

НОВОСТИ НАУКИ И
ТЕХНИКИ

<р>В атомной физике не
<р>Эксперимент МоEDAL ищ
<р>В популяциях африк
<р>Сложные задачи, для
<div class="note"><a href="http://
подробнее

НОВЫЕ КНИГИ

История философии для физиков и математиков



Книга о самых невообразимых животных. Бестиарий XXI века

ГЛАВНАЯ

ЕГЭ КАК КАТАЛИЗАТОР КРИЗИСА РОССИЙСКОГО ОБРАЗОВАНИЯ

Образование

Малинецкий Г.Г., Подлазов А.В.



Для единого государственного экзамена (егэ) в его нынешнем виде характерно совпадение декларируемых достоинств с системными недостатками.

Так, освобождение школьников от двойной экзаменационной нагрузки, достигаемое посредством совмещения выпускного и вступительного экзаменов, делает экзамен непригодным для выполнения обеих этих функций. А гипотетическое устранение образовательной коррупции, существующей на этапе перехода из школы в вуз, дается ценой ее общего усиления в результате создания новых коррупционных механизмов, действующих на школьном и вузовском уровнях. Обработка результатов егэ с помощью теории моделирования и параметризации педагогических тестов, которые позволяют объективно соотносить сложность разных вариантов (впрочем, на практике эта возможность не используется), значительно ухудшает способность экзамена дифференцировать поступающих в вузы и полностью лишает его способности характеризовать в целом качество школьного образования. Наконец, стандартизация требований к учащимся и подходов к оценке их успехов осуществляется за счет превращения сдачи егэ в отдельный предмет с одновременным вытеснением на периферию внимания многих важных знаний, умений и навыков.

В статье подробно рассматриваются эти и некоторые другие системные недостатки егэ и намечаются пути к их преодолению без отказа от единого экзамена. Поскольку в официальных документах, регламентирующие проведение егэ, не освещаются используемые методы моделирования и параметризации педагогических тестов, дается их описание с анализом недостатков – как неустранимых, так и привнесенных неграмотным использованием.

Постановка задачи

Единый государственный экзамен (егэ [1]), по всей видимости, в ближайшие годы будет той реальностью, в условиях которой предстоит существовать средней и высшей школам России. Отношение активной части общества к этой технологии было и остается, мягко говоря, неоднозначным. Как будет показано ниже, плюсы от внедрения егэ в его нынешнем виде во многом иллюзорны [2], тогда как минусы вполне осязаемы.

Отечественное образования находится в кризисе, столь глубоком, что он в долговременной перспективе угрожает самому существованию страны. Поэтому крайне важно понимать, как сказываются на перспективах системы

ПОИСК

ПРОЕКТЫ И ПРЕЗЕНТАЦИИ

- Инструментарий для планирования и решения социально-экономических проблем регионального развития и инновационной деятельности
- Центр математического моделирования и экспертного анализа ИПМ им. М.В.Келдыша РАН
- Рыжов В.А. Методики привлечения, оценки разработок и механизм отбора технологий на условиях, нужных инвестору

подробнее



Парнокопытные киты, четырехкрылые динозавры, бегающие черви...



подробнее

образования проводимые реформы, среди которых одну из первых ролей играет именно введение единого государственного экзамена, лишь усиливающего кризисные процессы, но никак не противостоящего им.

На наш взгляд, правильным был бы полный отказ от единого государственного экзамена, пагубным образом влияющего на образование. Однако, исходя из возможного, в данной работе авторы видят свою задачу не в обосновании необходимости отмены егэ, а в обсуждении его дефектов и в поиске путей их преодоления.

Сейчас события развиваются в сторону приспособления образовательной системы к егэ и выработки методов его саботажа. Такой результат не может радовать ни сторонников единого экзамена, ни его противников. Первые вряд ли удовлетворятся тем, что ни одна из декларируемых целей его введения так и не была достигнута [3]. А вторых пугают осложнения, к которым может привести болезнь под названием егэ, если не лечить ее, а лишь загонять вглубь.

Чтобы единый экзамен мог превратиться из очередного бездумного (а, возможно, и злонамеренного) начинания властей в полезный и приемлемый для общества инструмент, придется радикально изменить его формы и существенно уточнить его цели, а кроме того, привести первое хотя бы в минимальное соответствие со вторым.

Заметим, что в контексте реформы и развития системы образования тема егэ является, на самом деле, второстепенной. Вне всякого сомнения, гораздо важнее вопросы, связанные с тем, какое нам нужно образование, каковы его задачи, может ли их достижение быть измерено оценками и могут ли оценки служить мотивации ученика. Но эти вопросы, не имеющие простых и однозначных ответов, мы оставляем в стороне, сосредоточившись только на теме егэ, допускающей раскрытие в рамках отдельной работы.

Достижение в обществе взаимопонимания по этой узкой теме могло бы стать минимальной гарантией того, что и более сложные вопросы могут быть со временем успешно разрешены. А если взаимопонимание не будет достигнуто, то егэ вскоре добьет российское образование, и эти вопросы утратят свою актуальность.

Ты перестала пить коньяк по утрам?

Одним из острых вопросов, вставших в связи с переходом на егэ, является определение того, в какой мере обезличенные и унифицированные методы проверки применимы к тем или иным предметам.

Сложность организации человеконезависимого экзамена состоит в том, что он может включать в себя лишь такие вопросы, на которые существует однозначный, объективно верный ответ. Чем дальше дисциплина от математики, тем обременительней это условие. И если для русского языка, информатики или физики соблюсти его еще как-то можно, то для литературы, истории или обществознания оно вряд ли выполнимо, что на практике приводит к бессодержательности большинства егэшных заданий по этим предметам [4].

В настоящий момент никому не известно, для каких школьных предметов вообще возможна разработка содержательных заданий, пригодных для формальной проверки. Собственно говоря, нормальное введение единого государственного экзамена должно было начинаться с отработки способов его организации и проведения на математике, выделенной

наличием бесспорных критериев истинности. И лишь после достижения успеха с математикой можно было бы переходить к другим дисциплинам, постепенно выявляя и преодолевая специфические для них трудности.

Однако организаторы егэ решили наступать широким фронтом, не думая о сложностях, с которыми связана формализация экзамена почти по всем школьным предметам. Наличие множества различных концепций и трактовок, систем терминов и классификаций, приближений и уровней описания может стать причиной неоднозначного восприятия формулировки заданий. Более того, даже для однозначно трактуемого вопроса то, какой ответ на него следует считать правильным, может радикально зависеть от глубины знаний. А ведь при проведении массового экзамена возможно и даже неизбежно возникновение коллизий, когда экзаменуемый превосходит проверяющего по уровню знаний, что почти исключает выставление адекватной оценки. К аналогичному результату в случае единого экзамена ведет различное построение курсов или использование разных учебников.

Причем не следует думать, что это особенность только дисциплин, касающихся человека и общества. В наиболее острой форме она проявилась в ... химии, являющейся безусловным лидером по числу отловленных егэшных ляпов. Дело в том, что школьная химия – наука уже достаточно нестрогая, чтобы неквалифицированные авторы в изобилии создавали задания, которые понимаются и выполняются совершенно по-разному, в зависимости от того, чему и как учили. Но, с другой стороны, благодаря наличию объективных экспериментальных данных и точных физических теорий квалифицированный человек может отловить и аргументировано описать все глупости.

Для тех предметов, по которым проводится егэ, есть только два пути: либо составление продуманных заданий, что требует совершенно иного подхода к их разработке, либо унификация школьной программы до состояния единственно верного учения, что более соответствует нашей традиции.

К сожалению, пока события развиваются по второму пути. Происходит насаждение единомыслия, временами переходящего в безмыслие [5]. Чтобы успешно справляться с заданиями, предполагающими единственно верный ответ, экзаменуемый вынужден оперировать в той же системе взглядов, которой пользуются составители заданий. Если же он читал другие книги или его знания выходят за пределы школьной программы, он неизбежно оказывается в проигрыше.

Однако, не теряя надежды на то, что рано или поздно ситуация изменится, будем далее полагать, что описанная проблема уже решена.

Принципы и частности

Дефекты егэ уместно подразделить на системные и частные. Если исправление первых, которые и будут основным предметом дальнейшего анализа, требует существенного изменения формы проведения экзамена, то вторые могут быть легко устранены и в ее рамках.

Например, все варианты заданий, предлагавшиеся на экзамене, необходимо сразу после его проведения публиковать вместе с решениями и правилами начисления баллов, чтобы недовольные оценкой экзаменуемые понимали, где их могли «обсчитать». А чтобы они не боялись требовать своего, им следует гарантировать, что на апелляции оценка ни за одно

задание не может быть уменьшена даже при обнаружении пропущенных при его проверке недочетов. Кроме того, экзаменуемый должен сразу после экзамена получать на руки копии всех экзаменационных бланков, содержание которых проверяется компьютером. В силу того, что по ним не допускается апелляция, снова увидеть свои ответы уже невозможно и нет никакой возможности установить, кто допустил ошибку – экзаменуемый или компьютер.

Вся первичная информация по результатам проведения единого государственного экзамена должна находиться в открытом доступе. Любой желающий должен иметь возможность получить исчерпывающие сведения о выполнении каждого задания каждым экзаменуемым [6] (с указанием доставшегося ему варианта, а также учебного заведения, которое он окончил, и места, где проходил тестирование). Без публикации полной статистики егэ никакого доверия ему быть не может.

Еще одной очевидной и легко исправляемой глупостью является установка порогов прохождения итоговой аттестации после проведения экзамена. Это обесмысливает само понятие аттестации. Не вызывает сомнения, что минимальный балл, получение которого свидетельствует об освоении экзаменуемым школьной программы, должен устанавливаться до экзамена, причем не «от фонаря», а по формальным правилам, одинаковым для всех предметов.

Для удовлетворения этим и некоторым другим, столь же очевидным требованиям, необходима только политическая воля. А само возникновение этих дефектов егэ является следствием элементарного нежелания власти вступать с обществом в диалог.

Системные дефекты единого государственного экзамена существенно иные. Они обусловлены, в первую очередь, попытками придать экзамену множественные функции. Нельзя с помощью одного и того же инструмента проводить итоговую аттестацию выпускников и проводить конкурсный отбор абитуриентов, контролировать качество работы отдельного учителя и осуществлять мониторинг состояния системы образования в целом. **Указанные задачи взаимно несовместимы. Решать их все одновременно означает не решить ни одну.**

Другой бедой единого государственного экзамена является обработка его результатов с помощью методов, которые либо научно не обоснованы, либо применяются с грубейшими ошибками. Это не только самостоятельная проблема, но и показатель общего уровня проработанности всей процедуры егэ, в т.ч. и тех ее частей, практическая реализация которых не может быть соотнесена с положениями какой-либо теории.

Еще один огромный недостаток егэ связан с массовым использованием заданий-угадаек. Их наличие не только существенно искажает результаты экзамена, но и оказывает демотивирующее влияние на школьников. Чтобы отказаться от подобных заданий, необходимо поднять разработку контрольно-измерительных материалов на иной качественный уровень. То, с каким трудом изживается угадка [7], также свидетельствует о фундаментальных проблемах егэ, но уже не процедурного, а содержательного плана.

Отдельно стоит выделить то печальное обстоятельство, что ориентация учебного процесса на подготовку к единому экзамену радикально сказывается на содержании школьного образования. **Фактически из инструмента контроля егэ превращается в центральный элемент формирования образовательной политики, осуществляемой лицами, не**

имеющими для этого ни квалификации, ни полномочий.

Наконец, нельзя забывать и о том пагубном воздействии, которое оказывает егэ на сферы, не относящиеся непосредственно к подготовке к экзамену или к его проведению. Разрушается система профессиональной ориентации и довузовской подготовки школьников. Не используются или даже блокируются каналы обратной связи, необходимые для принятия управленческих решений в системе образования и получения объективной оценки их результатов. В массовое сознание внедряются искаженные представления о последствиях перехода на егэ. Усиливается образовательная коррупция, и появляются ее новые формы.

Рассмотрим все перечисленные системные дефекты егэ подробно.

За двумя зайцами

Экзамен, как и любой другой измерительный прибор, имеет вполне определенный рабочий диапазон, где его показания адекватны реальности, и за пределами которого прибор бесполезен. Даже возможность переключаться между разными диапазонами без изменения принципов измерения, реализованная в хороших приборах, не позволяет им осмысленно работать в нескольких диапазонах одновременно.

Невозможность объединения в одном приборе микроскопа и телескопа, хотя оба построены на оптических принципах, представляется достаточно очевидной. Впрочем, если нечто невозможно, это еще не значит, что не найдется желающих внедрить это нечто в повседневную практику, прекрасным примером чего и стал единый государственный экзамен. **Сейчас в нем объединены функции выпускного экзамена и вступительного. То есть, он используется для выявления одновременно и тех, кому нельзя выдавать аттестат зрелости, т.к. они не смогли освоить школьную программу, и тех, кому имеет смысл получать в проверяемой области профессиональную специализацию.** При этом обе задачи решаются из рук вон плохо. Единый измерительный прибор может справляться или не справляться со своей функцией в середине диапазона, но он оказывается заведомо неэффективен на обоих его концах – при оценивании как худших, так и лучших.

С одной стороны, основная масса экзаменуемых даже не пытается выполнять самые сложные задания, воспринимая их как заведомо находящиеся за пределами своих возможностей. Это означает лишь то, что эти выпускники изначально оцениваются не из полного балла, поэтому для них требования егэ оказываются завышенными. Ситуация дополнительно усугубляется тем, что для самых слабых выпускников результат экзамена по большинству предметов вообще определяется не выполнением заданий, а лишь угадыванием их ответов.

С другой стороны, относительно простые задания, нацеленные на проверку базовых *знаний, умений и навыков* (зунов) и сами по себе не представляющие проблемы для подготовленных школьников, не оставляют времени для качественного выполнения сложных заданий, над которыми нужно думать. Экзамен становится соревнованием по скорости и автоматизму, не позволяя абитуриентам продемонстрировать свои таланты, и с этой точки зрения, требования егэ неизбежно оказываются заниженными. И, опять же, ситуация дополнительно усугубляется тем, что уровень даже наиболее сложных заданий по каждому предмету недостаточен для того, чтобы проверить соответствие подготовки

абитуриентов требованиям сильнейших вузов.

Однако, сколько бы ни твердили учителя о невозможности совместить *тесты достижений* с *тестами способностей*, что потенциал человека не находится в прямой зависимости от объема его знаний, это никоим образом не препятствовало внедрению ЕГЭ в его нынешнем виде. Сформировался метод оценки, вроде бы, и не достижений выпускника, но, вроде бы, и не способностей абитуриента, а чего-то среднего, не являющегося, тем не менее, ни тем, ни другим. Это в равной степени губительно и для средней школы, поскольку учителям и ученикам задаются непонятные ориентиры, и для высшей, поскольку вузы вынуждены набирать неадекватно оцененных абитуриентов.

Абсурдность ситуации усиливается тем обстоятельством, что описанная проблема имеет тривиальное решение. Если достижения и способности нельзя проверить одновременно, их следует проверять порознь. **Единый экзамен не должен быть единственным!** Само слово «единый» следует понимать исключительно как общий для всей страны, но не как объединяющий в себе выпускной и вступительный.

Иными словами, по каждому предмету надо проводить вместо одного экзамена два, имеющих различную направленность. Первый, являющийся общеобязательным, проверяет наличие у выпускника тех базовых знаний, умений и навыков, которыми по завершению обучения в школе он не имеет права не обладать (образовательный стандарт). Второй, сдаваемый только поступающими в вузы, выявляет потенциал абитуриента, его умение самостоятельно мыслить, порождая идеи, которые не даны заранее, и анализировать материал, связывая причины и следствия, то есть активно применять весь арсенал знаний, полученных в рамках школьной программы.

Принципиально, что эти два экзамена не могут быть объединены в одном временном отрезке, ибо, с одной стороны, необходимо гарантировать независимость оценок за их выполнение, а с другой – они несовместимы по требуемому психологическому настрою экзаменуемого. Сдавать их вместе – это всё равно, как стрелять после бега (конечно, состязания по биатлону интересней, чем просто лыжные гонки или стрельба, но биатлонисты-то идут на это совмещение добровольно).

Кроме того, тесты достижений и способностей предполагают диаметрально противоположный взгляд на оригинальное мышление, на нетривиальные способы выполнения заданий. И если во втором случае умение найти неожиданное решение задачи или проявить более глубокое, чем предписано программой, видение проблемы, является высшим достижением, то в первом это есть попытка уклониться от требования продемонстрировать владение совершенно конкретными знаниями, которая должна караться.

Наконец, разделение экзаменов по времени в принципе позволяет проводить тест достижений существенно – скажем, на полгода – раньше теста способностей. При этом у школьников будет возможность на завершающем этапе обучения сосредоточиться на предметах, связанных с их будущей профессиональной специализацией, по которым им предстоит проходить тест способностей. А в случае провала теста достижений по какому-то предмету остается достаточно времени, чтобы подтянуть его и пересдать до окончания школы, дабы не остаться без аттестата.

Тем не менее, разумной процедуры единого государственного экзамена недостаточно для того, чтобы вузы согласились

признавать его результаты без выкручивания рук и угрозы антикоррупционного шельмования. И это не удивительно, ведь если взятки за поступление берет лишь какая-то часть вузовских преподавателей, то работать со студентами, не имеющими базовой подготовки и не умеющими учиться, не хочет никто. Нет никакой возможности на основе формальных критериев отличить те вузы, которые противятся введению егэ из-за стремления сохранить коррупционные доходы, от тех, которые видят в нем препятствие отбору удовлетворяющих их требованиям абитуриентов.

Сейчас разработка *контрольно-измерительных материалов* (кимов) для егэ осуществляется закрытыми структурами, аффилированными с Рособрнадзором. И хотя из года в год неудовлетворительный уровень наполнения кимов уже стал общим местом, никому толком неизвестно, что за люди являются их авторами, на каком основании именно они допущены к этой работе и какую ответственность несут за ее результаты.

Решить как проблему субъективного признания вузами результатов егэ, так и проблему объективно низкого качества кимов можно, передав их разработку преподавательскому корпусу. В случае второго экзамена – преподавателям высшей школы, которым и предстоит в дальнейшем работать с теми, кого оценивают с помощью егэ, в случае первого экзамена – школьным учителям, которые и закладывают основы проверяемых знаний, умений и навыков.

Поскольку результаты экзамена, так или иначе, свидетельствуют и о качестве работы самих учителей, нужна независимая от них экспертиза кимов, что и должно быть функцией Министерства образования и его служб в деле разработки кимов. Ее финансирование также следует вывести из его ведения, чтобы гарантировать независимость разработчиков. Зато ответственность последних должна стать персональной, то есть для каждого задания должно указываться лицо, отвечающее за его качество и корректность.

В этих естественных предложениях нет ничего революционного. К примеру, в научных фондах, выдающих гранты на проведение исследований, оценка деятельности и перспектив соискателей проводится представителями научного сообщества, а не чиновниками, на которых возложены лишь организационно-технические функции. Точно так же и в случае егэ решать, что и как требовать с выпускников и поступающих, могут только преподаватели средней и высшей школ, предметно знающие их конкретику. А на долю чиновников останется проверка соответствия кимов школьной программе и образовательному стандарту, ну, и, собственно, проведение егэ.

Тридцать восемь попугаев и одно попугайское крылышко

Если экзамен – измерительный прибор, то принципиально важным является то, каким образом получается результат измерения и в каких единицах он исчисляется.

На егэ каждое задание имеет определенную *номинальную стоимость*. Сумма номинальных стоимостей заданий, выполненных экзаменуемым, составляет его *первичный балл*, который, однако, носит только вспомогательный характер. В соответствии с некоторой нелинейной шкалой (о способе построения и особенностях которой речь пойдет далее) он пересчитывается в *тестовый балл*, который и является результатом сдачи единого государственного экзамена.

Тестовый балл исчисляется по стобальной шкале, поэтому далее происходит еще один раунд преобразований, связанный с определением ее диапазонов, соответствующих оценкам привычной «пятибалльной» [8] шкалы выпускного экзамена. Границы этих диапазонов различны для разных предметов и меняются от года к году. Переход в 2009 г. с «пятибалльной» шкалы на двухбалльную (зачет-незачет) ситуации не изменил, зато отчетливо выявил смысл манипуляций с границами диапазонов. Она имеет целью исключительно управление долей различных оценок, в первую очередь, – двоек, высокий процент которых вызывает широкий общественный резонанс.

С оценками выпускного экзамена ситуация не менее сложная. До недавнего времени здесь тоже присутствовал дополнительный раунд преобразований, превращавших тестовые баллы во вступительные. Правила этих преобразований устанавливались вузами самостоятельно, что, хотя и давало возможность как-то учесть специфику каждого из них, но окончательно запугивало картину, не позволяя абитуриенту получить однозначное представление о том, насколько успешен его результат. С 2009 г. тестовые баллы егэ по предметам, соответствующим выбранной абитуриентом специальности, просто суммируются, давая его *вступительный балл*. Но егэ – такая непродуманная конструкция, что если хвост здесь вылез, значит, нос, непременно, увяз. За некоторое прояснение правил пришлось заплатить полным обесмысливанием результата их применения. Очевидно, что и от специальности, и от вуза существенно зависит относительная значимость оценок по разным предметам. И просто суммировать их, это всё равно, что пытаться определить сумму денег в кармане, суммируя номиналы банкнот, не обращая внимания на то, в какой они валюте [9].

В действительности дела здесь обстоят намного хуже, поскольку даже при равной значимости двух предметов нет никаких гарантий, что, скажем, 70 баллов по одному предмету + 90 по другому – это то же самое, что 80 + 80 или 90 + 70. Если неизвестно, каким способом градуирована шкала прибора, то нельзя быть уверенным, что цена его делений одинакова по всей шкале. А в случае шкалы тестовых баллов егэ можно показать, что они заведомо неравноценны.

За десятилетие проведения единого государственного экзамена его организаторы так и не сумели внятно объяснить обществу, что же там считается, каким образом и почему именно так, а не иначе. Но само по себе это еще не есть проблема. В конце концов, измерительный прибор не обязан иметь простое и общепонятное устройство – он должен давать надежные и однозначно интерпретируемые показания. Но вот этого, как раз, и нет.

Номинально в основе процедуры пересчета первичных баллов в тестовые лежит *теория моделирования и параметризации педагогических тестов* (items response theory). Однако ее методы используются зачастую за пределами их применимости, а рекомендации грубо нарушаются буквально на каждом шагу. Немаловажно и то, что в официальных документах, регламентирующих проведение егэ, описание методов обработки его результатов вообще отсутствует. Вместо этого приводятся лишь ссылки на научные публикации неясного правового статуса.

Возьмем на себя труд по описанию и разбору используемых методов.

Тестология – in vitro & in vivo

Для проведения любых тестов необходимо большое число разных вариантов, которые неизбежно будут неодинаковы по сложности. Компенсация ее различий и является главной задачей пересчета баллов. Теория моделирования и

параметризации педагогических тестов решает эту задачу посредством математических моделей, выражающих вероятности определенных исходов выполнения задания экзаменуемым через некие характеризующие их параметры.

Краеугольным камнем теории является *однопараметрическая модель Раша*, формулирующаяся для заданий с *дихотомическим исходом*, т.е. оцениваемых из одного балла (успех или неуспех). Она сводится к гипотезе о том, что отношение *вероятности справиться* с некоторым заданием p к *вероятности не справиться* с ним $1-p$ равно отношению *знаний* экзаменуемого Z к *трудности* этого задания T . Две последние величины сами по себе неопределимы, и возможно оперировать только их отношением. Поэтому, как это часто делается в подобных случаях, переходят к логарифмической мере – *логитам* [10] знаний $z = \ln Z/W$ и трудностей $t = \ln T/W$, где W – некий *опорный уровень*. Его выбор, равно как и выбор основания логарифма, несущественен, т.к. линейное отображение диапазона логитов знаний на заранее определенный диапазон тестовых баллов элиминирует эти два параметра.

Сразу обратим внимание, что модель Раша обоснована только для гомогенных тестов, т.е. состоящих из однотипных заданий. В случае же столь сложной структуры тестов, которая нужна для выпускного или вступительного экзаменов (и тем более, в случае попыток их совмещения), невозможно гарантировать достаточно высокую корреляцию результатов выполнения отдельных заданий с результатами выполнения теста в целом, что автоматически выводит нас из области применимости модели. Более того, гомогенность становится невозможно даже проверить, поскольку, чем хуже устроен тест, тем в большей мере он проверяет не владение предметом, а готовность к сдаче тестов такого типа.

Кроме того, очевидно, что задания с выбором ответа из нескольких предложенных вариантов, используемые в части А егэ, радикально нарушают предположения модели. Вероятность успешного выполнения таких *заданий-угадаек* существенно ненулевая при сколь угодно низких знаниях, поскольку даже расставленные наугад крестики принесут какое-то количество баллов. Следовательно, результаты обработки результатов единого экзамена с помощью формул модели Раша неизбежно оказываются искажены.

Однако предположим, что этих проблем нет, и пойдем дальше.

Пусть имеется N экзаменуемых, характеризующихся логитами знаний z_i ($i = 1, 2, \dots, N$), и каждому из них предложен вариант, состоящий из K заданий с логитами трудностей t_j ($j = 1, 2, \dots, K$). В соответствие с моделью Раша, вероятность успешного выполнения j -го задания i -м экзаменуемым есть

$$P_{ij} = \frac{1}{e^{t_j - z_i} + 1}.$$

Результатом тестирования будет матрица исходов $\|a_{ij}\|$ размерности $N \times K$, состоящая из нулей и единиц. На ее основе с помощью метода наибольшего правдоподобия [11] можно получить систему уравнений

$$\begin{cases} \sum_{j=1}^K p_{ij} = \sum_{j=1}^K a_{ij} = b_i & i = 1, 2, \dots, N \\ \sum_{i=1}^N p_{ij} = \sum_{i=1}^N a_{ij} = c_j & j = 1, 2, \dots, K \end{cases},$$

решение которой дает неизвестные величины z_i и t_j .

Первая строка системы требует для каждого экзаменуемого совпадения суммы вероятностей успеха по всем заданиям с реальным количеством его успехов b_i , которое называется *первичным баллом экзаменуемого*. Вторая строка требует совпадения суммы вероятностей успеха на каждом задании по всем экзаменуемым с реальным количеством их успехов на нем c_j , которое называется *первичным баллом задания*. Результаты экзамена входят в уравнения только в виде сумм, определяющих первичные баллы. Поэтому экзаменуемые (задания), характеризуемые одинаковым первичным баллом, получают одинаковые оценки знаний (трудностей) [12].

Заметим, что приведенные уравнения не являются независимыми (сложение уравнений первой строки системы дает тот же результат, что и сложение уравнений второй строки). Это вполне ожидаемо, т.к. неизвестные z_i и t_j могут быть определены только с точностью до произвольной аддитивной постоянной, задание которой эквивалентно фиксации опорного уровня W , связывающего трудности и знания с их логитами. Как уже было сказано, его величина не влияет на конечный результат, поэтому можно считать, что полученная система уравнений однозначно разрешима.

Задания разных вариантов неизбежно различаются по трудности, поэтому в них включают т.н. *якорные задания*, общие для всех вариантов. Найденные как решения системы логиты подвергаются такому линейному преобразованию, чтобы трудности якорных заданий из разных вариантов совпали (или оказались максимально близки в смысле какого-либо критерия, если их точное совмещение невозможно). Это позволяет скомпенсировать неравноценность вариантов. Откорректированные таким образом логиты знаний линейно отображаются на какой-либо удобный диапазон значений (например, от 0 до 100) – это и будут тестовые баллы.

Здесь имеется очевидный подводный камень, на который умудрились напороться организаторы егэ. Дело в том, что модель Раша при определении ее параметров с помощью метода наибольшего правдоподобия некорректна для экзаменуемых, показавших *абсолютный результат* (всё или ничего), – для них она дает логиты знаний $z = \pm\infty$. Однако линейно отобразить неограниченный диапазон знаний на ограниченный диапазон тестовых баллов невозможно. Поэтому если кто-то осилил все задания или не осилил ни одного, процедура линейного масштабирования оказывается в принципе невыполнима. Тем не менее, официально принятые и опубликованные правила пересчета, действовавшие до 2008 г., требовали выполнения именно такой процедуры. Но чудес не бывает, и поэтому на практике преобразование осуществлялось не так, как того требовали правила.

Собственно говоря, тут существует единственный и очевидный способ побороть бесконечности. Нулевому результату сразу сопоставляется 0 тестовых баллов, стопроцентному – 100, а вот уже для остальных – *неабсолютных* – результатов оценки подготовленности подвергаются линейному преобразованию. Однако это элементарное уточнение нашло свое

отражение в правилах только в 2008 г., а до того применялось подпольно, что автоматически делает результаты егэ предшествующих лет юридически ничтожными. Понятно, что публично признать подобную накладку немислимо, поэтому ее исправление было официально обосновано устранением другого изъяна методики пересчета баллов.

До 2008 г. параметры линейного отображения определялись следующим образом (еще раз подчеркнем, что речь идет о практике, а не об официальных правилах). Постоянный сдвиг выбирался так, чтобы средний уровень подготовленности отображался на 50 тестовых баллов, а коэффициент растяжения – так, чтобы конец диапазона оценок подготовленности, дальше отстоящий от среднего, отображался на 1 или 99. При этом другой конец диапазона, который был ближе к среднему, мог отобразиться в тестовый балл, очень далеко отстоящий от своего края шкалы. Так, в 2007 г. на егэ по математике экзаменуемые, набравшие всего 1 первичный балл (из 37, т.е. 2,7%), получали сразу аж 12 тестовых, тогда как на егэ по русскому языку столько же оставалось до 100 тем, кто не добрал всего 2 последних первичных балла (из 60, т.е. 3,3%). Иными словами, было возможно возникновение большой дырки в диапазоне тестовых баллов, причем расположенной с непредсказуемой его стороны. И если первое просто плохо, то второе в условиях действия свидетельств егэ в течение двух лет недопустимо.

Поэтому правила 2008 г. предписывали выбирать параметры линейного отображения так, чтобы оно было по возможности симметричным. Минимальный неабсолютный результат (набран только один первичный балл) отображался на 6 тестовых баллов, а максимальный неабсолютный результат (не набран только один первичный балл) – на 94. Почему стоимость первого и последнего первичных была баллов принята равной именно 6 тестовым, да еще и сразу для всех предметов (задания по которым оцениваются из разного количества первичных баллов), осталось загадкой [13].

В 2009 г. в правила пересчета были тайно [14] внесены дополнительные изменения. Логиты знаний, попавшие в тот же диапазон, в котором они лежали в 2008 г. преобразовывались по той же формуле, что и год назад, а выходящие за пределы этого диапазона – по дополнительным формулам. Таким образом, отображение перестало быть линейным даже в области неабсолютных результатов, что еще дальше уводит применяемую процедуру от ее научного обоснования.

Кстати, следует заметить, что попытки сделать результаты егэ действительными в течение двух лет, тоже являются начинанием, научно не обоснованным. Оно требует использования якорных заданий для экзамена, проводимого в разные годы. Но это возможно лишь теоретически, а не практически, так как кимы прошлых лет становятся материалом для предэкзаменационной подготовки. В результате, входившие в их состав якорные задания оказываются знакомы новым экзаменуемым, так что субъективная трудность этих заданий уменьшается на неопределенную величину. Вообще говоря, проблематичным является соотнесение даже результатов егэ разных волн, разделенных неделями, – срок достаточный для ознакомления с предъявленными заданиями.

Все описанные выше глупости и несуразности, за исключением наличия в егэ задач-угадаек, в сущности, являются мелочами, свидетельствующими о неспособности егэшников разобраться в теории, на положения которой они, якобы, опираются. Однако тут имеется и по-настоящему серьезная, но тоже рукотворная проблема. В процедуру пересчета баллов была добавлена еще одна стадия, теорией непредусмотренная.

Дело в том, что после корректировки логитов с помощью якорных заданий, каждый вариант будет характеризоваться

своим диапазоном логитов знаний, а значит, и своими параметрами отображения на шкалу тестовых баллов. В результате возможно возникновение *инверсий* между первичным и тестовым баллом (ситуаций, когда меньший первичный балл превратится в больший тестовый балл за счет того, что получен на более сложном варианте). В условиях информационной закрытости системы единого государственного экзамена балльные инверсии, как нетрудно догадаться, должны вызывать бурю общественного негодования. Поэтому организаторы егэ, не готовые что-либо объяснять людям, изобрели дополнительную *стадию осреднения шкал*.

Сначала для каждого значения первичного балла b_i вычисляется среднее значение откорректированных z_i по всем экзаменуемым, набравшим такой первичный балл на всех вариантах, которое и считается для них *оценкой подготовленности*. И лишь только эти оценки преобразуются в тестовые баллы. Инверсий при этом не возникает.

Стадия осреднения, делая пусть не методы, но хотя бы результаты пересчета более понятными обществу, сводит на нет сам смысл применения этих методов. Основным их содержанием является возможность объективно соотносить между собой результаты выполнения разносложных вариантов. Но шкала пересчета, полученная на основе осреднения, преобразует первичные баллы в тестовые единообразно для всех экзаменуемых, т.е. уже без учета сложности вариантов. Наш Левша аглицкую блоху в очередной раз подковал, но танцевать та, как и следовало ожидать, перестала...

Заметим, что различия сложности вариантов в случае единого государственного экзамена оказываются куда больше, чем в случае традиционных вузовских экзаменов. Выравнивать сложность вариантов для традиционных вступительных экзаменов было сравнительно просто, т.к. разработчики заданий имели дело с небольшим числом вариантов. А для массового егэ необходимо большое число разных вариантов заданий, которые к тому же должны существенно варьироваться для разных часовых поясов [15], что делает задачу выравнивания их сложности практически неразрешимой. Впрочем, ее, скорее всего, никто и не пытается решать, поскольку, как принято считать, пересчет баллов компенсирует различия в сложности вариантов егэшных кимов.

Завершает этот паноптикум избранный подход к заданиям не с дихотомическим, а с *политомическим исходом*, т.е. оцениваемым не из одного, а из нескольких первичных баллов. Для обработки результатов выполнения таких заданий используется одно из самых неудачных расширений модели Раша, называемое *моделью частичного оценивания* (partial credit model) [16]. Последняя сводится к предположению, что баллы задания представляют собой *однобалльные ступеньки*, которые преодолеваются строго последовательно, причем вероятность преодоления каждой из них подчиняется модели Раша. При этом ее базовая формула

$$p/(1-p) = Z/T$$

заменяется (включаящим ее в себя как частный случай) более общим выражением

$$p_s/p_{s-1} = Z/T_s,$$

где p_s – вероятность набрать ровно s баллов при выполнении задания, а T_s – некая характеристика s -ой ступеньки.

Раскрывая это рекуррентное соотношение, получаем

$$\frac{p_s}{p_0} = \frac{Z}{T_1} \cdot \frac{Z}{T_2} \cdot \dots \cdot \frac{Z}{T_s},$$

где

$$\frac{1}{p_0} = 1 + \frac{Z}{T_1} + \frac{Z}{T_1} \cdot \frac{Z}{T_2} + \dots + \frac{Z}{T_1} \cdot \frac{Z}{T_2} \cdot \dots \cdot \frac{Z}{T_M},$$

а M – число ступенек в задании.

Как и в случае модели Раша, в модели частичного оценивания имеют смысл не сами по себе величины Z и T_s , а только их отношения. Однако если в модели Раша величина T характеризовала трудность получения балла, то в случае модели частичного оценивания величинам T_s нельзя придать аналогичный смысл трудностей получения s баллов (равно как и какой-либо иной внятный смысл). Но если невозможна содержательная интерпретация величин T_s , то становится неинтерпретируемой и величина Z , измеримая только в связке с ними. Модель частичного оценивания позволяет нечто измерять, но мы лишены возможности понять, что же именно она измеряет.

Более того, модель становится внутренне противоречивой при использовании в одном тесте заданий с разным числом ступенек. Поскольку согласно предположениям модели ступеньки преодолеваются строго последовательно, вероятность преодолеть s -ю ступеньку (набрать s или более баллов) не может зависеть от наличия последующих ступенек. Однако, как легко видеть, эта вероятность дается дробью, числитель и знаменатель которой есть полиномы от Z степени M , т.е. она зависит как от числа последующих ступенек, так и от их характеристик. Следовательно, смысл величины Z будет варьироваться при изменении числе ступенек в заданиях, используемых для ее измерения. Поэтому одновременное наличие в ЕГЭшных кимах заданий, оцениваемых из разного числа баллов, при пересчете по формулам модели частичного оценивания делает итоговую оценку величиной просто бессмысленной.

Наконец, третья проблема с этой моделью связана с тем, что задания, части которых могут выполняться строго последовательно, как она того требует, крайне неудобны с практической точки зрения. Обычно куда важнее проверить функциональную грамотность экзаменуемого и его умение выполнить некий комплекс взаимосвязанных действий, нежели выяснить, как далеко человек способен пройти по навязанной ему цепочке шагов. Поэтому устройство реально предлагаемых заданий с политомиическим исходом почти никогда не согласуется с исходными посылками модели частичного оценивания, что автоматически делает ее формулы неприменимыми для обработки результатов выполнения этих заданий.

Подводя промежуточный итог, можно сказать, что если модель Раша – это вполне разумный практический инструмент, который используется безграмотно и (как будет показано ниже) далеко за пределами области своей применимости, то

модель частичного оценивания – это плод абстрактного теоретизирования, для практического применения непригодный изначально. Единственная мыслимая причина обращения к этой экзотической модели – стремление сделать процедуру обработки результатов егэ максимально непонятной не только для широкой общественности, но и для большинства специалистов.

Не от той стенки гвоздь

Описанные выше проблемы со шкалированием вызваны непониманием организаторами егэ основ используемого ими математического инструментария. Однако есть сложности и совершенно иной природы, связанные с тем, в какой мере этот инструментарий соответствует решаемым задачам.

Тесты по решаемой ими задаче делятся на *критериально-ориентированные* и *нормативно-ориентированные*. Первые предназначены для проверки усвоения экзаменуемыми определенного материала (например, итоговая аттестация), а вторые – для сравнения экзаменуемых между собой (например, конкурсный отбор). В случае каждой из указанных задач применение модели Раша и ее обобщений приводит к проблемам (в первом случае – к непреодолимым, во втором – к преодолимым, но реально не преодолеваемым).

Одной из важных функций единого государственного экзамена как итоговой аттестации должно быть отслеживание изменений уровня школьного образования. **Однако используемая методика шкалирования результатов егэ делает их в принципе непригодными для этих целей. Шкала пересчета первичных баллов в тестовые строится таким образом, что егэ оказывается измерительным прибором, лишенным какого бы то ни было эталона, с которым соотносились бы успехи экзаменуемых.** Они сравниваются не с требованиями школьной программы, а только с успехами других экзаменуемых. В результате полностью утрачивается важная функция единого государственного экзамена как объективного измерителя положения дел в сфере образования.

Рассмотрим ситуацию, когда имеется две равночисленных группы экзаменуемых, таких, что знания представителей второй группы составляют фиксированный процент от знаний представителей первой:

$$Z_i^{(2)} = Z_i^{(1)} \cdot \text{const}$$

Пусть обеим группам предложен один и тот же набор заданий, достаточно полный для точной оценки знаний их представителей. Тогда логиты знаний представителей групп, полученные в результате независимой обработки результатов экзамена, будут различаться на некоторую постоянную величину:

$$z_i^{(2)} = z_i^{(1)} + \text{const}$$

Но как корректировка логитов по якорным заданиям, так и последующий пересчет логитов в тестовые баллы уничтожают постоянный сдвиг, разделяющий группы. В результате они, несмотря на разные уровни знаний их представителей, становятся неразличимы.

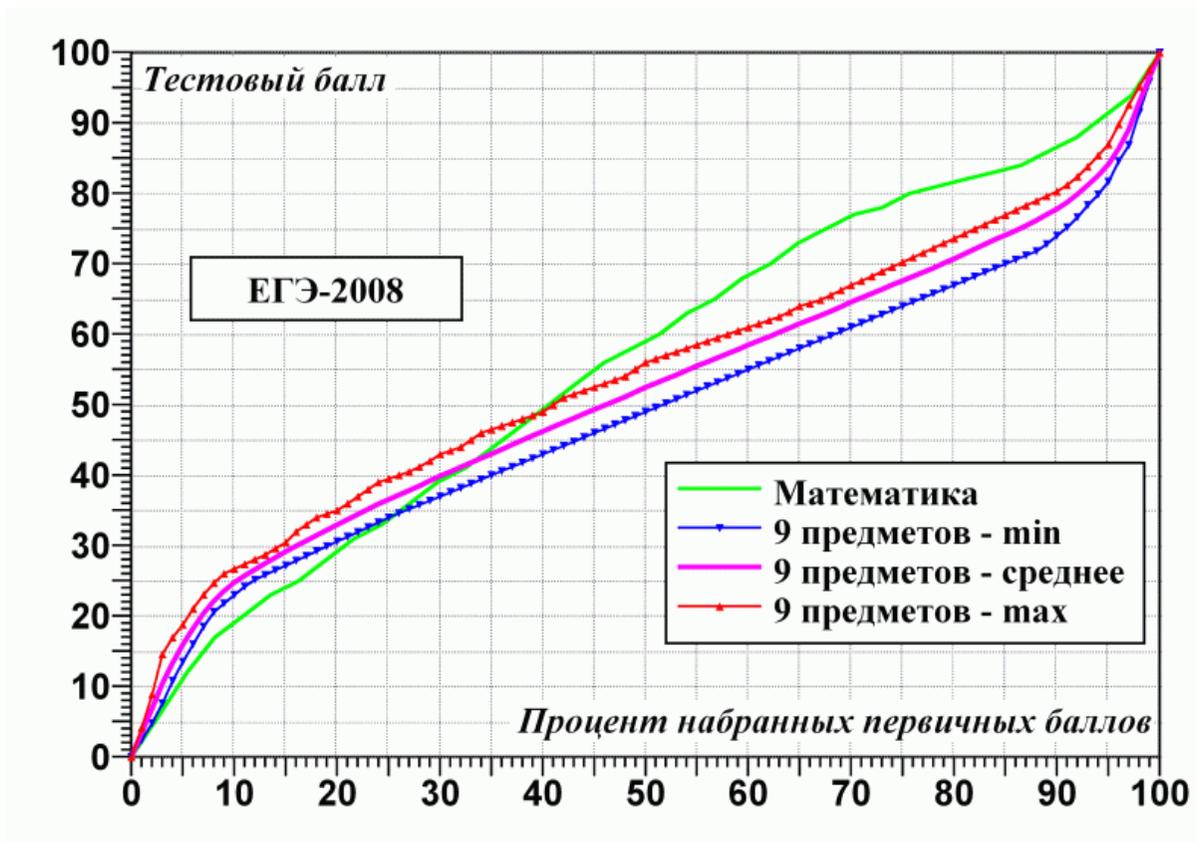
Если эти группы являются, скажем, школьными выпусками разных лет, то получается, что уровень образования может сколько угодно сильно измениться, что, однако, никак не отразится на результатах экзамена.

Из того, что модель Раша безнадежно плоха для моделирования и параметризации критериально-ориентированных тестов, к сожалению, еще не вытекает, что она окажется очень уж хороша для тестов нормативно-ориентированных. Для таких тестов ключевой характеристикой является *дифференцирующая способность* шкалы, т.е. то, насколько сильно изменяется итоговый результат экзаменуемого при приобретении/потере им одного первичного балла. Продифференцировав первое уравнение системы, находим

$$\frac{dz_i}{db_i} = \frac{1}{\sum_{j=1}^K p_{ij}(1-p_{ij})}$$

Легко видеть, что дифференцирующая способность шкалы оказывается максимальной на ее краях и минимальной в центре. В самом деле, все слагаемые в знаменателе правой части близки к нулю для экзаменуемых с самыми низкими (высокими) знаниями, для которых актуально получение самых первых (последних) первичных баллов [17]. Именно эти баллы и преобразуются в самое большое количество тестовых, обеспечивая сравнительно хорошую дифференцировку самых слабых (сильных) экзаменуемых. Но их немного и конкуренция между ними невелика, поэтому нет никакой необходимости дополнительно увеличивать разреженность их результатов. Высокая дифференцирующая способность нужна как раз на середине шкалы, где сосредоточена основная масса экзаменуемых, имеющих средние знания. Однако для них пересчет баллов с помощью модели Раша приводит лишь к увеличению скученности.

На рис. 1 показана зависимость тестового балла егэ от первичного по результатам проведения единого государственного экзамена в 2008-09 гг. На ее графике быстрый рост сменяется медленным, а потом снова – быстрым. Первые и последние тестовые баллы даются легче всего. Однако, как уже было сказано, для надежной дифференцировки экзаменуемых, оптимальной была бы прямо противоположная ситуация, при которой набор первых и последних баллов максимально затруднен (рост графика: медленно – быстро – медленно). Именно такова логика традиционных экзаменов, эффективность которых проверена многолетней практикой. Их варианты обычно включали по одному «утешительному» и одному «убойному» заданию, ориентированным на самых слабых и самых сильных, а также – большое количество заданий средней сложности, служивших целям дифференцировки основной массы поступающих.



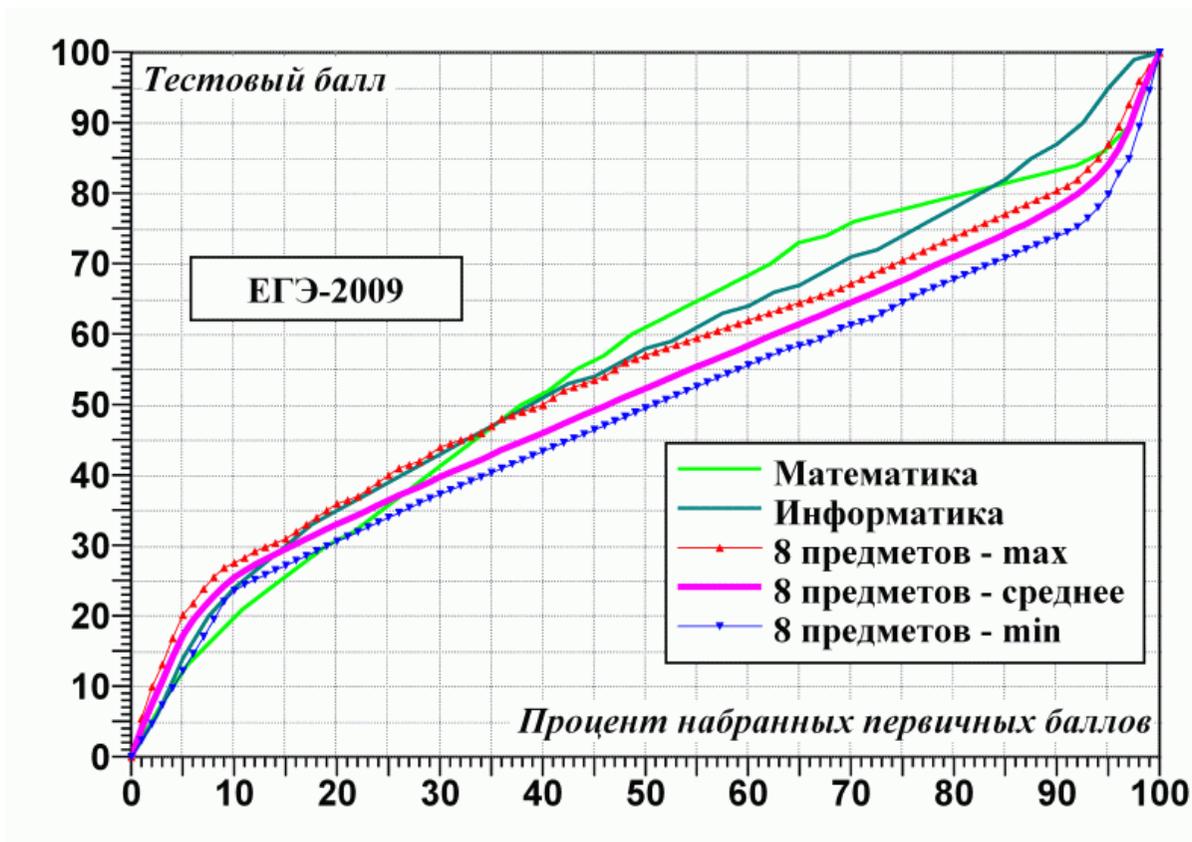


Рис. 1. Шкалы пересчета первичного балла егэ в тестовый (ОИП)

Для большинства предметов (физики, химии, биологии, географии, истории, литературы, обществознания, русского языка, а в 2008 г. еще и информатики) шкалы практически совпадают, располагаясь внутри узкого коридора со средней полувысотой в 3 балла. Они устроены так, что для первых и последних четвертей диапазона тестовых баллов их ценность, выраженная в первичных баллах, оказывается вчетверо ниже, чем для центральной половины.

Из общего ряда выбивается математика, для нее график пересчета близок к линейному. Причиной этого является сравнительная трудность получения по математике первых и последних первичных баллов. В 2009 г. из коридора в правой части графика выбралась и информатика, что обусловлено появлением в ее кимах нескольких субъективно сложных заданий, затруднивших получение последних баллов.

Идеальна ситуация, когда *интегральное распределение* экзаменуемых по результатам (зависимость от результата доли экзаменуемых, не сумевших его превзойти) имеет линейный вид без особо крутых или пологих участков. В этом случае приобретение (потеря) балла конкретным экзаменуемым будет означать, что он обгоняет (пропускает) примерно одинаковое число конкурентов вне зависимости от того, каков его результат [18]. И, соответственно, будет иметь хоть какой-то смысл привычное для нас вычисление суммарного вступительного балла от экзаменов по различным предметам.

На рис. 2 приведены примеры распределений участников егэ-2009 по набранным баллам. Легко видеть, что отклонение от линейного вида для тестового балла значительно больше, чем для первичного. Графики для тестового балла имеют слева и справа широкие участки, «выположенные» практически до горизонтали, вследствие чего средняя часть графика

очень сильно «выкручена». Части графика, на которую приходится 90% экзаменуемых, по разным предметам соответствует диапазон шириной лишь в 35÷50 тестовых баллов из 100, т.е. из-за пересчета баллов рабочая часть шкалы сокращается в 2÷3 раза.

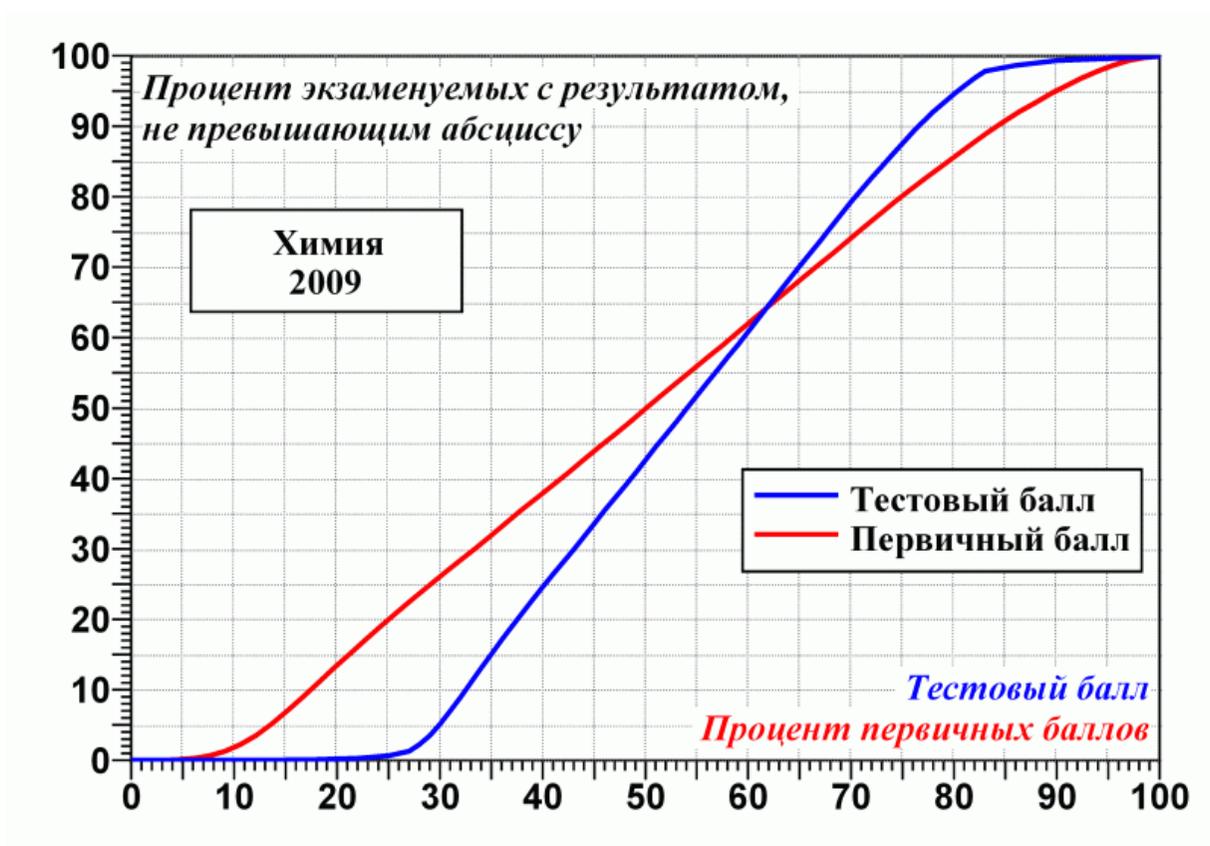




Рис. 2. Интегральное распределение участников егэ по химии и русскому языку по набранным тестовому и первичному баллам (ОИП)

Если вид графиков для первичных баллов еще близок к линейному, то графики для тестовых баллов уже имеют по краям отчетливо выраженные плато, а в середине – крутые подъемы. В результате более-менее линейные участки этих графиков оказываются узки (от 30 до 80 тестовых баллов по химии и от 35 до 75 – по русскому языку).

В этой связи следует обратить внимание на еще одну принципиальную особенность любого шкалирования: чем выше дифференцирующая способность шкалы, тем выше погрешность итоговой оценки (и наоборот). В случае модели Раша данное обстоятельство особенно наглядно.

Дисперсия балла, получаемого за задание с вероятностью выполнения p , есть $\sigma^2 = p(1-p)$. А дисперсия первичного балла b_i в предположении независимости отдельных заданий есть сумма их индивидуальных дисперсий:

$$\sigma_{b_i}^2 = \sum_{j=1}^K p_j (1 - p_j)$$

Заметим, что это выражение совпадает со знаменателем приведенной выше формулы для дифференцирующей способности шкалы. Учитывая пропорциональность стандартных отклонений дифференциалам соответствующих величин, получаем, что дисперсия логита знаний для модели Раша в точности равна дифференцирующей способности:

$$\sigma_{z_i}^2 = \frac{dz_i}{db_i}$$

Как мы видим, «хороших» участков у шкалы просто нет: там, где она хорошо дифференцирует, она плохо оценивает, а там, где хорошо оценивает, плохо дифференцирует. И чем дальше преобразование от линейного, тем сильнее будет проявляться это неустранимое противоречие.

На самом деле, ситуация не столь трагична. При использовании модели Раша единственно разумным является полный отказ от оценивания экзаменуемых, результат которых ниже (выше) некоторого порога. Им всем, а не только показавшим абсолютные результаты, независимо от точной величины первичного балла, следует директивно присваивать нулевой (максимальный) тестовый балл. Это всего лишь означало бы честное признание того, что предложенный тест (как вообще любой инструмент) имеет ограниченную область применимости. В результате, с одной стороны, были бы исключены участки шкалы, где погрешность определения тестового балла недопустимо высока, а с другой – растянулась бы средняя часть шкалы, что повысило бы ее дифференцирующую способность. Если ориентироваться на рис. 1, то применительно к егэ сказанное означает отказ от оценивания экзаменуемых, набравших менее 10% (более 90%) первичных баллов, что позволило бы вдвое растянуть рабочий участок шкалы на рис. 2. Однако, как нетрудно понять, это несовместимо с желанием егэшников измерить всех и сразу.

А ларчик просто открывался...

Проблема обработки результатов единого государственного экзамена не сводится к отдельным ошибкам, которые можно было бы ценой больших или меньших усилий исправить. **Она носит системный характер и связана с отсутствием внятного ответа на вопрос о том, что именно должна характеризовать количественная мера успехов экзаменуемого.** Для существующего формата егэ ответ на поставленный вопрос и не может быть дан в принципе, т.к. в случае выпускного и вступительного экзаменов это совершенно разные вещи. Однако если разделить единый экзамен на независимые выпускной и вступительный, то для каждого из них просто определяются и цели, и способы их достижения.

Цель вступительного экзамена – конкурсный отбор. Здесь компенсация различной сложности вариантов осуществляется элементарно, если в качестве экзаменационной оценки использовать *рейтинг-балл*. Рейтинг-балл экзаменуемого определяется как процент его конкурентов, выполнявших тот же вариант, что и он, и набравших первичный балл не меньший, чем набрал он. При этом автоматически обеспечиваются линейность интегрального распределения и

постоянная дифференцирующая способность на протяжении всей шкалы, а также гарантируется простая и понятная цена балла. Однако не менее важно, что решается и еще одна важная задача, выходящая за пределы тестологии.

Использование теории моделирования и параметризации педагогических тестов для обработки результатов вступительного экзамена неявно предполагает, что при отборе поступающих в вузы первостепенное значение имеет уровень достижений, а раз так, то достаточно поставить экзаменуемых в равные условия. Такая точка зрения представляется глубоко порочной. Принимать в высшие учебные заведения следует, в первую очередь, не тех, кто продемонстрировал хорошую подготовку, а тех, кто сможет учиться в вузе, т.е. ориентироваться необходимо на способности и мотивации, а не на текущие достижения [19].

Достижения человека зависят не только от его способностей и мотиваций, но и от социально-экономических условий, в которых он находится. Так, например, очевидно, что одинаковые результаты школьников из столицы и из глухой отдаленной деревни свидетельствуют, скорее всего, о совершенно разном уровне способностей и мотиваций. Данное обстоятельство может быть учтено, если каждый вариант вступительного экзамена предлагать на территории с более-менее одинаковыми социально-экономическими условиями, не смешивая столицы с провинцией, а города с селами. Тогда конкурировать между собой будут экзаменуемые, имеющие близкие, насколько это возможно, стартовые условия. Решение этой задачи социального плана также является весомым доводом в пользу использования рейтинг-балла на вступительном экзамене. Немаловажна и блокировка преференций, получаемых жителями более коррумпированных регионов страны.

В случае выпускного экзамена всё обстоит совсем по-другому. Его цель – проверка усвоения школьной программы. При этом нет никакой нужды сравнивать выпускников между собой. **Выпускной экзамен имеет смысл только в том случае, когда его результаты являются не относительными, а абсолютными.**

Выпускник должен иметь вполне определенный набор базовых знаний, умений и навыков. Поэтому единственно возможный способ определения номинальной стоимости задания – пропорционально количеству проверяемых им зунов. Владение каким их процентом продемонстрировал экзаменуемый, такова его оценка по стобальной шкале [20]. Остальное – от лукавого.

Такая организация выпускного экзамена имеет одно, несомненно, позитивное следствие: все его задания должны выбираться из *открытого банка*. Исключив процедуру пересчета баллов после экзамена, мы вынуждены заранее гарантировать равнотрудность его вариантов. Но для этого надо провести *калибровку* его заданий, т.е. убедиться в том, что количество начисляемых за их выполнение баллов, сообразно их сложности. А калибровка заданий невозможна без их предъявления для выполнения на контрольных, олимпиадах, пробных экзаменах или экзаменах прошлых лет [21]. Но однажды предъявленные задания неизбежно попадут в открытый доступ, в силу чего становятся бессмысленными попытки их скрывать.

Открытость банка заданий имеет массу положительных следствий. Во-первых, это сократит возможности организаторов его превращать оплаченные деньгами налогоплательщиков кимы в свою интеллектуальную собственность, страдая копирайтным законодательством всякого, кто пытается их обнародовать. Во-вторых, это позволит на ранней стадии

выявить и элиминировать некорректные задания за счет возможности для любого желающего проверить их и дать им оценку. Ну, а в-третьих, исчезнет проблема утечки экзаменационных материалов, поскольку лучшая защита секретной информации – отсутствие секретности или хотя бы минимизация объемов закрытой информации [22].

Описанные выше методики получения оценок для обоих видов экзамена устроены значительно логичнее и проще, чем инструментарий теории параметризации и моделирования педагогических тестов. Более того, они лишают модель Раша и ее расширения одного из их главных преимуществ, которым принято считать возможность построить не *порядковую*, а *метрическую* шкалу выставляемых баллов. Если для элементов порядковой шкалы имеют смысл только операции *сравнения* («хуже», «лучше», «одинаково»), то в случае метрической приобретает смысл и операция вычисления *разности*. Это означает, что для метрической шкалы *цена* каждого очередного балла одна и та же. То, чем исчисляется эта цена, и определяет логику построения шкалы.

Для шкалы, основанной на модели Раша, балл – это изменение логита знаний на определенную величину. Для шкалы, которая выше была предложена для выпускного экзамена, балл – это владение определенной долей школьной программы в рамках образовательного стандарта. Ну, а для шкалы, предложенной для вступительного экзамена, балл – это процент обойденных конкурентов. То есть, все эти три шкалы метрические. Однако есть и разница. Данное свойство, достоящееся первой из шкал просто в силу методики ее построения, далее никак не может быть использовано, поскольку вне тестологических моделей логиты знаний никакого смысла не имеют. Шкале вступительного экзамена метричность достается столь же легко, но зато в качестве бонусов она дает постоянную дифференцирующую способность экзамена и возможность частично скомпенсировать различия в социально-экономических условиях регионов, где обучались экзаменуемые. Обеспечение метричности шкалы выпускного экзамена требует определенных предварительных усилий, но они не пропадают втуне. Дополнительным преимуществом этой шкалы является то, что она является еще и *нормированной*, т.е. в ней естественным образом определено начало отсчета, что не позволяет ей «не заметить» изменений общего уровня образования.

Сколько баллов в стобалльной шкале

Одной из важных положительных черт егэ, которой традиционно уделяется немалое внимание при его пиаре, считается переход от традиционной «пятибалльной» шкалы к качественно более точной стобалльной. Однако действительно ли кимы единого экзамена позволяют производить оценивание подготовленности с такой высокой точностью?

Любой прибор имеет погрешность измерений, которая не уменьшается от измельчения градуировки его шкалы. Можно было пересчитывать оценки егэ и к тысячебалльной шкале – стали ли бы они от этого точнее? Мелкая градуировка спасает только от ошибок округления, но не от ошибок измерения.

Результат любого измерения подвержен неизбежным флуктуациям. Измерение подготовленности – не исключение. При одном и том же ее уровне экзаменуемые могут получить разные итоговые баллы, распределение которых характеризуется определенными значениями математического ожидания m и дисперсии σ^2 . И нет никаких оснований считать тех экзаменуемых, чей балл попадает в интервал от $m-\sigma$ до $m+\sigma$, различающимися по уровню подготовленности. В пределах

погрешности теста им всем можно приписать балл m . Это накладывает ограничение на количество *различимых градаций* оценки G , каждой из которых должен соответствовать диапазон разыгрываемых баллов шириной 2σ .

Естественным способом повышения точности является проведение множественных измерений – в данном случае – предъявление экзаменуемым тестов, состоящих из многих заданий. По мере роста числа разыгрываемых баллов K , дисперсия итогового балла σ^2 будет возрастать пропорционально K (в предположении, что баллы набираются независимо). Соответственно, максимально возможное количество различимых градаций описывается формулой

$$G = c\sqrt{K},$$

где c – некоторый коэффициент. Заметим, что это ограничение носит фундаментальный характер и не связано со структурой теста и методикой пересчета баллов при условии, что они корректны. Их некорректность может неопределенно сильно уменьшить величину G .

Потеря градаций оценки может быть связана как с эффективным уменьшением количества разыгрываемых первичных баллов K , так и с внесением в тестовый балл случайных искажений, которые раздувают дисперсию, уменьшая тем самым коэффициент c . При проведении ЕГЭ каждый из указанных способов был применен аж дважды.

Уменьшение числа разыгрываемых баллов обусловлено наличием в ЕГЭшных кимах как заданий-угадаек, за которые есть шансы получить баллы, ничего не зная, так и заданий с политомическим исходом, допускающих непоследовательное преодоление ступеней, в результате чего частичный балл за задание можно получить несколькими различными способами, тогда как модель предусматривает только один. И то, и другое приводит к систематическому завышению результатов экзамена по сравнению с реальным уровнем экзаменуемого, т.е. некоторое количество баллов, начисляемых просто так, фактически выводится из розыгрыша.

Искажения в тестовый балл вносятся как за счет различной сложности вариантов, которую не компенсирует усредненная шкала пересчета баллов, так и за счет использования заданий, оцениваемых из разного числа баллов, что лишает смысла измеряемые параметры модели частичного оценивания. В первом случае оказывается существенным, какой именно вариант выполнял экзаменуемый, во втором – на каких именно заданиях был набран первичный балл. С точки зрения общего итога, эти индивидуальные обстоятельства оказываются случайными факторами, не поддающимися учету.

Описанные практические ошибки, сокращающие количество различимых градаций итоговой оценки, являются в принципе устранимыми. Поэтому попробуем понять, как велико может быть количество градаций в том гипотетическом случае, когда они устранены, т.е. найдем ограничение величины G сверху.

Можно показать, что если тест состоит только из заданий с дихотомическим исходом (оцениваемых из одного балла), то значение c не превосходит величины $1,0\div 1,2$ (рассчитанной в предположении, что разброс логитов знаний составляет $5\div 7$ единиц или, что то же самое, различие экзаменуемых по знаниям не превышает $2\div 3$ порядков). Однако данная величина коэффициента на практике не достигается, чему есть несколько причин.

Во-первых, приведенная верхняя оценка c соответствует весьма специфической структуре теста, при которой все задания имеют равную или почти равную сложность. Для разумно организованных тестов, в которых задания существенно варьируются по сложности, коэффициент будет несколько меньшим.

Во-вторых, в любом реальном тесте разные задания, вообще говоря, измеряют разные сущности, что неизбежно повышает погрешность итогового результата и снижает значение c . И хотя для кимов единого государственного экзамена внутренняя согласованность теста весьма высока [23], указанное обстоятельство всё же нельзя сбрасывать со счетов.

В-третьих, необходимо учитывать наличие заданий с политоимическим исходом (оцениваемых более чем из одного балла). Части таких заданий не являются независимыми. В силу этого успешное выполнение одной части обычно повышает шансы на успешное выполнения остальных частей. И, наоборот, ошибка, допущенная в одной части, снижает шансы справиться с остальными. Наличие подобной корреляции увеличивает дисперсию балла, полученного за задания, оцениваемые из нескольких баллов, по сравнению с суммарным балом за соответствующее количество однобалльных заданий. А значит, количество различимых градаций итоговой оценки еще более сокращается.

Вряд ли возможно напрямую учесть влияние всех описанных факторов. Однако можно воспользоваться следующей косвенной оценкой. Традиционный экзамен по математике, структура которого была оптимизирована для обеспечения 4 различимых градаций итоговой оценки, обычно состоял из $5 \div 6$ заданий, выполнение каждого из которых могло оцениваться из $3 \div 5$ псевдобаллов: «+», «+/-», «+2», «-/+», «-» (с возможным отказом от одной или двух промежуточных ступеней). Это соответствует $K \approx 18 \div 25$. Такое количество разыгрываемых баллов обеспечивает 4 градации оценки при значении $c \approx 0,8 \div 0,9$, которое в свете сказанного выше представляется заслуживающим доверия.

Максимальный балл K ощутимо варьируется для ЕГЭ по различным предметам. Так, в 2008–10 гг. на ЕГЭ по математике разыгрываются $30 \div 37$ баллов (что соответствует 5 различным градациям оценки), по литературе и информатике – $39 \div 45$ баллов ($5 \div 6$ градаций), по физике – 50 баллов (6 градаций), по химии, биологии, географии, истории, обществознанию и русскому языку – $59 \div 69$ баллов ($6 \div 7$ градаций), по иностранным языкам – 80 баллов ($7 \div 8$ градаций) [24].

Таким образом, если не принимать в расчет дополнительное уменьшение величины G из-за безграмотного шкалирования и некорректной структуры тестов, то можно утверждать, что стобалльная шкала ЕГЭ реально содержит всего $5 \div 8$ различимых градаций оценки. Это, разумеется, больше, чем при проведении традиционного экзамена, однако чтобы довести это количество до декларируемых 100 градаций, нужно увеличить число предъявляемых заданий всего-навсего в несколько сот раз.

Следует ли на этом основании отказываться от стобалльной шкалы? Нет, не следует. Она удобна психологически, а ее явно избыточная точность помогает противостоять ошибкам округления. Однако восторгаться ею тоже особо не стоит.

Данный пример достаточно хорошо позволяет почувствовать разницу между масштабом ЕГЭшного пиара и тем положительным эффектом, который может быть достигнут хотя бы теоретически.

К сожалению, дела обстоят еще хуже. Упомянутые 5÷8 градаций оценки соответствуют всей шкале итоговых баллов, но целиком она никогда и нигде не используется. При конкурсном отборе в любой вуз за места в нем конкурируют абитуриенты с оценками, принадлежащими диапазону, который значительно уже 100 баллов. Абитуриенты с оценкой ниже определённого порога, либо вообще не допускаются к конкурсу [25], либо не имеют никаких шансов на поступление. Диапазон результатов, на котором происходит конкуренция за места в конкретном вузе, оказывается вдвое–втрое меньше полной ширины шкалы. Соответственно, реально используемое количество различных градаций оценки сокращается до 2÷3, при том, что традиционный вступительный экзамен, ориентированный по сложности на конкретный контингент поступающих, уверенно обеспечивал 3 различных «положительных» оценки (получившие двойку из конкурса выбывали). Таким образом, несмотря на значительно большее число заданий, **точность единого государственного экзамена чуть ниже, чем у традиционных вступительных экзаменов**, что есть прямое следствие совмещения функций выпускного и вступительного экзаменов. При этом обществу рассказывают про значительно возросшую точность.

Так и кончаются школьные годы...

Одной из официально декларировавшихся целей введения единого государственного экзамена была «разгрузка выпускников-абитуриентов через сокращение числа испытаний за счет совмещения выпускных экзаменов в школах и вступительных экзаменов в вузах». **Функции егэ как итоговой аттестации постепенно сходят на нет, результатом чего стала фактическая отмена выпускного экзамена** (надо полагать, это и имелось в виду под разгрузкой выпускников).

Чтобы выполнять функции итоговой аттестации, экзамен должен сдаваться всеми выпускниками [26]. Однако с 2009 г. егэ обязателен только по русскому языку и математике, а по остальным предметам единый экзамен сдается лишь по желанию. Естественно, те выпускники, которым эти предметы не нужны для поступления в вуз, их и не сдают [27]. Более того, они просто прекращают учить эти предметы, поскольку никаких альтернативных форм итоговой аттестацией после отмены выпускных экзаменов не введено.

Еще недавно как одно из важнейших достоинств егэ превозносилась его дифференцирующая способность, якобы, значительно более высокая, чем у оценок среднестатистического учителя. Однако с 2009 г. выпускной экзамен свелся к недифференцированному зачету, который реально проводится только по двум предметам.

Тем не менее, даже в таком виде, рассмотрение результатов егэшной аттестации дает богатую почву для размышлений.

Статистика знает всё

Прежде чем приступать к анализу результатов единого экзамена в его выпускной ипостаси, необходимо сделать одно важное замечание. Существуют две разных статистики результатов егэ. Одна публикуется в ежегодных аналитических отчетах «Результаты единого государственного экзамена» [28] Федерального института педагогических измерений (ФИПИ), а вторая размещается на Официальном информационный портале (ОИП) единого государственного экзамена в разделе «Статистика ЕГЭ» [29].

По некоторым позициям содержание этих источников разнится радикально. Более того, разночтения в содержании предварительной и итоговой информации ОИП порой выходят за пределы любых мыслимых уточнений, чему никто не потрудился дать хоть какое-то объяснение. Скорее всего, имеет место включение в окончательную статистику лиц, не являющихся выпускниками текущего года. По крайней мере, тот факт, что суммарное число сдаваемых добавленными лицами экзаменов (около 250 тыс. ежегодно) оставалось практически неизменным в 2007-09 гг., косвенно подтверждает данное предположение. Также есть основания полагать, что в статистике ФИПИ могли быть не учтены лица, сдававшие егэ досрочно. Кроме того, результаты егэ в аналитических отчетах ФИПИ приведены без учета передач русского языка и математики в 2009 г.

К сожалению, данных только какого-то одного источника не достаточно для анализа темы аттестации. Поэтому мы будем пользоваться обоими источниками, указывая их в скобках. Читателю не следует пугаться, встретив цифры, отличающиеся от известных ему, хотя возможные расхождения порой могут быть фантастичны. Так, например, доля лиц, получивших двойки по математике и русскому языку в 2009 г., составляет то ли 7,0% и 6,5% (ФИПИ), то ли 3,7% и 3,0% (ОИП), а по иностранному языку в 2008 г. – то ли 6,6% (ФИПИ), то ли 17,7% (ОИП). Впрочем, столь большой разброс – это пиковые ситуации. В большинстве случаев различия являются не очень существенными и не влияют на общую картину.

Я угадаю эту мелодию с двух нот

Часть А егэшных кимов представляет собой *тест-американку*, т.е. состоит из заданий с выбором одного правильного ответа из четырех (трех для некоторых заданий экзамена по иностранным языкам) предложенных вариантов. Соответственно, у каждого четвертого (третьего) задания правильный ответ может быть просто угадан. Причем наличие хотя бы фрагментарных познаний, позволяющих отбросить часть вариантов как заведомо неверные, существенно повышает вероятность угадывания верного. Более того, некоторые задания-угадайки сформулированы так, что возможно исключить все неверные варианты, не зная вообще ничего, а полагаясь лишь на здравый смысл и предположение, что правильный ответ должен быть ровно один.

Схожим образом построен и ряд заданий части В экзамена. Номинально они предполагают краткий ответ, однако фактически во многих случаях в ответе не оказывается ничего, что бы не содержалось в условии. Требуется или отметить несколько правильных ответов из приведенного набора, или упорядочить приведенные варианты в соответствие с некоторым критерием, или сопоставить друг другу элементы двух наборов [30], или выписать из текста какое-то слово или предложение. Во всех этих случаях сохраняется непренебрежимая вероятность угадывания правильного ответа, равно как и возможность существенно повысить ее, если знать хоть что-то.

Приведенный список угадаек с кратким ответом уместно дополнить еще одним типом заданий, требующих поставить приведенное иностранное слово в нужную форму. Здесь тоже делается выбор из сравнительно небольшого числа возможных вариантов. И хотя они и отсутствуют в условии, эта дополнительная сложность компенсируется тем, что половина заданий части А на экзамене по иностранным языкам предполагает выбор ответа всего из трех вариантов, а не из четырех.

Доля разыгрываемых баллов, приходящихся на задания–угадайки, кардинально сказывается на результатах итоговой аттестации. Ее анализ уместно проводить отдельно для 2009 г. и для предшествующих лет, поскольку с этого года егэ по русскому языку и математике стал всеобщим, а по прочим предметам – необязательным.

Эксперимент по введению егэ: 2006-08 годы

Школьные предметы очень сильно различаются по доле заданий-угадаек в части В (см. табл.). Их совсем нет в кимах по математике и почти нет – по литературе. В кимах по физике и информатике угадаек уже побольше, но они еще пребывают в меньшинстве, начиная доминировать в кимах по географии, истории, химии, обществознанию и русскому языку. Наконец, часть В кимов по иностранным языкам (в приведенной выше трактовке) и биологии полностью состоит из таких заданий.

На егэ по математике и литературе 2006-08 гг. задания-угадайки частей А и В в сумме давали не более 33% первичных баллов, тогда как для остальных предметов этот показатель составлял 58÷75% (см. табл.). В этом, по-видимому, и кроется причина того отрыва, с которым в эти годы математика и литература обгоняли другие предметы по проценту неудовлетворительных оценок. Показать высокий результат, только гадая, конечно же, невозможно, но разница между двойкой и тройкой на егэ во многом определяется именно угадыванием.

Можно приближенно считать, что процент двоек в 2006-08 гг. был прямо пропорционален проценту первичных баллов, приходящихся на *полноценные задания*, которые не допускают угадывания ответа. Зависимость показана на рис. 3. В рамках однопараметрической (без свободного члена) линейной регрессионной модели свыше 87% дисперсии доли двоек, полученных на егэ 2008 г., объясняется вариацией доли полноценных заданий. В 2007 и 2006 гг. коэффициент детерминации оказывается заметно ниже – соответственно, 49% и 29%, что, скорее всего, обусловлено неунифицированным способом пересчета баллов для разных предметов в эти годы.

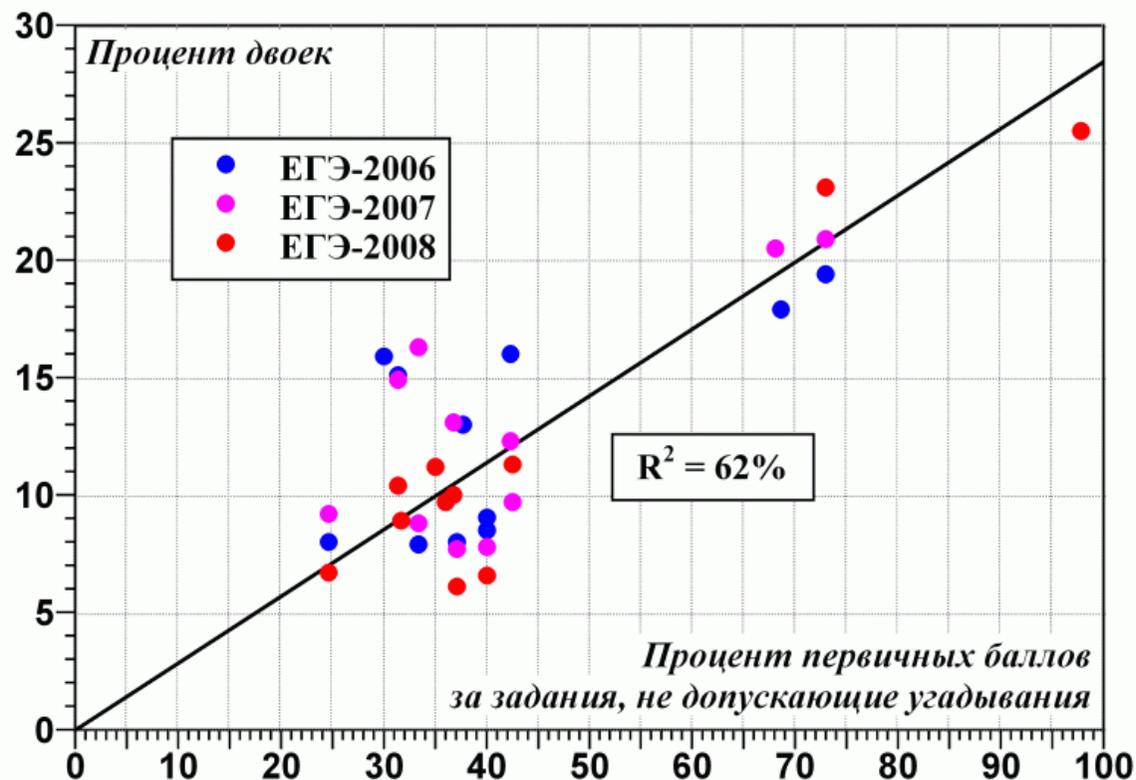


Рис. 3. Зависимость доли неудовлетворительных оценок от стоимости полноценных заданий (ФИПИ)

Сплошная линия – линейная регрессионная зависимость для данных за все три года, проведенная через начало координат. Точки в правой верхней части рисунка соответствуют егэ по математике и литературе, а в левой нижней части – по остальным предметам.

Особо следует отметить егэ по литературе, в кимах которого с 2008 г. не стало части А, где ранее разыгрывалось почти 30% первичных баллов. В результате этого литература смогла обойти математику по двойкам (самая правая и верхняя точка).

При объединении данных за три последних года коэффициент детерминации составляет 62% (см. рис. 3). Это достаточно много, чтобы утверждать, что сравнение результатов егэ по разным предметам позволяло характеризовать не различия в качестве их усвоения школьниками, а лишь различия в структуре кимов.

Егэ-2008. Результаты экзамена и особенности кимов

Предмет	Двоек (ФИПИ)	Первичных баллов на тройку	Стоимость заданий-угадаек	
			части А	частей А и В
Литература	25,5%	8,9%	0,0%	2,2%
Математика	23,1%	16,1%	33,3%	33,3%
Информатика	11,3%	27,5%	50,0%	57,5%

Иностранный язык	6,6%	31,0%	28,0%	60,0%
Обществознание	6,1%	30,6%	48,4%	62,9%
История России	10,0%	23,5%	47,1%	63,2%
Физика	9,7%	24,0%	60,0%	64,0%
Русский язык	11,2%	30,0%	51,7%	65,0%
География	8,9%	26,7%	51,7%	68,3%
Химия	10,4%	23,9%	44,8%	68,7%
Биология	6,7%	24,6%	52,2%	75,4%

Егэ в штатном режиме: 2009 год

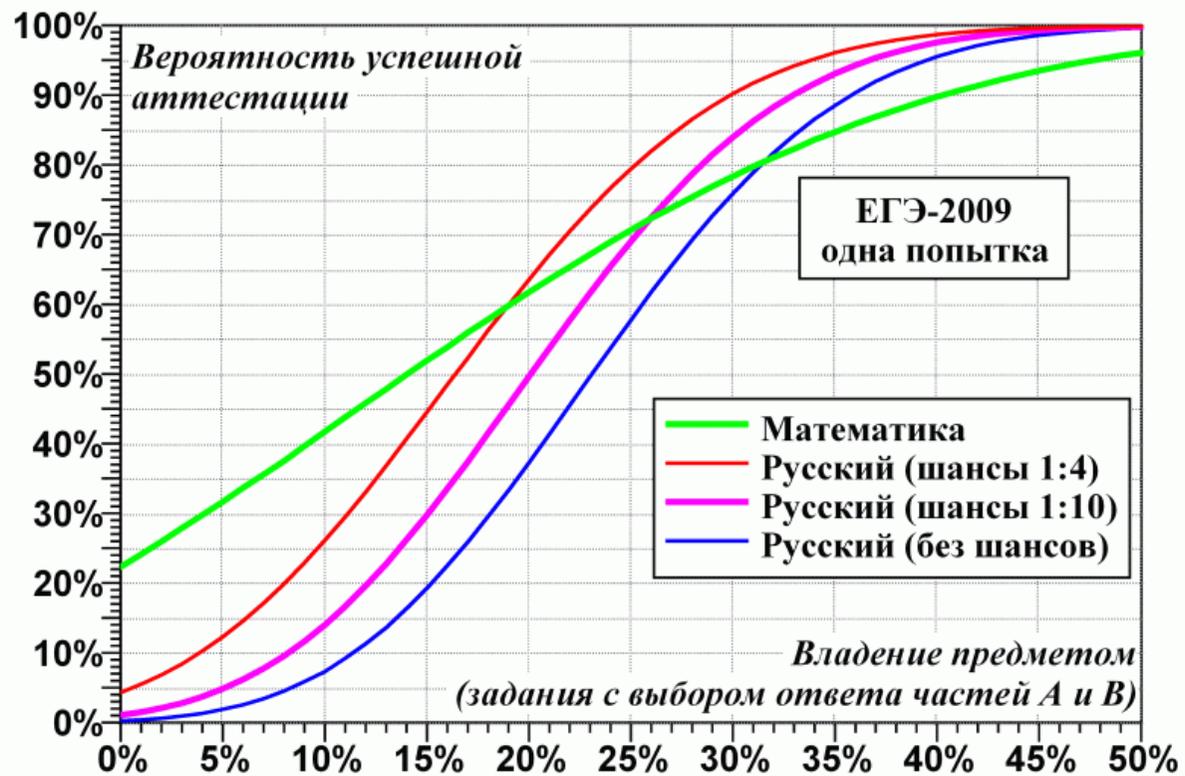
В 2009 г. егэ по русскому языку и математики сдавала вся страна, что, казалось бы, позволяет судить о состоянии дел в российском образовании хотя бы в разрезе этих двух предметов. Попробуем понять, что же означает подтверждение освоения «основных общеобразовательных программ среднего (полного) общего образования в 2009 году» по этим предметам, без чего нельзя было получить аттестат.

Контрольно-измерительные материалы по математике в том году содержали 10 заданий части А с выбором ответа из четырех предложенных вариантов. При этом достаточно было дать всего 4 верных ответа, чтобы избежать двойки. Если не знать совсем ничего и расставлять крестики наугад, то пройти такую итоговую аттестацию можно примерно в 2 случаях из 9. Однако если знать хотя бы что-то, вероятность успеха можно существенно повысить.

Определим *владение предметом* как долю тех заданий с выбором ответа, которые экзаменуемый может осилить, и будем считать, что иначе он попросту гадает. В реальности, конечно, не всегда имеет место четкое разделение на «знаю» или «гадаю», а часто происходит отбрасывание некоторых неверных вариантов ответа и случайный выбор между оставшимися вариантами. Однако математически эти ситуации однозначно связаны. Так, если экзаменуемый справляется с заданием с вероятностью p , а в противном случае случайно выбирает один вариант ответа из n предложенных, то в среднем это эквивалентно уменьшению числа возможных вариантов до

$$n' = \frac{n}{(n-1)p + 1}$$

Рис. 4 демонстрирует, как быстро вероятность подтвердить освоение общеобразовательной программы по математике, увеличивается по мере роста владения предметом. Видно, что крайне слабые знания обеспечивают довольно высокую вероятность успешной аттестации, которая достигает 50% при владении предметом на уровне всего 14%.



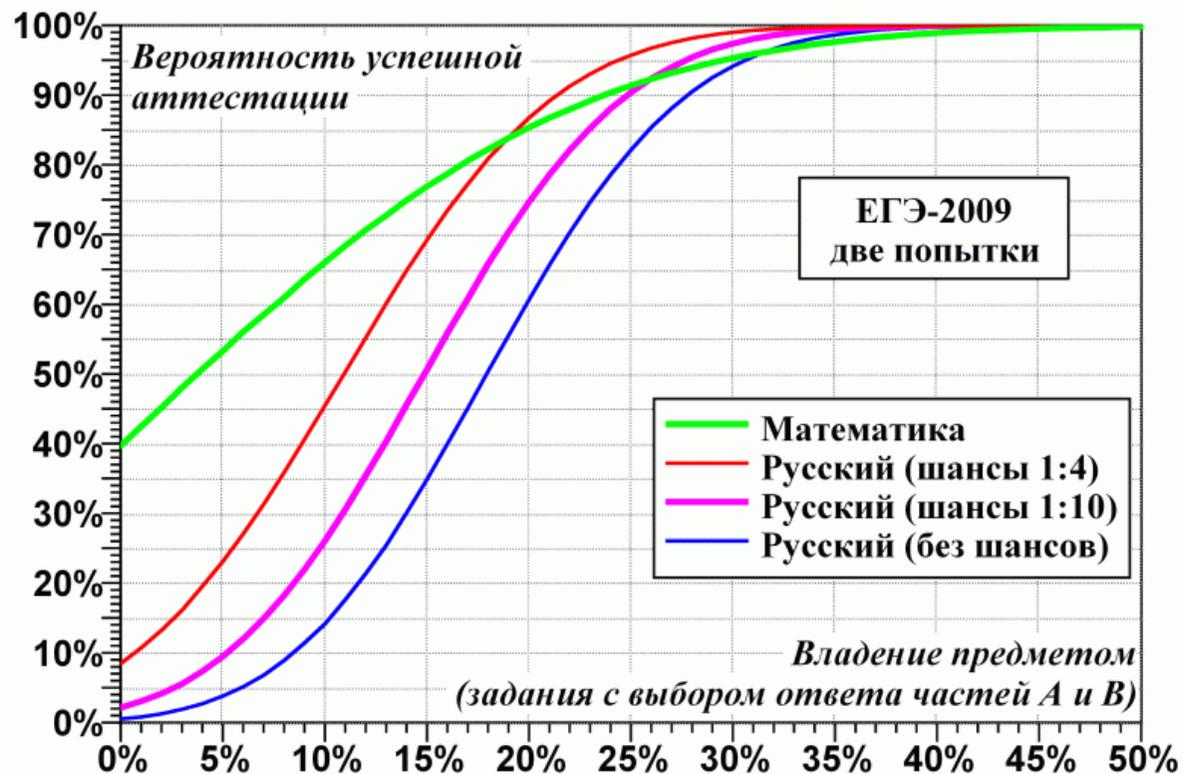


Рис. 4. Вероятность успешной аттестации по математике и русскому языку в зависимости от владением предметом

Для математики рассматривается только выполнение части А, а для русского языка – частей А и В с различными гипотезами относительно шансов на угадывание правильного ответа для заданий части В.

Учет возможности второй попытки связан с правом пересдачи предмета, освоение которого экзаменуемый не подтвердил с первого раза.

Ситуация с егэ по русскому языку не столь однозначна. Здесь почти нереально получить зачет, только расставляя наугад крестики, поскольку в 2009 г. было необходимо набрать 15 баллов, а в части А насчитывалось всего лишь 30 (однобалльных) заданий. Однако в кимах по русскому языку ответ на задания части В, в которой разыгрывалось еще 9 баллов, тоже выбирался из приведенного набора вариантов (в отличие от математики, где ответы в части В нужно получать самостоятельно) [31]. Эти варианты заведомо неравноценны, а их число непостоянно, в силу чего нельзя достоверно определить шансы на успех при угадывании ответов для заданий части В, но можно сформулированы следующие три гипотезы:

- *оптимистичная* – шансы угадать такие же, как и для части А, т.е. 1:4;
- *реалистичная* – шансы угадать примерно 1:10;

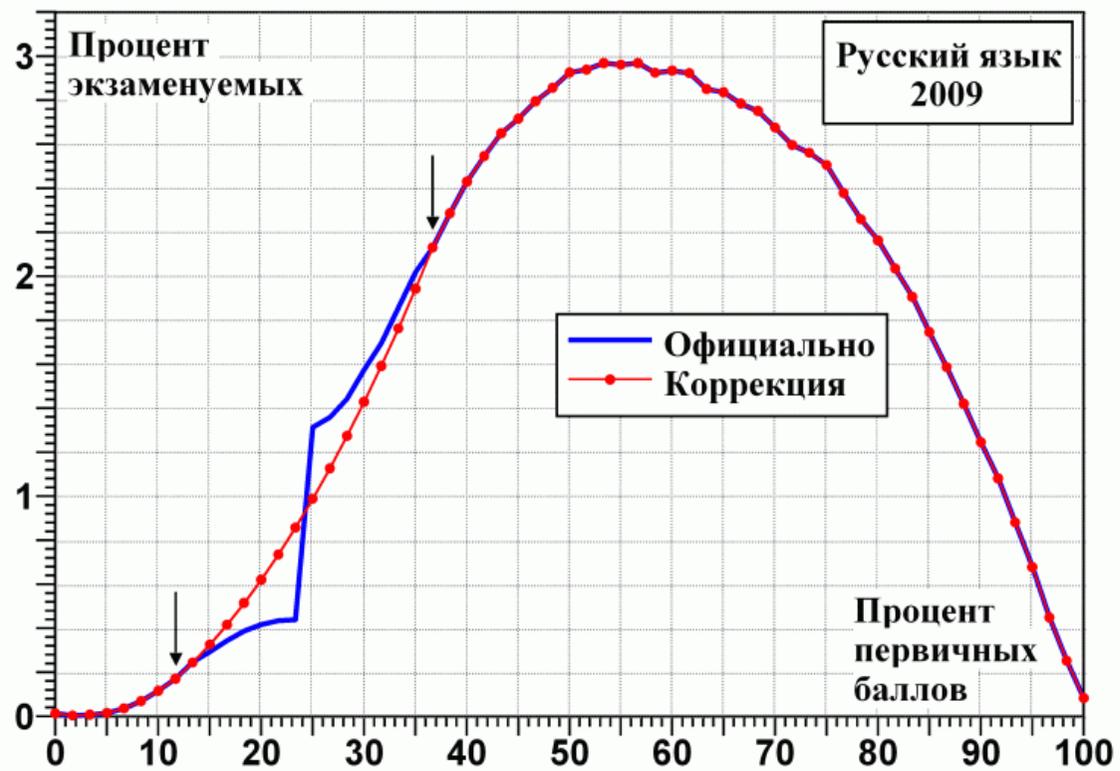
- *пессимистичная* – шансов угадать нет вообще.

Как видно из рис. 4, переход от оптимистичной гипотезы к пессимистичной эквивалентен изменению владения предметом примерно на $5\div 7\%$, т.е. особой разницы между крайними гипотезами нет, и разумно опираться на промежуточную реалистичную.

Зависимости для математики и русского языка качественно довольно похожи, но количественно всё же различны, поэтому в простейшем случае уместно ориентироваться на точку пересечения графиков. Из ее положения можно сделать общий вывод, что 70% вероятность подтверждения освоения основных общеобразовательных программ по этим предметам обеспечивается владением ими на уровне лишь в 25%. А уровень владения в 33,3% (здесь $n' = 2$, что соответствует пресловутому «попросим компьютер убрать два неверных ответа из четырех») поднимает вероятность успеха и вовсе до $80\div 90\%$.

Ценность подобной, с позволения сказать, аттестации близка к нулю. Однако организаторы егэ не были бы собой, если бы ограничились достигнутым. Правилами проведения единого государственного экзамена 2009 г. допускается однократная передача неудовлетворительной оценки. Возможность попытаться счастья еще раз существенно повышает шансы на благоприятный исход (см. рис. 4), что не удивительно в ситуации, когда результат сдачи в основном зависит не от знаний, а от везения. За счет повторного бросания костей подтвердить освоение программы по математике при нулевых знаниях удастся уже в 2 случаях из 5. Владение русским языком или математикой на уровне 25% обеспечивает аттестацию с вероятностью свыше 90%, а на уровне 33,3% – свыше 97%.

Передача проваленного егэ имела одно неожиданное пиар-следствие. На рис. 5 приведены распределения экзаменуемых по первичному баллу. Аналоги этих графиков, широко растиражированные в печати, принято трактовать как свидетельство фальсификации результатов экзамена. И впрямь, если выпускникам, которым чуть-чуть не хватает до тройки, набросить недостающие баллы, то график распределения слева от порога просядет, а справа – приподнимется. Однако к точно такому же эффекту приводит и легальная передача двоек. Так что дело здесь не столько в том, что результаты егэ были сфальсифицированы, сколько в том, что сам егэ изначально является фальсификацией измерений в сфере образования. Понятно, что если есть ненулевые шансы пройти аттестацию только за счет везения, то остается всего лишь сделать достаточное количество попыток. Коррекция графиков на рис. 5 позволяет оценить долю исправленных за один раз двоек на уровне 28% по русскому языку и 34% по математике (ОИП) [32]. Надо полагать, что если бы разрешили попробовать сдавать экзамены еще по несколько раз, то двоек не осталось бы вовсе.



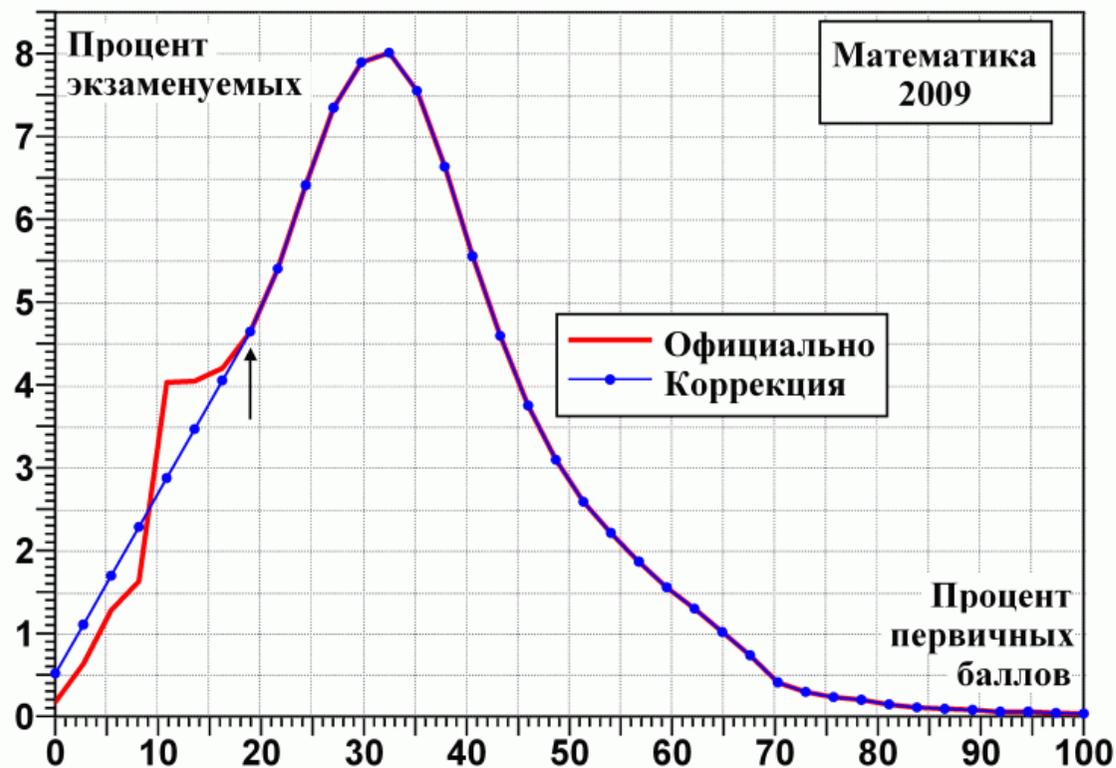


Рис. 5. Распределение участников егэ по русскому языку и математике по набранному первичному баллу (ОИП)

Оба графика демонстрируют практически вертикальный взлет, приходящийся на точку, отделяющую двойку от тройки. Такое поведение свидетельствует о чудесном превращении некоторого количества двоечников в троечников. Замена «испорченных» участков графиков гладкими кривыми дает 4,2% двоек для русского языка и 5,6% для математики против официальных значений в 3,0% и 3,7%, соответственно.

В качестве корректировочных кривых взяты парабола (для русского языка) и прямая (для математики), проведенные через опорные точки, помеченные стрелочками, и сохраняющие интеграл. Гладкая стыковка графиков позволяет предполагать удовлетворительное качество реконструкции.

Всё течёт туда, где ничего не меняется

Завершая обсуждение пригодности единого государственного экзамена для единообразной аттестации наиболее слабых выпускников, следует лишней раз обратить внимание на патологическое стремление организаторов егэ сбалансировать предметы по проценту двоек. Казалось бы, простейшим решением здесь было уравниванием доли заданий-угадаек в кимах по разным предметам.

Однако егэшники не ищут легких путей даже в деле очковтирательства. Поэтому ставка была сделана на снижение порогов получения выпускной тройки по наиболее проблемным предметам (см. табл. выше). Так, в 2008 г. для этого на

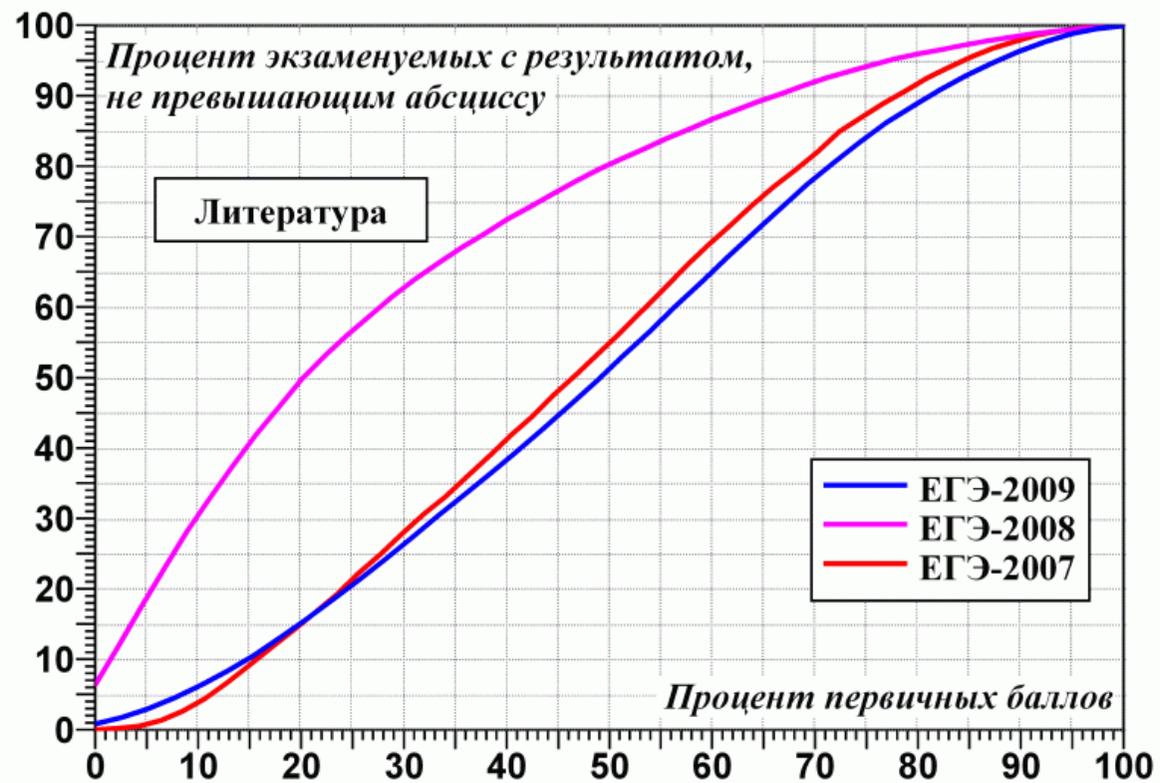
егэ по литературе было достаточно набрать 9%, а по математике [33] – 16%. В тоже время на егэ по остальным предметам порог получения выпускной тройки составлял в том году 24÷31% первичных баллов. Впрочем, уменьшение минимальных требований по математике и литературе не особо помогло – отношение суммарной стоимости заданий-угадаек к троечному порогу для этих предметов всё равно оставалось самым низким, а процент двоек – самым высоким.

В 2009 г. порог для тройки по математике был опущен в полтора раза, что, наконец-то, позволило свести количество двоек до общественно приемлемого уровня. А вот с литературой в том году произошла вещь в высшей степени нетривиальная. Снижать порог тройки здесь было уже невозможно [34]. И порог был поднят до уровня, даже более высокого чем у математики (15% разыгрываемых первичных баллов против 11%). При этом процент двоек парадоксальным образом тоже сократился.

При более внимательном анализе ощущение чуда только усиливается. В 2008 г. экзаменуемые справлялись с 34÷52% заданий части В кимов по литературе, а в 2009 г. – с 62÷79% (ФИПИ). Эта часть не претерпела никаких существенных изменений ни по структуре, ни по содержанию. Тем не менее, самое простое ее задание оказалось в 2008 г. субъективно сложнее, чем самое сложное в 2009 г.

Ситуация начинает проясняться, лишь если принять во внимание сокращение с 8,4% в 2008 г. до 6,2% в 2009 г. доли выпускников, сдававших егэ по литературе [35] (ОИП). Кроме нее уменьшение доли сдающих егэ, хотя и не столь значительное, продемонстрировали только география [36] (с 4,5% до 3,6%), математика (с 103,8% до 94,2%) и русский язык (с 109,4% до 99,4%), а для остальных предметов этот показатель возрос и порой весьма значительно. То, что произошло с егэ по литературе, лучше всего описывается притчей «купи козу – продай козу»: как только выпускной экзамен по этому предмету стал не обязательным, выпускники шарахнулись от него, как черт от ладана [37].

Изменение субъективной сложности егэ интересно не только в контексте подтверждения освоения предмета, но и само по себе. Судить о субъективной сложности экзамена можно по виду графика интегрального распределения экзаменуемых по первичному баллу. По сравнению с 2008 г. субъективная сложность единого государственного экзамена в 2009 г. уменьшилась только для русского языка, математики и литературы, вернувшись для всех трех предметов на уровень примерно соответствующий 2007 г. Но если в первых двух произошло лишь незначительное уменьшение сложности, связанное, скорее всего, с приспособлением образовательной системы к формату егэ, то в случае литературы, как можно видеть из рис. 6, оно было гигантским. По всем остальным предметам субъективная сложность егэ изменилась незначительно. Систематический ее рост на демонстрирует только информатика (см. рис. 6), для которой количество сдающих увеличивается наибольшими темпами. Ее сдавали 1,7% выпуска в 2007 г., 3,0% – в 2008 г. и 7,2% – в 2009 г. (ОИП).



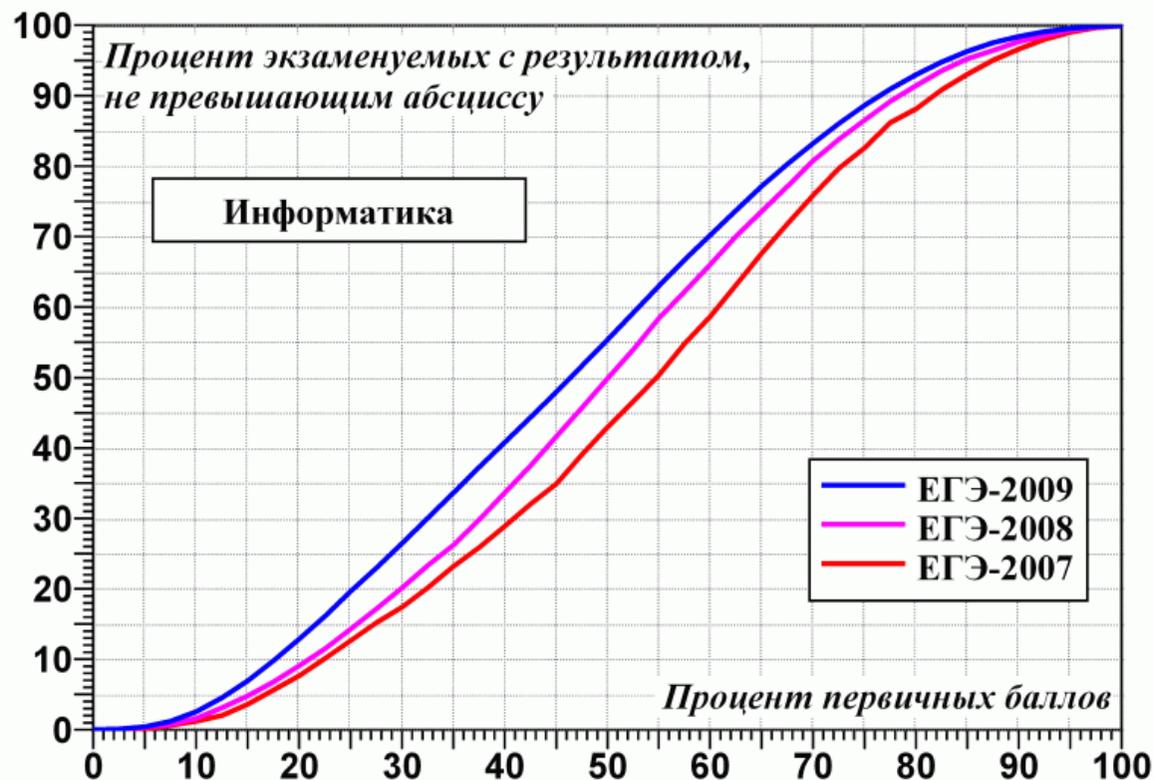


Рис. 6. Интегральное распределение участников егэ по литературе и информатике по набранному первичному баллу (ОИП)

Смещение кривой влево и вверх означает, что экзамен стал субъективно сложнее, а вправо и вниз, что – проще. Экзамен по литературе отыграл в 2009 г. резкий рост сложности в 2008 г., вызванный исключением части А. Экзамен по информатике, напротив, от года к году становится сложнее.

Впрочем, даже не имея графиков экзамена иногда возможно составить представление об изменении его субъективной сложности. В 2010 г. происходит эпохальное событие: из экзамена по математике полностью удалены задания-угадайки. Аналогичное изменение структуры кимов по литературе двумя годами ранее привело к резкому росту субъективной сложности егэ (см. рис. 6). И можно предсказать, что с математикой случится ровно то же самое. Это вызовет всплеск числа двоек, который не удастся демпфировать никаким снижением порога тройки.

Любые характеристики егэ, важные для его функционирования в качестве выпускного экзамена, варьируются год от года и предмет от предмета самым причудливым образом – в силу действия как объективных, так и субъективных факторов. Как при этом экзамен может не только считаться объективным измерителем чего бы то ни было, но и даже называется единым, понять решительно невозможно.

И ставит, и ставит им градусники

Поклонники егэ, кажется, пребывают в искренней уверенности, что одно его введение позволит что-то серьезно изменить в системе образования, причем непременно к лучшему. Однако о механизмах, посредством которых измерения в образовательной сфере, могли бы повлиять на ее состояние, они не очень задумываются. И хотя диагностика, не предполагающая лечения, выглядит странно и дико, это еще не означает, что она не оказывает влияния на исследуемый объект.

Состояние сложных систем, обладающих способностью к рефлексии, действительно может изменяться в результате одних только измерений. В случае егэ имеются два канала воздействия на процесс обучения, связанных, соответственно, с учителями и с учениками.

Уже упоминавшаяся идея оценивать с помощью егэ не только подготовку учеников, но и труд учителей, в дополнение к усилению коррупции в сфере образования негативно сказывается и на его уровне.

Качество преподавания измеряется прогрессом ученика, а вовсе не его результатами как таковыми, поскольку они зависят и от индивидуальных особенностей последнего, и от множества сопутствующих социально-экономических факторов, повлиять на которые учитель не в состоянии. Сложно сказать, возможно ли в принципе унифицировано и обезличено судить о работе учителя, но заведомо неадекватный способ оценки ее результатов вынуждает школу искать пути приспособления к новым требованиям. Одним из них является жульничество при тестировании, другим – вынужденное изменение целей преподавания.

Раньше учитель имел возможность держать тех, кто не может и не хочет учиться, на некотором минимуме, а на выпускных экзаменах просто накидывать им лишний балл. Это было, конечно, очковтирательством, но зато позволяло сосредотачивать основные усилия на тех детях, которых можно чему-то научить. Теперь учителю, лишенному возможности повлиять на результаты экзамена, приходится либо принимать участие в коррупционных схемах, целью которых является перераспределение уже отнюдь не учебного времени, либо тянуть всех учеников одновременно, натаскивая их на решение тестов, чтобы улучшить оценку своей работы. При этом ее продуктивность резко сокращается, т.к. слабые ученики незначительно прибавляют в достижениях, а сильные лишаются возможностей развивать способности.

Кроме того, неизбежен отказ от преподавания тех сторон предмета, которые не проверяются в рамках единого государственного экзамена, чтобы уделить больше времени проверяемому. Здесь каждый легко может придумать множество примеров.

Егэ по литературе не является обязательным (и вряд ли таковым когда-нибудь станет, поскольку разрыв между тем, что составляет задачи и содержание литературы в школе, и тем, что спрашивается по этому предмету на егэ, совершенно запределен)? Значит, программные часы литературы передаются учителями-словесниками на усиление русского языка, егэ по которому обязателен.

Геометрические и текстовые задачи, присутствующие в егэ по математике, в течение ряда лет влияли только на

вступительный балл, но не на выпускной, по которому оценивается качество работы учителя? Значит, эти разделы школьной математики, составляющие ее суть, остаются за бортом [38].

Навыки постановки эксперимента и обработки его результатов на егэ проверить нельзя? Значит, лабораторный практикум по физике и химии приносится в жертву натаскиванию на экзамен (и это даже очень удобно, если принять во внимание дефицит оборудования и реактивов).

Навыки устной речи, умение слушать и вычленять смысловые единицы из речевого потока тоже, вроде бы, больше не нужны? Значит, и их наработкой можно пренебречь при преподавании литературы, истории, обществознания.

Использование справочной литературы и источников на экзамене не допускается, чтобы облегчить разработку кимов, наполненных вопросами, апеллирующими к знанию, а не к пониманию? Значит, работе со справочниками и поиску источников можно больше не учить.

Устных экзаменов больше нет? Значит, умение вести диалог и последовательно отстаивать свою точку зрения, жизненно необходимое для гуманитарных дисциплин, более не актуально.

Etc, etc, etc...

Еще один немаловажный фактор, воздействующий на учителя, проистекает из низкой предметной квалификации разработчиков кимов. Неумение составить изящные содержательные задания они компенсируют расстановкой разного рода ловушек, предъявлением изощренных требований к оформлению и другими приемами, замещающими проверку предметных знаний проверкой обстоятельств, к предмету никакого отношения не имеющих. В результате существенную часть учебного времени приходится тратить на освоение методов преодоления рукотворных препятствий, а не на изучение самого предмета. Фактически, сдача егэ превращается в новый предмет, а точнее, спектр предметов, не предусмотренных школьной программой, но отнимающих время, необходимое на ее освоение. Единый государственный экзамен не столько является инструментом контроля и измерения в сфере образования, сколько фактором, определяющим его форму и содержание.

Ориентиры ученика также меняются вследствие того простого обстоятельства, что форма контроля определяет и форму учебной деятельности. То, что не контролируется, становится необязательным. А егэ в принципе не позволяет проверить целый ряд знаний, умений и навыков.

Егэ требует знания определенного набора фактов, но практически не требует их анализа, умения видеть их взаимосвязь и понимать структуру изучаемой дисциплины. Тем более он не позволяет выявить навыки рефлексии и объяснения своих знаний и умений, неплохо проверяемые традиционными устными экзаменами. Соответственно, школьники, обучение которых «заточено» под егэ, скорее всего, не смогут стать учителями или исследователями. Ровно то же самое можно сказать и в отношении умения воплощать теоретические знания в практику, проверяемого экспериментальными заданиями опять же в рамках устного экзамена.

Наконец, егэ, становящийся единственной целью учебного процесса, прямо подавляет естественную любознательность

ученика, стимулируя его ограничивать свои знания четко заданным набором того, что может войти в тест. Ориентация на ЕГЭ означает утрату навыка развернутого ответа, умения конспектировать или изменять формат представления информации. Знания и информация из важной цели обучения превращаются в средство сдачи тестов.

Отдельно следует отметить такую жуткую особенность ЕГЭ, как форма организации его тестовой части, выполняя которую, экзаменуемый должен в каждом задании поставить крестик против правильного ответа из заданного набора. При этом он пытается отбросить заведомо неверные ответы, уловить логику чередования правильных ответов или угадать их – делать всё, что угодно, но только не рассуждать. А любому человеку, занимавшемуся преподаванием, прекрасно известно, что стоит ученику позволить гадать, он мгновенно и необратимо прекращает думать.

Не стоит, однако, полагать, что тестовые формы проверки несовместимы с нормальным учебным процессом. Вовсе нет. Они по-своему и удобны, и полезны, если исключить элемент угадывания. Для этого достаточно всего-навсего заменить задания с одним правильным ответом (американка) на задания с заранее неизвестным числом правильных ответов, которые экзаменуемый должен отметить все (и только их), чтобы задание было зачтено (мексиканка). Как вариант, можно собирать задание из нескольких связанных по смыслу пунктов, в каждом из которых необходимо выбрать один правильный ответ, но начислять балл лишь тогда, когда по всем пунктам даны верные ответы. Второй тип заданий несколько менее интеллигентен, чем первый, зато проще в плане разработки и проверки. В любом случае, элемент угадывания был бы практически исключен.

Тот факт, что даже этой элементарной вещи до сих пор не было сделано, не оставляет сомнений в том, что ни преодоление негативного влияния ЕГЭ на образовательную систему, ни, тем более, его использование для каких-то позитивных изменений в ней не воспринимаются проводниками ЕГЭ как цели, заслуживающие внимания.

А ведь устранение других перечисленных разрушительных последствий введения ЕГЭ требует существенно более сложных мер.

Сомнительные приобретения и несомненные потери

Являются ли проделанный разбор дефектов единого государственного экзамена каким-то откровением? Ни в коем случае! Они, в массе своей, хорошо известны не только специалистам, но и всем людям, которым небезразлична судьба отечественного образования. Более того, многие из описанных проблем предсказывались экспертами задолго до перехода на ЕГЭ.

Подаются ли описанные дефекты исправлению? В большинстве своем, да, поддаются.

Будет ли это сделано? Вряд ли.

ЕГЭ и коррупция

Одной из публично декларируемых целей введения единого государственного экзамена является борьба с коррупцией в сфере образования. Однако, как будет показано ниже, переход на ЕГЭ не ослабляет, а лишь усиливает коррупцию и

структурирует ее, перенаправляя теневые финансовые потоки. И странно было бы ожидать, что бенефициары новых коррупционных схем добровольно откажутся от своих выгод, отвоеванных в тяжких боях с бенефициарами схем былых.

Излишние возможности должностного лица по принятию решений, которые принято считать причиной коррупции, в действительности являются лишь одним из ее механизмов. И никакие антикоррупционные меры, независимо от того являются ли они подлинными или притворными, не могут сократить платежеспособный спрос, имеющейся в той сфере, к которой они применяются. Бесполезно бороться с коррупцией без устранения ее причин, а только лишь путем усложнения коррупционных схем. Рост любых транзакционных издержек компенсируется за счет повышения стоимости услуг, и коррупционные услуги не являются исключением. Чем сложнее обходить закон, тем большую долю платежеспособного спроса покрывают именно они, и тем меньшая доля, соответственно, приходится на легальные услуги.

Платежеспособный спрос в сфере образования весьма велик, и его объем не может ни с того, ни с сего взять и уменьшиться. И если люди, не имея возможности приобретать нормальные образовательные услуги или не видя в этом смысла, находят приемлемым приобретение поступления, оценок и дипломов, то всегда найдутся те, кто продаст им требуемое.

Какие бы ни проводились образовательные реформы, те средства, которые родители способны тратить на образование своих детей или на связанные с ним взятки, они всё равно будут продолжать тратить. А воздействию поддается только соотношение этих двух компонент в расходах. Пока нет никаких оснований полагать, что переход на ЕГЭ может каким-либо образом усилить собственно образовательную компоненту, а значит, неизбежно будет усиливаться компонента коррупционная. А вследствие массовости ЕГЭ и его исключительно высокого социального статуса механизмы коррупционного перерождения идут особенно быстро.

ЕГЭ ведет к замене горизонтальных коррупционных механизмов вертикальными. Непрозрачная для власти система взяток, существовавшая на этапе перехода из школы в вуз, ныне вытесняется легко контролируемыми денежными насосами, действующими как на уровне школы, так и на уровне вуза.

На уровне школы коррупционные механизмы, вызываемые к жизни введением ЕГЭ, связаны с использованием его результатов еще и для оценки качества обучения в школах и в регионах. Из-за этого как преподаватели, так и местные власти оказываются активно заинтересованными не противодействовать нечестной сдаче экзамена. Скорее, наоборот, они готовы всячески потворствовать нарушениям или даже участвовать в них, но зачастую не бескорыстно, ибо далеко не все готовы идти на нарушение закона только из идейных соображений, ничего не получая за риск. К тому же, чтобы нарушения не выплывали наружу, необходимо как-то договариваться с проверяющими – за счет средств, собранных с детей и их родителей. Ну, а проверяющие, вынужденные отстегивать выше, будут всячески стимулировать этот сбор средств. В результате завышение показателей из цели махинаций быстро превратится в общую норму, отклонение от которой будет безжалостно наказываться. Не зря же проведение госэкзамена было передано в регионы на уровень местных органов управления образования – дань должна собираться на местах.

Кроме того, не будем забывать, что и у проверяемых, и у проверяющих, а также их друзей, родственников и знакомых есть дети, которым тоже приходится сдавать ЕГЭ. Поэтому неизбежно заключение пактов о ненападении (мы не ловим ваших, а

вы не ловите наших), по крайней мере, в отношении некоторых экзаменуемых. Это отяготит коррупционные схемы, адаптированные к контролю со стороны вышестоящих инстанций, блатными отношениями, действующими при контроле со стороны соседей (вывоз учащихся для сдачи егэ в другой район или приглашение для надзора учителей из другой школы), но не избавит от обязанности платить наверх.

В эту игру начинают включаться и ученики, и их родители. Для них егэ – это беда, с которой надо справиться, по возможности, быстро и эффективно. И если преференции для поступления в вузы покупаются в индивидуальном порядке, то вопрос с прохождением итоговой аттестации может решаться и коллективно. Оптом – дешевле, чем в розницу, а организовано – проще, чем дикарями. Экзамен же единый! Поэтому родители, интересующиеся у работников школ, по сколько скинутесь, чтобы у детей не было проблем со сдачей егэ, уже воспринимаются как норма.

На уровне вуза спектр новых коррупционных механизмов богаче. Нужно понимать, что если вуз «гнилой» и в него принимают за взятки, то коррупция там не исчезнет и после того, как его нечистых на руку преподавателей лишат возможности манипулировать результатами вступительных экзаменов. В «гнилом» вузе качественно учить не могли, а, соответственно, и не начнут, даже если прием вдруг станет честным. В нем станут зарабатывать на неотчислении уже набранных студентов, сделав невозможной сдачу сессий без посещения дополнительных платных занятий (что из-за объективно низкого уровня подготовки студентов, набранных по егэ, не представляет особой проблемы). Разумеется, деньгами, полученными таким путем, надо будет делиться с контролирующими инстанциями, чтобы у тех не возникало лишних вопросов.

Другая форма коррупции на уровне вуза связана с получением эксклюзивных разрешений тем или иным способом обходить егэ при наборе студентов в силу творческой специфики вуза, его особого значения для страны и иных столь же субъективных причин. Как вариант могут применяться и непубличные методы уклонения от егэ. Можно, скажем, не проверять у некоторых абитуриентов соответствие указанных ими баллов тем, которые они реально набрали. А можно оттеснять абитуриентов со средними вступительными баллами, официально объявляя высокий проходной балл, а реально зачисляя «правильных» людей с низкими баллами... Способов, как полулегальных, так и откровенно криминальных, много. Но возможность их спокойно использовать дается, конечно, не даром. И даже вполне приличные вузы, изначальная цель которых заключается в возможности набирать способных студентов, будут вынуждены изыскивать и возможность зарабатывать на приеме, чтобы было на что выкупать право независимой оценки абитуриентов.

Не следует, однако, думать, что чисто экзаменационная компонента коррупции исчезает с переходом на егэ. Она лишь перемещается с собственно сдачи егэ на апелляцию. При проверки части С имеется масса возможностей начислить меньше баллов, чем объективно заслужил экзаменуемый. И даже самый толковый ребенок не сумеет вернуть свои баллы, если ему заявляют, например, что раз задание выполнено не так, как в официальном образце, присланном для проверки работ, значит, неправильно. А как действовать в ситуации, когда возвращают один–два балла и говорят, что больше нельзя, сообразит не любой взрослый. Это ведь не так пугающий сторонников егэ устный экзамен, на котором «валят», – это уже разборки совсем не детского уровня. Поэтому успех на апелляции будет сопутствовать лишь тем, кто либо придет на нее в сопровождении опытного репетитора или иного предметного профессионала (на которого должна быть своевременно оформлена доверенность от родителей), либо, не мучаясь, «занесет», куда следует. Впрочем, эти два

алгоритма можно и объединить, сразу нанимая в качестве представителя ребенка лицо, аффилированное с апелляционной комиссией, – просто и надежно.

Бессмысленно пытаться предсказать и описать все формы, которые может принять коррупция. Народ всегда оказывается изобретательней власти, ибо, как мудро заметил один известный российский политик, «тех, кто пишет законы, – тысячи, а тех, кто думает, как их обойти, – миллионы». Так, в 2009 г., ко всеобщей неожиданности, золотой жилой стали разные исключения из правил приема – для сирот, инвалидов и иных льготных категорий, в которые массово (и безвозмездно) стали записываться абитуриенты. В результате огромное количество сегодняшних студентов получило этот статус вне конкурса. А на нескольких факультетах МГУ им. М.В. Ломоносова и вовсе не оказалось ни одного студента, поступившего на общих основаниях, – все места были заняты льготниками... Несомненно, данная лазейка будет рано или поздно взята под контроль, но, как известно, вода дырочку найдет, и можно не сомневаться, что вскоре будут протоптаны новые народные тропы в обход установленных властями шлагбаумов.

В заявления руководства Рособнадзора о борьбе с фальсификацией результатов егэ может поверить только тот, кто не в курсе, что стобалльный результат с гарантией прохождения любых проверок стоит на порядок дороже тех же ста баллов, но без всяких гарантий [39]. А если знать цены, становится очевидно, что борьба ведется не против коррупции, а против демпинга на рынке коррупционных услуг. И когда Л.Н. Глебова объясняет воронежцам, что в показанных в их регионе результатах, по всем предметам превышающих средние по стране, она видит «отсутствие контроля и питательную среду для развития коррупции» и имеет «основания считать, что... имели место коллективные решения с вовлечением в этот процесс большого количества людей» [40], понимать это можно только как «Не по чину берете, ребята!».

Главным результатом введения единого государственного экзамена становится выстраивание коррупционной вертикали власти в системе образования. При этом совершенно неважно, сформировался ли этот результат путем утряски множества частных интересов или он и был истинной целью проводников системы егэ. Можно исходить из представлений о невидимой руке бюрократического рынка или из теории заговора – результат получается один и тот же. Злоупотребления из осуждаемых и стихийных становятся санкционированными и строго обязательными. Когда эта систем окрепнет, уничтожить ее будет невозможно, даже отменив егэ.

Борьба невежества с несправедливостью

Весь социальный пафос, призванный преподнести обществу егэ как средство борьбы с коррупцией и обеспечения равных возможностей получения образования, представляет собой не более чем дымовую завесу, скрывающую истинные процессы. Тем не менее, результаты вдыхания этого дыма поистине удивительны и заслуживают отдельного разбора.

Если внимательно посмотреть, как разделилась активная часть общества по отношению к егэ, то можно подметить следующую нетривиальную закономерность. Группу его последовательных противников составляют, в первую очередь, школьные учителя, вузовские преподаватели, опытные репетиторы, квалифицированные научные работники, выпускники лучших вузов и школ с углубленным изучением предметов, родители, стремящиеся дать своим детям качественное образование, – словом, все те, кто много и хорошо учился или учил. А в группу восторженных сторонников егэ

преимущественно входят люди, затронутые образовательным процессом лишь по касательной, но ратующие за социальную и иную справедливость, иногда понимаемую ими в отрыве от всяких реалий, на уровне абстрактных лозунгов.

Временами дискуссия и вовсе переходит из области социологии в область психологии, чтобы не сказать, психиатрии. И тогда становится очевидно, что спор между противниками и сторонниками ЕГЭ – это чаще всего спор между людьми разного опыта, а вовсе не разных взглядов или интересов. Это спор между теми, кто надеется сохранить лучшее, что мы имели в системе образования, и теми, кто пытается именно лучшее разрушить. Те, кому в учебе выпала удача, ориентируются на хорошее, понимая, что оно возможно и воспроизводимо, а те, кому удача не выпала, – на плохое, стремясь отыграться за свои беды на других. Психологически это понятно – каждый человек стремится к воспроизведению своего опыта, каким бы он ни был. Но когда личные комплексы (да еще и отягченные криминальной компонентой) становятся основой для принятия решений государственного масштаба, этого ни понять, ни принять нельзя.

В обществе нет почвы для конструктивного диалога по вопросу ЕГЭ – к нему относятся по-разному не в силу различия во взглядах или интересах, а в силу различия в глубине понимания вопроса. Маловероятно, что столь отчетливый раскол, линия которого пролегает между экспертами и теми, кто не хочет прислушиваться к их точке зрения, мог возникнуть без целенаправленного манипулирования общественным мнением, призванного скрыть некие неблагоприятные цели.

У нас в стране нет организованных, осознающих себя сил, которые бы стремились противодействовать образовательной коррупции. В ее искоренении объективно заинтересовано общество в целом, субъективно – отдельные личности, но ни общество не может консолидироваться, ни эти личности – найти общий язык. И одной из важнейших причин неспособности объединиться перед лицом коррупции как раз и является выбор заведомо негодных методов борьбы с нею, помноженный на их агрессивную пропаганду. При переходе на ЕГЭ никто не интересовался ни нуждами системы образования, ни мнением участников образовательного процесса. Стоит ли удивляться, что непродуманные и насильственно навязанные преобразования не только не сплотили общество перед лицом давно назревших проблем, а, напротив, расколоты его?

Еще один важный элемент ЕГЭшного пиара заключается в педалировании предоставляемой им возможности выявления талантливых детей в провинции, у родителей которых нет средств, чтобы отправить чадо сдавать экзамены в крупный город. Это классический пример полуправды, которая хуже лжи. И дело даже не в том, что без ответа остается пустяковый вопрос, откуда появятся у такого студента деньги, чтобы после поступления несколько лет существовать в крупном городе, пока он не выучится и не встанет на ноги. Куда важнее другое. Для поддержки талантов, обитающих в медвежьих углах, где, к сожалению, нет надежды на нормальное школьное образование для всех, нужны не общие, а очень частные меры: заочное и дистанционное обучение, гранты и стипендии для одаренных детей, приглашение их на олимпиады и в учебные лагеря. Вряд ли ЕГЭ может помочь реализации этих мер.

Нужно обеспечивать равные возможности по получению образования школьникам, которые хотят учиться. Вместо этого предполагается предоставить якобы равные возможности по поступлению тем из них, кто уже как-то выучился. При этом

вопрос, как им это должно удаваться, видимо, является, с точки зрения проводников системы егэ, второстепенным.

А судьи кто?

Сколько бы ни велась речь о непоправимо низком качестве егэ, независимую и авторитетную оценку тому, насколько хорошо или плохо он выполняет функции вступительного экзамена, могут дать только вузы. Только они являются единственными потребителями «услуги» по оценке абитуриентов.

Можно предложить два принципиально разных механизма получения вузовской оценки единого государственного экзамена, связанных, соответственно, с его замещением и дополнением.

Первый механизм обусловлен выбором вузом того, что ему лучше – проводить собственные вступительные испытания, неся связанные с их организацией затраты, или довериться результатам единого экзамена, рискуя получить студентов, уровень подготовки которых оценен неправильно. Принципиально, что признание вузами результатов егэ может как-то свидетельствовать о его адекватности только в том случае, когда вузы свободы в выборе форм конкурсного отбора. Однако такой свободы вузам – как государственным, так и негосударственным – не предоставляется.

Примитивное предположение, что данное ограничение обусловлено борьбой с коррупцией, критики не выдерживает. Во-первых, как показано выше, переход к приему по егэ образовательную коррупцию увеличил, а не уменьшил. Во-вторых, наиболее активно и последовательно против введения единого государственного экзамена выступали именно руководители благополучных вузов, в которых уровень коррупции сравнительно низок. Ну, а в-третьих, необходимость для негосударственных вузов признавать результаты егэ вообще не может быть объяснена антикоррупционными мотивами, т.к. в этих вузах просто нет бюджетных мест, преимущественный доступ к которым мог бы обеспечиваться взятками.

Тем не менее, прием на основе егэ обязателен для всех вузов, имеющих государственную аккредитацию, в т.ч. и для негосударственных [41]. Вряд ли такую практику возможно объяснить иначе, чем желанием ни в коем случае не допустить объективной оценки качества единого государственного экзамена. А такое желание лучше любой критики свидетельствует о том, насколько неудовлетворительно это качество.

Косвенно данное предположение подтверждается и тем, как проходило внедрение единого экзамена. Он распространялся по регионам, а не по вузам. Предположим, что эта система так хороша для конкурсного отбора, как нас пытаются убедить ее проводники. Тогда разумно было бы на стадии эксперимента по введению егэ предоставлять вузам право принимать по его результатам, а не обязывать их поступать так с выпускниками из тех регионов, где егэ уже введен. Расширение списка вузов, добровольно признающих результаты единого государственного экзамена, повышало бы его авторитет и способствовало дальнейшему его распространению. Требования по исправлению каких-либо недостатков, которые вузы могли бы выдвигать как условие признания егэ, позволяли бы совершенствовать систему его проведения. Ничего этого, однако, не произошло. Экзамен просто был навязан высшей школе России, но качественная услуга в навязывании не нуждается.

Второй механизм независимой оценки качества единого государственного экзамена связан с проведением вузами в дополнение к нему собственных вступительных испытаний. Сейчас такие испытания разрешено проводить не всем вузам и, что особенно важно, только по тем предметам, по которым не проводится егэ. Единственный смысл последнего ограничения как раз и состоит в том, чтобы скрыть истинное состояние дел с оценением уровня подготовки абитуриентов. Необходимость сделать процедуру единого экзамена человеконезависимой привела к тому, что он оказывается принципиально непригодным для проверки значительной части предметных знаний, умений и навыков, а его формализация разрушила систему профориентации школьников.

Следует разрешить всем вузам проводить собственные вступительные испытания в форме экзамена по профилирующим предметам [42], что сняло бы сразу множество проблем.

Во-первых, статистическое сопоставление результатов единого и вузовского экзамена по одному и тому же предмету позволило бы выявлять ситуации, когда первый или второй существенно искажают картину. Декорреляция результатов на уровне вуза была бы свидетельством того, что с вузовским экзаменом что-то не так, а декорреляция на уровне страны, – что что-то не так с единым (и именно боязнь обнаружения последнего обстоятельства, а вовсе не борьба с коррупцией, как нам представляется, является истинной причиной недопущения вузовских экзаменов).

Во-вторых, проведение собственных вступительных испытаний позволяет учесть специфику вуза, помогает ему отобрать «своих» студентов. А добавление во вступительный балл результатов дополнительного экзамена по профилирующему предмету, позволит выделить его на фоне прочих.

В-третьих, формат вузовского экзамена может быть существенно отличен от формата единого экзамена. Это возвратит интерес к формам учебной деятельности, которые егэ обрек на вымирание, и позволит проверять те знания, умения и навыки, для проверки которых егэ непригоден. Идеальным вариантом стало бы проведение вузовского экзамена в устной форме. Основным упреком в адрес устного экзамена всегда была его необъективность. Однако, с одной стороны, этой беде устные егэшные апелляции подвержены не в меньшей степени, чем традиционные устные экзамены, а с другой – возможность статистического сопоставления результатов единого и вузовского экзамена не зависит от формы проведения последнего.

Ну, и наконец, в-четвертых, егэ лишает абитуриента права на ошибку. Плохо сдав единый экзамен, он лишается возможности что-либо изменить до следующего года. Проведение вузовских экзаменов отчасти вернуло бы эту возможность.

Царь-зеркало

Как известно, Царь-пушка никогда не стреляла, а Царь-колокол никогда не звонил. Единый государственный экзамен, несомненно, достоин продолжить этот ряд как зеркало, которое ничего и никогда не отражало. Школьник получил на егэ какие-то баллы, которые влияют на его судьбу, но на самом деле ни о чем не свидетельствуют. Обществу предъявлена какая-то статистика, которая его взбудоражила или успокоила, хотя и она тоже не значит ровным счетом ничего...

Организация и проведение единого государственного экзамена требуют огромных усилий, но результаты оказываются в государственном масштабе ничтожны. А ведь в ходе егэ собирается огромный массив ценнейшей информации, которая далее толком никак не обрабатывается и не исследуется.

Если результаты экзамена объективны, то их можно и должно использовать для выработки и принятия обоснованных управленческих решений. Где основные болевые точки системы образования? Какое влияние оказывают те или иные принимаемые меры на его качество? Каким школам и регионам необходимо оказывать помощь? Никто не пытается этого узнать с помощью егэ. И не только потому, что нет доверия к нему как к инструменту измерений, но и потому, что само его введение не было продумано.

Напомним, что изначально единый государственный экзамен замышлялся лишь как дополнение к ныне забытой системе *государственных именных финансовых обязательств* (гифо), предназначавшейся для мягкой отмены 3 части 43 статьи Конституции РФ[43]. Потом выяснилось, что для достижения этой цели не обязательно придумывать столь сложную систему, а достаточно просто недофинансировать вузы, сокращая в них тем самым количество бюджетных мест. Тогда о гифо забыли, а егэ остался как система, не выполняющая ни одной государственной задачи. В результате объем информации о происходящем в системе образования, который ныне дает единый экзамен, качественно меньше, чем те искажения, которые он вносит в работу этой системы.

Помимо уровня принятия решений информация, которую позволяет собрать единый экзамен, может быть востребована на уровне учителей и методистов. Им важно понимать, какие разделы программы являются проблемными, как уместно перераспределить учебное время, на чем заострить внимание, каковы типичные ошибки, выявляемые при независимой проверке. Школа не получает этих сведений – в лучшем случае ей достаются сборники рассекреченных заданий егэ с объяснением алгоритмов их выполнения. Это подталкивает учителя к натаскиванию, а не к совершенствованию своей работы.

Выскажем в этой связи еще одно соображение общего плана. Любой экзамен (тестирование, аттестация) выполняет три различных функции. Первая – дать преподавателю возможность понять, насколько хорошо его ученики усвоили пройденный материал. Вторая – помочь ученику увидеть целостную картину пройденного, подтолкнуть его к систематизации и уточнению своих знаний. Третья – оценить достижения или перспективы ученика. Значение последней компоненты закономерно возрастает для экзаменов, сдаваемых на стыке школы и вуза, однако это не означает, что от двух первых следует полностью отказаться. Экзамен всегда должен в той или иной степени быть механизмом обратной связи, иначе он перерождается, переставая служить целям обучения и подчиняя себе учебный процесс.

Мы выбираем, нас выбирают...

Переход егэ в шпательный режим в 2009 г. породил вал заявлений, подаваемых абитуриентами сразу во множество вузов, что затрудняло работу приемных комиссий и не позволяло составить какое-либо представление о реальном конкурсе. Однако это лишь следствия еще одной серьезной деформации системы образования, вызванной введением единого экзамена. Суть ее в уничтожении существовавшей системы профессиональной ориентации школьников без создания каких-либо

альтернатив.

Опыт 2009 г. показал, что документы могут не просто подаваться одновременно на несколько разных специальностей, а на специальности, между которыми нет вообще ничего общего, кроме списка предметов, по которым необходимо сдавать егэ. Люди просто не знают, кем они хотят стать. Единый государственный экзамен превращает для абитуриента вузы и факультеты, каждый из которых имел свое лицо, в одну безликую высшую школу.

Традиционный вступительный экзамен по любому предмету в каждом вузе был немножечко свой. Поэтому абитуриент, выбирая конкретный вуз, сразу ориентировался на его специфику и готовился к поступлению на основе его экзаменационных материалов прошлых лет. Вуз, в свою очередь, тоже был заинтересован в том, чтобы найти «своего» студента, который не разочаруется в сделанном выборе. Поэтому издавались учебные материалы в помощь поступающим, организовывались лекционные курсы и кружки для школьников, проводились дни открытых дверей, создавались «придворные» школы с углубленным изучением профильных предметов и заочные школы для жителей удаленных районов.

Унификация экзамена разрушила этот алгоритм взаимной притирки абитуриента и вуза. Первому до самого момента подачи документов больше не нужно задумываться о том, куда пойти учиться. Однако и второму теперь нет смысла заботиться о довузовской подготовке школьников. Ее система, скорее всего, будет в ближайшее время свернута. Для провинциальных вузов она просто утрачивает смысл, поскольку все плоды трудов – лучших абитуриентов – у них отбирают столичные вузы. Но и последним становится незачем думать о возвращении своего контингента, поскольку к их услугам вся страна, – где-нибудь да возрастят.

Кому-то может показаться, что старая система была не столь уж и хороша. Традиционный экзамен – это мощный инструмент, предоставляющий большие возможности, которые можно использовать как во благо, так и во зло. И ничто не мешало вузу практиковать такой формат экзамена, к которому просто невозможно подготовиться без репетиторов, аффилированных с этим вузом. Основной смысл такого экзамена – обеспечить их работой.

Это действительно имело место, однако было совсем не таким распространенным явлением, как хотели бы нас убедить сторонники егэ. Массовое репетиторство было вызвано к жизни совсем не недобросовестностью отдельных вузов, а общей деградацией школьного образования. Поэтому раньше репетиторы не столько натаскивали, сколько учили и доучивали, ликвидируя разрыв между уровнем, обеспечиваемым средней школой, и теми требованиями, которые предъявляла высшая.

Теперь же репетиторы именно натаскивают на егэ. Причем благодаря его введению репетиторство только ширится (разные части кимов требуют, вообще говоря, разного натаскивания, чем порой занимаются разные люди, а вызываемое им дополнительное падение школьного уровня только повышает спрос на их услуги). Но никаких реальных знаний, занимаясь с репетитором, будущие студенты больше не получают и, прорываясь в вуз, оказываются просто не готовыми к получению высшего образования. Соответственно, их приходится спешно доучивать (не суть важно – самому вузу за счет часов учебного плана, или тем же репетиторам во внеурочное время), что, однако, сделать теперь намного труднее, поскольку упущены и время, и, что самое главное, возможность отобрать обучаемых. Для многих специальностей

начинать готовить человека только с момента поступления уже поздно, начинать надо до, а лучше – задолго до поступления. Именно эту возможность и уничтожает егэ.

Особенно печальна судьба школьных олимпиад, на протяжении десятилетий бывших незаменимым элементом отечественного образования. Их смысл, в первую очередь, не соревновательный, а просветительский – возбудить интерес школьников к предмету, дать им возможность понять свои возможности и почувствовать свои силы, ободрить способных независимой оценкой, подарить призерам умные книжки... Необходимость найти хоть какую-то альтернативу егэ ныне превратила олимпиады в замаскированные экзамены. Не возьмемся судить, насколько хорошо олимпиады справляются с этой задачей, но профориентации школьников они больше не служат.

Выводы

Фундаментальный порок единого государственного экзамена связан с тем, что его пытались приспособить для достижения одновременно целого ряда совершенно различных целей. Он был призван стать и итоговой аттестацией для выпускников, и вступительным испытанием для абитуриентов, и инструментом мониторинга системы образования, и средством контроля объективности школьных оценок. Однако при решении любой конкретной задачи универсальный инструмент неизбежно проигрывает специализированному (так, гусь летает, но значительно хуже орла, поет, но значительно хуже соловья, плавает, но значительно хуже пингвина, и бегает, но значительно хуже страуса).

В частности, итоговая аттестация и вступительное испытание оказываются несовместимы по целому ряду причин, перечисленных в таблице.

Несовместимость итоговой аттестации и вступительного испытания

	Итоговая аттестация	Вступительное испытание
Тип испытания	Тест достижений (взгляд в прошлое)	Тест способностей (взгляд в будущее)
Цель испытания	Выявление худших	Выявление лучших
Участие в экзамене	Обязательное	Добровольное
Оценка подготовленности	Абсолютная (сравнение с некоторым стандартом)	Относительная (сравнение с другими экзаменуемыми)
Тип теста (по решаемой задаче)	Критериально-ориентированный (проверка усвоения учебного материала)	Нормативно-ориентированный (ранжировка или сравнение испытуемых)
Содержательный смысл балла	Доля усвоенных знаний, умений и навыков из обязательного набора	Доля обойденных конкурентов
Необходимая шкала оценок	Нормированная (определены и цена каждого балла, и начало отсчета)	Метрическая (определена цена каждого балла, но не начало отсчета)

Обсуждать возможности совершенствования единого государственного экзамена без разделения его на независимые выпускной и вступительный бессмысленно, поскольку любые возможные улучшения в одной части означают ухудшения в другой.

Тему антикоррупционной роли егэ в его нынешнем виде вообще нельзя воспринимать всерьез. «Егэ против коррупции» –

это примерно то же самое, что «рок против наркотиков». И до тех пор, пока система егэ не будет сделана прозрачной и понятной, пока ее организаторы не начнут активный диалог с обществом, ситуация не изменится.

У проблемы егэ, кроме уровня функционирования средней и высшей школ, есть еще один уровень, о котором нельзя забывать при ее анализе. Это уровень национальных интересов и национальной безопасности. Все мы – ученики, учителя, их родственники и знакомые – прежде всего, граждане России. И нашей стране сегодня, как воздух, требуются квалифицированные кадры, чтобы удержать техносферу, созданную усилиями нескольких предшествующих поколений.

Цена квалификации специалиста, его профессиональных и морально-волевых качеств, а также цена его ошибки неизмеримо возросли. Результаты действий нескольких человек, обслуживающих, контролирующих, руководящих опасными объектами, могут иметь глобальный характер. Ускоряя деградацию российского образования, егэ ставит под вопрос будущее страны.

Невозможно организовать отдельную качественную подготовку кадров для критически важных профессий, если всех прочих превращают в коекакеров. Система образования – живое существо, которое не удастся разделить на две части, живущие независимо друг от друга. Это как сообщающиеся сосуды: будет понижаться уровень в одном, неизбежно упадет и в другом.

Кроме того, ущерб, которым может обернуться падение уровня в критически важных профессиях, вполне соизмерим с тем ущербом, которым обернется падение квалификации представителей профессий менее экстремальных, но зато более многочисленных – учителей, врачей, инженеров, – им мы повседневно доверяем свою жизнь и жизни своих близких. Да и просто деградация общего уровня общества – рядовых обывателей – сулит интегральные потери вполне соизмеримые с последствиями самых тяжелых катастроф. Вспомним, что максимального экономического успеха в XX веке добились те нации, которые в XIX веке имели наибольший процент грамотного населения.

Перед нашей страной сейчас стоит важнейшая задача новой модернизации, связанной со сменой технологических укладов. Привычные отрасли промышленности, еще недавно бывшие основой мощных экономик, сейчас отходят на второй план, освобождая место только возникающим отраслям, с созданием и развертыванием которых будет связано развитие тех стран, которые найдут свое место в мире завтрашнего дня. Это – шанс для России. Но чтобы его реализовать, необходимы массовые квалифицированные научные и управленческие кадры. Готовить их из жертв егэ – не стоит и пытаться.

Ситуация усугубляется тем, что мы попали в невиданную демографическую яму, уменьшающую всего за десятилетие вдвое число выпускников и поступающих. Каждый из них, по сути, штучный товар. Но сегодняшний абитуриент – это поручик Киже. Ходят бумаги, а людей никто не видел ни до поступления, ни во время. Кого учить?

Чтобы единый государственный экзамен перестал представлять опасность для общества и государства, необходимо, как минимум:

1. ограничить сферу применения егэ дисциплинами, возможна постановку содержательных вопросов, имеющих однозначный ответ;

2. разделить по каждому предмету один экзамен на два – общеобязательный тест достижений, без сдачи которого невозможно получение аттестата зрелости, и необязательный тест способностей, используемый только для поступления в вуз;
3. разрешить вузам проводить собственное вступительное испытание по профилирующим предметам, результаты которого будут суммироваться с результатами егэ;
4. отказаться от заданий с выбором одного ответа;
5. передать разработку кимов преподавательскому сообществу;
6. привести систему начисления баллов в соответствие с целями экзаменов;
7. отказаться от оценивания труда учителей по результатам егэ.

Отношение к егэ – это отношение нашего общества к своему будущему. Это экзамен на гражданскую зрелость.

[1] Поскольку эта аббревиатура уже вполне субстантивировалась и обросла производными, здесь и далее она пишется строчным буквами – точно так же, как аббревиатуры «вуз», «зун», «ким» и «гифо».

[2] По всей видимости, единственным более-менее позитивным аспектом егэ на сегодняшний день является восстановление в стране единого образовательного пространства.

[3] Это прямо констатируется в опубликованном в декабре 2009 г. отчете «Итоги деятельности Комиссии по изучению и оценке результатов эксперимента по введению единого государственного экзамена, созданной распоряжением Председателя Совета Федерации от 30 декабря 2008 года №605рп-СФ».

[4] Ни один вопрос, исчерпывающий ответ на который можно найти в справочнике, нельзя считать содержательным. Разумеется, для успешного выполнения заданий экзаменуемому нужно обладать некоторым объемом элементарных знаний на уровне правил, определений, дат, терминов, фактов и т.п. Однако демонстрация этих знаний не может и не должна быть предметом оценки. Мы по-настоящему знаем лишь только то, что умеем. Поэтому проверять следует не знание правил, а умение их применять, не знание определений, а способность соотносить их с реальными объектами, не знание дат, а понимание логики и последовательности событий, не знание терминов, а навыки их осмысления, не знание фактов, а готовность видеть за ними систему...

[5] Вопиющим примером являются задания по истории, при выполнении которых экзаменуемый должен выбрать наиболее предпочтительную точку зрения из нескольких предложенных. Для неоднозначной темы такой выбор означает ее профанацию, ведь отказ от рассмотрения сложной проблемы во всей полноте и противоречивости ведет к формулировке относительно нее заведомо ошибочного суждения.

[6] Подобные данные, разумеется, должны быть анонимными. Сейчас же, если случаются утечки первичных результатов, то в сеть попадают данные с фамилиями и паспортными данными экзаменуемых, но при этом почему-то без указания номеров доставшихся им вариантов.

- [7] Пока она исключена из кимов только по литературе (с 2008 г.) и по математике (с 2010 г.).
- [8] Реально эта шкала является четырехбалльной, т.к. предполагает лишь четыре оценки (неудовлетворительно, удовлетворительно, хорошо и отлично), поэтому ее название взято в кавычки.
- [9] Если у вас в кармане 200 долларов, 300 рублей и 500 юаней, то следуя логике вычисления егэшного вступительного балла, на вопрос «Сколько денег?» следует отвечать «Тысяча», а на недоуменное «Тысяча чего?» радостно уточнять «Тысяча денег!»
- [10] Единицей измерения логитов в русскоязычной литературе является... «логит», что не только нелепо, но и неграмотно. Во-первых, изначальное значение слова «логит» – название функции, посредством которой z или t выражаются через p . А во-вторых, единица относительной величины при использовании натуральных логарифмов называется (в честь их изобретателя) Непер. А поскольку применительно к анализу успехов и неудач сие название приобретает и вовсе сакральный смысл (в студенческом фольклоре «непер» – единица невезения), манкировать им просто-таки грешно.
- [11] Метод определения неизвестных параметров модели, основанный на требовании максимальности некоторой их функции, в качестве которой обычно берется произведение вероятностей получить именно те значения измеряемых величин, которые и наблюдаются.
- [12] Особо обратим внимание на ошибочность распространенной точки зрения, согласно которой первичные баллы за более сложные задания имеют больший вес при вычислении тестового балла. Имеет значение только количество баллов, но не то, за задания какой сложности они получены.
- [13] Не исключено, что ориентиром послужила 12-балльная дырка в математике, которую просто решили поделить поровну между концами диапазона тестовых баллов.
- [14] Насколько известно авторам, правила 2009 г. в открытых источниках не публиковались. Во всяком случае, их содержание не индексировано ни одним поисковым сервисом рунета.
- [15] Так, в 2009 г. было подготовлено 85 оригинальных вариантов кимов по русскому языку и математике и от 12 до 41 – по остальным предметам. Причём реально предлагаемые варианты получались комбинированием заданий из оригинальных вариантов, что требует компенсации различий в сложности не только вариантов как целого, но и в сложности их отдельных заданий.
- [16] Иногда можно встретить и другой (весьма странный) перевод названия – модель с промежуточными категориями ответов.
- [17] Эти баллы, как правило, набираются на заданиях наименьшей (наибольшей) сложности, что, видимо, и породило упоминавшееся выше ошибочное представление о разной цене баллов за задания различной сложности.

- [18] А в нынешней ситуации егэ – это такая линейка, цена делений которой произвольным образом меняется от измерения к измерению.
- [19] Как показал недавно проведенный анализ, среди выпускников Московского физико-технического института наибольших успехов добивались преимущественно те, кто при поступлении набрал 18÷20 баллов (все знают) и 14÷15 баллов (мало знают). Представители второй группы, не имевшие возможность получить в школе хорошее образование в ряде регионов, брали трудолюбием, энергией, целеустремленностью. При приеме по егэ они просто оказались бы за бортом.
- [20] Заметим, что если далее необходимо представить результаты экзамена в оценках «пятибалльной» шкалы, то для этого должно использоваться единое (для всех предметов) и неизменное (во времени) правило пересчета. Его смысл заключается исключительно в формализации размытых понятий «отлично», «хорошо» и «удовлетворительно». Логичной и удобной представляется следующая градация: владение на уровне 90% и более – пятёрка, 75% и более – четвёрка, 50% и более – тройка, иначе – двойка. Эти требования могут показаться излишне жесткими, но не будем забывать, что в случае выпускного экзамена речь идет не обо всей программе по предмету, а лишь об образовательном стандарте, т.е. о той ее части, не овладеть которой выпускник не имеет права. И понятно, что он не может быть аттестован, не осилив хотя бы половину предложенных заданий.
- [21] Собственно говоря, наработка банка калиброванных заданий для выпускного экзамена и должна была быть основной задачей эксперимента по введению егэ. Теперь, когда он функционирует в штатном режиме, решать эту задачу будет намного труднее.
- [22] Сейчас варианты егэ рассылаются по пунктам его проведения с соблюдением сложных мер, направленных на сохранение секретности, которые всё равно не помогают. А если задания берутся из открытого банка, то и скрывать почти нечего. Можно, например, передавать на места сразу много вариантов, а непосредственно перед началом экзамена сообщать сгенерированные компьютером номера тех из них, которые следует предложить выпускникам. А можно и просто пересылать варианты в пункты проведения экзамена в электронном виде, а там – распечатывать.
- [23] Мерой внутренней согласованности теста является коэффициент Кронбаха α . Он определяется долей дисперсии результатов выполнения всего теста, которая не объясняется суммарной дисперсией результатов выполнения его отдельных заданий. Иными словами, α служит количественным выражением того, насколько одно и то же измеряют разные задания. Для удобства коэффициент Кронбаха нормируется на диапазон от 0 до 1. Для егэшных кимов $\alpha \approx 0,9$, что считается хорошим показателем. Однако им не следует особо обольщаться, поскольку одним из основных объектов измерения егэ является готовность к сдаче егэ, что может маскировать низкую согласованность в плане проверки предметной подготовки.
- [24] Заметим, что до 2009 г. на егэ иностранным языкам разыгрывалось даже 100 баллов (что теоретически должно соответствовать 8÷9 градациям оценки), однако дополнительные 20 баллов приходились на задание по говорению, сдававшееся устно, что неизбежно влекло значительно более высокую корреляцию между частями этого задания, чем при письменной сдаче. Поэтому нет оснований учитывать это увеличение количества градаций.

[25] Поступать в вузы могут только лица, подтвердившие основных общеобразовательных программ среднего (полного) общего образования по предметам, соответствующим выбранному направлению подготовки (специальности). То есть, чтобы быть допущенным к конкурсу, надо показать результат, как минимум, в 20÷39 тестовых баллов (в зависимости от предмета). Однако многие вузы вводят и свои пороги, более жесткие, чем уровень аттестационной тройки. Например, в МФТИ они в 2010 году установлены на уровне 51 балла по русскому языку и 61 балла по математике, физике и информатике.

[26] Как вариант, выпускники могли бы сдавать определенное число обязательных экзаменов, имея свободу только в выборе предметов для некоторых из них. Но обязательность прохождения аттестации остается ключевым моментом.

[27] Любопытна статистика сдач необязательных предметов в 2009 г. В среднем один выпускник принимал участие в сдаче 1,42 необязательного экзамена. Обществознание сдавали 46,8% выпускников, физику – 23,1%, историю – 20,1%, биологию – 17,8%, иностранный язык – 8,9%, химию – 8,2%, информатику – 7,2%, литературу – 6,2%, географию – 3,6%. По большинству предметов егэ сдавала столь малая доля выпуска, что результаты экзамена не несут никакой информации об уровне всех школьников.

[28] Отчеты ФИПИ.

[29] Статистические данные о проведении ЕГЭ. За время, прошедшее после написания первой редакции настоящего текста, эта страница была уничтожена, поэтому ее содержимое воспроизведено здесь.

[30] Любопытно, что если требуется взаимно однозначное сопоставление элемента одного набора элементу другого и за каждое попадание начисляется по баллу, то сопоставление, выполненное наугад, будет приносить в среднем ровно 1 балл вне зависимости от длины предложенных наборов. Примерно так устроены некоторые задания в кимах для иностранных языков – с той разницей, что в один из наборов включен лишний пункт, не имеющий пары. Это чуть-чуть уменьшает средний результат гадания – на $1/(n+1)$ балла, где n – число элементов набора, имеющих пару.

[31] На самом деле, в кимах по русскому языку есть еще редкостная по «питательности» часть С. Человек, для которого этот язык не совсем чужой, обязательно получит за нее некоторое количество баллов. Участники егэ по русскому языку в 2009 г. выполнили часть А на 67%, В – на 31%, а С – на 45% (ФИПИ), т.е. в пересчете на балл задания части С проще заданий части В. Эта ситуация анекдотична и не имеет аналогов среди других предметов. Однако поскольку формально часть С все-таки не содержит заданий-угадаек, мы здесь не принимаем в расчет уникальную возможность ее частичного выполнения даже при нулевом уровне предметных знаний (т.е. за счет лишь бытовых навыков).

[32] Напомним, что данные ОИП, скорее всего, учитывают и лиц, которые не проходили итоговую аттестацию в 2009 г., поскольку окончили школу раньше. Поэтому доли выпускников, исправивших свои оценки по математике и русскому языку, в действительности больше приведенных значений, т.к. исчисляются от меньшей базы. Однако разыскать достоверную информацию ни по количеству исправленных двоек, ни по категориям сдававших егэ нам не удалось.

[33] Речь в данном случае идет именно о математике в целом, а не только об алгебре.

[34] Ни по одному предмету ни в одном году не было такого низкого порога тройки и ни по одному предмету ни в одном году не было такой высокой доли двоек (ФИПИ), как по литературе в 2008 г. Сочетание этих двух взаимоисключающих абсолютных рекордов как нельзя более полно характеризует способность ЕГЭ что-либо измерять в системе образования.

[35] Доли рассчитаны как отношение числа сдававших ЕГЭ к числу выпускников в тех регионах, где он проводился (а не к полному выпуску).

[36] ЕГЭшная география может показаться чем-то похожей на литературу – доля сдающих ее мала и уменьшается. Однако есть принципиальные различия. Во-первых, география характеризуется одной из самых высоких долей разыгрываемых первичных баллов, приходящихся на задания-угадайки, в силу чего ее сдача не вызывает такого ужаса, как сдача литературы. А во-вторых, если содержание кимов по географии все-таки имеет какое-то отношение к этому школьному предмету, то кимы по литературе в действительности ориентированы на совсем другую дисциплину (особенно в части В) – литературоведение, – которую необходимо изучать отдельно.

[37] Очень любопытна здесь динамика гендерного состава: доля юношей среди сдающих литературу составляла 26÷28% в 2005-07 гг., 39% – в 2008 г. и 17% – в 2009 г. Таких резких бросков соотношения полов (сначала в 1,8 раза в одну сторону, потом в 3,1 раза в другую) не демонстрировал никакой предмет ни в каком году.

[38] Сейчас делается отчаянная попытка выправить ситуацию. Так, новый формат кимов по математике предполагает повышение с 2010 г. доли геометрических задач в разыгрываемом первичном балле с 16,2% до 26,7%, а текстовых – с 2,7% до 10%. Однако, как бы не было поздно. Во многих школах преподавание геометрии в старших классах уже де факто прекращено. Да и способность видеть за формулами и преобразованиями содержательный смысл, давно уже вымывавшаяся из практики преподавания математики в школе, в связи с ориентацией на ЕГЭ, кажется, необратимо утрачена.

[39] Насколько известно авторам, в одном из регионов центральной России тарифы в 2009 г. составляли, соответственно, 400 и 50 тыс. руб. Разумеется, липовые стобалльные сертификаты, относящиеся к элитным коррупционным услугам, – лишь верхушка айсберга. Однако она позволяет получить верное представление и об остальных секторах рынка коррупционных услуг, порожденного введением ЕГЭ.

[40] Чернаков А. Глава Рособнадзора Любовь Глебова: «Честно ЕГЭ сдали не везде»// Известия. 02.09.2009.

[41] За последний год в России были закрыты 22 негосударственных вуза. Формальные поводы различны, но, по слухам, основная причина – нежелание принимать по ЕГЭ.

[42] Заметим, что именно так устроен в 2010 году конкурс на большинстве факультетов МГУ им. М.В. Ломоносова и СПбГУ – прием осуществляется на основе результатов ЕГЭ и одного собственного экзамена. Однако эта возможность обеспечена за счет присвоения столичным университетам эксклюзивного статуса «уникальных научно-образовательных комплексов, старейших вузов страны, имеющих огромное значение для развития российского общества». Иными словами, разумное решение проблемы было превращено в привилегию, что фактически является прививкой против его

распространения на остальные вузы.

[43] Каждый вправе на конкурсной основе бесплатно получить высшее образование в государственном или муниципальном образовательном учреждении и на предприятии.

Ключевые слова

реформа образования, кризис, единый государственный экзамен, дефекты егэ, образовательная коррупция, два экзамена, пересчет баллов, дифференцирующая способность, образовательный стандарт, модель Раша, partial credit model, шкалирование, логит

Институт прикладной математики им. М.В. Келдыша РАН, Сайт С.П.Курдюмова "Синергетика"
Сайт разработан при поддержке Российского гуманитарного научного фонда, проект **№ 07-03-12134-в**

