

Теория

ПЕДАГОГИЧЕСКИЙ АНАЛИЗ ЕДИНОГО ГОСУДАРСТВЕННОГО ЭКЗАМЕНА (ЕГЭ)

В. Беспалько

В июне 2006 года средства массовой информации сообщили всему миру, что внедрение ЕГЭ в России отложено до 2009 года. Очень правильное решение: есть ещё время подумать... и отказаться от этого экзамена вовсе.

Содержание лекции:

Откуда взялся ЕГЭ?
Что собой представляет ЕГЭ?
Что от ЕГЭ хотят?
Что ЕГЭ может дать?
Что с ЕГЭ делать?
Тесты, экзамены и демократия.
Заключение.

Откуда взялся ЕГЭ?

Он не свалился с неба. У американцев вводят финальный тест, и мы не лыком шиты: и у нас будет финальный тест! Его попросту слепо и некритично списали с печально известных творений корпорации ETS, так называемых **тестов-достижений SAT** и **фи-**

нального теста, введённого администрацией президента Буша в американских школах. Как и в Америке, под него откачали массу денег из бюджета, но в Штатах к этому времени накопился столетний опыт тестирования и существовала многолетняя и многоопытная организация ETS, на заделы которой и опираются финальные тесты в школах США. В России же, где в отличие от Америки уже давно сформировалась достаточно развитая педагогическая наука, руководители образования не нашли ничего лучшего, как попросту отбросить этот вековой научный задел и на пустом месте создать несколько федеральных организаций для решения сложнейшей проблемы измерения и оценки качества знаний учащихся и в целом — качества образования в стране. Неудивительно поэтому, что, поскольку заимствование выполнено столь близко к оригиналу, ему присущи все недостатки оригинала. Кроме того, авторы российского варианта SAT — ЕГЭ при его применении исказили основную установку оригинала — **добровольность теста SAT и автономность финального теста** — и сумели придать этому педагогически несовершенному продукту статус **обязательного централизованного государственного экзамена**. С поистине

российским размахом сделана попытка внедрения повсеместно того, что в стране рождения, США, получает всё более негативную оценку и педагогов, и общественности и от чего всё больше отмежёвываются школы.

Основные причины этой тенденции таковы:

а) отсутствие заранее поставленной и гарантированно достижимой цели обучения, поэтому ЕГЭ, как и SAT, уподобляется «телеге, поставленной впереди лошади». Действительно, что проверяется? Известно: дети ходят в школу и там что-то изучают, но что и особенно как — остаётся диагностично неопределённым и отданным для оценки на откуп учителю. Не случайно американские школьники всё чаще демонстрируют своё неприятие финального теста, правильно утверждая, что его требования не соответствуют школьному уровню обучения, к которому они привыкли;

б) отсутствие чётких исходных позиций в конструировании материалов экзамена, необычность обстановки и процедур его осуществления, а также неясность способа оценки результата. Авторы SAT, учитывая эти слабости своего создания, стараются завуалировать словесной эквилибристикой пустоту его психолого-педагогических основ и присваивают

Теория

Откуда взялся ЕГЭ?

ПЕД диагностика
ПЕД диагностика

Откуда взялся ЕГЭ?

испытуемым **абстрактную** оценку по шкале в несколько сотен баллов, которую и не пытаются соотносить со школьной шкалой оценки. Авторы ЕГЭ пошли дальше своих предшественников и «изобрели» множество противоречивых шкал перевода из абстрактной 100-балльной шкалы в привычную пятибалльную. И этим неожиданно для самих себя **полностью опровергли необходимость такого экзамена** вообще, поскольку ту же самую информацию об успеваемости учащихся всегда сообщал учитель на основе своих субъективных оценок;

в) запредельная стоимость и иллюзорная объективность громоздкой «машины» ЕГЭ, призванной выполнять самую рутинную функцию всего учебного процесса — контроль качества усвоения. Эта функция по всем педагогическим канонам всегда доверяется учителю, и вопрос состоит лишь в том, чтобы дать ему в руки **объективный и надёжный инструмент диагностики** и от него непосредственно получать необходимую информацию о состоянии образования в стране безо всяких посредников и преобразователей. К сожалению, таким инструментом не обладает ни ETS, ни его неудачный «клон» ЕГЭ.

Примечательно, что со времени принятия Правительством РФ исторического реше-

ния № 119 от 16 февраля 2001 года «Об организации эксперимента по введению Единого государственного экзамена» не состоялось ни одной научной конференции или совещания о психолого-педагогических основаниях этого тоталитарного мероприятия, о целях и задачах эксперимента, о его методике и прочих атрибутах серьёзного педагогического исследования. Слово «эксперимент» обязывает к основательности подхода к этому важному исследовательскому методу не только в физике или химии, но и в педагогике. Вместо этого авторы идеи ЕГЭ и поддержавшие их администраторы образования как загипнотизированные пошли по следам американских руководителей образования из администрации Буша с их финальным тестом по математике и английскому языку и стали внедрять своё наспех подготовленное создание в широкую российскую практику. Апологеты ЕГЭ пошли дальше своих американских наставников. Если в США финальный тест — это прерогатива штата, то у нас — масштабно, на всю Россию, если в Штатах — только по математике и английскому, то у нас — по всем предметам!

Претенциозная, чисто политическая программа администрации Буша «NO CHILD

LEFT BEHIND (NCLB)» исходит из ложной, бизнес-подобной идеи о том, что стоит предъявить требование (тест) и показать в одной руке кончик доллара, а в другой — кончик кнута, как уровень образования в стране немедленно возрастёт до небес. Надо абсолютно не представлять себе природу образовательного процесса, чтобы поверить в подобную выдумку. Но наши образовательные администраторы, похоже, в неё поверили, несмотря на то, что на страницах американской прессы всё реже акцентируется внимание общественности на пресловутом NCLB и финальном тесте — его порождении. И не случайно также всё чаще сообщается о протестах учащихся, заданных неправомерным финальным тестом. Протестует также самая мощная учительская организация Америки «Национальная ассоциация образования» (NEA).

Очень симптоматично в этом отношении только что появившееся в американской печати сообщение под таким красноречивым заголовком: «**NEA to Challenge «No Child Left Behind**», что означает: «NEA протестует против закона «Не допустить отставания ни одного ребёнка». На состоявшейся 3 июля 2006 года Ассамблее NEA только трое из 9000 делегатов возражали против обра-

щения к правительству об отмене закона. Основной тезис Ассамблеи состоит в том, что «хорошие школы и умелые учителя — это замечательно, но “патологическое” увлечение правительства тестированием успехов учащихся и наказанием провалившихся школ подрывает систему образования...». Президент NEA Рег Вивер высказал мысль о том, что закон подвергнется серьёзному обсуждению сразу после окончания президентского срока Д. Буша, который и породил этот закон.

Как и широковещательные акты прежних президентов о радикальном реформировании американского образования, уйдёт в небытие вместе с Бушем и его крестник NCLB. Причиной этому, как и во всех предшествующих опытах (1983; 2000), будет его неизбежное фиаско ввиду **полной научно-педагогической несостоятельности этого мероприятия.**

Научно несостоятельным, как мы это покажем ниже, является также создание и применение ЕГЭ.

Что собой представляет ЕГЭ?

Исходная идея ЕГЭ была бы прогрессивной, если бы содержала корректную методику объективной оценки качества знаний учащихся и попросту,

Теория

Откуда взялся
ЕГЭ?

ПЕД диагностика
ПЕД диагностика

**Что собой
представляет
ЕГЭ?**

без излишнего фанфаризма, предлагала бы её для применения в школах и вузах. И даже постановление правительства по этому поводу было бы вполне уместным, способствуя прогрессу образования в стране. Почему бы и педагогике вслед за литературой не продекларировать:

**Я хочу, чтоб
к штыку
приравняли перо!**

Введение такой экзаменационной системы дало бы в руки всех, кого интересует качество знаний учащихся, объективный инструмент его выявления, измерения и оценки. И вовсе нет необходимости централизовать этот процесс и посылать тесты и контролёров из центра на места, чтобы замерить качество знаний школьников в любой Тьму-Таракани. Местным учителям, владеющим объективной методикой контроля, можно вполне доверить это рутинное мероприятие, и они выполнят его лучше столичных курьеров. Главное, чтобы они не боялись показать истинную картину в собственном классе. Ведь не посылаем же мы гири и весовщиков в Тьму-Таракань для взвешивания получаемого оттуда груза. Когда груз прибывает к нам, мы взвешиваем его эквивалентными гирями и на подобных весах. Почему же для проверки зна-

ний учащихся надо и гири, и весы везти из центра, чтобы взвесить качество знаний учащихся? Пусть эти знания взвешивают на местах, а когда учащиеся станут приезжать, например, в вузы, пусть уже там, если хотят, проверяют правильность «взвешивания» знаний, но обязательно с помощью таких же гирь и таких же весов. И всё будет просто и буднично, без колокольного звона и милицейского оцепления.

О раскрытии содержания ЕГЭ позаботились методисты практически всех учебных предметов: опубликованы так называемые кодификаторы элементов содержания предметов, сборники примерных тестов для экзамена, баллы, шкалы, оценки, проценты и множество постановлений, приказов и директив. **Но все вместе они не могут заменить одной-единственной отсутствующей до сих пор странички с изложением педагогических основ ЕГЭ**, а этим начинаются и этим кончаются все его беды.

Вероятно поэтому, **и это самое важное**, в разработке контрольных измерительных материалов для экзамена (КИМов) в применяемых шкалах оценки и их интерпретации многочисленные авторы демонстрируют полный разноречивый из-за отсутствия исходных научно-педагогических принципов постро-

ения тестовой лестницы, способов конструирования тестов и методов оценки результата. Разве можно считать научно-педагогическим обоснованием тестов такое утверждение:

Экзамен состоит из тестов трёх уровней, которые делятся на:

Часть 1 — задания обязательного уровня сложности.

Часть 2 — задания повышенного уровня сложности.

Часть 3 — задания высокого уровня сложности.

И понимаете эти «части» и «уровни сложности» как хотите! Для экзамена, претендующего на объективность, да ещё в государственном масштабе — это полное разрушение такой претензии в самом его исходном пункте. Предлагаемая классификация тестов подобна мнимой классификации тестов, принятой в SAT («лёгкие, средней трудности и трудные»), со всеми её негативными педагогическими следствиями. В системе SAT тесты различаются не по сложности, а по трудности. Путать эти категории непозволительно, как нельзя путать длину и глубину: сложность — это по ступени абстракции, а трудность — по уровню усвоения. В системе SAT различение категорий трудности теста не имеет критерия, но хотя бы соотносится с числом провалившихся на тесте студентов. Тест при таком его различении будет ме-

нять свою трудность или, что то же самое, свою измерительную объективность при встрече с разными по подготовленности студентами. Можно себе представить, что бы творилось в нашей жизни, если бы вес гири постоянно менялся при взвешивании разных по природе грузов. А вот в образовании США всё тихо и всё меряют непонятно чем, непонятно как и получают никто не знает что. Однако все довольны и все дружно сетуют на низкое качество образования. В Америке это выгодно: напугать очередного президента плохим качеством образования, подсунуть ему очередной широчувательный лозунг и откачать из бюджета миллиарды долларов на образовательную бедность.

В ЕГЭ даже и этого нет. Просто называйте тест «сложным» и меряйте!

Не проясняет картины и следующее «научное» разъяснение различия в тестах.

Вопросы и задания состоят из трёх частей (выше было «три уровня»):

а) тестовые — т.е. задания с выбором варианта ответа из четырёх предложенных (это, надо думать, «обязательный уровень сложности»);

б) задания с кратким свободным односложным (одно слово или число) ответом (это, надо думать, «повышенный

Теория

Что собой представляет ЕГЭ?

ПЕД диагностика
ПЕД диагностика

**Что собой
представляет
ЕГЭ?**

уровень сложности» и уже не тестовые, а какие?);

в) задания с развёрнутым обоснованным ответом (задания «высокого уровня сложности» и тоже не тестовые?).

Эта «педагогика» напоминает популярные в советском заочном («заушном», как называл его Аркадий Райкин) образовании анекдотичные критерии оценки знаний студентов-заочников:

— явился на экзамен — 3 балла;
— открыл рот — 4 балла;
— произнёс звук! — 5 баллов.

Ниже будет показано, что ЕГЭ недалеко ушёл от этих «критериев». Но вернёмся к приведённым выше «определениям» сложности тестов.

Во-первых, не «задания состоят из трёх частей», а тест-лестница состоит из трёх частей. И чем же отличаются эти части? Судя по многим примерам из тестовых сборников, по сути, ничем. Например, вот задания по математике из «Типового варианта ЕГЭ»:

Часть 1

Упростите выражение
 $4 - 5\sin^2 x - 5\cos^2 x$.

Варианты ответов: 1) 9; 2) -1; 3) 3; 4) 5.

Для чего здесь варианты ответов — непонятно. Выбирать здесь нечего, надо решать и в итоге получить число. Следовательно, тест относится ко второй части («повышенной сложности»).

Часть 2

Найдите значение выражения
 $\cos 75^\circ (\sin 100^\circ \cos 25^\circ - \cos 100^\circ \sin 25^\circ)$

Тоже повышенная сложность.

Часть 3

Найдите все значения параметра a , при которых множество решений неравенства (приводится неравенство) содержит число, равное сумме квадратов корней уравнения (приводится уравнение).

То же самое, но позапутаннее: как минимум две задачи в одном тесте.

Все три задачи являются педагогически эквивалентными деятельности второго уровня усвоения, независимо от наличия или отсутствия вариантов ответа, и все формулируются на третьей ступени абстракции. Следовательно, все они равнотрудные и равносложные и никакой лестницы не образуют.

В заданиях по химии в основном используются тесты первого уровня. Вот типовой тест из Части 1:

Взаимодействуют между собой

- 1) этанол и водород;
- 2) уксусная кислота и хлор;
- 3) фенол и оксид меди (II);
- 4) этиленгликоль и хлорид натрия.

А этот из Части 2:

И серная кислота, и гидроксид бария способны реагировать с

- А) гидроксидом калия;
- Б) гидроксидом аммония;
- В) цинком;
- Г) водородом;
- Д) оксидом магния;
- Е) силикатом натрия.

Ответ: _____.

(Запишите соответствующие буквы в алфавитном порядке.)

Здесь нарушено требование простоты теста: две задачи в одном тесте.

И наконец, в Части 3 появляется тест второго уровня:

Даны водные растворы: хлорида железа (III), иодида натрия, бихромата натрия, серной кислоты и гидроксида цезия.

Приведите уравнения четырёх возможных реакций между этими веществами.

Здесь опять нарушено требование простоты теста: много задач в одном тесте.

Вот примеры тестов по литературе.

Часть 1:

Определите жанр произведения, из которого взят фрагмент.

- 1) Роман.
- 2) Притча.
- 3) Очерк.
- 4) Рассказ.

Этот из Части 2:

Какому традиционному жанру классической поэзии близко стихотворение С. Есенина «Спит ковыль. Равнина дорогая...»?

- 1) Ода.
- 2) Элегия.
- 3) Баллада.
- 4) Сонет.

А этот из Части 3:

Каково авторское определение жанра пьесы А.С. Грибоедова «Горе от ума»?

- 1) Трагедия.
- 2) Драма.
- 3) Комедия.
- 4) Водевиль.

Все тесты первого уровня построены на содержании второй ступени абстракции. Не потому ли наши учащиеся так «любят» уроки литературы?

Такое положение вещей объясняется тем, что при разъяснении различий в тестах акцент сделан на их внешней форме, а не на содержании деятельности учащегося по решению теста, а поэтому неудивителен полный разноречивый полный разноречивый в конструкции КИМов.

Различение тестов только по их внешней форме ведёт к ошибкам и созданию псевдо-тестов, когда по форме это тест одного уровня, а по содержанию — другого. Такими ошибками изобилуют КИМы всех предметов ЕГЭ.

Теория

Что собой представляет ЕГЭ?

ПЕД диагностика
ПЕД диагностика

**Что собой
представляет
ЕГЭ?**

Так, в КИМах по физике, как и по математике, все три части состоят из тестов второго уровня, имеются ли в них варианты ответов или нет, требуется ли «односложный» или «развёрнутый» ответ. В то же время КИМы по химии во всех их частях, несмотря на все ухищрения авторов, тяготеют к тестам первого уровня. Тесты по литературе — это наиболее близкий клон SAT преимущественно с тестами первого уровня.

В ходе ЕГЭ учащиеся набирают баллы за выполненные тесты по абсолютно необоснованной 100-балльной шкале. По элементарным педагогическим понятиям за шкалой должно стоять какое-то содержание измеряемого признака. Например, чем характеризуются знания учащегося при 20 баллах? А при 40? А поскольку баллы в ЕГЭ складываются, то при 40 баллах знаний должно быть в два раза больше, чем при 20. Позже данные 100-балльной шкалы по произвольным критериям соотносятся с привычной, но столь же произвольной 5-балльной шкалой. Это соотношение производится таким «хитрым» образом, что мы снова возвращаемся к пресловутым «нормам оценки» и мнимой «высокой» (90% — 100%) успеваемости учащихся. К чему же было весь этот

огород городить, чтобы наконец-то прийти к той же субъективной 5-балльной оценке, которую не задумываясь выставил бы сам учитель? Собственно «высокие» предметные комиссии тем только и занимались, что не переводили 100-балльную шкалу в пятибалльную, а подгоняли первую под вторую. В результате, сделав огромный и дорогостоящий пируэт, пришли к тому, от чего ушли. И мы так и не знаем: хорошее или плохое у нас образование в стране. Говорят, «лучшее в мире».

Несмотря на разноречивую «сложность» тестов по разным предметам, экзамен тем не менее приводит к тому необъяснимому факту, что по всем предметам учащимися неизменно достигается в среднем 50-балльный показатель успешности преодоления теста. По всем правилам логики и педагогики средние баллы должны быть выше в тестах с преобладанием тестов «обязательного» уровня и ниже — с преобладанием тестов «повышенного» и «высокого» уровня «сложности». Но этого каким-то чудодейственным образом в ЕГЭ не происходит.

Что от ЕГЭ хотят?

От него хотят совсем «немногого»: чтобы показал реальное

качество общего среднего образования в России и стимулировал бы его совершенствование. Но по всем правилам любого эксперимента, прежде чем его разворачивать, надо договориться о его целях, процедурах и корректных способах оценки результата.

Для ЕГЭ это означает, что прежде чем начинать разработку тестов и проводить экзамены, следовало бы диагностично обозначить цели изучения предметов в целом и их отдельных разделов в терминах теории диагностического целеполагания. Тогда будет ясно, что мы проверяем и какое образование хотим. На сегодняшний день этого не сделано и школа учит, как может, едва достигая по всем предельно перегруженным предметам показателя $K1 \sim 0,2-0,3$. А тесты ЕГЭ ищут в школе несуществующий второй уровень усвоения и, конечно же, его не находят, но это хорошо вуалируется суммарным баллом общей оценки теста. Что же мы тогда проверяем тестами ЕГЭ? По-видимому, как и в SAT, проверяется, могут ли учащиеся преодолеть искусственно придуманный тест, которым и замещается реально достижимая цель усвоения предметов. Следовательно, целью обучения становится не вполне определённое усвоение предмета, а преодоление наду-

манного теста. Я говорю «надуманного», поскольку никакой объективной педагогической теории измерения и оценки знаний учащихся за ним не стоит. Кроме того, такой тест, который не направлен на цель обучения, не считается с реальными возможностями обучения. Поясним сказанное. Каждый учитель, готовясь к уроку, формулирует цель этого урока и подбирает дидактические средства для наилучшего достижения этой цели. Затем он проводит запланированный урок и проверяет, достигнута ли цель урока. Во всей этой последовательности действий учителя важнейшим фактором является способ постановки цели: диагностично или недиагностично.

Если цель поставлена диагностично (чего в современной школе нет), то степень её достижения проверяется надёжной батареей содержательно и функционально валидных тестов. В этом случае по результату тестирования выносятся объективное суждение как о качестве знаний учащихся, так и о качестве выполненного учебного процесса.

Если же цель поставлена недиагностично (что и свойственно современной школе), то по качеству выполнения теста учащимися можно только судить о том, как выполнен данный тест, но не каково качест-

Теория

Что от ЕГЭ хотят?

ПЕД диагностика
ПЕД диагностика

Что от ЕГЭ хотят?

во знаний учащихся и насколько качественным был процесс обучения. Это всё равно, что при желании измерить длину забора мы станем измерять высоту столба на противоположной стороне улицы. Ведь учитель всегда и с полным правом может сказать, что поставленная им цель обучения и уровень требований теста не совпадают: он формирует научное мировоззрение учащихся (или их предметное сознание), а знания для него — только средство достижения цели и качество их усвоения по критериям теста его не интересует. Это высказывание учителя совершенно справедливо, потому что обучающие и тестирующие при отсутствии диагностической цели говорят на разных языках о целях обучения. Почему об этом ещё не заявляют учителя, можно легко догадаться: они попросту не понимают, что любое тестирование знаний их учащихся без предварительно диагностично сформулированной цели обучения неправомерно. Кроме того, из-за нелепой секретивности тестирования учитель не может согласовать даже ту размытую цель обучения, которая предполагается предстоящим тестом со своей текущей учебной работой. Как для учителя, так и для учащихся тест чаще всего представляет собой изощрённую ловушку, подготов-

ленную «хитрыми» составителями теста. О том, что ЕГЭ изобилует такими ловушками, говорит тот странный факт, что на составление тестов объявляется конкурс тестоизобретателей. Что же поощряется в конкурсе? Изощрённость ловушек. В то время как при ориентировке на требования содержательной и функциональной валидности теста тесты не изобретаются, а строятся соответственно чёткой методологии тестирования.

Каков выход из положения для загнанных в ловушки учителя и ученика?

В лучшем случае отставить систематическое изучение предмета и натаскивать учащихся на прототипы теста, а в худшем — ловчить уже получившими известность нелегальными способами, благо рекламы этих способов хоть отбавляй.

Что ЕГЭ может дать?

Тест ЕГЭ может только сообщить его авторам, как он сам преодолевается учащимися, но никакой существенной информации о качестве учебного процесса он сообщить не может по следующим причинам:

а) поскольку он не опирается на какие-либо чётко сформулированные критерии качества обучения, с которыми можно сопоставить результаты

тестирования. Деление невесть откуда взятой 100-балльной шкалы на четыре произвольных интервала (произвольных ещё и в каждом предмете), чтобы иметь все оценки пятибалльной шкалы, извращает само понятие «качество обучения»;

б) тестирование является предельно ненадёжным по объёму выборки;

в) оценка ответов испытуемых осуществляется по конечному результату часто многооперационных задач, не отражая процесса их решения.

Но если цель не поставлена диагностично, то можно проводить любые манипуляции с полученными результатами тестирования, чтобы показать желаемую, а не действительную картину качества образования в стране. Знакомясь с результатами тестирования и их интерпретацией, у меня возникло твёрдое мнение, что вся затея ЕГЭ и созданные для её реализации федеральные институты, центры и службы только тем и занимаются, что вуалируют действительность. Из чего это видно? В первую очередь из принятых шкал перевода балльных сумм в пятибалльную шкалу оценки (см. результаты ЕГЭ за 2004–2006 гг.). Как объяснить тот странный факт, что каждый предмет (это при едином-то экзамене!) имеет свою шкалу

перевода баллов в пятибалльную шкалу, изобретаемую и утверждаемую предметной комиссией(!), а не единственно закономерную, выводимую из естественных этапов процесса обучения? И что это за шкала, которая варьирует для одного и того же предмета в разных классах и в разные годы эксперимента? Ведь шкала — это отображение объективного критерия и сопряжённой с ним процедуры измерения. Это стандарт, который применяется во всех случаях однородных измерений. Как же его можно, даже по велению высокой комиссии, «гнуть» и «мять» от предмета к предмету и внутри предметов? Представьте себе, что перевод фунтов в килограммы варьировал бы при переезде из Америки в Россию и в России взвешивали бы другими килограммами по приказу какого-либо министерства! Этого, всем понятно, делать нельзя, а в образовании, как видите, можно.

При анализе результатов тестирования, как уже подчёркивалось, обнаруживается не менее странный факт: почему-то все учащиеся по всем предметам дают одну и ту же успеваемость, характеризуемую средним баллом в 50 единиц (?), хотя тесты по разным предметам имеют разную трудность (условно от первого до третьего уровня) и выполняются

Теория

Что ЕГЭ может дать?

ПЕД диагностика
ПЕД диагностика

**Что ЕГЭ может
дать?**

они с разным успехом. По-видимому, манипуляции со шкалированием позволяют получить любой средний балл. Так, даже по алгебре, где максимальное число баллов, которое может набрать учащийся, равно 37, а пятёрка ставится в интервале 19–30 баллов, средний балл равен 49,89, а 148 учащихся набрали 100 баллов!

Но это не единственная загадка и арифметический парадокс. По химии, к примеру, можно набрать максимально 67 баллов, но отдельные учащиеся, как нас уверяют отчёты, набирают по 100 баллов. Вот уж действительно ХИМИЯ! Не отстают от них и литераторы: у них 100 баллов набирают из 81 возможных. Ну просто Ворошиловские стрелки!

Почти в дюжине предметов, по которым проводится ЕГЭ, по-разному конструируются тесты одного и того же «уровня сложности», по-разному начисляются баллы за выполненные тесты и по-разному выставляются оценки. В опубликованных в разное время таблицах перевода из 100-балльной в 5-балльную шкалу существенно варьируют интервалы начисления оценок, да и интервалы выбраны совершенно произвольно, без какой бы то ни было опоры на имеющиеся психолого-педагогические критерии качества усвоения. Так, только в 2006 году

Федеральным центром тестирования опубликованы **три** переводных таблицы:

- *общая*, в которой, к примеру, *по русскому языку*, испытуемому ставится *двойка*, если он набрал в тестах менее 30 баллов, *тройка* — при 31 балле, *четвёрка* уже при 50 баллах, а *пятёрка* при 67 баллах (в 2004 году было соответственно: 31, 32, 51 и 70 баллов);
- *в другой шкале по русскому языку (для 9-го класса)* двойка выставляется при менее 12 баллах, тройка — при 13 баллах, а четвёрка уже при 28 баллах и пятёрка при 53 баллах;
- *шкала для 11-го класса* соответственно: двойка — при менее 13, тройка — при 14, четвёрка — уже при 34, а пятёрка — при 64 баллах.

Что это? Разный русский язык в разных классах или левая рука не знает, что пишет правая? Кроме того, что это за пятёрка по русскому (родному!) языку, допускающая 36% ошибок у выпускника средней школы? И что это за «зрелость» выпускника школы, у которого по русскому языку тройка? Он получит её, если лишь в 13% случаев в состоянии правильно пользоваться грамотой родного языка! Четвёрку учащийся получит, если будет ошибаться и спотыкаться в 66% случаев, а пятёрку в 9-м классе — в 47% и пятёрку в 11-м классе,

если допустит 36% возможных ошибок!

Учитывая 25-процентную вероятность случайного угадывания в тестах первого уровня («обязательного уровня сложности»), даже шимпанзе может оказаться грамотеем! Но это будет слишком поспешным суждением, если не познакомиться с самим тестом по русскому языку. Демонстрационный вариант его опубликован в Интернете. Он **подготовлен Федеральным государственным научным учреждением «Федеральный институт педагогических измерений»**. **Согласован с Научно-методическим советом ФИПИ по русскому языку и утверждён руководителем Федеральной службы по надзору в сфере образования и науки**. И что же представляет собой этот тест? Является ли он тестом **по русскому языку**, т.е. предмету, который призван научить учащихся **правописанию** на родном языке? Нет, тест выходит далеко за рамки своего предмета и глубоко вторгается на территорию, занятую в учебном плане предметом под названием *Литература*. Авторы теста по русскому языку поднимаются до таких высот, которые достижимы разве что студенту филологического, литературного или другого специального литературного творческого института. Чтобы не быть голослов-

ным, приведу пример теста так называемого обязательного уровня сложности:

Прочитайте текст и выполните задания А7–А12.

(1)... (2) Вся другая информация — как звуки, так и изображения — для обработки на компьютере должна быть преобразована в числовую форму. (3) Аналогичным образом на компьютере обрабатывается и текстовая информация: при вводе в компьютер каждая буква кодируется определённым числом, а при переводе на внешние устройства по этим числам строятся соответствующие изображения букв. (4) Это соответствие между набором букв и числами называется кодировкой символов. (5) Все числа в компьютере представляются с помощью нулей и единиц, а не десяти цифр, как это привычно для людей. (6)...

А7 *Какое из приведённых ниже предложений должно быть **первым** в этом тексте?*

- 1) Персональные компьютеры — это универсальные устройства для обработки информации.
- 2) Компьютер может обрабатывать только информацию, представленную в числовой форме.

Теория

Что ЕГЭ может дать?

ПЕД диагностика
ПЕД диагностика

**Что ЕГЭ может
дать?**

3) Вся информация, предназначенная для долговременного пользования, хранится в файлах.

4) Информация в компьютере хранится в памяти или на различных носителях, например на гибких и жёстких дисках.

A8 *Какое из приведённых ниже предложений должно быть шестым в этом тексте?*

1) Прежде всего компьютеры обычно работают в двоичной системе счисления.

2) Однако компьютеры обычно работают в двоичной системе счисления.

3) Кроме того, компьютеры обычно работают в двоичной системе счисления.

4) Иными словами, компьютеры обычно работают в двоичной системе счисления.

A9 *Какие слова являются грамматической основой во втором (2) предложении текста?*

1) Информация — звуки, изображения.

2) Информация должна.

3) Информация должна быть преобразована.

4) Информация преобразована.

A10 *Укажите верную характеристику второго (2) предложения текста.*

1) Простое.

2) Сложносочинённое.

3) Сложное бессоюзное.

4) Сложноподчинённое.

A11 *Укажите правильную морфологическую характеристику слова ПРЕОБРАЗОВАНА из второго (2) предложения текста.*

1) Действительное причастие.

2) Страдательное причастие.

3) Прилагательное

4) Деепричастие.

A12 *Укажите значение слова КОДИРУЕТСЯ в предложении 3.*

1) Воспроизводится в определённой последовательности.

2) Постоянно повторяется.

3) Записывается в виде текста.

4) Преобразуется с помощью определённых условных обозначений.

Что это? Тест по подготовке редактора для издательства или проверка знаний школьника по правописанию?

Дальше — лучше: в единственном тесте «высокой сложности» учащемуся уже прямо предлагается написать сочинение не менее чем в 150 слов как рецензию на одну страничку из К.Г. Паустовского, состоящую из 200 слов. Чтобы понять, что экзаменаторы (так и хочется написать экзекуторы) хотят от учащегося, достаточно познакомиться с критериями оценки этого сочинения.

№	Критерии оценивания ответа на задание С1	Баллы
I	Содержание сочинения	
К 1	Формулировка проблем исходного текста	
	Экзаменуемый (в той или иной форме) верно сформулировал одну из проблем исходного текста. Фактических ошибок, связанных с пониманием и формулировкой проблемы, нет	1
	Экзаменуемый не смог верно сформулировать ни одну из проблем исходного текста	0
К 2	Комментарий к сформулированной проблеме исходного текста	
	Сформулированная экзаменуемым проблема прокомментирована. Фактических ошибок, связанных с пониманием исходного текста, в комментариях нет.	2
	Сформулированная экзаменуемым проблема прокомментирована. Фактических ошибок, связанных с пониманием исходного текста, в комментариях нет, но допущена 1 фактическая ошибка в комментариях, связанных с пониманием проблемы исходного текста.	1
	Сформулированная экзаменуемым проблема не прокомментирована, или допущено более 1 фактической ошибки в комментариях, связанных с пониманием проблематики текста, или прокомментирована другая, не сформулированная экзаменуемым проблема, или в качестве комментариев дан простой пересказ текста или его фрагмента, или в качестве комментариев цитируется большой фрагмент исходного текста	0
К 3	Отражение позиции автора исходного текста	
	Экзаменуемый верно сформулировал позицию автора исходного текста по прокомментированной проблеме.	1

Теория

ПЕД диагностика
ПЕД диагностика

	Фактических ошибок, связанных с пониманием позиции автора исходного текста, нет	
	Позиция автора исходного текста экзаменуемым сформулирована неверно, или позиция автора исходного текста не сформулирована	0
К 4	Изложение экзаменуемым собственного мнения по проблеме	
	Экзаменуемый этически корректно сформулировал своё мнение по одной из проблем, поставленных автором текста (согласившись или не согласившись с позицией автора), аргументировал его (привёл не менее 3 аргументов), опираясь на жизненный или читательский опыт. Фактические ошибки в фоновом материале отсутствуют	3
	Экзаменуемый этически корректно сформулировал своё мнение по одной из проблем, поставленным автором текста (согласившись или не согласившись с позицией автора), аргументировал его (привёл 1–2 аргумента), опираясь на жизненный или читательский опыт. Фактические ошибки в фоновом материале отсутствуют	2
	Экзаменуемый сформулировал своё мнение о проблеме, поставленной автором текста (согласившись или не согласившись с позицией автора), но не привёл аргументы, и/или экзаменуемый сформулировал своё мнение по одной из проблем, поставленных автором текста (согласившись или не согласившись с позицией автора), этически некорректно, и/или допущено не более 1 фактической ошибки в фоновом материале	1
	Мнение экзаменуемого лишь формально заявлено (например: «Я согласен/не согласен с автором»), или вообще не отражено в работе, и/или допущено более 1 фактической ошибки в фоновом материале	0

II	Речевое оформление сочинения	
К 5	Смысловая цельность, речевая связность и последовательность изложения	
	Работа экзаменуемого характеризуется смысловой цельностью, речевой связностью и последовательностью изложения: – логические ошибки отсутствуют; – в работе нет нарушений абзацного членения текста	3
	Работа экзаменуемого характеризуется смысловой цельностью, связностью и последовательностью изложения, но – допущено не более 1 логической ошибки; – в работе имеется 1 нарушение абзацного членения текста	2
	В работе экзаменуемого просматривается коммуникативный замысел, но имеется не более 2 логических ошибок, и/или имеются 2 случая нарушения абзацного членения текста	1
	В работе экзаменуемого просматривается коммуникативный замысел, но имеется более 2 логических ошибок, и/или имеется более 2 случаев нарушения абзацного членения текста	0
К 6	Точность и выразительность речи	
	Работа экзаменуемого характеризуется точностью выражения мысли, разнообразием грамматических форм, уместным использованием средств выразительности	2
	Работа экзаменуемого характеризуется точностью выражения мысли, разнообразием грамматических форм, но средства выразительности используются неуместно	1
	Работа экзаменуемого отличается бедностью словаря, однообразием грамматического строя речи и неуместным использованием средств выразительности. Средства выразительности не используются	0

Теория

ПЕД диагностика
ПЕД диагностика

Ш	Грамотность	
К 7	Соблюдение орфографических норм	
	орфографических ошибок нет (или 1 негрубая ошибка)	2
	не более 1 ошибки	1
	более 1 ошибки	0
К 8	Соблюдение пунктуационных норм	
	пунктуационных ошибок нет (или 1 негрубая ошибка)	2
	не более 2 ошибок	1
	более 2 ошибок	0
К 9	Соблюдение языковых норм	
	грамматических ошибок нет	2
	не более 2 ошибок	1
	более 2 ошибок	0
К 10	Соблюдение речевых норм	
	допущено не более 1 речевой ошибки	2
	не более 3 ошибок	1
	более 3 ошибок	0
Максимальное количество баллов за всю письменную работу (К 1 – К 10)		20

И далее:

При оценке грамотности (К 7–К 10) следует учитывать объём сочинения. Указанные в таблице нормативы оценки разработаны для сочинения объёмом в 150–300 слов.

Если объём сочинения в два раза меньше нижнего предела нормы (т.е. в сочинении менее 70 слов), то такая работа не засчитывается вовсе (0 баллов), задание считается **невыполненным**.

При оценке сочинения объёмом от 70 до 150 слов число допустимых ошибок всех четырёх видов **уменьшается**: 2 балла по критерию К 7 ставится при отсутствии ор-

фографических ошибок (или если экзаменуемый допустил 1 негрубую ошибку); 1 балл по критериям К 8 – К 10 может быть поставлен, если в работе по каждому из этих критериев допущено не более 1 ошибки.

2 балла по критериям К 7 – К 10 за работу объёмом от 70 до 150 слов вообще не ставится.

Если сочинение представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, за такую работу по всем аспектам проверки (К 1 – К 10) ставится 0 баллов.

И никакого эталона! Как же применять эти критерии при отсутствии эталона? О какой объективности оценки может идти речь?

К примеру, какой нелепостью является критерий: «Если объём сочинения в два раза меньше нижнего предела нормы (т.е. в сочинении менее 70 слов), то такая работа не засчитывается вовсе (оценивается нулём баллов), задание считается **невыполненным**».

А как же быть с известным утверждением, что «краткость — сестра таланта»? Сколько талантов будет погребено под обломками этого диковинного теста по русскому языку?

Удивительно ли, что учащиеся показывают удручающе низкую успеваемость по этому содержательно инвалидному тесту по русскому языку на ЕГЭ (93% фактически не усвоенного материала). Под видом проверки правописания в русском языке тестируются их публицистические и литературные способности, развитие которых этим предметом не предусмотрено. И не только содержательно не валидное тестирование. Оно и функционально не валидное: приведённые примеры тестов уже дотягивают до третьего-четвёртого уровней (Эксперт — Творец), хотя и содержат варианты ответов для выбора. Каков источ-

ник этих ошибок в подготовке тестов по русскому языку? Источник налицо: это знаменитый TOEFL, педагогическая дефектность которого очевидна.

Итак, хотелось бы отметить тот формализм, с которым авторы тестов по различным предметам стремятся во что бы то ни стало соблюсти те формы тестов, которые им были заданы по инструкции: «с выбором ответа из четырёх вариантов, с однозначным ответом, с развёрнутым ответом». Эти формальные, а не содержательные требования связывают и сковывают авторов тестов часто неприсущими конкретному предмету формами. Так, по русскому языку от века устоявшейся формой контроля усвоения правописания были диктант или изложение, но никогда — сочинение. В рамках полчасового диктанта или изложения, которые являются прекрасным тестом второго уровня, можно исчерпывающе и надёжно проверить успешность усвоения учащимся правописания в общеупотребительном, не изощённо усложнённом языке — русском или иностранном.

Предлагаемый же в ЕГЭ тест по английскому языку ни в какое сравнение с TOEFL не идёт: он намного, с большим отрывом превзошёл его по нарушению всех требований

Теория

Что ЕГЭ может дать?

ПЕД диагностика
ПЕД диагностика

Что ЕГЭ может дать?

валидного построения тестов. Тест английского языка не валиден ни содержательно, ни функционально, в нём, как и в русском языке, за пределами нарушено требование простоты теста, он чудовищно перегружен искусственно усложнённым текстом и проверяет больше качества памяти, внимания, аналитико-синтетических способностей, чем рутинное знание обыденного английского языка в объёме и номенклатуре, требуемого от иностранца. Не поэтому ли так убога успеваемость учащихся по английскому языку по результатам ЕГЭ?

Та же картина произвола в построении тестов и в оценке знаний учащихся характерна и для других предметов. Так, двойка по географии по разным шкалам может быть выставлена соответственно, если число баллов менее 35, 22 или 17. Дешевеет география! По химии: в разных шкалах при 30, 27 и 24 баллах учащийся получит двойку. Химики колеблются: у них в разных классах химия то «дешевле», то «дороже», как и география. Шкала для перевода 100-балльной оценки в 5-балльную для иностранного языка опубликована только в общей шкале. Интересно, что требования к знанию иностранного языка намного выше, чем для знания родного. Для получения пятёр-

ки по русскому языку, даже по самой строгой, общей шкале достаточно набрать 67 баллов, а по иностранному — не меньше 89 баллов. Это по постановлению предметной комиссии Министерства образования, которая игнорирует опубликованные Федеральным центром тестирования шкалы! Кстати, ни одна министерская комиссия ни по одному предмету не выдвигает столь высоких требований для пятёрки. Даже «королева» школьного учебного плана — математика — довольствуется в общей шкале 72 (комиссия постановила — 75) баллами для присвоения знаниям испытуемого пятёрки.

Как объяснить такой размах различий в конструкциях тестов и критериях оценки, который характерен для всех учебных предметов? Ведь не могли же создатели ЕГЭ не видеть этих бросающихся в глаза нелепостей? Думаю, что все эти различия сделаны преднамеренно, чтобы подогнать результаты тестирования под пятибалльную шкалу оценки таким образом, чтобы не исказить традиционную министерскую отчётность о состоянии образования в стране! После такой подгонки появились и тройки, и четвёрки, и пятёрки в соотношениях, соответствующих традиционной «наробразовской»

отчётности. Но если в традиционной наробразовской отчётности никто и никогда не понимал, что такое тройка и что такое пятёрка, удивляясь, в частности, что медалисты проваливаются на экзаменах в вузы, то ЕГЭ при всей его педагогической дефективности всё же показал истинную цену школьной оценки и стопроцентной успеваемости. И в этом факте его единственная полезность для образования. Но захотят ли её заметить и правильно её использовать? Этот вопрос возникает по той простой причине, что многочисленные экзаменаторы и не менее многочисленный персонал Министерства образования и науки, Федеральных институтов, центров и служб не обратили внимания на тот краеугольный психолого-педагогический факт, что общий средний балл по всем предметам, колеблющийся вокруг **50**, означает, что по существу **ни один учебный предмет вообще не усвоен учащимися!** И знания практически всех учащихся (за исключением, может быть, «верхних» 3–5%) по всем предметам должны быть оценены двойкой. Опираясь на изложенную в предшествующих трёх лекциях педагогическую теорию измерения и оценки знаний учащихся, очевиден вывод: **по всем школьным предметам учащиеся**

остаются недоучками. Средний балл 50 означает, что $K\alpha = 0,5$, тогда как минимальный критерий завершенности обучения — это $K\alpha = 0,7$. От 70 баллов только и должна начинаться шкала перевода оценки. И не суммарно, а по каждому уровню отдельно: менее 70% баллов — двойка! Пятёрка — при более 90% баллов.

Ведь только после 70% усвоения предмета на данном уровне усвоения можно говорить о последующем устойчивом совершенствовании деятельности (на данном уровне). Заканчивая обучение до значения 0,7, мы обрекаем учащегося на выполнение изученной им деятельности в будущем с систематическими и неистребимыми ошибками и навсегда оставляем его безграмотным в данном предмете и на данном уровне! Таким образом, сфабрикованные шкалы перевода баллов в оценки, хотя и утверждены в высоких инстанциях различными комиссиями (Федеральная служба по надзору в сфере образования и науки), своими тройками-пятёрками искусственно вуалируют действительный катастрофически низкий уровень успеваемости в современной общеобразовательной школе.

Насколько это плохо? Ровно настолько, насколько мы в состоянии разобраться в кар-

Теория

Что ЕГЭ может дать?

ПЕД диагностика
ПЕД диагностика

Что ЕГЭ может дать?

динальном и кажущемся самоочевидным вопросе, на который пока нет ответа: *Что такое хорошее и что такое плохое образование?* Ведь вполне возможно, что кто-то из ответственных за образование господ скажет, что именно такое образование нам и нужно «на данном этапе его развития». Если же приведённые выше данные будут восприняты как тревога, то предстоит ответить на два следующих естественных вопроса: *Кто в этом виноват? И что делать?*

Требуется также ответ на важнейший вопрос, *кто эти 5–7% учащихся, которые умудряются выбираться из дебрей ЕГЭ и получать мифическую пятёрку?* Ответ на этот вопрос может перевернуть всю традиционную парадигму структуры и содержания образования в стране.

Для планирования и проведения ЕГЭ (если он будет сохранён) важно учитывать принципиальные возможности традиционного учебного процесса, общепринятого в современной общеобразовательной школе. Между требованиями хитроумно изобретённых тестов и возможностями традиционного учебного процесса образуется если не пропасть, то существенный разрыв, и этот разрыв только и обнаруживает тест. Самые высокие принципиальные возможности традиционного

обучения — это α_1 и $K\alpha = 0,3–0,5$, тогда как во многих предметах тесты ЕГЭ требуют от учащегося деятельности на втором уровне (α_2). Откуда ему взяться? Но это не смущает авторов ЕГЭ: получая низкие баллы за разные тесты, они суммируют их при полном отсутствии аддитивности и затем делят искусственно созданный столбальный интервал совершенно механистически на четыре почти равные части (см. «Шкалы переводов...») и ставят тройки и четвёрки там, где по всем педагогическим критериям должны стоять прочные нули. Пятёрки ставятся там, где должны стоять прочные двойки и в редких случаях тройки. Что это? Новый этап пресловутой процентомании или выражение педагогического невежества? Если не невежества, то **кого мы обманываем и зачем?** Пусть этот вопрос проанализирует Федеральная служба по надзору в сфере образования и науки — это её епархия, а выход из положения может подсказать только педагогическая наука.

Что с ЕГЭ делать?

Скажу сразу, не тая: немедленно его прекратить!

Эта рекомендация следует из всего сказанного выше, поскольку так называемый экс-

перимент с ЕГЭ показал следующие его неисправимые недостатки:

а) эмпиризм и противоречивость его исходных педагогических принципов;

б) неинформативность его показателей;

в) громоздкость и сложность его осуществления;

г) неоправданно высокую стоимость проведения экзамена;

д) моральную дефектность его форм: секретивность, централизацию, диктатуру.

Но чем тогда его заменить, чтобы удовлетворить здоровый интерес нашего Министерства образования в достоверной и оперативной информации о состоянии образования в стране? Заменить его можно очень понятной, простой и дешёвой процедурой, которая состоит из трёх элементов. Состав этой процедуры следующий:

1) надо, наконец, диагностично определить цели общего среднего образования как в целом (на выходе педагогической системы), так и на этапах обучения (по предметам); бесцельное образование не может быть ни хорошим, ни плохим — оно просто никакое;

2) утвердить сопряжённую с целью критериально-ориентированную систему оценки успехов учащихся;

3) разработать валидные тесты успешности усвоения по

всем предметам и по всем УЭ каждого предмета.

Весь этот материал затем передать школам для практического использования.

После того как учителя школы и профессора вузов будут вооружены названным выше инструментарием, они могут и должны проводить диагностику качества обучения и периодически сообщать о положении дел в министерство, не боясь при этом никаких санкций и репрессий, даже если результаты обучения будут близкими к нулю. Применив к полученному массиву данных соответствующие статистические процедуры, министерство получит наиболее точную информацию о состоянии образования в стране, а используя научно-педагогические методологии, сумеет управлять этим состоянием в правильном направлении.

При этом нет никакой необходимости тестировать учащихся по трём уровням (хотя никаких уровней во многих КИМах и не просматривается) и применять суммарную 100-балльную оценку, где никакой аддитивностью и не пахнет. В ЕГЭ эту грубейшую ошибку делают все, включая математиков, физиков, химиков, которым по статусу следовало бы обратить внимание на эту нелепость. В отсутствие аддитивности уровневых оценок ито-

Теория

Что с ЕГЭ делать?

ПЕД диагностика
ПЕД диагностика

Что с ЕГЭ делать?

говое тестирование, каким является ЕГЭ, и итоговая оценка знаний является фикцией. Итоговый контроль знаний должен проводиться только по тому уровню, который является целью обучения. Для чего, к примеру, тестировать по первому уровню, если цель обучения второй уровень? То же самое можно сказать относительно тестирования по третьему уровню, если цель обучения — первый или второй уровень. Избыточная информация здесь уже ни к чему, поскольку ничего уже исправить нельзя. Кроме того, целенаправленный экзамен сократил бы в три раза затраты времени, денег и нервов ученика.

Для устранения этих несуразностей разработчикам не следует идти по пути американских эмпириков и создавать тесты и методы их обработки на конкурсной основе (кто остроумнее запутает учащихся!), а воспользоваться теорией педагогических измерений, достаточно хорошо проработанной в отечественной педагогической науке как базой будущих реформистских проектов.

И наконец, последнее — совершенно нелепа секретивная процедура контроля качества усвоения знаний учащимися. В ней нет абсолютно никакой необходимости, если строить объективный контроль

усвоения содержательно и функционально валидными тестами. Демократизм образования состоит в том, что учащемуся предоставляется не только содержание некоторого предмета для изучения, но и средства (это могут быть тесты) самоконтроля качества его усвоения. Эти средства должны быть открыты для всех, кто пожелает ими воспользоваться: для учащихся, учителей, родителей, вузовских приёмных экзаменаторов — всех, кто соблюдает содержательную и функциональную валидность их применения. Нынешняя атмосфера секретности и таинственности искусственно нагнетается с подачи пресловутой частной корпорации ETS, которая таким путём охраняет свои огромные барыши на одурачивании своих педагогически невежественных клиентов. Но в России ЕГЭ принадлежит государству и нагнетание всё той же туманной атмосферы непредсказуемости и репрессивности экзамена может быть объяснено только стремлением людей, имеющих отношение к этой сомнительной затее, таким путём самоутвердиться и придать себе иллюзорный ореол значимости.

И ещё одно, чисто этическое замечание. ЕГЭ оскорбляет безответного российского учителя и его учащихся недоверием: экзамен готовят, прини-

мают и оценивают не те, кто учит. Но, как показывает первый опыт ЕГЭ, сразу же буйным чертополохом вокруг столь тайного экзамена расцвели коррупция и криминал, не оставляющие камня на камне от мнимой засекреченности этого экзамена. С этим можно справиться только открыв все карты и сделав оценку знаний учащихся рутинной работой самого учителя, которому надо и доверять, и, конечно, иногда проверять.

И наконец, последнее, что немаловажно учитывать в сегодняшней российской криминальной обстановке — это возможность покупки любого аттестата или диплома, от школьного до докторского, и необходимость поэтому постоянной перепроверки реальности того образования, на которое претендует предъявитель диплома. Для этого надо располагать простой, надёжной и объективной методикой диагностики качества знаний и умений абитуриента, претендующего на поступление в вуз или на некоторую должность, такой методикой диагностики мастерства, которой легко могут воспользоваться как школьный учитель, так и кадровик любого предприятия или учреждения, включая государственное, для которых подлинность образования не пустяк.

И конечно, преждевременно вводить автоматическое зачисление в вуз по результатам ЕГЭ, поскольку эти результаты, во-первых, ненадёжны, а во-вторых, с большой вероятностью могут быть фальшивы.

Ну а что будет, если руководители образования не отменят ЕГЭ, а будут упорно его «совершенствовать» и далее пугать им школы, учителей и учащихся? **Не изменится в нашем народном просвещении ровным счётом ничего**, поскольку, как показал тот же ЕГЭ, хуже ему уже становится некуда и будет ли ЕГЭ или его не будет — ничего в образовании не изменится. Просто за школу обидно: задержали её, бедную, разные самородки-«инноваторы», вместо того чтобы, опираясь на современную психолого-педагогическую науку, провести полноценную её реформу на уровне XXI века.

Тесты, экзамены и демократия

Вопрос состоит в том, в каком духе мы хотим воспитывать наших детей: в духе демократии или тоталитаризма и диктатуры. Мы только что одной ногой вышли из тоталитаризма и о его пережитки всё ещё спотыкаемся на каждом шагу. От тоталитарного режима не избавилась и школа. Его выражением является и специфическое по-

Теория

Что с ЕГЭ делать?

ПЕД диагностика
ПЕД диагностика

**Тесты,
экзамены
и демократия**

нимание «образовательного стандарта», и «единого образовательного пространства», и введение централизованного единого государственного экзамена. Все эти категории при различном их понимании могут оказаться либо тоталитарными оковами и репрессивными средствами, либо демократическими свободами и средствами, стимулирующими развитие. О категориях образовательного стандарта и единого образовательного пространства речь пойдёт позже, а вот об условиях демократичного использования тестов поговорим сейчас.

Демократичное использование тестов возможно только при определённых условиях, и тогда экзамен становится обычным, деловым и бесстрессовым отчётом учителя и учащихся о проделанной ими совместной учебной работе.

Первое условие — это обеспечение содержательной валидности тестов. Для этого надо разработать и поместить в учебнике все тесты, отображающие цель обучения и охватывающие **все** изученные учебные элементы. Учение по такому учебнику позволяет ученику вести систематический самоконтроль качества своей учебной работы, а учителю осознанно управлять ею. И если по результатам такого тестирования окажется, что цель не достигнута, то можно будет

вполне осмысленно искать причину. И с этой задачей должны иметь дело участники процесса обучения, а не посторонние его оценщики.

Второе условие — это во всех случаях проверки и оценки знаний учащихся включать в экзаменационные испытания только опубликованные в учебнике тесты или аналогичные им. Это условие соответствует требованию содержательной валидности экзамена. Известное стремление экзаменаторов отклоняться от содержательной валидности теста и включать на экзаменах тесты по ранее не изученным учебным элементам, оправдываясь тем, что ученик должен уметь ориентироваться в неожиданных ситуациях, используя усвоенные умения в новых условиях, ничем не оправдано. Если такое должно случиться в будущей практике ученика, то соответствующей должна быть цель обучения (третий уровень) и в учебнике должны быть также представлены тесты третьего уровня по соответствующим учебным элементам. Если тесты третьего уровня включаются в экзамен, то испытуемому надо разрешать пользоваться любой литературой и помощью экспертов. В то же время следует подчеркнуть, что тесты третьего и четвёртого уровня совершенно неуместны в об-

шеобразовательной школе и могут использоваться только при итоговой аттестации в **высшем** профессиональном образовании. Наконец, если в учебниках будет полный набор тестов по предмету для текущего самоконтроля, это может поставить под вопрос необходимость экзамена как такового вообще. Аттестовать ученика по результатам текущего тестирования может сам учитель, поскольку её (аттестации) правдоподобность будет проверена вступительным экзаменом в профессиональном учебном заведении или при приёме на работу, если работа требует такого образования.

И третье, кажущееся наиболее противоречивым с точки зрения традиционного образования условие: предоставление учащемуся полной свободы пользования во время экзамена всеми возможными источниками информации: учебниками, монографиями, конспектами, шпаргалками, что моделирует будущую самостоятельную деятельность испытуемого, когда использование различных источников никак и ничем не будет ограничено. Регулировать степень достигнутого учеником мастерства можно с помощью параметра автоматизации (время на выполнение теста), минимального ($K_t < 0,5$) на первом

этапе усвоения мастерства и постепенно повышаемого ($K_t > 0,5$) на последующих этапах обучения.

И последнее. На приёмных экзаменах в вузы должны использоваться только тесты из того набора, которые представлены в школьном учебнике или подобные им, чтобы обеспечить полную преемственность в образовательной системе. Вузы могут вводить дополнительные психологические испытания абитуриентов по способностям к тому виду деятельности, который характерен для данного вуза, оставляя неизменной процедуру и валидность общеобразовательного экзамена.

Заключение

Подводя итог изложенному выше о теории и методике измерения и оценки знаний учащихся, принимая во внимание педагогический анализ типовых тестовых систем (SAT; TOEFL и ЕГЭ), можно сделать следующие общие выводы:

1) Используемые в настоящее время в России и за рубежом тестовые системы являются **инвалидными и недемократичными**. Они не опираются на какие бы то ни было явно выраженные научно-педагогические принципы измерения и оценки знаний и умений учащихся, их процедуры

Теория

Тесты, экзамены и демократия

ПЕД диагностика
ПЕД диагностика

Заключение

имеют выраженный репрессивный характер. Анализ показывает, что они построены на чисто эмпирической («здравый смысл») основе, что делает их чрезвычайно противоречивыми как в конструировании тестов, так и в их применении к оценке знаний учащихся.

2) Требования к знаниям учащихся в существующих системах не выражены в явной и конструктивной форме, не отражают диагностично поставленных целей обучения, чем не способствуют целенаправленной и осознанной подготовке тестовых материалов, с одной стороны, и учащихся — с другой, к экзамену.

3) Как правило, известные тестовые системы предъявляют учащимся в одном наборе функционально разные тесты (тесты разного уровня), а неправомерный суммарный показатель качества их исполнения создаёт иллюзорное представление учащихся и общественности об их мастерстве в изучаемых предметах.

4) Такая иллюзия может быть причиной глубоких разочарований студентов при встрече с реальностью. Не редки случаи, когда успешно сдав SAT, TOEFL или ЕГЭ, студент из-за академических трудностей должен покинуть вуз.

Именно на этой почве разыгралась трагедия в одном из американских вузов, когда студент из Нигерии, исключённый из вуза за неуспеваемость, застрелил декана факультета и своего профессора в припадке беспомощности перед непреодолимыми требованиями. А ведь этот студент получил удовлетворительный балл по TOEFL, иначе он не был бы принят в американский вуз

Публикации

Директивные и методические материалы о ЕГЭ опубликованы в печати и Интернете правительственными учреждениями и Федеральным центром тестирования.