

От заданий в тестовой форме - к тестовым заданиям

Вадим Аванесов

testolog@mail.ru

Опубликовано в ж. «Педагогические Измерения» № 4, 2010 года.

Публикуется в новой редакции 2019 г.

Введение

Название этой статьи может показаться несколько парадоксальным, поскольку трудно сразу вникнуть в тайны различий между заданиями в тестовой форме и тестовыми заданиями. Для постижения подлинных смыслов различных заданий и других сопряжённых понятий приходится постепенно погружаться в тайны языка педагогических измерений, или по-старому, тестологии.

Так получилось, что в образовании тесты стали занимать заметное место. При профессиональной экспертизе используемых тестов выясняется, что настоящих тестов в образовании нет, или почти нет, а то, что везде называется тестом, оказывается совсем не тестом или псевдотестом, в котором вместо тестовых заданий применяются не вполне удачно сформулированные вопросы. Совсем нет собственно тестологической (метрической) информации о качестве применяемых «тестов», а без этого трудно поверить, что мы имеем дело с настоящим тестом. Достаточно привести пример т.н. КИМов ЕГЭ, где метрическая информация о качестве (точнее, о некачественности) используемых методов засекречена с первого дня.

Между тем, этические нормы тестологической службы США, Великобритании, Австралии и многих других стран Запада и Востока запрещают применять тесты, если нет опубликованной информации об их качестве. Таким образом, в этих странах личность защищена от негодных

методов получения информации. В России вместо тестов оказываются КИМы ЕГЭ т.п. Хотя вокруг нас говорят преимущественно о тестах. Тестом называют даже отдельное задание, хотя тест – это система заданий равномерно возрастающей трудности, позволяющая качественно и эффективно измерить уровень и структуру подготовленности каждого испытуемого. Отсюда вытекает важная задача - научить каждого дифференцировать тесты от псевдотестов и, если можно так сказать, не тестов.

Автору этой статьи часто приходится объяснять, что сделать хорошее задание труднее, чем поставить правильный вопрос, а точный вопрос сформулировать труднее, чем дать точный ответ. Поэтому в логике не случайно говорят о культуре вопроса, а в педагогических измерениях - о культуре формулирования заданий в тестовой форме.

Чему учит история?

Элементы истории уже затрагивались в публикациях ряда авторов. До XX века в России тестов не было. Вместо тестов проводились экзамены. В циркуляре по Санкт-Петербургскому учебному округу “О производстве испытаний зрелости без послаблений”, сообщалось, что во многих гимназиях испытания зрелости производятся крайне снисходительно, с весьма большими послаблениями. А это обстоятельство, в свою очередь, весьма неблагоприятно отзывается на всем ходе учебного дела, приучая воспитанников поверхностно относиться к усвоению преподаваемых им наук, в надежде на ожидаемые послабления на испытаниях.

В этом же циркуляре определён главный принцип определения содержания контрольных заданий: “не обременяя требованиями мелких и

второстепенных подробностей... необходимо удостовериться в знании существенно-важного в каждом предмете»¹.

Идея принудить лучше учиться угрозой проведения экзамена имеет корни в старой немецкой педагогике. Однако международный педагогический опыт, особенно современный, показывает, что принуждение в образовании приносит больше отрицательных результатов, чем положительных. Эти результаты предвидел замечательный русский педагог и хирург Николай Пирогов. «Я почти ежедневно убеждаюсь из опыта, что экзаменационное направление ... не приносит никаких благих результатов. Оно вредно, возбуждает склонность в учащихся учиться для экзамена, а не для науки», - писал он². Ещё в древности было известно, что... «ни одну науку не следует изучать рабски» так как «...насильственно внедренное в душу знание непрочно».

Попытки научного обоснования качества заданий были сделаны в начале XX-го века. Это произошло во Франции. Авторы первого теста, созданного для измерения интеллектуальных способностей детей, А. Бине и Т. Симон провели эмпирическую проверку заданий, которые предполагалось включить в тест. Для оценки пригодности заданий авторы использовали два основных критерия:

1) эмпирическую меру трудности каждого задания, которую они определяли по доле или, умножением долей на сто, по проценту правильных ответов в группах детей разного возраста.

2) информацию о степени совпадения результатов теста с мнением преподавателей.

Для достижения сопоставимости данных, полученных ими с данными других исследователей и минимизации ошибок измерения, А. Бине

¹ См. Н.Я. Пирогов. Избр. пед. соч. М.1952. - С. 219-220).

² Циркуляр по Санкт-Петербургскому учебному округу. №5, Май 1901г., с. 218.

и Т. Симон прилагали к тесту стандартную инструкцию по проведению тестирования. Интересен метод выявления дифференцирующей способности, заданий, которым пользовались эти авторы. Результаты ответов испытуемых, на каждое задание, представлялись в виде точек на плоскости, где по оси абсцисс откладывались значения возраста, а по оси ординат - доля правильных ответов, в каждой возрастной группе. Усредняя полученные точки и затем, соединяя их линией, они получали графический образ, на основании которого делали выводы об адекватности задания для того или иного возраста³.

Позже, в конце XX-го века появились компьютерные программы RUMM 2010 (сейчас 3020) и Winsteps, которые позволяют получать графические образы каждого задания. Если графический образ удовлетворяет требованиям настоящего теста, то задание с таким графиком включается в тест, если не удовлетворяет, то в тесте таком заданию места нет.

Так, постепенно, мы подошли к тому, что математика, информатика, и педагогические измерения стали настоящими помощниками педагога, который стремится делать настоящие тесты. Можно сказать и больше: без знания элементов теории педагогических измерений и без технологий делать высококачественные тесты стало делом невозможным. Это и есть главный урок истории. Усвоить это урок может только тот, кто хочет и может извлечь пользу из исторического знания. Остальных история ничему не учит.

О новой роли заданий в тестовой форме

³ Подробнее см. Аванесов В.С. Методологические и теоретические основы тестового педагогического контроля. Дисс. соиск. уч. степ. докт. пед. наук. С-Пб. гос. ун-т. 1994, 339с.

В системе среднего и высшего образования сейчас начали активно внедряться новые образовательные технологии, основанные на достижениях современной науки. Девиз педагогики XXI –го века «От обучения всех - к образованию каждого! Каков подлинный смысл этого девиза? Обучение всех человечество пыталось реализовать групповыми формами, где, по прежним меркам, лучше себя проявляла классно-урочная форма образовательной деятельности.

Но прошло время, и мы убедились, что при классно-урочной форме образование нацелено на всех сразу, но ни на кого - конкретно. Вузы стали выпускать специалистов сильно различающегося уровня подготовки, что зависит не только от качества преподавания, но и от индивидуальных способностей и учебной мотивации отдельных студентов. При классно-урочной форме организации занятий в центре образовательного процесса находится не индивид и индивидуальность, без которой подлинное образование каждого невозможно, а педагог и класс в целом.

Как это ни покажется странным, учить всех оказывается легче, чем научить одного. Здесь примерно такое же соотношение, как в медицине: лечить вообще не трудно, трудно вылечить конкретного больного. Не случайно в медицине сложилась такая ситуация: лечат все, вылечивают немногие. И в системе образования очень схожая ситуация: учат многие и многих, а научают хорошо единицы, и только отдельных учащихся или студентов.

Для реального претворения девиза «От обучения всех - к образованию каждого» в жизнь нужно по-новому взглянуть на роль заданий в тестовой форме. Успехи образования во многих странах мира связаны с изменением отношения именно к заданиям, и к вопросам повышения их качества. Если раньше качество образования почти полностью зависело от профессиональной подготовки педагогов, то теперь, по мере развития

образовательных технологий и Интернет, заметную роль в развитии образования стали выполнять задания в тестовой форме. Они формулируются так, чтобы:

- компьютер сумел быстро и чётко предъявить задание испытуемым;
- смысл задания был понятен каждому;
- компьютер смог безошибочно распознать правильные ответы и выставить оценки так, чтобы большему уровню знаний соответствовали и более высокие оценки;
- в конце тестирования получаемые исходные баллы испытуемых соответствовали требованиям теории педагогических измерений. А это значит, у более подготовленного испытуемого должен быть и более высокий тестовый балл.

Задания стали выполнять не только контролирующую, но также развивающую и обучающую функцию. Причём в новых образовательных технологиях роль развивающей и обучающей функций оказывается выше, чем контролирующая. Наиболее эффективны для обучения развивающие задания в тестовой форме. Эти задания не отменяют важную роль речевых упражнений, письменных форм контроля, коллоквиумов, практикумов и прочих хорошо зарекомендовавших себя форм обучения и контроля.

Задания выполняют синергетическую функцию объединения усилий педагога и обучаемого. Метафорически задание можно представить как мост между учащимся и педагогом, открывающим возможность для их взаимодействия, объединяющим их усилия в учебном и воспитательном процессе. Если держаться и далее упомянутой метафоры, то качество и эффективность учебного процесса в существенной мере зависят от качества моста. Чем лучше задания, тем (при прочих равных условиях)

прочнее и весомыми могут быть результаты взаимодействия учащегося и педагога.

Лекционные и прочие формы педагогической деятельности представляют только одну, преподавательскую сторону педагогического процесса. Выполнение заданий испытуемыми – вторая, не менее, а часто и более важная сторона педагогического процесса. Основные виды заданий - задача, вопрос, упражнение, творческое задание, задание в тестовой форме, тестовое задание, учебная проблема, курсовая и дипломная работа, а также и другие.

В наши дни, наряду с задачами и развивающими вопросами, заметную роль приобрели задания в тестовой форме. Обучение без заданий - не эффективно, а нередко и вредно. Отсутствие или недостаток опыта самостоятельного решения учебных задач переносится и на неумение выпускника школы или вуза успешно решать затем и жизненные задачи.

С этой точки зрения, сложившаяся практика планирования, учёта, оценки и оплаты педагогического труда по количеству проведенных лекций или уроков, без реального учёта количества и качества выполненных учащимися заданий, представляет собой анахронизм, тормозящий улучшение образовательной деятельности.

Эффективное задание

Эффективное задание позволяет учащимся за единицу учебного времени добиться большего эффекта, сравнительно с другими педагогическими заданиями: быстрее усвоить новые элементы знаний, правильнее их применять, улучшить воспроизведение знаний, повысить точность изложения и др. В наше время эффективность заданий можно повысить, если удаётся:

- точно отобрать и сформулировать содержание задания. Формулировка даётся не в форме вопроса, а в форме высказывания, краткий ответ на которое превращает высказывание в истинное или ложное.

- так сформулировать задание, чтобы его мог предъявить испытуемым компьютер;

- чтобы компьютер сумел быстро распознать правильный ответ от неправильного и быстро выставить исходные тестовые баллы.

Именно в вопросах производства и применения эффективных заданий российское образование стало сильно отставать от общемировых тенденций. Это одно из самых слабых мест российской педагогики и педагогической науки. Отсюда – актуальность проблемы измерения эффективности и качества педагогических заданий.

Главная метрическая ошибка в т.н. ЕГЭ по русскому языку

В общем случае результаты всех испытуемых по настоящим тестам должны иметь, чаще всего, так называемое нормальное, симметричное распределение результатов. В итоге измерения посредством настоящего теста должна получиться шкала измерения с равными или с примерно равными интервалами. Последнее требование сейчас выполняется очень редко, потому что для этого надо провести так называемое шкалирование результатов испытуемых посредством упомянутых выше компьютерных программ.

Неоднократно уже приходилось писать, что теория педагогических измерений не допускает возможность использования одной и той же системы заданий для разных целей, и особенно, для разных уровней подготовленности испытуемых. При педагогических измерениях такое использование повышает в разы погрешность получаемых оценок и резко

снижает качество измерений. Именно это сейчас происходит в т.н. «Едином государственном экзамене» по русскому языку. В математике и в других учебных дисциплинах этот государственный нонсенс был преодолен лишь к 2009 году. А до того времени образовательные власти упирались, как могли. Сейчас «Единым» в смысле объединённого аттестационного для выпускников школ и вступительного для абитуриентов вузов остался только один ЕГЭ – по русскому языку. Для того, чтобы убедиться в его ошибочности в смысле очень высокой погрешностей измерения, достаточно посмотреть на гистограмму распределения реальных результатов ЕГЭ по русскому языку.

Посмотрим на реальный пример проверки знаний по русскому языку в ЕГЭ 2016 и 2017 годов⁴. Об этом на гистограммах свидетельствует очень завышенные результаты так называемого модального класса значений баллов ЕГЭ по русскому языку. Завышение времени и, кроме того, облегчение множества заданий позволило повысить общее количество баллов испытуемых, создать впечатление о высоком и очень высоком знании выпускниками школ русского языка. Этот приём может вызвать в сознании некоторых чиновников головокружение от успехов и надежду на получение государственных наград. Многократно уже было сказано, что реально уровень владения русским языком, математикой и многими другими учебными дисциплинами снижается.

⁴ Педагогические измерения №3, 2017. С. 9

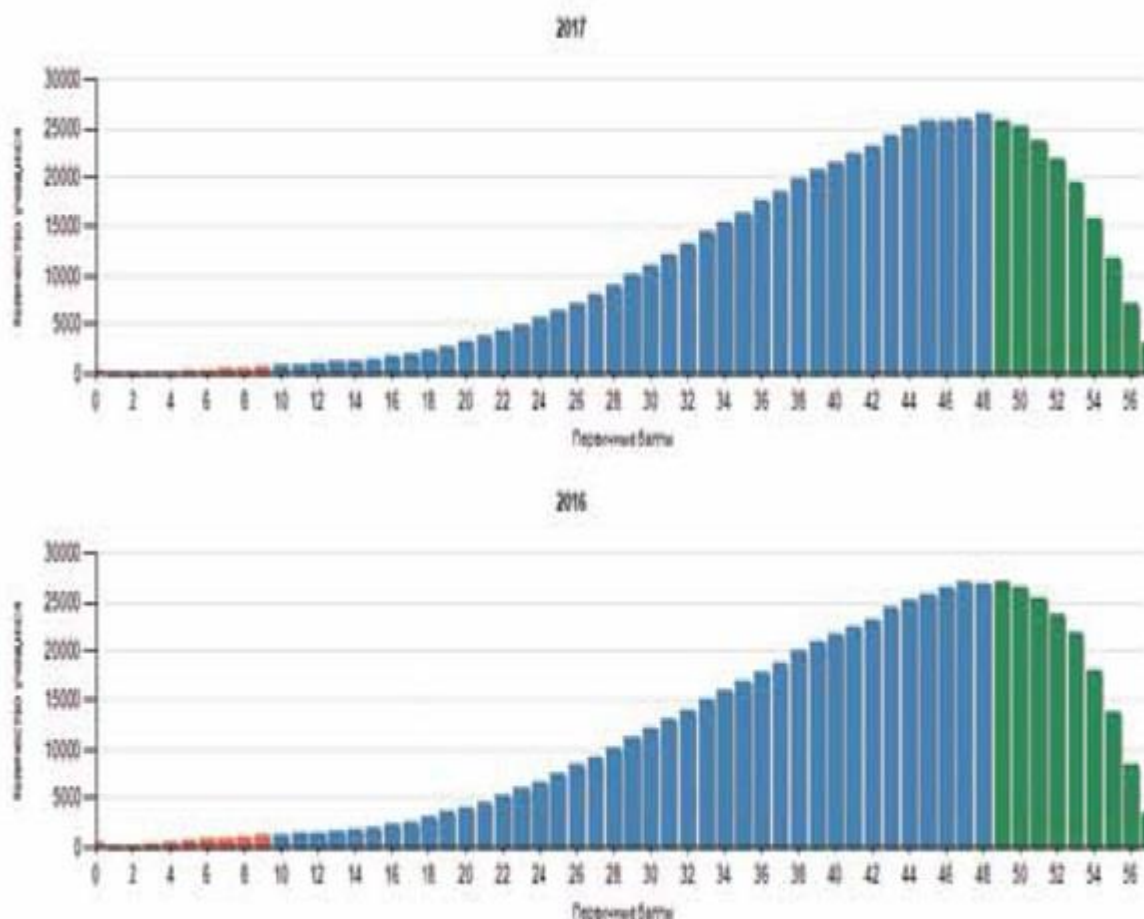


Рис.1 и 2. Гистограммы распределения результатов ЕГЭ по русскому языку в 2017 и 2016 годах.

Приведённые гистограммы указывают так же на то, что многократно обещанное совершенствование т.н. КИМов ЕГЭ по русскому языку в 2016 и 2017 годах в реальности не произошло.