**ОБ ИСТОРИИ ВОЗНИКНОВЕНИЯ И ИСПОЛЬЗОВАНИЯ ТЕСТИРОВАНИЯ КАК ДИАГНОСТИЧЕСКОГО МЕТОДА**

*Шатрова Наталья Владимировна учитель обществознания*

*МБОУ «Отрадненская средняя общеобразовательная школа №2»*

Сфера школьного образования – одна из важнейших областей жизни общества.

И, вполне естественно то, что образовательный процесс должен являть строго организованную систему, отвечающую современным тенденциям развития и изменения социума.

Познавательная деятельность учащихся должна быть контролируемой. Это необходимо как для определения степени соответствия приобретенных знаний и умений поставленной учебной цели, так и для управления познавательной деятельностью.

В настоящее время система педагогического контроля приобретает серьезные изменения. Введение ФГОС и использование Единого Государственного Экзамена придает проблеме надежной оценке знания большую значимость. Ведь практика последних лет часто дает неоднозначные результаты. Случается, что отметка, полученная учащимися на ЕГЭ, значительно отличается от той, которая год за годом выставлялась по данной дисциплине в школе, причем обычно не в лучшую сторону. Результат же Единого Государственного Экзамена не должен быть непредсказуемым, он должен являться естественным продолжением текущей оценки знаний учащегося.

Проблема объективности и субъективности оценивания знаний учащихся является весьма актуальной до сегодняшних дней.

По прежнему имеет место в учебном процессе переоценивание и недооценивание уровня подготовки учащихся, субъективность при анализе уровня интеллектуального развития. Зачастую предметом оценивания становятся не усвоенные знания и умения, а особенности учащихся (физические, психологические).

Российская пятибалльная педагогическая система оценивания, к сожалению, имеет свои несовершенства. Она фактически сводится к трех уровневой. 4 (хорошо) – средний уровень, 5 (отлично) – выше среднего, 3 (удовлетворительно) – ниже среднего. Практически не используется оценка «1», а оценка «2», в особенности, как итоговая, очень нежелательна.

Использованием Единого Государственного Экзамена, современная отечественная педагогика, конечно, хочет добиться повышения объективности педагогического контроля. И одним из перспективных методов этого направления является тестирование.

Исследование значения слова «тест» отсылает нас к латинскому языку. Testa (лат.) – ваза из глины. Этим словом называли небольшие емкости из обожженной глины, которые использовались алхимиками для проведения опытов.

«Словарь-справочник по психологической диагностике» сообщает нам о смысловом значении термина «тест», характерном для русского языка XIX века. В тот период этот термин использовался в двух значениях. Во-первых, как испытательная присяга, которую каждый вступающий в общественную должность должен приносить, дабы доказать, что служит вере и отечеству. Во-вторых, как плоский плавильный сосуд для выделения олова из золота и серебра. Большинство же современных ученых рассматривают происхождение этого слова от английского (test – проба, испытание, исследование).

Сущность теста, как метода исследования тех или иных вопросов достаточно неоднозначна. Но при всем существующем многообразии его определений, необходимо его рассмотрение в трех аспектах: во-первых, как инструмент теоретического исследования; во-вторых, как метод эмпирического исследования; в-третьих, как средство получения качественной оценки.

Однако, достаточно часто понятие «тест» отождествляется с понятием «тестовое задание». А последнее часто употребляется в смысле «вопрос». Тест – это инструмент теоретического исследования, представляющее собой стандартизированное испытание, в основе которого лежит набор заданий, которые позволяют оценить исследуемые качества на основе использования статистических методов.

В качестве структурных компонентов теста выделяют: во-первых, тестовое задание – задачу, предназначенную для выполнения и имеющую тестовую форму; во-вторых – эталон, как образец правильного решения, который служит для сравнения достигнутого уровня с планируемым. При этом необходимо учитывать, что тест предполагает определенные критерии качества. Наиболее важными из них являются валидность и надежность.

Валидность (англ.) – действительный, обоснованный. Этот термин является трудно переводимым на русский язык и зачастую трактуется очень широко. В российских исследованиях он соотноситься с точностью измерений и используется как понятие, выражающее «надежность» теста, адекватность, действенность. Отталкиваясь от этого, тест считается валидным, если он измеряет то, для чего он предназначен. Однако, рассматривая понятие валидность, необходимо обратить вниманиееще на один существенный момент – относительный характер валидности. А.Н. Майоров в своей работе «Тесты школьных достижений: конструирование, проведение, использование» пишет: «В общем виде понятие валидности раскрывает, насколько полученные результаты тестирования соответствуют объективной реальности, а поскольку отсутствует способ точного определения объективной реальности, то валидность – понятие относительное, так как тесты, как и любой измерительный инструмент, никогда не являются абсолютно - эквивалентными реальному положению вещей. Соответственно, валидность устанавливается относительно характеристики, признака, величины, принимаемой (а не истинно являющейся), за объективную.

Современная методология демонстрирует нам несколько видов валидности.

1. Очевидная валидность.

Предполагает внешнее представление о тесте со стороны испытуемого, его понимание, что этот тест должен измерить. Чем более тест выглядит измеряющим то, для измерения чего он создан, тем выше, его очевидная валидность. Хотя, не всякому тесту присущ этот вид валидности. Это связано с целями исследования. Но, большинство тестов не призвано скрывать свои реальные цели, поэтому, очевидная валидность выступает либо как нейтральное качество, либо как позитивное.

2. Содержательная валидность.

Характеризуется степенью репрезентативности содержания заданий теста по отношению к измеряемой характеристике. Н.Н.Нохрина отмечает: «Она определяется на основе экспертных методов. Наиболее распространение нашли такие способы экспертной оценки, как метод коллективной оценки, метод средневзвешенной оценки, метод ранжирования и метод парного сравнения».

3. Функциональная валидность связана с установлением определенных уровней сложности. Она характеризуется степенью соответствия содержания заданий теста по отношению к выявляемому уровню усвоения опыта деятельности. Тестовые задания предполагают решения, путем использования способов деятельности определенного уровня сложности (способы деятельности более низкого уровня здесь не приемлемы).

4. Прогностическая валидность связана с точностью и обоснованностью суждений о диагностируемой переменной. Для этого изучают корреляции между данными, полученными в ходе исследования и полученными спустя какое-то время.

5. Критерии валидности.

Предполагает выбор критерия оценки результатов работы.

В настоящее время каждому работнику системы образования, для успешной работы с педагогическими заданиями в тестовой форме, необходимо иметь представление о критериях качества теста, его структуре и основах тестологии. Причем, это полезно не только для практикующих подобную форму контроля, но и для занимающихся разработкой тестовых заданий. История развития тестирования показывает, что тесты уровня обучимости являются важным способом оценки педагогической деятельности, с помощью которого результаты обучения могут быть достаточно объективно измерены, обработаны и использованы в педагогической практике.

Предыстория тестирования уводит нас вглубь веков. Имеются сведения, что уже с III тысячелетия д. н.э. в станах Древнего Востока

(Египет, Вавилон, Индия, Китай), использовались системы испытаний интеллектуального характера, предназначенные для отбора претендентов на высокие посты и правительственные должности.

В трудах античных мыслителей, таких как Гераклит, Аристотель, Гиппократ, Гален, Теофраст приведены диагностические показатели особенностей личности. Так Гераклит (IV-V в. д. н.э.) впервые указал на разницу между душой взрослого и ребенка, соотносятся познавательные способности со степенью влажности души.

Афинянин Платон (V в. д. н.э.), разрабатывая теорию идеального государства, размышляет о социальной принадлежности граждан, соотнося социальный и психологический статус человека. Он ставит вопрос соотношения педагогического метода с психологическим статусом и особенностями ребенка и говорит о занятиях, которые помогают распознавать интеллектуальные и психологические особенности человека. В «Государстве» он пишет: «…надо искать людей, которые всех доблестнее стоят на страже своих взглядов и считают, что для государства следует делать все, по их мнению, наилучшее. Даже в их детские годы, предлагая им занятия, надо наблюдать, в чем кто из них бывает особенно забывчив и поддается обману. Памятливых и не поддающихся обману надо отбирать, а кто не таков, тех отвергнуть».

Гиппократ (V в. до н.э.) выдвинул предположение, согласно которому , различие в поведении и здоровье людей связаны с определенной комбинаторикой основных составляющих человеческого тела. В его трактате «О природе человека» мы можем прочитать: «Тело человека содержит в себе кровь, флегму, желчь желтую и черную: они составляют природу тела, через них [человек] испытывает боль или здоров. Он совершенно здоров, когда они соразмерны между собой в отношении смеси, силы и качества и совершенно между собой перемешаны. Боль же [человек] испытывает, когда одно из них оказывается в большем или меньшем количестве или же отделится в теле и не будет смешанно со всеми».

Античная философия демонстрирует многочисленные попытки систематизации различий между людьми и пытается создать адекватные методики для выявления индивидуальных особенностей. Рассмотрение истории философии и науки демонстрирует множество попыток решения этих вопросов и в дальнейшем. Однако, все эти методы исследования нельзя назвать тестами в современном понимании этого слова. Первый этап становления тестирования.

Известно, что тестология появилась на стыке психологии, социологии и педагогики, и призвана заниматься вопросами разработки тестов для объективного контроля знаний, умений, навыков и представлений.

И первичное применение тестов в школьной практике часто связывают с именем Френсиса Гальтона, английского психолога и антрополога.

Гальтон первым применил статистический анализ в биологии и психологии. Он полагал, что между различными (физическими и психическими) свойствами индивидов имеются корреляции, создал статистические методы их определения. Его работа «Исследование человеческих способностей и их развитие» является исследованием человеческого интеллекта. Именно в ней мы встречаем описание основ психологического тестирования.

Как отмечает Т.М. Фоменко: «Важнейшим вкладом Ф.Гальтона в развитии теории тестов было определение трех основных принципов, которые используются и по настоящий день.

1) применение серии одинаковых испытаний к большому количеству испытуемых;

2) статистическая обработка результатов;

3) выделение этапов оценки».

Впервые опубликованные в 1883 году, результаты его исследований послужили основой для французского психолога Альфреда Бине, который вместе со своим помощником Теодором Симоном осуществил вначале XX в. настоящий переворот в измерении умственных способностей.

В 1904 году во Франции, в связи с введением всеобщего образования, министр образования занялся проблемой «отсталых» детей. Была создана специальная комиссия для изучения этого вопроса, в нее вошли Бине и Симон. Они предложили серию стандартных заданий повышающейся сложности для школьников. Они хотели выяснить, что такое «отсталый» ребенок и путь к этому видели в методе тестирования. Этот тест должны были

Этот тест должны были решить 75% детей со средними интеллектуальными способностями. Однако, данный тест имел и свои минусы. Он не позволял точно определить степень понимания и сообразительности учащихся. Именно поэтому, восприятие теста психологами и педагогами различных учебных заведений было неоднозначным. Но, бесспорно, заслуга данного теста в том, что он положил начало оценке интеллектуальных возможностей человека.

До конца своей жизни Бине работал над усовершенствованием своих тестов. В 1911 году его работу продолжил американец Льюис Терман. Основных нововведений по сравнению с тестами Бине было два: введение в качестве показателя по тесту коэффициента интеллектуальности (IQ) и введение статистической тестовой нормы. Его тесты (тесты Стенфорда-Бине) и сегодня используют для разного рода задач. Исследование Бине сформулировали понятие «умственного возраста» или характеристики интеллектуального развития индивида на основе его сравнения с уровнем интеллекта – других людей того же возраста.

Второй этап становления тестирования относят к 20-60 годам прошлого столетия и связывают с деятельностью американца В.А. Макколла, который разделил тесты на педагогические и психологические. Основной задачей педагогических тестов являлось измерение успешности учащихся в определенных дисциплинах за некоторый период обучения, а так же успешность применения определенных методов преподавания и организации.

Создателем первого педагогического теста считается американский психолог Эдуард Ли Торндайк. Его называют основоположником педагогических измерений. Первый педагогический тест был связан с решением арифметических задач. И надо отметить, что первоначально, именно в США педагогические тесты для проверки знаний, навыков и умений учащихся по отдельным предметам нашли широкое распространение.

Внедрение тестовой диагностики в России в 20-е годы XX века часто связывают с возникновением новой науки педологии. Педология направление в психологии и педагогике, возникшее под влиянием эволюционных идей. В педологии ребенок рассматривался комплексно, во всех своих проявлениях, в постоянном развитии и в различных, в том числе социальных, условиях, а целью ставилось помочь развитию всех его потенций.

Л.С. Выготский определял педологию как науку о целостном развитии ребенка.

Школьные психологические службы в то время, в силу понимания предмета педологии, имеют комплексный характер, принимая в себя черты медико-биологической, социологической и психологической службой одновременно.

В Москве после 1925 года при педагогическом отделе Института методов школьной работы, была создана тестовая комиссия для разработки стандартизированных тестов для советской школы. Для этой работы были задействованы известные российские ученые: М.С. Бернштейн, П.П. Блонский, А.П. Болтунов, С.Г. Геллерштейн, Г.И. Залкинд и др. Но, поскольку, педология претендовала на комплексный подход к изучению ребенка, то она занималась не только теорией, но и практикой. Она стала заниматься и проблематикой оценки результатов обучения и воспитания.

Начало 30-ых годов прошлого века становится временем массового внедрения тестов в народном образовании. К несчастью, их использование не всегда целесообразно и профессионально. Как замечает Н.Н. Нохрина в своей работе «Тест как общенаучный диагностический метод»: «Массовые тестовые обследования не подкреплялись серьезной проверкой качества инструментария, решение о переводе некоторых учащихся в заведение для умственно отсталых детей принимались на основе коротких тестов без учета других факторов, влияющих на результаты проверки. К тому же, применявшиеся в массовой педологической практике тесты умственной одаренности, в основном Бине-Берита, были слабо адаптированы к местным условиям и не вполне отвечали задачам, для решения которых они предназначались.

Однако, в конце 20-х годов в педологии возникает разрыв между теорией и практикой. Разработчики теории педологии, изучая причины отставания развития школьника, полагали необходимым комплексное изучение ребенка. Судить о прич

Однако, в конце 20-х годов в педологии возникает разрыв между теорией и практикой. Разработчики теории педологии, изучая причины отставания развития школьника, полагали необходимым комплексное изучение ребенка. Судить о причинах отставания можно проанализировав данные, полученные в результате длительного и целенаправленного наблюдения, анкетирования, тестирования, и антропометрического измерения. На практике все происходило иначе. Вместо комплексного программированного исследования массово внедрялись тесты, как оперативные и «простые» средства диагностики. Повсеместно в исследованиях стала царить бессистемность и, как следствие, ошибки в педологических диагнозах.

В 1936 году принимается постановление ЦК ВКП (б) «О педологических извращениях в системе Наркомпроса»\*. Последствия этого события очень серьезны. Педология характеризуется как лженаука, накладывается запрет на применение тестов и анкет. Исследователь Н.Н. Нохрина отмечает: «Метод тестирования был признан орудием дискриминации учащихся и «изгнан» из советской школы». В дальнейшем в течение многих лет тестирование официально не признавалось объективным методом и не находило практического применения, хотя сама идея педологии как «комплексной человедческой науки» была перспективна и актуальна.

Это постановление, конечно, обладало большой значимостью. Началось сворачивание разработок диагностических методик и отказы от систематической психо-педагогической диагностики. Таким образом, в отечественной науке применение тестов было приостановлено. Однако, в ряде западных стран тестирование не прекращает своего развития.

В 30-50-е годы особой популярностью пользуются прогностические тесты, они внедряются в систему образования. Особую актуальность в то время приобретают вопросы, связанные с разработкой тестов педагогами – практиками.

Третьим этапом в развитии тестирования в отечественной науке является период с начала 60-х годов и до конца 70-х годов. Развитие программирования, реформы по преподаванию ряда предметов, способствовали возобновлению тестологической деятельности, совершенствования тестирования. Ведутся разработки по внедрению компьютеров в эту сферу. Разрабатывается принцип разветвленного программирования, сущность которого заключается в следующем: если испытуемый предоставляет правильное решение, то дальше ему даются более трудные задания, и наоборот. При этом отмечается, что этот подход дает возможность учащимся выявить свои умственные способности. В этот же период, возобновляются исследования в области тестов психолого-педагогической направленности. Здесь, конечно, анализируется многолетний опыт, накопленный в зарубежных странах.

Четвертый этап развития тестирования в России приходится на начало 80-х годов. Его выделение, как самостоятельного этапа, связано с началом интенсивного использования компьютеров, как в процессе тестирования, так и при обработке полученных результатов. В ряде западных стран тестирование становится практически ведущей формой контроля. Англия, Нидерланды, Япония, Дания, Канада, Австралия, США занимаются организацией массового тестирования. В этих странах развивается теория и практика тестирования, создаются службы разработки тестов.

В настоящее время тестирование является одним из основных методов диагностического контроля ЕГЭ в России. Единый государственный экзамен - централизованно проводимый в Российской Федерации экзамен в школах и лицеях. Примечательно то, что являясь выпускным экзаменом из школы, он при этом становится и вступительным испытанием для абитуриентов ВУЗов.

Как известно, ЕГЭ предполагает применение однотипных заданий и единых методов оценки качества выполненной работы на всей территории страны. Единственной формой выпускных экзаменов в школе и основной формой вступительных экзаменов в ВУЗы ЕГЭ стал в 2009 году.

Создателем ЕГЭ принято считать Владимира Алексеевича Хлебникова - доктора технических наук, директора Федерального Центра тестирования при Министерстве образования и науки РФ. Он создал отечественную технологию ЕГЭ. Впервые эксперимент по введению ЕГЭ был проведен в 2001 году в республиках Чувашия, Марий Эл, Якутия, а так же в Ростовской и Самарских областях. В последующем году ЕГЭ был проведен в 16 регионах страны, в 2003 в 47 субъектах РФ, а в 2004 в 65.В 2008 его сдавали свыше миллиона учащихся во всех регионах.

Задания для Единого государственного экзамена разрабатываются ФИПИ (Федеральным институтом педагогических измерений). Для основной массы предметов задания подразделяются на тестовые, предполагающие один вариант ответа из четырех предложенных, задания, предполагающие краткий ответ, состоящий из одного или нескольких слов и несколько заданий с развернутым ответом.

Дискуссии вокруг ЕГЭ не утихали долгие годы. Особое обострение полемики приходилось на 2008 год. Именно в этот период к ЕГЭ подключаются все регионы. Негативное восприятие ЕГЭ во многом было связано с тем, что учащиеся стали получать оценки ниже тех, которые им ставили в период обучения в школе.

На мой взгляд, система ЕГЭ несомненно является перспективной для нашей страны. Учащиеся достаточно быстро к ней привыкли. Однако, мы продолжаем приобретать опыт, а он имеет положительные и отрицательные аспекты. Элиминация ошибочного происходит ежегодно. Но для того, чтобы сделать процедуру и результаты экзамена качественными, необходима работа с основами тестирования как диагностического метода применительно к каждой школьной дисциплине, как к области обладающей своим особым предметным содержанием и методиками преподавания.

Согласно Е.Е. Шестернинову (статья "Педагогические тесты: надежный способ оценки и повышения уровня знаний") сложность оценивания в условиях современности усугубляется некоторыми глобальными тенденциями, свойственными мировой педагогической практике.

1. В XX веке произошел переход от "классической культуры" к глобальной, толерантной, мультикультурной. Сложно говорить об относительно стройной, иерархической системе знаний о мире. К сожалению, современный человек значительную часть знаний получает не в системе рационального образования, не столько через систему образования, сколько через средства массовой коммуникации.

2. В сфере образования - особенно во второй половине XX века отчетливо прочвляется тенденция технократизма. Произошло перемещение центра тяжести в сторону функциональных характеристик личности, в сторону профессионализма, дефицит которого уже невозможно компенсировать простым усердием. Интенсивно развивается информатизация и компьютеризация общества. Тесты же традиционно связываются с техническим подходом к контролю знаний.

3. Изменяется положение педагога в системе образования. Меняется его роль. Если раньше он в основном выполнял функцию накопителя и распространителя научной информации, то теперь его центральная задача - управлять познавательной активностью учащихся и контролировать ее результаты.

По замечанию Шестернинова, постепенно учитель начинает осознавать явный приоритет письменной формы оценки знаний перед устной. Причины: "Повышается объективность и надежность оценивания, уменьшается стресс у учащихся, снижается субъективизм учителя, связанный с симпатиями и антипатиями, возможностью негативной отметкой наказать ученика за различные проступки. Суммирование результатов текущего (рубежного) и итогового контроля стимулирует ученика к систематической работе в течение года, а использование индивидуального рейтинга как основного показателя успехов в обучении порождает состязательность в учебе, положительно влияет на мотивацию учащихся, сводит до минимума случайность в оценивании и т.п."

Важная цель ЕГЭ - попытка улучшения качества образования за счет объективизации контроля и повышение мотивации на успешное его прохождение. Естественно, что дети должны быть хорошо подготовлены, тогда они сдадут экзамены успешно, и их результаты можно будет сравнивать.

Готовить к экзамену должны, конечно, учителя. И это большой труд, учитывая еще и то, что педагогам приходится постоянно перестраиваться, организовывая свою работу инновационным образом. Возникает очень много задач, ведь ученик должен уметь решать не только тесты, в нем должно развиваться творческое начало, умение находить нестандартные решения. Поэтому, конечно, несмотря на очевидные плюсы тестирования как диагностического метода, тестовые методики в формате ЕГЭ, именно на учителей возлагают огромную ответственность за результат и качество школьного образования.

Список литературы

Абасов З.А. От оценки учителя - к самооценке ученика //Директор школы 1999 №2 с.63-70

Декарт Р. Рассуждения о методе, чтобы верно направлять свой разум и отыскивать истину в науках// Философское наследие т.1 Москва 1989

Единый государственный экзамен: проблемы и перспективы. Выступления и дискуссии круглого стола// Университетское управление. 2001. №4(19). С. 51-64

Кириллова Светлана. Из истории тестирования //Первое сентября. 2003. №28

Лебедев А.В. Фрагменты ранних греческих философов//Памятники философской мысли ч.1. Москва 1989

Майоров А.Н. Тесты школьных достижений: конструирование, проведение, использование// СПб 1997

Платон// Философское наследие т.3 Москва 1993

Тимаева В.С. Тестирование как метод контроля качества усвоения учебного материала учащимися// Первое сентября. 2010 №9

Струкчинская Е.Н., Л.С. Выготский о педологии и смежных с нею науках// Рукописный журнал Общества ревнителей русской философии 2002 №5

Шестернинов Е.Е. Педагогические тесты: надежный способ оценки и повышения уровня знаний// Директор школы. 2003. №9 с.6-13

Интернет ресурсы:

Нохрина Н.Н. Тест как общенаучный диагностический метод// <http://www.isras.ru/files/File/Socis/2005-1/nohrinann.pdf>