

# Педагогическое тестирование в России конца XX — начала XXI века

Вадим Сергеевич Аванесов,  
доктор педагогических наук,  
testolog@mail.ru

Для педагогического тестирования XXI век в России оказался трудным. Трудно было и до этого, в XX веке, в СССР этот научный метод долго не признавали, нередко осуждали. Причина одна: при правильном использовании тесты показывали реальные результаты, а не те, которые хотели видеть. В 1936 году Постановлением Совета Народных Комиссаров СССР тесты были запрещены под предлогом их буржуазности.

Но сейчас политическая ситуация в стране и мире иная. Конец XX века был отмечен заметным развитием тестирования в России. Но в 2016 году руководство Рособнадзора решило удалить задания в тестовой форме из т.н. Единого государственного экзамена (ЕГЭ)<sup>1</sup>. Что будут ещё удалять из ЕГЭ ради его спасения и что станет с педагогическим тестированием в России? Вот вопросы, на которые пока нет ответа.

В данной статье рассматриваются проблемная ситуация, вопросы истории, теории и практики педагогического тестирования в России. Приводятся основные результаты исследований, уточнённое название теории, даны определения понятий.

*Ключевые слова:* теория педагогических тестов и практика педагогического тестирования, тест, тестовое задание, задание в тестовой форме

*Какая-то подделка под науку ...  
присвоила себе название знания и ставила  
возврату последнего почти  
неодолимые преграды.  
Ж.-Ж. Руссо*

## Проблемная ситуация

Педагогическое тестирование — главный метод педагогической диагностики и педагогических измерений. Поэтому не случаен интерес читателей журнала к этому методу. В конце XX — начале XXI веков в системе советского образования был краткий период подъёма интереса к педагогическому тестированию. Появились публикации по данной проблеме, диссертации, началось повышение квалификации вузовских и школьных педагогов. Повышение педагогической квалификации проводилось по решению бывшего Госкомитета СССР по высшему образованию, специально созданной кафедрой, по учебной программе педагогического тестирования, написанной автором этой статьи.

<sup>1</sup> Сергей Крайцов. ЕГЭ как стимул к развитию. <http://www.vedomosti.ru/opinion/articles/2016/06/17/645705-ege-kak-stimul-razvitiyu#/galleries/140737488877577/normal/1>.

Но вот к самому концу XX века в России возникла идея создания нового экзамена, внешне похожего на педагогическое тестирование, но, по сути, не являющегося таковым. Такая идея возникла в недрах бывшего тогда Министерства образования РФ, под предлогом борьбы с коррупцией и для создания условий равного доступа в вузы.

В начале XXI века Правительство РФ начало внедрять этот экзамен в жизнь. В нём использовалось, в числе других методов, подобие тестовой формы. Работа велась, некачественно, без научного проекта или программы, без опоры на теорию тестов и на подготовленных к тому времени в стране нескольких тысяч начинающих тестологов.

Деньги на проведение в России ЕГЭ вначале дал займы Мировой банк. И этим, вероятно, может объясняться многое. Большие деньги на проведение ЕГЭ потрачены и из госбюджета. Естественно задать вопрос — каков результат? Самый общий, для начала, ответ таков: плюсов очень мало (если они вообще есть и доказуемы), а минусов очень много.

Власти рассматривали ЕГЭ как главную надежду в борьбе с коррупцией и как средство выравнивания доступа к высшему образованию, в период усиления сверхизбыточного неравенства в доходах граждан. Но коррупцию ЕГЭ не победил, а наоборот, усилил. А идею равного доступа к высшему образованию в условиях сверхизбыточного неравенства населения по доходам можно было считать верной лишь частичной, и лишь для тех выпускников, которые могли быть хорошо подготовлены. Но могли ли все они учиться в принимающем их вузе, при очень низкой стипендии — этот экономический вопрос инициаторы ЕГЭ не исследовали. В этой связи идея равного доступа становится

отчасти из разряда социально привлекательных мифов. У автора этой работы была даже такая публикация — Мифология ЕГЭ<sup>2</sup>.

Свершившуюся в России подмену научного метода тестирования новой, неизвестной миру формой т.н. единого государственного экзамена мало кто заметил. Поскольку в этом экзамене декларировалось применение тестов. Из кабинетов бывшего тогда министерства образования РФ было запущено в оборот название «тесты ЕГЭ», которые тогда собирались применять в этом экзамене<sup>3</sup>.

Вероятно, из-за свалившихся неожиданно денег органы управления образованием решили делать ЕГЭ сами, по-своему разумению, преимущественно с опорой на педагогов-предметников. Всё в этом деле им казалось понятным. Но их потенциала для разработки качественных тестов было недостаточно. В таком деле нужно было задействовать, в центре и на местах, потенциал и тестологический. К тому времени он уже был создан в стране. Но это их не интересовало. Стремление чиновников того времени опереться при создании ЕГЭ на собственные силы, без какого-либо проекта или научной программы, и без привлечения к работе тестологов, оказалось для страны ошибкой.

Не случайно качество созданных тогда, как называл официоз, «тестовых материалов» было низким, не улучшаемым из-за ошибочных представлений, положенных в разработку ЕГЭ. А потому там сочли за лучшее засекретить информацию, касающуюся статистических распределений, получаемых в результате массового применения отмеченных материалов. Однако весь брак скрыть не удалось. Кое-какие данные несколько лет назад удалось разыскать в Интернет и проанализировать их научными методами.

<sup>2</sup> Аванесов В.С. Мифология ЕГЭ. Газета «Первое сентября» — 10.07.2001. — № 50.

<sup>3</sup> «...Организаторы ЕГЭ берут на себя часть ответственности за эту сумятицу, поскольку именно мы запустили в оборот слово «тест» применительно к ЕГЭ. Болотов В.А. ЕГЭ: промежуточные итоги. Вопросы образования. — 2004. — № 2. — С. 155–167.

Опубликованные результаты и выходы оказались для создателей ЕГЭ неутешительными<sup>4</sup>.

Разъяснения автора этой статьи о том, что такого рода метод не может быть *единым* по метрическим требованиям, *государственным* по этическим и политическим требованиям и в форме *экзамена*, по технологическим требованиям, отвергались, как говорили раньше, с порога.

Такая позиция помогала органам управления образованием упорно вести свою ошибочную линию, устраняя лишь некоторые текущие перегибы, но порождая при этом новые ошибки. Например, отменили обязательность сдачи всех выпускных экзаменов, которые раньше требовалось сдавать по Закону об аттестации выпускников школ. Обязательными и достаточными для итоговой государственной аттестации выпускников школ оставили только два предмета — русский язык и математику.

Таким образом тихо расстались с ошибочной идеей *единых* экзаменов по всем предметам, которые давали фактическую основу для принятия решения об аттестации выпускников школ по итогам сдачи всех государственных экзаменов. Эту основу ликвидировали. Детям и их родителям это новшество показалось большим облегчением, а потому оно некоторым понравилось. И не все понимали, что вслед за этим многие дети перестанут изучать те предметы, по которым, в данных условиях, не надо сдавать экзамены.

Явно наблюдаемое разрушение среднего школьного образования

пошло именно с этой ошибки ведомственной образовательной политики. Все остальные ЕГЭ, кроме двух названных, стали только приёмными экзаменами в вузы. Букву «Е» из аббревиатуры ЕГЭ можно было бы после этого удалить, но тогда крах единого экзамена стал бы очевидным всем и сразу. Пострадало бы и благозвучие. А потому новым разделённым (не единым) экзаменам было решено оставить прежнее название «ЕГЭ».

В нормально функционирующих образовательных технологиях вообще никакие экзамены не нужны. Метод динамического (текущего) автоматизированного оценивания снимает множество проблем современного образования. В этом смысле курс на проведение единого государственного экзамена представляется анахронизмом XIX века.

Было немало и других ошибок, о которых автор этой статьи неоднократно писал. Но бюрократический молот теперь уже не единого государственного экзамена продолжал упорно долбить по российской школе, ничего и никого не слушая.

## Главное событие 2016 года

Совсем недавно было озвучено мнение руководства Минобрнауки и Рособнадзора об удалении «тестов» из ЕГЭ. Хотя причины вновь возникшего отрицательного отношения образовательных властей к тестам в России толком не объясняются, похоже, что массовая критика

<sup>4</sup> Аванесов В.С. Единый государственный экзамен в фокусе научного исследования. // Педагогические измерения. — 2000. — № 1.

Аванесов В.С. Являются ли КИМы ЕГЭ методом педагогических измерений? // Педагогические измерения. — 2009. — № 1. — С. 3–26. 2 ред. С табл. и графиками <http://viperson.ru/wind.php?ID=563869&soch=1>.

Аванесов В.С. Проблема демаркации педагогических измерений. // Педагогические измерения. — 2009. — № 3. — С. 3–37.

Аванесов В.С. Ошибочные цели — плачевные результаты // Педагогические измерения. — 2010. — № 4.

Avanesov V.S. Consequences of the EGE in RUSSIA. KEDI Journal of Educational Policy. Special Issue: pp. 89–100. Guest Editor — Prof. Такая позиция Hoi Suen. Penn. State Univ. V. 3, 1, 2006. <http://testolog.narod.ru> и др.

применения тестоподобных форм в ЕГЭ стала, наконец, действительно угрожать проведению этого государственного экзамена.

Хотелось бы заметить, что собираясь сейчас удалить «тесты» из ЕГЭ, руководители органов управления образованием хотят удалить то, чего в ЕГЭ нет, и никогда там не было. Там были лишь тестоподобные задания, которые не отвечали критериям качества тестовых заданий, ни по форме, ни по содержанию, ни по статистическим требованиям. Примеры ошибочных заданий в т.н. «контрольно-измерительных материалах» (КИМах) ЕГЭ приводились в наших статьях<sup>5</sup>.

Гегель говорил, что в истории всё повторяется как бы дважды. А К. Маркс добавил к словам Гегеля: первый раз это случается в виде трагедии, второй раз — в виде фарса.

## Проблема

Между тем, мы уже живём в XXI веке — веке интеллектуальной конкуренции, образовательных технологий и научно-технического прогресса. Именно эти факторы определяют основную стратегию развития многих стран, потому что ни одна страна не сможет быть более развитой, чем развита её система образования. А современное образование не может нормально развиваться без тестовых технологий, как это происходит во многих странах мирах.

Улучшению образования мешают не только возросшие ошибки ведомственной образовательной политики, но также и накопившиеся трудности функционирования самой системы образования. Многолетняя практика ослабленного внимания к вопросам научной организации учебного процесса привели к недостаточной материально-технической оснащённости учебного про-

цесса, нехватке учебной техники и технологий. Кроме того, это привело к снижению стимулов повышения педагогического мастерства учителей к разработке новых учебных материалов, к выпуску огромного числа негодных учебников, к раздробленности учебных дисциплин и однообразному стилю работы школ, характерному для канонизированной ещё со средних веков классно-урочной формы обучения.

В исторически сложившейся системе преподавания много времени занимает словесная, объяснительно-иллюстративная работа учителей и недостаточно внимания уделяется научно организованной самостоятельной работе учащихся, применению новых образовательных технологий, где самообразование и самоконтроль играют ведущую роль. Отставание в новой организации учебного процесса становилось устойчивым в условиях усиливающегося дефицита финансовых ресурсов.

Образовательный бюджет сильно перекосялся в сторону финансирования непедагогического ЕГЭ, допуск на который проводится с элементами полицейских проверок содержимого карманов экзаменуемых школьников, применения других непедагогических форм слежения за испытуемыми. Не случайно в абсолютном большинстве стран, особенно демократических, с развитой системой образования, такого рода непедагогичных государственных экзаменов нет.

В постсоветской России госэкзамены остались со времён правления И. Сталина. Они ассоциировалось тогда с идеей усиления, таким образом, роли советского государства в системе бесплатного, для граждан, образования. С той поры госэкзамены стали формой психологического воздействия на нерадивых студентов, а затем и школьников, превратились в дополнительное средство

<sup>5</sup> Аванесов В.С. Единый государственный экзамен в фокусе научного исследования. // Педагогические измерения. — 2000. — № 1.

запугивания учащихся и принуждения их к учёбе. Но для нормальной системы образования метод принуждения — самый плохой, разлагающий всю систему, приводящий её в кризис.

В бедственном положении оказались малокомплектные школы, часть которых либо уже закрыта, либо хотят закрыть. Их лучше пере-профилировать в школы дистанционного обучения. Но для этого нужны качественные учебные материалы, которых в России очень не хватает.

К сожалению, для таких школ не были своевременно заготовлены квантованные учебные тексты с заданиями в тестовой форме, а также новые обучающие и контролирующие программы, необходимые для организации в школах электронного изучения предметов. В этих школах уже давно можно было бы наладить современные формы централизованного дистанционного обучения и контроля, если бы деньги из образовательного бюджета тратились именно на нужды образования, а не на усиление бюрократического контроля. Как справедливо отмечается в интересной статье Г.В. Безродной, Д.А. Севостьянова и М.В. Чельцова, преувеличение роли контроля в системе образования вызывает инверсию целей, а это сказывается разрушительно на самой этой системе<sup>6</sup>.

Уникальной помехой для становления тестовой культуры в системе российского образования является стремление государственных органов управления самим заниматься разработкой и применением контролирующих форм, попытки создавать методы, внешне похожие на тестовые, но далёкие от них по

сути, качеству, теоретическим основам и по эффективности. Примеры таковых государственных самоделок в сфере образования — т.н. «тесты» и «КИМы» ЕГЭ, Федеральный интернет-экзамен системы профессионального обучения (ФЭПО), «показатели» мониторинга вузов<sup>7</sup> и др.

Такая странная, настойчивая и некачественная методическая деятельность министерства образования и науки, а также Росособнадзора, при наличии Российской Академии образования, вызывает много вопросов и порождает много проблем с состоянием качества образования. В стране нет организованной силы, которая могла бы остановить министерскую методическую самодеятельность.

Было сделано ещё несколько ошибок. Требования к аттестации выпускников школ посредством КИМов ЕГЭ сделали сильно заниженными<sup>8</sup>. Было время, когда аттестовали всех, кто только приходил и ставил свою фамилию в пустых, незаполненных бланках ответов на задания государственных экзаменов. И после этого пришедших и просто поставивших в бланках свои фамилии стали считать аттестованными за школьный курс.

Так возникли «зияющие», по аналогии с лексикой А. Зиновьева, «успехи» российского образования: по результатам такого вот государственного экзамена. Школу «успешно» заканчивали теперь не менее 95 процентов выпускников. С поступающими в вузы абитуриентами поступили иначе. Для них была приготовлена третья часть ЕГЭ, с трудными и сверхтрудными заданиями. После чего резко возросли количество и стоимость репетиторских

<sup>6</sup> *Безродная Г.В., Севостьянов Д.А., Чельцов М.В.* Педагогический контроль: инверсия целей и средств. <http://www.odn2.ru/index.php/biblioteka/34-uchebnyj-protsess/289-pedagogicheskij-kontrol-inversiya-tselej-i-sredstv>.

<sup>7</sup> *Аванесов В.С.* Неэффективность российских вузов: реальная и мнимая. <http://viperson.ru/wind.php?ID=657524&soch=1>.

<sup>8</sup> *Иванов А.* Пути выхода из катастрофы ЕГЭ. <http://gazeta-licey.ru/educ/school/28916-svoboda-ne-uchit-ne-uchitsya>. 29 янв. 2015 г.

услуг по таким заданиям. Бизнес и здесь получил свою подпитку. О доходах с продажи книжной продукции можно было косвенно судить по тому, что на полках магазинов главной стала литература не по изучению школьных предметов, а по сдаче ЕГЭ. Те, кто думает, что это одно и то же, ошибаются. Объём материалов для изучения много больше материалов для сдачи. Здесь всегда имеет место редукция.

По мере такого процесса внедрения государственных экзаменов общая деградация качества массового образования приобрела вполне зримые очертания. В ответ на массовую критику ЕГЭ обещали реформировать. Как и положено, обещанного ждали три года. А на четвёртый, вместо реформы, получили запрет применения тестовой формы в ЕГЭ. Теперь вместо тестовых форм вновь предлагается использовать традиционные экзамены. Относительно экзаменов было много раз сказано, что это нетехнологичная форма оценки знаний, имеющая древнюю историю, превышающую четыре тысячи лет<sup>9</sup>. В наше время массовое применение формы государственных экзаменов экономически не эффективно. Поэтому пришла пора менять функции, формы и методы работы органов управления образованием.

Со времён перестройки жизнь в стране постепенно менялась, возникало всё больше открытых зон, но в сфере образования начался, похоже, противоположный процесс: возникли новые запреты и ограничения, возросла бумажная карусель отчётности, введено очень вредное для развития образования засекречивание статистических распределений результатов ЕГЭ. Трудно поверить, что эта линия российского государства. Но тогда кому и зачем всё это нужно? Всё это похоже на неизбеж-

ные перегибы ведомственной образовательной политики, о которых своевременно предупреждал покойный академик Российской Академии образования Эдуард Днепров.

Итак, проблема такова: есть общемировой процесс развития теории тестов и методик педагогического тестирования. Но в России управленцы создают и укрепляют свой собственный непедagogический метод тотального, принудительного оценивания учащихся и студентов. Такая ситуация, в силу отрицательных следствий, стала ухудшать качество образования. Особенно мешает подмена тестов т.н. КИМами, экзаменами, субъективным оцениванием и т.п. И всё это вместо тестов. Впрочем, терминологические подмены такого рода — явление исторически не новое<sup>10</sup>.

Стало очевидным, что без ограничения методической самостоятельности органов управления, изменения их функционала, движение вперёд стало невозможным. Отсюда следует актуальность проблемы организации педагогического тестирования, соответствующего требованиям XXI века.

## Язык ЕГЭ

В логической литературе термином называют слово или группу слов, выражающих понятие о предмете. Смысл термина — выражаемое им понятие. Одно из требований логики и методологии любой науки — определенность и однозначность её терминологии. Сейчас это требование обычно не выполняется. Такое состояние терминологии есть результат не только того, что исследователи не договорились относительно словоупотребления. Хотя и это тоже может иметь место, поскольку все они говорят, как при строитель-

<sup>9</sup> Аванесов В.С. Из глубины веков. <http://testolog.narod.ru>.

<sup>10</sup> Какая то подделка под науку ... присвоила себе название знания и ставила возврату последнего почти неодолимые преграды. Руссо Ж.-Ж. Сочинения /Пер с фр. Н.И. Кареев и др. Сост. и ред. Т.Г. Тетенькина. Калининград: Янтарный сказ, 2001. — 416 с.

стве Вавилонской башни, на разных научных языках. Казалось бы, говорят на одном, русском языке, но тексты малопонятны, в них много терминов, некорректно переведённых с английского языка на русский.

А.А. Зиновьев писал о замутнении смысла терминов. С похожим явлением мы многократно встречаемся при чтении предметно-методических отчётов по проведению ЕГЭ. Там пишут о надёжности КИМов ЕГЭ, о распределении «тестовых» баллов в КИМах ЕГЭ. И это при том, что там нет ни тестов, ни надёжности результатов, как об этом принято писать в настоящем тестовом отчёте. Вся эта лексика вытекает из отождествления КИМов ЕГЭ с тестами.

Результатом является лексическая смесь, препятствующая пониманию сути ЕГЭ, КИМов и, конечно, тестов. Оправданием для применения отмеченной лексической смеси является мнение сторонников ЕГЭ о том, что КИМы являются методом педагогических измерений. Но надо повториться ещё раз, что это совсем не так<sup>11</sup>.

Единые государственные экзамены декларировались как методы, которые по замыслу, должны были обладать следующими характеристиками: соответствие цели, справедливость, честность, доверие общественности к результатам, действенность и экономическая эффективность, прозрачность контрольно-оценочных процедур, положительное влияние результатов контроля на образовательную практику<sup>12</sup>. Но эти характеристики оказались недостижимыми в КИМах ЕГЭ.

Название «КИМы ЕГЭ» очень широкое, под него можно подвести любой «материал» — задачи, вопросы, упражнения и т.п. Создание КИМов можно охарактеризовать как деятельность, лишь *похожую* на

измерение, и судя по открытым публикациям, не соответствующую научному процессу. В то время как создание теста — это разработка научно обоснованного метода педагогического измерения.

Из США к нам пришла традиция называть «тестами» любые задания с выбором одного правильного ответа, а то и просто любые наборы вопросов. В российских образовательных учреждениях эта традиция имеет, похоже, ещё большую распространённость, хотя уже многие знают, что тест — это не отдельное задание, а система заданий возрастающей трудности, представленных в тестовых формах. Знают, но ничего поделать не могут. Слишком сильной оказалась привычка называть тестами любой набор задач, вопросов, или упражнений, подготовленный для контроля знаний студентов или школьников.

Обыденный смысл слова *test* охватывает любой метод проверки, научный, так и не научный. Научное же мышление рассматривает тест не как элементарный перевод общего смысла иностранного слова, а как обоснованный метод, как результат, полученный данным методом.

## Название науки о тестах

В науке нет вечных определений понятий, истинных на все времена. Не случайно научные понятия классики философии и логики называют «текучими», с меняющимися определениями, испытывающими постоянное давление времени, изменчивой практики и научного прогресса.

Отсюда проистекает и постоянная необходимость критики понятийного аппарата любой, можно

<sup>11</sup> Аванесов В.С. Являются ли КИМы ЕГЭ методом педагогических измерений?

Педагогические измерения — 2009. — № 1.

<sup>12</sup> Болотов В.А., Ефремова Н.Ф. Система оценки качества российского образования. <http://www.den-za-dnem.ru/page.php?article=150>.

сказать, науки. Время заставляет нас периодически обращаться к прежним теоретическим разработкам по определению понятий и по концепциям тестовой теории, с тем, чтобы внести в них уточнения и изменения, адекватные современным научным требованиям.

Для нормального становления науки важную роль играет правильное название её самой, потому что, как писал Э. Сепир, вещи, качество, события вообще воспринимаются так, как они называются<sup>13</sup>. В XX веке в числе приемлемых использовались такие названия науки о тестах как «тестология»<sup>14</sup>, «теория педагогических измерений»<sup>15</sup>, теория педагогического контроля<sup>16</sup> и другие.

Это явление не уникальное для новых наук. Например, на стыке вычислительной техники и лингвистики родилась наука, которая несколько раз меняла своё название: сначала она называлась математической лингвистикой, потом структурной лингвистикой и вычислительной лингвистикой. Наконец за ней прочно укрепилось её современное название — компьютерная лингвистика<sup>17</sup>.

Проблему осложняет положение теории педагогического тестирова-

ния. В наше время существенный признак тестирования — это педагогически обоснованный отбор содержания теста, точный подбор формы заданий, система основных педагогических понятий этой теории<sup>18</sup>, получение данных об испытуемых и заданиях на шкале натуральных логарифмов.

Идея необходимости перехода от тестирования к системе педагогических измерений и обратно к научно обоснованному тестированию возникла как результат сравнительного исследования вопросов методологии педагогических измерений<sup>19</sup>, практики государственного тестирования в России<sup>20</sup>, введения российского Единого государственного экзамена (ЕГЭ)<sup>21</sup>.

После нескольких лет поисков различных названий для исследуемой теории автор этой статьи считает наиболее подходящим название *теория педагогических тестов*. Это частная теория относительно более широкой теории педагогических измерений, куда кроме последней можно отнести вопросы теории шкалирования и рейтинга.

За рубежом в середине XX века выходили книги под названием «Теория педагогических и психологических тестов». Авторы этих книг

<sup>13</sup> Сепир Э. Избранные труды по языкознанию и культурологии. — М.: Прогресс. 1993. — С. 227.

<sup>14</sup> Аванесов В.С. Понятийный аппарат педагогической тестологии // Педагогическая диагностика. — 2002. — № 2. — С. 33–37.

<sup>15</sup> Аванесов В.С. Основы педагогической теории измерений // Педагогические измерения. — 2004. — № 1. — С. 15–21.

<sup>16</sup> Аванесов В.С. Основы научной организации педагогического контроля в высшей школе. Пособие для слушателей учебного центра. — М.: МИСиС, Исследовательский центр. — 1989.

<sup>17</sup> Боярский К.К. Введение в компьютерную лингвистику. Учебное пособие. — СПб: НИУ ИТМО. — 2013. — 72 с.

<sup>18</sup> Аванесов В.С. Научные проблемы тестового контроля знаний. — М.: Иссл. Центр, 1994. — 135 с.

<sup>19</sup> Аванесов В.С. Вопросы методологии педагогических измерений // Педагогические измерения. — 2005. — № 1. — С. 3–27. <http://testolog.narod.ru/Theory34.html>

<sup>20</sup> Аванесов В.С. Централизованное тестирование лучше Единого Государственного Экзамена // «Развитие тестовых технологий в России. Тезисы докладов Всероссийской научно-методической конференции / Под ред. Л.С. Гребнева. — М.: Центр тестирования Министерства образования РФ, 2002. <http://testolog.narod.ru>.

<sup>21</sup> Аванесов В.С. Единый Государственный Экзамен в фокусе научного исследования // Педагогические измерения. — 2006. — № 1. — С. 3–31. <http://testolog.narod.ru/Education39.html>.



считали, вероятно, различия между тестами этих двух наук не существенными. Однако эта логика могла относиться только к статистическому аппарату обеих теорий. Что касалось целей, задач, содержания и применения, то различия, конечно, существенны.

В русскоязычной научной литературе название «тестология» возникло в 20-х и 30-х годах двадцатого столетия. В СССР конца 30-х годов XX века теорию педагогических и психологических тестов называли «буржуазной тестологией». Из этого можно видеть, что в истории исследуемой проблемы без политики не обходилось.

С начала XXI века стало утверждаться название «Теория педагогических измерений». Сейчас интересующую нас теорию лучше называть теорией педагогических тестов (ТПТ).

## Миссия педагогического тестирования

Слово «миссия» имеет смысл предназначения, главной цели или функции, ради которых используется этот вид деятельности. Главная миссия педагогического тестирования — повышение *объективности* результатов проверки учебных достижений учащихся и студентов.

Так получается, и не только в России, что объективность получаемых тестовых оценок достигается в случаях, когда тестирование проводится независимо от государства. Тесты и государство — несовместимые понятия, что подтверждается историей тестов<sup>22</sup>. Это очень важное условие для мировой практики тестирования было проигнорировано при проведении ЕГЭ в России. Объективность достигается в условиях открытости банков тестовых заданий, компьютерных программ выбора равно трудных вариантов

заданий для каждого испытуемого, независимой экспертизой содержания и формы тестовых заданий, надёжностью и валидностью тестовых результатов.

Объективность связана со справедливостью. Последняя предполагает необходимость коррекции баллов в государственных экзаменах у тех испытуемых, условия обучения которых не были равными с условиями обучения других. Главное системное требование перед началом тестирования — создание равных условий для изучения школьных и вузовских предметов. Мы слишком далеки от выполнения этого социального требования, в полной мере. Начинать массовый контроль лучше с создания равных условий для обучения. И только тогда имеют смысл равные условия для контроля. Но когда приоритет отдаётся государственному контролю, то возникает инверсия, которая начинает свою разрушительную деятельность.

## Актуальность педагогического тестирования

Нынешняя практика тестирования даёт немало фактов для разных выводов — как против тестов, так и в их пользу. Голоса «против» звучат иногда громче голосов «за». Важно понять простую истину: сами по себе тесты — полезный метод, но их применение в практике может оказаться как полезным, так и вредным. Многое зависит от цели, качества, соблюдения принципа добровольности, возможности потенциального или реального ущерба личности.

Актуальность тестового метода объясняется некоторыми его преимуществами перед другими методами. Кратко напомним о пяти основных преимуществах тестов:

1. Высокая *научная обоснованность* настоящего теста;

<sup>22</sup> Аванесов В.С. История тестов. Viperson.ru.

2. *Технологичность* тестовых методов;

3. Объективно фиксируемая и открыто проверяемая *точность измерений*;

4. *Одинаковость правил измерения* и *адекватной интерпретации* тестовых результатов;

5. *Сочетаемость тестовой технологии* с другими современными образовательными технологиями.

Можно надеяться, что в последующие годы XXI века руководители сферы образования вряд ли смогут без риска для карьеры отрицать, или не принимать во внимание отмеченные научно-технологические преимущества тестов и тестовых заданий.

### Два основных предмета педагогического тестирования

Педагогическое тестирование имеет два основных предмета: *теорию педагогических тестов* и *практику педагогического тестирования*. Оба этих предмета основной сферой применения имеют учебный процесс в школах и вузах, с основной деятельностью под названием «педагогическая диагностика». Результаты педагогической диагностики обеспечивают информацию для обратной связи с учебным процессом. Эти результаты необходимы для очень важной педагогической деятельности, от которой зависит будущее страны — для автоматизированной и обычной педагогической *коррекции* приобретаемых знаний всех изучающих науки.

Раньше два упомянутых предмета тестирования автором этой статьи рассматривались в рамках более широкой *теории и практики педагогических измерений*. Учёные знают о преимуществах выделения более узкого предмета исследования. Его легче познать и его можно точнее описать.

## Основные понятия педагогического тестирования

### Педагогическое задание

Под этим именем можно представить многие педагогические формы, используемые для упражнений учащихся, проверки и самопроверки уровня подготовленности. В результате теоретического анализа было определено, что *задание* — это педагогическая форма, нацеленная на проверку усвоения учащимися изучавшихся ими тем и разделов курса.

Каждое задание имеет как бы свою миссию, создаётся для определённой цели и имеет свои характеристики. Главная миссия заданий — побуждение учащихся к активной самообразовательной деятельности, к развитию учащихся и овладению необходимыми знаниями, умениями и навыками до требуемого уровня компетенции.

### Содержание педагогических заданий в тестовой форме

Содержание заданий в тестовой форме тесным образом связано с содержанием образовательного процесса и зависит от последнего. В заданиях нет, и не может быть содержания, которое не было бы запланировано в учебных курсах.

Содержание заданий определяется как отображение содержания элементов учебной дисциплины во множестве заданий. Такое отображение осуществляется в результате педагогического анализа содержания учебной программы, выделения в ней всех тем, понятий, положений, принципов и, где надо, формул, а затем — поэлементной редукции всего перечисленного в содержание заданий в тестовой форме. Редуцирование осуществляется посредством логического анализа знаний и представления таковых в тестовой форме. При этом форма выступает как

способ связи, упорядочения элементов содержания. Сейчас чаще других применяются задания с выбором нескольких правильных ответов. Они позволяют проверить способности к дивергентному мышлению.

Содержание заданий в тестовой форме может охватывать полное множество предлагаемых знаний, что особенно важно при организации самостоятельной работы в текущем учебном процессе, а также при разработке супертеста<sup>23</sup> для проведения текущего мониторинга.

Решая задания, студенты (школьники) получают возможность последовательно и без спешки, в удобное для них время, проверить уровень своей подготовленности по всем темам и разделам курсов содержания которых было включено в учебную программу. Эта программа обеспечивается тысячами заданий в тестовой форме. Тем самым создаётся одна из двух главных педагогических составляющих электронных учебников. Вторая главная составляющая – это квантованные учебные тексты.

Например, студенты медицинских вузов должны знать названия всех костей и мышц, изучаемых в курсе анатомии, а также обширный лексикон физиологии, названия тысяч болезней и лекарств. Поэтому и число заданий в тестовой форме для самостоятельной работы должно быть порядка нескольких тысяч. Содержание заданий в тестовой форме определяют преподаватели учебной дисциплины.

### **Содержание заданий формулируется как можно точнее, короче и яснее**

*Точность* содержания обеспечивается использованием терминов, элементов искусственного языка, формул, исключением метафор и неадекватной лексики.

*Краткость* достигается тщательным подбором слов, символов, графиков, позволяющих минимумом средств добиваться максимума ясности смысла задания. Для достижения краткости в каждом задании лучше спросить о чём-нибудь одном. Хорошо, когда задание содержит не более одного придаточного предложения, а лучше, чтобы его не было вовсе.

*Ясность* содержания достигается тогда, когда все испытуемые правильно понимают подлинный смысл задания. С этой целью из заданий полностью исключаются повторы слов, малопонятные, редко употребляемые, а также не изучавшиеся в курсе символы и иностранные слова, затрудняющие восприятие сути задания.

В заданиях, например, по физике, полезно использовать рисунки интересующего процесса. Рисунки повышают понимаемость задания, позволяют избежать многословия, способствуют образному представлению, что помогает также и быстрой восприятия, а это существенно при применении тестовых форм.

Анализ содержания курсов обычно относится к компетенции преподавателей, в то время как анализ содержания заданий – один из ключевых вопросов теории педагогических измерений.

Общие принципы разработки содержания заданий рассматриваются в теории педагогических измерений, в то время как содержание учебных дисциплин относится больше к преподаваемой науке, знание которой проверяется.

### **Педагогический тест**

Система заданий в тестовой форме, возрастающей трудности, позволяющая качественно оценить структуру и измерить уровень знаний.

*Педагогическими* называются тесты, которые разрабатывают педа-

<sup>23</sup> Аванесов В.С. Супертест // Педагогические измерения. – 2014. – № 2. – С. 3–13.

гоги для объективной оценки результатов обучения и, надо теперь добавлять, самостоятельного изучения. Поскольку значение самостоятельного изучения школьных и вузовских курсов кардинально меняет всю деятельность образовательных учреждений. Настолько меняет, что можно говорить о совершающемся на наших глазах кардинальном переходе от массового обучения — к самостоятельному изучению, при компьютерной поддержке. Тех, кто не умеет учиться сам, теперь вряд ли хорошо научат, поскольку образовательная деятельность стремительно меняется от обучения к изучению. В России с этим есть проблемы из-за практического отсутствия подходящих квантованных учебных текстов и технологичных заданий в тестовой форме.

### Содержание педагогического теста

В тестовом задании внимание привлекает, в первую очередь, содержание и форма. Содержание определяется как отображение фрагмента учебной дисциплины в тестовой форме, форма — как способ связи, упорядочения элементов задания.

Содержание теста — это оптимальное отображение содержания образования в системе тестовых заданий. Слова «оптимальное отображение» предполагают необходимость отбора такого контрольного материала, ответы на который с высокой вероятностью (больше 95%) свидетельствовали бы об уровне подготовленности каждого учащегося.

Требование оптимальности отображения подводит к необходимости периодического пересмотра цели и смысла педагогической деятельности. До недавнего времени практика общего среднего образования сводилась к овладению известным перечнем знаний, умений и навыков; в методической литературе последние иногда, по первым буквам, называют ЗУНами.

Оптимизация содержания является ведущей идеей традиционного теста, а в ещё большей степени — адаптивного теста, где ведущая идея — минимумом числа заданий, за короткое время, быстро, качественно и с наименьшими затратами измерить знания как можно большего числа учащихся.

Эта идея близка по смыслу задаче повышения эффективности педагогической деятельности при использовании массовых форм контроля знаний.

### Критерии отбора содержания теста

1. Соответствие содержания теста целям тестирования.
2. Значимость проверяемых знаний в общей системе знаний.
3. Взаимосвязь содержания и формы.
4. Содержательная правильность тестовых заданий.
5. Репрезентативность содержания учебной дисциплины в содержании теста.
6. Соответствие содержания теста уровню современного состояния науки.
7. Комплексность и сбалансированность содержания теста.
8. Системность содержания.
9. Вариативность содержания.
10. Соответствие уровня трудности содержанию цели тестирования.

### Тестирование

Это применение тестов для получения информации об интересующих свойствах личности, о результатах обучения и изучения предметного и метапредметного содержания образования. Это содержание обычно предписывается в самых общих терминах, в государственных образовательных стандартах. Такое определение относится сразу к психологическому и педагогическому тестированию.

## Педагогическое тестирование

Определяется как научно обоснованная *практика* применения педагогических тестов для объективной оценки уровня и структуры подготовленности учащихся и студентов.

## Система педагогического тестирования

Эта система может оказаться научно обоснованной, не обоснованной и даже антинаучной. Исходные результаты тестирования теперь уже не рассматриваются как педагогические измерения; это лишь первый этап сбора данных для последующего логарифмического шкалирования. Можно сказать и так: тестирование — это незаконченное педагогическое измерение. В новейшей литературе только шкалированные результаты тестирования признаются полноценными педагогическими измерениями.

## Форма тестовых заданий

Форма тестовых заданий представляет собой специфический способ связи содержания заданий с наиболее подходящей для этого содержания формой. Кроме того, в тестовых технологиях иногда применяются задания открытой формы и задания на установление правильной последовательности. Задания на установление соответствия используются сейчас мало, из-за их громоздкости и возможности замены заданиями с выбором нескольких правильных ответов.

Вне тестовых форм ни тест, ни его содержание не существуют. По форме все известные в теории и практике тестовые задания можно разделить на четыре группы. Вот их точные названия: задания с выбором одного или нескольких пра-

вильных ответов, задания открытой формы, задания на установление соответствия и задания на установление правильной последовательности. Тысячи примеров заданий различных форм, по многим, если не всем учебным дисциплинам, читатель найдёт в книгах автора<sup>24</sup>. Пример логически сбалансированного учебного задания в тестовой форме, с выбором одного правильного ответа:

### 1. ЕДИНИЦА СЧИТАЕТСЯ ЧИСЛОМ

- 1) простым
- 2) составным
- 3) простым и составным
- 4) ни простым, ни составным

Задания с выбором нескольких правильных ответов, которым, за редким исключением, не нашлось места в ЕГЭ, во многих случаях оказываются лучше заданий с выбором одного правильного ответа.

В таких заданиях, наряду с одним правильным ответом, есть несколько неправильных, но правдоподобных. Последние называют дистракторами (от англ. *to distract* — отвлекать). Число дистракторов в заданиях с выбором нескольких правильных ответов может колебаться, обычно, от одного до 10–12.

## Применение языка и методов статистической науки для разработки педагогических тестов

Центральной моделью статистической теории педагогических измерений обычно является т.н. нормальное распределение результатов. Существует много причин, по которым методы для массового оценивания ориентируются именно на это распределение. Отклонение от такой модели нередко вызывает обоснованные подозрения в некачественности метода или данных. В первой части ЕГЭ плохо сформулированные задания способствуют феноме-

<sup>24</sup> Аванесов В.С. Форма тестовых заданий. — М.: Центр тестирования. 2006.

ну угадывания правильного ответа, что заметно отклоняет результаты от нормального распределения и даёт основания усомниться в качестве результатов. Сходный эффект наблюдается при списывании. Последние удалось уменьшить за счёт усиления непедagogических мер, похожих на полицейские.

Самые распространённые в педагогической практике тестирования статистические характеристики тестов и тестовых заданий — это доля правильных и неправильных ответов, коэффициенты корреляции между ответами испытуемых по заданию и исходными тестовыми баллами испытуемых, показатели средней тенденции (средняя арифметическая, мода и медиана), показатели вариации исходных тестовых результатов. При разработке теста нередко вычисляются пять показателей вариации — размах данных, сумма квадратов отклонений от средней арифметической, дисперсия, стандартное отклонение, коэффициент вариации.

Чем больше значение коэффициента корреляции, тем выше *дифференцирующая способность задания*. В последние годы вместо понятия «дифференцирующая способность задания» автор этой статьи чаще применяет понятие «различающая способность задания». Это связано с тем, что слово «дифференцирующая» в зарубежной теории педагогических и психологических измерений используется для нескольких целей.

Кроме того, для исходных тестовых баллов рекомендуется определять значения коэффициентов асимметрии и мер эксцесса. Это позволяет сразу увидеть некачественность полученных результатов, точки зрения их соответствия или несоответствия основной аксиоме тестовой теории: уровень трудности заданий должен соответствовать уровню подготовленности испытуе-

мых. Пока в Интернете можно было найти хоть какое-то распределение результатов ЕГЭ, мы неизменно убеждались в нарушении создателями методов ЕГЭ этой аксиомы.

### Критерии приемлемости тестовых результатов

История теории педагогических тестов содержит много тысяч наименований трудов, нацеленных на разработку критериев объективности, эффективности и качества тестовых результатов. Всего насчитывается четыре основных критерия оценки результатов тестирования. В литературе часто их относят к тестам, но это неверно. Правильнее говорить к результатам тестирования. Это критерии объективности, надёжности, валидности и эффективности<sup>25</sup>.

В XXI веке можно будет надеяться, что в России появится профессиональный стандарт на разработку и применение педагогических тестов, который должен будет поставить заслон всем методам, не имеющим научно подтверждённых данных по названным критериям.

### Основная сфера применения тестовых форм в XXI веке

В XXI веке основная сфера применения тестовых форм станет учебный процесс, а также разработка заданий и квантованных учебных текстов для банков цифровых образовательных ресурсов (ЦОР). Что может обеспечить качественными материалами учебный процесс во всех образовательных организациях. И в этом деле российское образование сильно отстаёт от стран-лидеров качественного образования.

Успешная модернизация образовательной деятельности в стране

<sup>25</sup> Статьи автора по всем критериям представлены в российском научно-методическом журнале «Педагогические измерения» и на сайте автора [testolog.narod.ru](http://testolog.narod.ru).

невозможна без системы цифровых образовательных ресурсов, ассоциированной с уже имеющимися мировыми электронными образовательными ресурсами. Создание и объединение собственных образовательных ресурсов с другими позволит облегчить доступ студентам и преподавателям к лекциям известных педагогов и учёных, быстро тиражировать передовой педагогический опыт, довести до каждого видеоматериалы учебных экспериментов, а также создать задания и тексты для адаптивных технологий обучения и для контроля уровня подготовленности. Создание вузовского фонда цифровых образовательных ресурсов позволит наладить централизованную систему самостоятельной работы всех студентов страны, вести полный учёт изученного материала и числа решённых заданий каждым. То же самое относится и к школе.

При разработке цифровых образовательных ресурсов главное — это короткие и понятные студентам тексты, снабжённые заданиями в тестовой форме, разного уровня трудности. Разработке таких заданий педагогов надо специально учить. Студенты могут во второй половине дня самостоятельно изучать такие задания и учебные тексты. Где надо, задания надо решать с применением компьютера и компьютерных программ. Это сильно экономит учебное время. Что позволяет индивидуализировать учебный процесс, сделать его продуктивным, адаптивным, а значит, интересным и эффективным.

## Выводы

1. Россия является единственной в мире страной, где в качестве альтернативы научному методу тестирования созданы т.н. КИМы ЕГЭ. Научная ценность последних никак и нигде эмпирически не подтверждена. Опубликованными исследованиями подтверждена лишь их ошибоч-

ность и неприемлемость для применения в сфере образования.

2. В России педагогическое тестирование находится сейчас в трудном положении по причине ошибочного отношения к этому методу руководителей Минобрнауки и Рособнадзора. Поэтому педагогическое тестирование нуждается сейчас в поддержке со стороны верховной власти, учёных, педагогов, общественных и профессиональных объединений.

3. Неудачное творение — ЕГЭ — не поддаётся улучшению в силу конструктивных ошибок, заложенных при его создании. Поэтому этот экзамен рекомендуется закрыть. Вместо него в России лучше создать независимую, самокупаемую, автоматизированную, общественно-профессиональную службу тестирования в сфере образования, типа Educational Testing Service в США.

4. Попытки удаления из ЕГЭ заданий, похожих на тестовые, не улучшат качество этого экзамена, и не спасут его. После удаления большого числа заданий ЕГЭ станет ещё менее репрезентативным, по числу заданий, относительно объёма проверяемого содержания учебного предмета и установленных государственных образовательных стандартов. Этот экзамен может тогда стать более некачественным по содержанию, длительным по времени и нетехнологичным. Возрастут и финансовые запросы на проведение такого рода нетехнологичного экзамена. С заданиями, похожими на тестовые, создалась тупиковая ситуация, когда их нельзя ни применять, ни удалять. Гораздо проще закрыть сам ЕГЭ.

5. Хорошо бы прислушаться к мнению учёных, которые пишут о проблемной ситуации с ЕГЭ. Вот одно из самых аргументированных и интересных мнений, с которым могут согласиться *все* граждане РФ. «Никакие исправления *частных* ошибочных действий не могут помочь, если неверно избрано *общее* направление развития. Для оценки и уст-

ранения этих принципиально ошибочных действий необходимо прибегнуть к системному анализу происходящих преобразований, сделать надлежащие выводы и не повторять ошибок, которые давно учтены и проанализированы другими»<sup>26</sup>.

**6.** Все средства государственного бюджета лучше тратить на нужды образования в самих образовательных организациях. Электронный контроль и коррекцию знаний целесообразно сосредоточить исключительно в этих организациях, и применять эти очень полезные формы

лучше в процессе самообразования и самокоррекции знаний.

**7.** Нуждаются в исследовании этические вопросы и последствия массового нарушения педагогической этики при проведении непедагогического ЕГЭ в образовательных организациях.

**8.** Образовательная политика российского государства нуждается в изменении, особенно в части устранения ЕГЭ. При нынешнем уровне развития образовательных технологий многие проблемы образования могут вполне решаться в текущем учебном процессе, без этого экзамена.

<sup>26</sup> Безродная Г.В., Севостьянов Д.А., Чельцов М.В. Педагогический контроль: инверсия целей и средств. Статья представлена по адресу: <http://www.odn2.ru/index.php/biblioteka/34-uchebnyj-protsess/289-pedagogicheskij-kontrol-inversiya-tselej-i-sredstv>.