

Проблема становления системы педагогических измерений

Вадим Аванесов
testolog@mail.ru

Опубликовано в ж. «Педагогические Измерения» №1, 2008 г.
Обновлённая редакция 2019 г.

Аннотация

Название и содержание данной статьи находятся в заметном противоречии с реальной российской практикой создания систем контроля знаний. В официальных документах по созданию таких систем педагогические измерения не упоминаются.

Между тем, тестирование – наукоёмкая сфера общественно-профессиональной деятельности. В ней без развития культуры и теоретического фундамента педагогических измерений все практические системы контроля знаний обречены быть затратными, некачественными, нежизнеспособными и вредными для дела образования. Этот вывод подтверждается неудачным опытом внедрения в России двух практических систем централизованного тестирования (ЦТ) и единого государственного экзамена (ЕГЭ). Другая важная причина непригодности ЦТ и ЕГЭ – бюрократическое управление этими системами, их необеспеченность научными проектами и научно-общественной поддержкой.

Идея перехода от ЦТ и ЕГЭ к системе педагогических измерений, развиваемая в данной статье, возникла как результат исследования вопросов методологии и теории педагогических измерений, научного анализа существующей образовательной политики и практики.

Ключевые слова: система, педагогические измерения, оценка, экзамены, тест, тестирование, ЕГЭ

Введение

Цель настоящей статьи – поставить и раскрыть проблему становления научной системы объективных педагогических измерений. В России эта система оказалась практически не существующей. Это произошло из-за укоренившейся практики подмены таких систем элементарными словесными или цифровыми оценками, централизованным тестированием, некачественным и затратным единым государственным экзаменом. Для полноценного становления российской системы педагогических измерений необходимо рассмотреть перечень ключевых вопросов: выставление оценок, возможности тестового метода, опыт ЦТ и ЕГЭ. Этот опыт был построен на ошибочной идее оценки испытуемых двух различных контингентов – выпускников школ и абитуриентов вузов - единым набором заданий.

Автор неоднократно отмечал, что в массовом процессе аттестация выпускников школ – это, в основном, вопрос проверки усвоения образовательного

минимума, а вопрос приёма в вузы – это совсем другая проблема - профессионального отбора, требующего качественных тестов с высокой различающей способностью, на интересующих интервалах континуума подготовленности. А потому единого (универсального) метода для использования в системах оценки выпускников школ и отбора абитуриентов вузов различной профессиональной направленности нет. И не может быть, если посмотреть на эту проблему на научной основе. Содержание теста для отбора в вузы зависят от требований профессии к способностям испытуемых, а трудность теста - от квоты приёма. Это мнение чиновники не разделяют.

А потому в централизованном тестировании и в ЕГЭ образовательная бюрократия проводила в жизнь политически привлекательную, но метрически абсурдную линию на создание единого метода аттестации выпускников и приёма абитуриентов всех вузов. В едином государственном экзамене в роли такого универсального метода стали использоваться т.н. контрольно-измерительные материалы, качество которых оказалось ниже всякой критики.

Кто виноват и что делать? – вот вопросы, на которые пришлось искать ответы в настоящей статье.

Три системы оценки знаний

На сегодняшний день сложились три основные системы оценивания уровня подготовленности учащихся и студентов. Слово "оценивание" здесь употреблено как аналог английского понятия "evaluation". При использовании этого понятия полезно иметь в виду два аспекта употребления: первый - в смысле полученного результата, словесной или цифровой оценки, и второй аспект – в смысле процесса формирования получаемой оценки. В науке понятие "оценка" имеет тоже два основных смысла - как суждение о ценности (или значимости) и как приблизительная характеристика некоторой величины.

Первая система оценивания знаний - это традиционная словесная градация меры успешности в обучении, а затем и пришедшая ей на замену более удобная пятибалльная цифровая шкала. Обе используются практически неограниченно во всех образовательных учреждениях России. Пятибалльная шкала применяется для проведения текущего и итогового контроля, имеет своим главным преимуществом простоту и привычность. Это объясняет ее широкую и длительную, по времени, распространенность. Но пятибалльная градация и шкала имеют и недостаток - субъективность, большую зависимость от того, кто оценивает и кого оценивают.

Вторая система, апробированная в России – это централизованное тестирование (ЦТ). Там итоговая оценка возникала как результат процесса тестирования. Она была внешне похожа на западную систему объективного независимого тестирования, но только внешне. Потому что тестирование оказалось государственным по сути, а качество тестов, технология их разработки, степень научного обоснования до западного общественно-профессионального уровня явно не дотягивали.

Централизованное тестирование начиналось в России как кооперативная коммерческая деятельность. Коммерческим было, соответственно и понимание этой деятельности. Самое интересное в этом деле - ЦТ поначалу не было ни централизованным, ни тестированием, а лишь одной из форм платной массовой проверки знаний.

Постепенно тестирование стало развиваться, обычные вопросы начали заменяться высказываниями, улучшились тестовые формы, стали применяться и публиковаться статистические методы анализа результатов. Автор этой статьи в меру сил способствовал развитию ЦТ, особенно в кризисный период. Но помехи со стороны чиновников Министерства образования делали работу в ЦТ практически невозможной. В стране не было тестового процесса.

Тестовый процесс представляет единство теории педагогических измерений с практикой разработки и применения тестов. Существенное место в этом единстве занимают тестовые задания. В новых образовательных технологиях ключевую роль начинают занимать задания в тестовой форме, которым присущи свои признаки. Тестовый процесс является частью второго процесса - педагогических измерений¹. Существенные характеристики тестового процесса – направленность на развитие личности и общества, на улучшение образовательной деятельности, на гуманизм и культуру.

С окончательным утверждением ЕГЭ в качестве единого метода оценивания уровня подготовленности выпускников школ и абитуриентов, работа центра тестирования в качестве самостоятельного субъекта деятельности была приостановлена. Теперь ЦТ стало технологической службой ЕГЭ.

Третья система массового оценивания - единый государственный экзамен. Подлинное авторство схемы ЕГЭ не установлено, а желающих признаться нет. По

¹ Подробнее о процессе и содержании педагогических измерений см. в работах автора, опубликованных в журнале «Педагогические измерения» за 2004 и 2005 годы. А также на сайте автора: <http://testolog.narod.ru>

одной версии ввести ЕГЭ предложило Министерство образования РФ², по другой – идея ЕГЭ прорабатывалась в центре стратегических исследований, по третьей – в администрации Президента, и оттуда попала в Правительство РФ. Эта маленькая тайна будет, возможно, раскрыта по истечении какого-то времени.

В распоряжении Правительства Российской Федерации 10 июля 2001 года №910-р объявлялось, что ЕГЭ вводится с целью обеспечения объективности и унификации итоговой аттестации и вступительных испытаний системы профессионального образования³. Идея объективности выглядела привлекательно, но в связке с ошибочной идеей универсального метода сразу же теряла в доверии. Потому что такая универсалия противоречила теории и практическому опыту педагогических измерений.

Документально удалось установить, что истоки ЕГЭ просматриваются в планах ввести дифференцированную плату за высшее образование. Это предполагалось осуществить посредством введения т.н. ГИФО⁴. Планом действий Правительства Российской Федерации на Министерство образования была возложена обязанность организовать в 2001-2003 годах эксперимент по переходу на нормативно-подушевое финансирование учреждений высшего профессионального образования, с использованием государственных именных финансовых обязательств (ГИФО)⁵. На основе результатов вводимого для этой цели единого государственного экзамена.

Вузы и местная власть быстро осознали – чем выше оценки испытуемых по ЕГЭ, тем больше денег они получают из бюджета на образовательные нужды. Однако в Министерстве финансов быстро разобрались в этом дефекте ЕГЭ и отказались выделять дополнительные суммы. В результате ГИФО рухнули из-за тенденции чрезмерного завышения оценок по ЕГЭ. Таким образом, план правительства по ГИФО оказался проваленным из-за некачественного ЕГЭ.

Но сам ЕГЭ остался живым и невредимым. Этот экзамен определённо не тест и не тестирование, как это поначалу преподносилось педагогическому обществу⁶. Но это и не экзамен в обычном понимании этого слова. В ЕГЭ применяются т.н. контрольно-измерительные материалы неизвестного, точнее, засекреченного

² <http://www.r-komitet.ru/obraz/EGE.htm>

³ <http://www.r-komitet.ru/obraz/EGE.htm>

⁴ О ГИФО см. в статьях сайта <http://testolog.narod.ru>

⁵ <http://www.r-komitet.ru/obraz/EGE.htm>

⁶ Аванесов В.С. И не тест, и не экзамен // Университет и школа. № 1-2, май, 2002 г. С.26-27. См. также <http://testolog.narod.ru>

качества. Сокращённое название пресловутых «материалов» – это КИМы. В первой и второй части КИМов используются внешне тестоподобные задания эмпирически неподтверждённого качества, а в третьей части содержатся экзаменационные задания запредельной трудности, за выполнение которых экзаменаторы выставляют повышенные баллы.

По данным автора этой статьи, по многим территориям РФ эти баллы коррелируют отрицательно с баллами по первой и второй части. Интерпретировать этот факт можно по-всякому. Но очевидно, что при такой структуре метода складывать баллы первых двух частей с баллами третьей части КИМов не разумно. Сумма баллов испытуемых при таком сложении становится похожей на среднюю температуру по больнице.

В ЕГЭ была поставлена внешне справедливая, но метрически ошибочная задача - создать один универсальный набор вопросов и задач (КИМ), пригодный для государственного отбора на все специальности всех вузов и для аттестации всех выпускников школ. Министерство образования и науки знает об этой ошибке, но её устранением заниматься не хочет или не может, поскольку это граничило бы с невыполнением политической воли, проявленной свыше.

Проведение в России единого государственного экзамена и государственной аккредитации вузов дало мощное начало росту нового витка коррупции в сфере образования. В итоге Россия оказалась отброшенной уже на 143 месте по уровню некоррупционности⁷. По данным ООН, в России ежегодная общая сумма взяток за поступление в ВУЗы составляет 520 миллионов долларов⁸.

ОСОКО

Сейчас уже можно говорить ещё об одной, четвёртой системе оценивания, которая пока не внедрена, но уже широко обсуждается на различных совещаниях. Она называется *общероссийской системой оценки* качества образования (ОСОКО) и мыслится её создателями как не имеющая аналогов в мире. Фактически ОСОКО призвана установить не только полный контроль уровня подготовленности уровня подготовленности всех учащихся и студентов, но и тотальный чиновно-

⁷ Аванесов В.С. Комплексный подход к противодействию коррупции в образовательной сфере//Государственная политика противодействия коррупции и теневой экономике в России. Материалы научной конференции. М. Центр проблемного анализа и государственно-управленческого проектирования. М.: <http://testolog.narod.ru/Obrazov16.html>

⁸ Научный эксперт, 2007. С. 717-723. <http://www.epochtimes.ru/content/view/10999/2/>

административный контроль за работой всех образовательных учреждений России, во всех краях, областях, и даже в автономных республиках.

«Предлагаемая нами схема такова – сообщает А.А. Фурсенко⁹ относительно ОСОКО:

- для всех выпускников школ обязательна сдача ЕГЭ по установленным предметам, один из которых, безусловно, русский язык;
- экзамены по другим предметам могут сдаваться в форме ЕГЭ или традиционно;
- для поступления в вуз необходима сдача ЕГЭ по тем предметам, по которым установлены вступительные испытания в выбранный выпускником вуз.

А.А. Фурсенко добавил, что в перспективе целесообразно, в рамках развития общероссийской системы оценки качества образования, вести также разработку и апробацию независимых и объективных процедур оценки образовательных достижений учащихся по завершении обучения на ступенях начального и основного общего образования (после 4-го и 9-го кл.)»¹⁰.

Эти слова министра можно понимать так: сейчас опять будет введён ЕГЭ, а в перспективе целесообразно разрабатывать другие, независимые методы, но обязательно в рамках новой тотальной государственной системы ОСОКО. Тогда возникает вопрос: – какие методы и независимые от кого собирается ввести министерство образования и науки?

Любопытны упоминания об «объективных процедурах оценки». В русском языке понятие «процедура» ассоциируется с *оформлением* какого-либо дела, или с *лечебным* мероприятием¹¹. Оценивание – это отображение уровня развития свойства личности числами. И это процесс, который бывает объективным только тогда, когда оценки испытуемых не зависят от набора заданий теста, а оценки параметров заданий не зависят от уровня подготовленности испытуемых¹². Достижение такого рода идеала объективности возможно только в хорошо организованном тестовом процессе. Зная реальное отношение министерства к настоящему независимому тестированию,

⁹ Тезисы к выступлению А.А. Фурсенко на заседании Государственного совета РФ 24 марта 2006 года http://www.school.edu.ru/dok_min.asp?ob_no=29402

¹⁰ Там же

¹¹ Словарь русского языка. Т.3., М.: АН СССР, 1959. с.730..

¹² Wright B. Sample-free test calibration and person measurement. ETS Invitational Conference on Testing Problems. MESA Research Memorandum. <http://www.rasch.org/memo1.htm>

легко понять, что массовый процесс объективных педагогических измерений в России придёт не скоро.

Круговорот понятий

Культурное исследование проблем педагогических измерений немыслимо без создания системы научных понятий. Потому что любое измерение в общественных науках начинается с концептуализации, с поиска признаков, образующих то или иное понятие, попадающего в фокус измерения. На русском языке такой системы научных понятий педагогических измерений сейчас нет, хотя отдельные попытки её создания уже предпринимались, в том числе и автором этой статьи¹³. Здесь предстоит серьёзная расчистка понятийного поля, сильно засорённого плохими переводами английской лексики случайными людьми в этой специфической области знаний.

Анализ использованных во всех трёх системах понятий показывает удивительный результат. Если в них отставить, на время, в сторону все поясняющие определения в таких словосочетаниях, как «пятибалльная оценка», «централизованное тестирование», «единый государственный экзамен» и система оценки качества образования», то останутся такие ключевые слова: *оценка, тестирование, экзамен* и ещё раз *оценка*. Используемое два раза ключевое слово *оценка*, в начале и в конце, указывает на своего рода понятийный круг или *круговорот понятий*.

Оппоненты могут возразить, что нельзя отрывать ключевые понятия от определяющих их слов. Давайте на минуту согласимся с ними, объединим понятия и определяющие слова, упорядочим и сопроводим их кратким комментарием. Получится примерно такой текст: в России на смену ранее существовавшей субъективной пятибалльной *оценке* пришло объективированное централизованное *тестирование*. Затем тестирование, как более прогрессивный метод педагогических измерений, министерские работники заменили баллами псевдонаучного единого государственного *экзамена*, неизвестного качества. Теперь на смену некачественным оценкам ЕГЭ может прийти ещё более некачественная *оценка* в общероссийской системе оценки качества образования.

¹³ Аванесов В.С. Тест как педагогическая система // Педагогические Измерения, №1 2007г.

Пятибалльная шкала оценивания

В середине тридцатых годов XIX века была осознана неадекватность системы оценки знаний задачам, поставленным перед советской школой. В Постановлении Совета народных Комиссаров и ЦК ВКП(б) от 3 сентября 1935 г. утверждалось, что установленная Наркомпросами система оценки успеваемости не дает представления о фактических знаниях учащихся и ведет, на практике, к понижению уровня учебы. А потому учебному отделу ЦК ВКП(б) поручалось "разработать, обязательно для всех школ СССР, нормы оценки успеваемости учащихся, с тем, чтобы один и тот же уровень знаний одинаково оценивался во всех школах".

Ввиду неспособности учебного отдела ЦК партии решить эту методическую задачу, поручение так и осталось невыполненным. Не решили её и все последующие поколения руководящих деятелей союзного, а позже и российского образования. На этом, собственно, партийные опыты по улучшению качества оценивания учащихся закончились. Если не считать введение цифровой пятибалльной оценки, или иначе, цифровых отметок, как числовых эквивалентов оценок качества подготовленности учащихся.

Цифровая пятибалльная оценка была введена в советскую школу в качестве начальной меры улучшения дел. Это было сделано Постановлением Совета Народных Комиссаров РСФСР от 10 января 1944 года. Постановление было принято, как там писалось, "в целях более четкой и точной оценки успеваемости" и замены "словесной системы" оценки успеваемости - отлично, хорошо, посредственно, плохо и очень плохо – соответствующей цифровой пятибалльной системой 5, 4, 3, 2 и 1. Интересна деталь – постановлением предписывалось ввести цифровую пятибалльную систему отметок немедленно, со следующего дня - 11 января 1944 г.- от момента публикации.

В других странах сложилась традиция отмечать единицей отличные знания, а пятеркой - плохие. Из этого видно, что одинаковым оценкам можно ставить в соответствие различные цифры-отметки. Интересны суждения об отметках известного педагога Ш.А.Амонашвили: "в отметках мало педагогики...в них больше власти. Отметки - это костыли хромой педагогики"¹⁴.

Основной недостаток пятибалльной шкалы оценок касается ее слабой дифференцирующей способности. Она позволяет провести грубую классификацию студентов только на четыре группы - отличников и тех, кто учится хорошо, удовлетворительно и неудовлетворительно. Более тонкое различие, особенно

¹⁴ Амонашвили Ш.А. Педагогическая симфония. Ч.3. Екатеринбург, Изд-во Уральского Ун-та, 1993. - 216с.

нужное на приемных и выпускных экзаменах, в экспериментальной работе и для целей эффективного управления качеством обучения, пятибалльная шкала не даёт, и дать не может. Отсюда понятна актуальность вопроса совершенствования педагогических оценок на основе измерений.

Пятибалльная шкала оставляет мало простора для вариации оценок по их значениям. Последнее обстоятельство затрудняет внедрение статистических методов обоснования преимущества новых подходов в обучении и воспитании, препятствует качественной проверке учебников, учебных пособий, а также экспериментальной проверке эффективности организационных и воспитательных мероприятий.

Для выставления итоговых оценок по пятибалльной шкале применялась, и всё ещё продолжает применяться в России, экзаменационная форма проверки знаний.

Экзамены

Проведение экзаменов имеет длительную историю, превышающую четыре тысячи лет¹⁵. Труды по истории экзаменов представлены у ряда авторов¹⁶.

Экзамен – это традиционная и нетехнологичная форма педагогического контроля уровня подготовленности, в процессе которой делаются попытки оценить у испытуемого, в основном, знания, умения и навыки. Экзамен характеризуется высокими затратами времени, субъективизмом, не технологичностью. В массовых процессах экзамены стараются применять меньше по причине экономической неэффективности. Экзамены обычно содержат вопросы и задачи.

Экзамены применяются в России, в программах международного исследования PISA, на Тайване¹ и в других стран Юго-Восточной Азии. Преимущества этого метода - возможность оценить речевые способности учащихся, провести углубленную проверку знаний по избранным темам, в небольших учебных группах. В условиях малых групп лучше обеспечивается связь между содержанием обучения и содержанием экзамена. Но для массового процесса измерения уровня подготовленности экзамены не годятся, ввиду их повышенной трудоёмкости, субъективности, не технологичности и дороговизны.

¹⁵ DuBois P.H. A History of Psychological Testing. P.3. Boston. Allyn and Bacon Inc., 1970. -152p; .Аванесов В.С. Из глубины веков. <http://testolog.narod.ru>

¹⁶ См. напр. Кадневский В.М. Генезис тестовых методов и тестов в мировом цивилизационном процессе. ПИ, №1, 2008г.

При использовании экзаменов для приёма в вуз они сами нередко становятся инструментом коррупции. Это вторая причина, по которой экзамены принципиально не применяются в странах, имеющих развитые системы образования.

Третьим недостатком экзамена является несравнимость оценок, полученных студентами разных институтов, по одной и той же дисциплине, и отсюда - принципиальная невозможность корректно установить - в каком вузе процесс обучения поставлен лучше? Ответ на этот вопрос можно дать только в случае использования стандартной тестовой программы с применением технических средств контроля. Поэтому с управленческой точки зрения полезность тестовых оценок заметно выше экзаменационных.

Можно вывести, что в наше время эффективное измерение подготовленности больших масс испытуемых посредством экзаменов экономически неэффективно. Этот вывод относится и к единому государственному экзамену, представляющему эклектическую смесь тестоподобных заданий и конкурсных, заведомо трудных задач. В то время как для получения качественных измерений нужно использовать настоящие тесты.

Тестирование

Самым распространённым в мире массовым методом оценки знаний является *тестирование*. Оно применяется в США, странах Латинской Америки, Европы, Казахстане, Азии и в странах Юго-Восточной Азии. По сравнению с научно-обоснованными тестами, все другие методы оценки знаний оказываются менее эффективными, менее точными, более субъективными и более подверженными коррупции. Не случайно президенты Казахстана и Украины, по примеру многих государств, не разрешают в своих странах использование иных методов массового контроля знаний, кроме тестов.

Необходимо отметить, что не всякий тест лучше экзамена, а только тот, который разрабатывается и применяется на основе теории обучения и теории педагогических измерений. Тест лучше применять при массовой проверке знаний по той или иной учебной дисциплине в школе, у абитуриентов, на отдельных студенческих потоках, факультетах и в институтах. С целью повышения эффективности таких масштабных проверок рекомендуется использование технических средств контроля, стандартных бланков и использование ЭВМ, что позволяет в короткий срок собрать информацию, обработать ее и выдать результаты.

Тестовый метод имеет и некоторые отрицательные последствия с точки зрения цели и содержания учебного процесса. Если целью является сдать учебный предмет, то учащиеся обычно учат только тот материал, который включён в итоговый тест. А это содержание примерно тридцати заданий. Тестирование имеет также социально-психологические издержки, особенно в случаях, если тестовая оценка становится решающей для формирования жизненной и профессиональной траектории развития молодёжи¹⁷.

Централизованное тестирование

В последние двадцать лет для контроля знаний стали использоваться компьютеры, сканеры, программирование, автоматизацию, образовательные и тестовые технологии. Всё это получило широкое распространение на Западе и с неизбежностью должно было каким-то образом внедриться в России. Формой внедрения оказалось централизованное тестирование (ЦТ).

Централизованное тестирование определялось тогда как добровольный, платный процесс оценки уровня достижений выпускников школ¹⁸. По итогам такого тестирования испытуемым выдавался сертификат. В сертификате указывалось количество баллов, полученных испытуемым по каждому предмету. Результаты тестирования засчитывались многими школами в качестве выпускных экзаменов в школах и ВУЗах в качестве результатов вступительных испытаний. В некоторых вузах сертификаты Централизованного тестирования принимались во внимание при поступлении только на платную форму.

В приказах Министерства образования утверждалось, что создан Всероссийский Центр тестирования, имеющий свои региональные представительства и выдающий, за плату, государственные сертификаты поступающим в вузы. Слово «государственный» из приказов затем исчезло, а «негосударственный» так и не появилось, что создавало впечатление неопределённости фактического статуса центра тестирования.

Попытка определения ЦТ как образовательной услуги порождала цепочку вопросов. Первый из них - юридический статус ЦТ. На конференции 1999 года автор этой статьи отметил спорный статус ЦТ: оно продолжало оставаться ни

¹⁷ Это т.н. high-stake tests. Подробнее смотрите, например: Аванесов В.С. Единый государственный экзамен в фокусе научного исследования// Педагогические Измерения № 1, 2006, стр. 3-31. См. также на сайте автора <http://testolog.narod.ru>

¹⁸ http://www.kai.ru/info/admission/files/FAQ/faq4_ege_ere_ct.htm

государственным, ни частным¹⁹. Буквально на следующий день после публикации тезисов докладов ЦТ прямо на конференции был объявлен государственным учреждением, в соответствии с только что подписанным приказом. Вместе с тем, приказами Минобразования государственное ЦТ выдавало 10% своих «государственных» сертификатов частной фирме "Телетестинг". На вопросы - почему 10%, и на какой юридической, лицензионной, методической, договорной или финансово-компенсационной основе это делалось - ответов не было.

Важный вопрос – качество тестов. Оно не соответствовало требованиям, предъявляемым к тесту как системе параллельных заданий равномерно возрастающей трудности, позволяющей оценить структуру и качественно измерить уровень подготовленности испытуемых¹. Настоящий тест эффективно применим в массовых процессах ввиду кратковременности и экономической эффективности. Хорошо, если время педагогического тестирования не превышает сорока минут.

Если рассматривать ЦТ как деятельность государственного органа по регулированию вопросов отбора абитуриентов в вузы, то данное истолкование входит в противоречие с Конвенцией по борьбе с дискриминацией в области образования (Париж, 14 декабря 1960 г.). В соответствии с этой Конвенцией Россия обязана не воздвигать барьеры перед своими гражданами на пути получения образования, не закрывать никому доступ к образованию любой ступени или типа; не ограничивать образование какого-либо лица...

Низкое качество тестов – главная нерешённая проблема в работе центров тестирования в России и других странах бывшего СССР. Основная причина производства и применения некачественных тестов – монополизация тестовой деятельности под эгидой министерства, отсутствие мотивов и возможности приглашения учёных работать в центрах тестирования. Не случайно в большинстве центров тестирования стран СНГ в условиях бюрократического управления практическим тестированием учёным-педагогам места не оказалось. В службе ETS в США и в Европе это не так. Там есть ученые. В России же и странах СНГ неизменно подтверждается общее удручающее состояние практического тестирования.

Хорошо подготовленный и этично применяемый тест исключает возможности возникновения коррупции. Этими, а также двумя другими критериями - объективности и справедливости - объясняется широкое применение тестов в развитых странах, имеющих, одновременно, самые низкие уровни коррупции.

¹⁹ Аванесов В.С. Тест как педагогическая система//ПИ, №1, 2007, стр. 33-55; см. также сайт <http://testolog.narod.ru>

ЕГЭ & КИМы

Система ЦТ продержалась до 2007 года. После чего её полностью заменила система единого государственного экзамена.

В теории педагогических измерений утверждается: качество измерения тем выше, чем лучше уровень подготовленности испытуемого соответствует уровню трудности заданий. Объективность педагогических измерений возникает как следствие интеграции методов обоснования надежности и валидности тестовых результатов. Поскольку в ЕГЭ нет ни тестов, ни других методов педагогических измерений, то объективность там должна была появиться, вероятно, из воздуха или из благих пожеланий работников бывшего министерства образования.

С выбором ЕГЭ прежнее российское правительство допустило в 2000 году управленческую ошибку, но критику в свой адрес слушать не захотело. Теперь эта ошибка дополнена другой ошибкой. В п.27 постановления Правительства РФ №71 от 14 февраля 2008г. написано, что приём в государственные и муниципальные вузы по программам бакалавриата и подготовки специалистов будет теперь осуществляться на конкурсной основе по результатам единого государственного экзамена (ЕГЭ).

Одно из следствий этого постановления - отказ в приёме в вуз тех лиц, кто закончил школу в прошлые годы и не имеют сертификат ЕГЭ. Таковых может набраться по стране порядка десятков тысяч человек. Фактически эти граждане принуждались сдавать ЕГЭ. Другие возможные следствия постановления Правительства № 71 - это вероятность возникновения судебных расследований низкого качества баллов ЕГЭ, по которым государство принимает решения относительно судеб каждой личности. Не случайно проведение ЕГЭ сопровождается ежегодными расследованиями массовых завышений результатов и множества ухищрений, допускаемых испытуемыми, членами комиссий и организаторами проведения ЕГЭ. Результаты ЕГЭ фабрикуются в массовом порядке и систематически, особенно на некоторых территориях²⁰.

Автор этой статьи уже называл ряд причин, почему ЕГЭ опасен для страны²¹
Полезно их повторить:

²⁰ Факты расследования *массовых* фальсификаций и коррупции, представлены в публикации Е. Мельник и О. Герасименко: «ЕГЭ получил хорошие расценки». http://www.gazeta.ru/2007/04/05/oa_235842.shtml, 5 апреля 2007г. мн. др.

²¹ Аванесов В.С. Доживёт ли единый государственный экзамен до 2009 года? <http://testolog.narod.ru>

- разгул коррупции вокруг ЕГЭ;
- превращение ЕГЭ в средство обогащения чиновников и бизнесменов за счёт бюджета, абитуриентов и их родителей;
- безразмерная погрешность оценок ЕГЭ. Недопустимая метрическая погрешность ЕГЭ исключает принятие каких-либо осмысленных управленческих решений на основе ЕГЭ. И по сей день министерство образования не решается обнародовать рейтинг областей, краёв и территорий по качеству образования на основе имеющихся данных.
- редукция учебной работы по всем учебным дисциплинам только к проверяемому содержанию ЕГЭ;

Далее можно говорить о потере учащимися способности писать, читать, развивать речь и мышление в случае, если ЕГЭ станет, как предусмотрено законом, основным методом итоговой аттестации.

Более семи лет автор этой статьи пишет о некачественности и нелегитимности ЕГЭ²². Наконец, прислушались ко второму тезису, но упорно игнорировали первый. В итоге Госдума узаконила этот экзамен как ключевой метод реформирования образования. Но теперь каждый двоечник может попытаться опротестовать в суде свои баллы, потому что погрешность оценивания в ЕГЭ невообразимо высокая из-за его ошибочной конструкции. В результате граждане и государство попали в опасный для страны капкан некачественных результатов.

Не случайно ЕГЭ отвергается большинством населения страны. Согласно опросу Всероссийского центра изучения общественного мнения в 2007 году популярность ЕГЭ среди россиян резко снизилась. Если в июне 2006 г. 31% респондентов предпочли бы, чтобы их дети или внуки сдавали выпускные экзамены в виде ЕГЭ, то в июне 2007 г. - только 22%. Уменьшилась доля тех, кто видит в ЕГЭ те или иные положительные стороны²³. Был придуман экзамен, в котором первая и вторую часть представляли задания с устаревшей тестовой формой. К этому добавили конкурсные задачи запредельного уровня трудности для третьей части. Получился многоцелевой, невиданный в мире, гибрид трёх форм – тестовой, вопросной и задачной. Первым среди высших руководителей страны ошибочность ЕГЭ признал

²² Первая критическая статья на эту тему: Аванесов В.С. «Единый Государственный экзамен, или эта вещь будет посильнее Фауста Гёте» // Учительская Газета, №49, 28 ноября 2000 года. Результаты оказались намного разрушительнее, чем можно было тогда себе представить.

²³ www.regnum.ru/news/885466.html 14:44 15.09.2007.

Председатель Совета Федерации С.М. Миронов²⁴. Поначалу странный гибрид дал повод руководителям бывшего Министерства образования утверждать об оригинальности избранного ими пути. После жестокой критики утверждения такого рода затихли, а информация о качестве гибридных материалов оказалась засекреченной. Она и теперь является теперь главной тайной Министерства образования и науки. Однако по имеющимся фрагментарным данным, качество «материалов» ниже всякой критики. Возможно, что именно этим объясняется отсутствие внятных публикаций о качестве КИМов.

Главный метрический недостаток КИМов – это попытка оценить уровень подготовленности ординарных выпускников школ и, одновременно, уровень подготовленности абитуриентов престижных вузов одним и тем же набором вопросов и задач. Теоретически ясно, что точность измерения одним и тем же тестом уровня подготовленности столь различных контингентов может быть только низкой.

С принятием новой версии закона «Об Образовании» этот недостаток закреплён теперь законодательно. Там сейчас написано: «Единый государственный экзамен представляет собой форму объективной оценки качества подготовки лиц, освоивших образовательные программы среднего (полного) общего образования, с использованием заданий стандартизированной формы (контрольных измерительных материалов), выполнение которых позволяет установить уровень освоения ими федерального компонента государственного образовательного стандарта среднего (полного) общего образования²⁵».

Если бы объективность КИМов-ЕГЭ сверялась по Закону, принятому нынешним составом Госдумы, то «жить бы стало лучше, жить стало веселее». Но объективность метода порождается не законом, а свойствами метода. Объективность метода КИМов-ЕГЭ может быть удостоверена только вследствие открытого анализа матриц реальных результатов, достижения уже упомянутой независимости параметров испытуемых от параметров заданий, опоры на научные критерии надёжности и валидности результатов. А как раз этого в ЕГЭ как не было, так и нет!

А потому упомянутый фрагмент Закона РФ об «объективности» метода ЕГЭ представляет собой казус, который на века войдёт в мировую юридическую науку как несомненное свидетельство избыточного вторжения государства в научно-методическую деятельность по оцениванию личности.

²⁴ ЕГЭ как широкомасштабная ошибка: Обзор СМИ Приамурья. www.regnum.ru/news/772724.html 04:01 27.01.2007

²⁵ <http://www.akdi.ru/gd/proekt/099384GD.SHTM>

ЦТ было лучше ЕГЭ²⁶

ЦТ можно определить как форму деятельности государственного органа по эффективной и технологической оценке подготовленности выпускников школ и абитуриентов вузов. В других странах тестированием занимается не государство, а независимые общественно-профессиональные корпорации типа ETS²⁷.

Несколько лет технологию ЦТ пытались приспособить к нетехнологичному ЕГЭ. Но сравните: для тестов обязательны публикация меры надежности и валидности результатов, научно-теоретическая обоснованность, требования логики и тестовой формы²⁸, публикация отчетов о подлинных данных, открытый научный анализ каждого задания и его параллельных вариантов, детальные отчеты о доходах и расходах, выполнение кодекса тестологов²⁹. В ЕГЭ ничего из перечисленного нет. Да и можно ли сравнить, по научному потенциалу, известные теории педагогических измерений с псевдонаучным ЕГЭведением?

ЦТ лучше ЕГЭ также в финансовом и в правовом отношениях. ЦТ прибыльно, ЕГЭ затратное. ЦТ добровольно, ЕГЭ принудительно.

С юридической точки зрения ЕГЭ нарушает автономию вузов, установленную Законом. «Государству следует всегда помнить, – писал В. Гумбольдт, – что оно не может и не должно подменять университеты в их деятельности; каждый раз вмешиваясь, оно создаёт препятствия».

Утверждают, что ЕГЭ улучшит образование в школе. Но вот мнение разработчика первого поколения государственных образовательных стандартов В. Фирсова: никаких проблем школы ЕГЭ не решит³⁰! Утверждают также, что ЕГЭ улучшит высшее образование. На самом деле, все наоборот: введение ЕГЭ в его нынешнем системно-конфликтном варианте разорвёт устоявшиеся комплексы «школы-вузы», подорвет и без того ухудшающееся качество образования.

²⁶ Аванесов В.С. Централизованное тестирование лучше единого государственного экзамена. Тез. докл. Всер. Научно-методической конференции «Развитие тестовых технологий в России». /Под ред. Л.С. Гребнева.- М.: Центр тестирования Министерства образования РФ. 2002». М.: 21-22 ноября 2002г. с.71-73. <http://testolog.narod.ru/Education27.html> .

²⁷ Educational Testing Service. Princeton, N-J.

²⁸ Аванесов В.С. Композиция тестовых заданий. Издания 1996, 1998 и 2002 гг.

²⁹ Code of Fair Testing Practices in Education. (1988) Washington, D.C.: Joint Committee on Testing Practices. (Mailing Address: Joint Committee on Testing Practices, American Psychological Association, 1200 17th Street, NW, Washington, D.C. 20036.).

³⁰ Учительская газета, 17 сентября 2002 г.

Считается также, что ЕГЭ переборет коррупцию. И это в условиях многолетней блокировки Закона о борьбе с коррупцией? Много говорят и о противниках ЕГЭ. Это главным образом, дети и их родители. Но это миллионы граждан, и разве допустимо рассматривать их как противников ошибочной схемы. Всё обстоит наоборот. Но главный противник ЕГЭ - это он сам, с присущей ему совокупностью признаков:

- 1) фактическая нелегитимность эксперимента по коммерциализации образования посредством эксперимента ЕГЭ-ГИФО;
- 2) отсутствие гипотез и системы критериев успешности (неуспешности);
- 3) претензия на принудительную тотальность, стремление оценить одним «КИМом» всех выпускников и абитуриентов различных вузов и специальностей;
- 4) нетехнологичность; субъективность на всех этапах оценивания;
- 5) системная конфронтационность;
- 6) фактическая засекреченность (недоступность) реальных результатов, что сделано вопреки Закону.

Не случайно ЕГЭ отвергается большинством населения страны. Согласно опросу Всероссийского центра изучения общественного мнения в 2007 году популярность ЕГЭ среди россиян резко снизилась. Если в июне 2006 г. 31% респондентов предпочли бы, чтобы их дети или внуки сдавали выпускные экзамены в виде ЕГЭ, то в июне 2007 г. - только 22%. Уменьшилась доля тех, кто видит в ЕГЭ те или иные положительные стороны.

Поэтому переход от ЦТ к ЕГЭ – это отступление из прошлого века в позапрошлый век.

Тесты или ЕГЭ?

В России в качестве средства достижения обеих целей был выбран метод, предложенный бывшим министерством образования - единый государственный экзамен (ЕГЭ). Выбор обеспечивался целенаправленной подменой понятия тест понятием «экзамен».

Вот пример обоснования того, почему в России вместо *тестов* надо использовать «механизмы» *экзамена*.

Цитата: «Особое значение при создании национальной системы *экзаменов* имеет разработка требований к обеспечению качества при организации и проведении экзаменов, разработке контрольных измерительных материалов, на этапе обработки и шкалирования результатов и их использования, а также создание механизмов обеспечения качества системы. В США идет активная разработка механизмов

обеспечения качества экзаменов. С этой целью разрабатываются и внедряются стандарты по различным аспектам *тестирования*. Наиболее признанными считаются «Стандарты для педагогического и психологического *тестирования*» (Standards for Educational and Psychological Testing), принятые совместно тремя организациями: Американской Ассоциацией исследователей в области образования (American Educational Research Association), Американской Ассоциацией Психологов (American Psychological Association) и Национальным комитетом по Измерениям в Образовании (National Council on Measurement in Education), а также «Стандарты качества и справедливости ETS», разработанные специалистами Службы образовательного *тестирования* США (ETS)³¹.

Получается так: в США – тестирование, но оно является «механизмом экзамена».

В результате такой и похожей на это бесконечной словесной эквилибристики вместо качественного тестирования в России уже восьмой год проводится некачественный и дорогостоящий единый государственный экзамен.

В Казахстане создано и функционирует Единое Национальное Тестирование (ЕНТ), способное выдать итоговые результаты в день проведения тестирования, причём по 3 - 5 коротким тестам сразу. Все зарубежные консультанты, в числе которых на протяжении многих лет выступал и автор этих строк, советовали Правительству Казахстана применять исключительно тестовый метод. В результате Н. Назарбаев утвердил тесты как главный метод массовой оценки знаний.

Успеху ЕНТ способствуют, в частности, приобретение высокоскоростных оптосканов, разработанных в ETS - тестовой службе США. У казахстанских коллег, как и в других странах СНГ, нуждается в существенном улучшении качество тестов, а база тестовых заданий вообще требует принципиального обновления, если не замены. Это необходимо сделать из-за появления организованных групп профессиональных репетиторов, внимательно отслеживающих содержание заданий каждый год³².

И на Украине был сделан стратегический выбор в пользу создания Национальной независимой (внешней) системы тестирования. И это был правильный выбор. Целью становления Национальной системы *тестирования* поставлена борьба с коррупцией.

³¹ Развитие национальной системы экзаменов: опыт РОССИИ, СНГ И США. Материалы и тезисы докладов Международной конференции, 19 – 24 апреля 2003 года, Москва 2003. - С.6.

³² В Казахстане необходимо ежегодно обновлять базу данных ЕНТ <http://www.gazeta.kz/art.asp?aid=100270>

Статус тестирования в различных странах

В США, Казахстане и Украине, а также во многих других странах принят статус Национальной системы тестирования. В западном истолковании понятие «Национальное», означает, что тестирование является общественным. Оно создаётся из потребностей развития общества, и развивается под контролем общественных организаций. В определении статуса Национального тестирования важно понять соотношение роли государства и общества. Первое и главное место принадлежит обществу. Цивилизованному государству не следует заниматься тестированием каждого ребёнка или студента. Тем более, это не следует делать принудительно. В документах министерства образования любят использовать словосочетание государственно-общественное управление, где государственное управление мыслится главным, а общественное управление – декоративным.

Как известно, в США центральное правительство не имеет права проводить тестирование в том или ином штате. И это правильно, потому что удерживает государственные органы от соблазна расширения властных полномочий и от коррупции. Но государство там стоит на страже в случае нарушения прав гражданина в виде необъективно выставленной оценки, полученной на основе некачественного метода.

Вот почему в настоящем Национальном тестировании есть и вторая важная составляющая. Это профессионализм тех, кого общество привлекает к проведению Национального тестирования. Именно профессионализм обеспечил существенный успех общественно ориентированной, негосударственной службы тестирования в США. Полное название этой успешно управляемой службы - Educational Testing Service (ETS). В США наряду с системой тестирования есть и общественный орган - Национальный Совет по измерениям в сфере образования (National Council on Measurement in Education).

Система педагогических измерений

Педагогические измерения – это научная теория, сформировавшаяся на стыке педагогики, психологии, статистики, математики, логики и философии. Это наука о разработке тестов, других методов и показателей эффективности образовательной деятельности вообще. Педагогические измерения можно также определить как процесс отображения числами уровней проявления интересующего латентного качества личности.

Предмет педагогических измерений – не только разработка тестов, но и шкалирование тестовых результатов испытуемых (Measures of achievement), проведение рейтинга и мониторинга (в части разработки показателей текущих и итоговых учебных достижений). И, кроме того, - построение системы качественных показателей образовательной деятельности³³.

Большое значение в педагогических измерениях имеют также и методы определения качества заданий, шкалирования заданий по уровню трудности и различающей способности - то, что на английском языке называется Item Calibration. Измерением учебных достижений (person measures) и шкалированием заданий (item calibration) на интервальной шкале практическое тестирование почти не занимается, поскольку это научная работа, требующая целенаправленной метрической и программно-вычислительной подготовки.

Становлению системы педагогических измерений фактически препятствуют идеи проведения единого государственного экзамена. Без чётких разграничений понятий - экзамены, тестирование и педагогические измерения - различие систем может показаться неясным.

В наше время существенный признак педагогических измерений – это педагогически обоснованный отбор содержания заданий теста, точный подбор форм заданий, подходящих для того или иного содержания, система основных педагогических понятий этой теории³⁴, получение данных об испытуемых и заданиях на интервальной шкале. Это достигается использованием соответствующих алгоритмов трансформации исходных результатов практического тестирования посредством компьютерных программ типа RUMM 2020, Winsteps и др.

Многолетнее отставание педагогических измерений от потребностей практики тестирования и систематическое отсутствие учёных в штате большинства центров тестирования на пространстве СНГ превращает практику тестирования в зону постоянного риска получения ненадёжных и невалидных результатов. Этот вывод легко подтверждается экспертизой работы любого центра тестирования. Вывести тестирование из зоны перманентного риска получения ошибочных оценок личности может только стабильная опора на науку – теорию педагогических измерений. Но для этого нужны научные подразделения в центрах тестирования, которых практически нет.

³³ Аванесов В.С. Основы теории педагогических измерений//ПИ, №1, 2004г.

³⁴ Аванесов В.С. Научные проблемы тестового контроля знаний. М.: Иссл. Центр, 1994. – 135с.

Проблему осложняет положение самой теории педагогических измерений. Все известные в этой теории методы обоснования качества тестовых результатов неизменно оказывались статистическими и математическими. О том, что мы имеем дело именно со статистическими теориями психологических и педагогических измерений, было чётко сказано в классической работе³⁵ конца шестидесятых годов прошлого века. Так сложилось исторически, таким был идеал научности тестовых результатов. Однако попытки построения собственно педагогической теории педагогических измерений, других идеалов научности не прекращались ни в прошлом, ни в наше время. Известны попытки создания педагогической теории педагогических измерений в США³⁶.

Идея необходимости перехода от тестирования к системе педагогических измерений возникла как результат сравнительного исследования вопросов методологии педагогических измерений и тестирования³⁷, научного анализа практики централизованного тестирования³⁸, введения российского Единого государственного экзамена (ЕГЭ)³⁹ и личного опыта взаимодействия с практиками–работниками центров тестирования. Последние чаще стремятся выполнять указания сверху, нежели рекомендации учёных. К тому же многие остепенённые практики считают себя учёными не в меньшей степени, чем те, кто работают в науке. Такого рода менталитет характерен и для небольшого числа чиновников многих министерств.

В качестве педагогической основы теории педагогических измерений автором этой статьи предлагаются:

³⁵ Lord, F.M., Novick, M. Statistical Theories of Mental Test Scores. N-Y, 1968.

³⁶ Например, в работе *Ebel, R.L. Measuring Educational Achievement*. Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1965. – 481pp.

³⁷ Аванесов В.С. Вопросы методологии педагогических измерений// Педагогические Измерения, №1, 2005г. С. 3-27. <http://testolog.narod.ru/Theory34.html>

³⁸ Аванесов В.С. Централизованное тестирование лучше Единого Государственного Экзамена// «Развитие тестовых технологий в России. Тезисы докладов Всероссийской научно-методической конференции / Под ред. Л.С. Гребнева.- М.: Центр тестирования Министерства образования РФ. 2002. <http://testolog.narod.ru>

³⁹ Аванесов В.С. Единый Государственный Экзамен в фокусе научного исследования// Педагогические Измерения, №1, 2006г. С.3-31. <http://testolog.narod.ru/Education39.html>

- концепция пяти этапов тестового педагогического процесса⁴⁰ и система основных понятий педагогических измерений⁴¹;
- формулировки целей педагогических измерений⁴²;
- результаты исследования содержания⁴³, форм⁴⁴ и принципов композиции тестовых заданий⁴⁵;
- математические модели⁴⁶ и аксиоматика теории педагогических измерений⁴⁷.

Система педагогических измерений включает в себя всю научную составляющую разработки и применения тестов для получения качественных тестовых результатов и последующего шкалирования, проведение рейтинга, мониторинга и качественных показателей образовательной деятельности. Основным предмет педагогической теории измерений – разработка качественных тестов для измерения уровня и структуры *подготовленности* учащихся и студентов. Такие тесты используются почти во всех современных образовательных технологиях.

Для того, чтобы педагогические измерения рассматривались как основа формирования объективной системы оценивания образовательной сферы наиболее эффективные средства противодействия коррупции. Но они сами должны отвечать некоторым требованиям:

Во-первых, они не должны быть государственными. В РФ и в государствах Средней Азии тестирование и ЕГЭ *de facto* является **государственным**. Это достаточное условие для того, чтобы контроль знаний время от времени подвергался риску искажений.

⁴⁰ Аванесов В.С. Пять этапов педагогических измерений. <http://testolog.narod.ru/Theory58.html>

⁴¹ Аванесов В.С. Определение исходных понятий. 3 редакция, 7 апреля 2007. <http://testolog.narod.ru/Theory46.html> ; Аванесов В.С. Понятийный аппарат теории педагогических измерений// <http://testolog.narod.ru/Theory26.html>

⁴² Аванесов В.С. Основы педагогической теории измерений// №1, 2004 г. С.15-21.

⁴³ Аванесов В.С. Форма тестовых заданий. М.: Центр тестирования, 2006.

⁴⁴ Аванесов В.С. Форма тестовых заданий. М. Центр тестирования, 2006. – 137 стр.

⁴⁵ Аванесов В.С. Теоретические основы разработки заданий в тестовой форме. Уч. пособие для профессорско-преподавательского состава высшей школы. М. МГТА, 1995. -95с. Аванесов В.С. Композиция тестовых заданий. М. Центр тестирования. 2003. 217 стр.

⁴⁶ Аванесов В.С. Математические модели педагогического измерения. Научное издание. М.: Иссл. центр проблем качества подготовки специалистов, 1994. - 26с.;

⁴⁷ Аванесов В.С. Педагогическое измерение латентных качеств //Педагогическая диагностика, №4, 2003г.

Во-вторых, измерения не должны быть средством коммерциализация образования. В условиях интенсивного расслоения населения это ведёт к фактическому ограничению доступа к качественному образованию, и далее - к усилению невежества молодежи, к конфронтации и политической нестабильности – и всё это в стране, имеющей большую армию, ядерное оружие и ракеты.

В-третьих, педагогические измерения могут быть объективными только тогда, когда правильно определен их научный *предмет*. В России, Казахстане и в Узбекистане - это уровень учебно-дисциплинарных знаний, в США и Израиле – способности к овладению программами высшего образования. В Киргизстане намечен переход к оценке способностей посредством адаптированного американского варианта SAT-2⁴⁸. Но школа остается всё ещё ориентированной на знания. Возникает вопрос - будут ли результаты тестирования честными для тех школьников, родители которых не в состоянии нанять репетитора для нового теста?

Выводы:

1. Для стабильного и качественного развития независимой российской общественно-профессиональной системы педагогических измерений надо осуществить перевод всех имеющихся в стране центров тестирования в статус независимых центров системы педагогических измерений. И не по названию. А по сути. Это создаст психологические и организационные предпосылки усилить научный потенциал центров, заметно повысить качество педагогических измерений, заменить менее качественное тестирование более качественными педагогическими измерениями.

2. ЕГЭ надо закрыть, и навсегда. К этой мысли, похоже, склонился, наконец, и сам основатель ЕГЭ, В.А. Болотов. В своём интервью по случаю десятилетия ЕГЭ на вопрос корреспондента он ответил коротко и ясно: «через десять лет никакого ЕГЭ не будет...»⁴⁹.

Вместо ЕГЭ в стране надо организовать настоящее добровольное независимое педагогическое тестирование уровня подготовленности школьников и абитуриентов. Для этого понадобится сеть новых центров тестирования, опирающихся на научные основы педагогических измерений.

⁴⁸ Scholastic Aptitude Test

⁴⁹ Юлия Варшавская. Через 10 лет никакого ЕГЭ не будет.
https://mel.fm/yege/9830471-bolotov_ege?utm_referrer=https%3A%2F%2Fzen.yandex.com

3. Для становления национальной системы педагогических измерений требуется взаимодействие граждан, педагогической общественности, научных и технических специалистов. Государство должно помогать такому взаимодействию.

4. Нужно принять юридические акты, обеспечивающие ответственность за разглашение правильных ответов к заданиям и индивидуальных результатов без согласия граждан, прозрачность организации тестирования, открытость и доступность, для исследователей и независимых экспертов матриц тестовых результатов (без указания фамилий), по всем учебным дисциплинам и регионам, за все годы. Обязательна также публикация отчетов по разработке тестов по итогам финансовой деятельности центров.

5. Вместо задуманного перехода от ЦТ к ЕГЭ, и далее к ОСОКО, лучше приступить к созданию в России независимой системы педагогических измерений.
