



Теория образования и обучения

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

ИСТОРИЯ ПЕДАГОГИЧЕСКОЙ ТЕОРИИ ИЗМЕРЕНИЙ¹

В статье рассматриваются вопросы истории и теории педагогических измерений, уточняются цели, задачи, объект и предмет этой теории, определяются результаты развития педагогической теории измерений.

Теория педагогических измерений включает в себя методические правила проведения тестирования, и, кроме того, три важных направления научной деятельности: формирование языка этой теории, теорию формы и вопросы отбора содержания тестовых заданий. Подлинное педагогическое измерение появляется после тестирования испытуемых, в процессе шкалирования их результатов, а также в процессе шкалирования заданий по их существенным свойствам.

Исторически сложилось так, что теории педагогических измерений предшествовала разработка вопросов теории оценивания². Оценивание можно определить как эксперт-

ный процесс выставления каких-либо баллов по заранее предписанным правилам. Автор этой статьи полагает научный статус оценивания ниже статуса педагогического измерения.

История возникновения общей теории измерений в общественных науках обстоятельно исследована Б. Райтом³. Он был уверен, что создаваемая им теория пригодна для всех общественных наук, включая педагогику. Так оно и случилось, и для того времени это был правильный ход мысли. Но рано или поздно в науках появляются тенденции расширения и разделения, выделения собственного предмета. А потому естественным образом возникла проблема отделения педагогической теории измерений от общей теории измерений в психологии, социологии и в других общественных науках.

¹ Статья печатается с сокращениями. Полный текст в журнале «Педагогические Измерения», № 3, 2012.

² *Smallwood, M.L.* A Historical Study of Examinations and Grading Systems in Early American Universities. Cambridge: Harvard Univ Press. Harvard Studies in Education v.24, 1935; *Caldwell O.W., Curtis S.A.* Then and Now in Education: 1845–1923. Harcourt, 1924. 400 p; *McArthur D.L.* Educational Assessment: A Brief History /McArthur D.L. (Ed). Alternative Approaches to the Assessment of Achievement. Kluwer Academic Publishers. Boston, 1987. 268 p.

³ *Wright B.D.* A History of Social Science Measurement. University of Chicago, MESA Psychometric Laboratory; *Wright B.D.* Fundamental measurement in social science and education. / Research Memorandum № 33a, MESA Psychometric Laboratory, University of Chicago, March 30, 1983.

Теория педагогических измерений, как научная сторона процесса тестирования, стала развиваться с небольшим опозданием от теории тестов. Начало теории педагогических измерений положил Эдвард Торндайк^{4,5}. В своих трудах он затрагивал широкий круг философских, психологических и педагогических вопросов измерений, а также профессионального отбора.

Новаторство Э.Л. Торндайка проявилось в том, что он ранее других понял необходимость развития не только теории обучения, но также и двух других теорий — учения (самостоятельного изучения) и теории педагогических измерений. С появлением электронного обучения значимость этих двух теорий была осознана глубже, чем это было раньше.

Со времён Э. Торндайка в педагогическом сознании западных педагогов укоренялась мысль о том, что основная сфера применения тестовых форм — это не столько педагогический контроль, сколько самоконтроль в процессе собственного учения. В России эта мысль получила, можно сказать, противоположное истолкование, что связано с иным типом государственного устройства и систем управления образованием, таким типом, где роль бюрократии преувеличена, а роль науки и учёных минимизирована.

Исторически сложилось так, что первыми, больше ста лет назад, появились труды по психологической теории измерений.

В 1936 году теоретики этого направления создали журнал высокого научного уровня под названием «Psychometrika». Сейчас он издаётся и распространяется по всему миру. Затем стало развиваться общее направление «Психологическое и педагогическое измерение»⁶. Язык и методы этого научного направления считались, видимо, по инерции, общими, равно применимыми в психологии и педагогике.

Позже последовало издание журнала по педагогическим измерениям⁷, где были сделаны попытки выделить собственный педагогический предмет исследования и создать собственные педагогические методы. Успеха, однако, эти попытки не имели из-за отсутствия общей методологии педагогических измерений.

Три теории измерений

Разработка педагогической теории измерений проходила на основе имеющихся теорий. Наиболее известны три теории измерения:

1. Статистическая или, иначе, классическая теория тестов (classical test theory, СТТ). Так она называлась в начале XX века. Сейчас слова «классическая теория» можно рассматривать как метафору. В подтверждение такого мнения можно привести два аргумента.

Во-первых, эта теория имеет точное, а не нынешнее метафорическое название. На

⁴ Thorndike E.L. Theory of Mental and Social Measurements. Science Press. 1904.

⁵ Thorndike, E.L. Educational Measurements of Fifty Years Ago / J. of Educational Psychology, 1913, 6: 551–552.

⁶ Educational & Psychological Measurement. Published by Sage Publications <http://www.sagepublications.com>

⁷ Journal of Educational Measurement. Published by The National Council on Measurement in Education (NCME).

английском языке она называется true-score model of measurement. Поэтому её лучше называть не теорией, а моделью измерений.

Во-вторых, эта модель — перенос идей теории физических измерений на измерения в психологии и педагогике самого начала XX века. Отсюда упрощённая, мало реалистичная для наших дней, аксиоматика. Её точнее называть моделью истинного и ошибочного компонентов измерения. Разумеется, эта модель уже обросла тысячами исследований подлинно теоретического характера. Поэтому название «теория» реально существует, оно приемлемо, за исключением случаев сравнения этой модели с другими моделями педагогического измерения.

2. Вторую теорию можно очень условно перевести на русский язык как математическая теория измерений (Item Response Theory, IRT)⁸.

3. Третья теория — больше чем теория. Это одновременно ещё методология и метрическая система. Потому она получила название метрической теоретической системы Г. Раша (Rasch Measurement, RM)⁹. Когда эту модель измерения называют теорией, то и другие модели можно называть теориями, как это случилось, например, с IRT. Хотя другие, позже появившиеся модели измерения, так и остались моделями. Да и более

точное название IRT — это Latent Trait Model of Measurement.

Все перечисленные — относительно общие теории, применимые не только в педагогике, но и в психологии, социологии, в медицинских и иных исследованиях. Но общие теории не давали и не могли дать адекватные ответы на специфические педагогические проблемы. Нужна была специфическая педагогическая теория.

В педагогической теории измерений выделены три предмета исследования: педагогические понятия теории, принципы отбора педагогического (предметного) содержания заданий и логические принципы формулирования педагогических тестовых заданий.

Существенные признаки педагогических измерений — это педагогически обоснованный отбор содержания теста, точный подбор формы заданий для того или иного содержания, система основных педагогических понятий теории¹⁰. Удивительно, что измерения педагогических латентных свойств, в том числе относящихся к педагогической науке, до сего дня некоторые зарубежные авторы считают задачами психометрики, с чем согласиться невозможно.

Новая история теории педагогических измерений

Началом целенаправленных исследований по созданию именно педагогической теории измерений в России можно считать решение бывшего Комитета по высшей школе о создании Учебного центра при Исследо-

⁸ Аванесов В.С. Item Response Theory: Основные понятия и положения. Статья первая // Педагогические Измерения № 2, 2007. С. 3–28; Аванесов В.С. Истоки и основные понятия математической теории педагогических измерений (Item Response Theory). Статья вторая // Педагогические Измерения № 3, 2007. С. 3–36.

⁹ Аванесов В.С. Применение тестовых форм в Rasch Measurement // Педагогические Измерения № 4, 2005. С.3-20.

¹⁰ Аванесов В.С. Научные проблемы тестового контроля знаний. М.: Иссл. Центр, 1994. 135 с.

тельском центре бывшего Госкомитета СССР по высшей школе.

Конечно, некоторые авторы могут возразить, что и раньше были публикации, тяготеющие больше к вопросам теории педагогических измерений, чем к вопросам практики тестирования. Но там всё-таки было больше тяготения к теории тестов. Здесь в качестве отправного пункта принят факт создания специализированной кафедры педагогических измерений в упомянутом Учебном центре. Насколько известно, в истории СССР, а затем и России, создание специализированной кафедры в научном учреждении случилось впервые.

Подходящих учебных материалов на русском языке тогда не было, а то, что имелось в печати, для учебного курса не годилось. Такие материалы пришлось создавать в процессе работы кафедры. Началась целенаправленная работа по исследованию возможностей тестовых форм для улучшения качества учебного процесса в вузах. Первым заметным итогом той работы стало учебное пособие для профессорско-преподавательского состав вузов¹¹. Именно в том пособии были сформулированы и раскрыты вопросы истории, теории, методики и практики педагогического контроля.

В те годы не было принято писать о тестировании и педагогических измерениях. У многих ещё хранился в памяти запрет на применение тестов. Отсюда и вынужденная редакционная замена непривычных понятий «тестирование» и «измерения» на

приемлемые тогда понятия «проверка» и «контроль».

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 создания качественных показателей образовательной деятельности. Основной 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.

Исходные понятия теории педагогических измерений

В тот период было начато решение несколько проблем педагогических измерений¹². Первой из них стала проблема разработки исходных понятий теории¹³. Понятия педагогической теории измерений целенаправленно начали формулироваться в период написания автором этой статьи диссертации на соискание учёной степени доктора педагогических наук. Для того чтобы избежать обычной путаницы в толковании сути рассматриваемых научных положений, необходимо было с самого начала дифференциро-

¹¹ Аванесов В.С. Основы научной организации педагогического контроля в высшей школе. Пособие для слушателей учебного центра Гособразовании СССР. М. МИСиС, 1989 г.

¹² Аванесов В.С. Методологические и теоретические основы тестового контроля. Дисс... докт. пед. наук. СПб госуниверситет, 1994. 339 с.

¹³ Аванесов В.С. Определение исходных понятий. <http://testolog.narod.ru/Theory46.html>; Аванесов В.С. Понятийный аппарат теории педагогических измерений// <http://testolog.narod.ru/Theory26.html>

вать два главных понятия: педагогическое тестирование и педагогические измерения.

Педагогическое измерение — это преимущественно научная деятельность, результатом которой является методология и теория организации практического тестового процесса, такого, который позволяет в итоге получить объективированные результаты на интервальной, а возможно, и на пропорциональной шкале¹⁴.

Педагогический тест — это научно обоснованный и технологичный метод исследования результатов учебной деятельности посредством системы заданий возрастающей трудности, позволяющей качественно оценить структуру и измерить уровень подготовленности испытуемых.

В учебном процессе преимущественно применяются не тесты, а *задания в тестовой форме*. Если последние сделаны правильно, с учётом формальных, содержательных, логических и семантических требований. Но правильная разработка заданий в тестовой форме встречается редко. То, что сейчас иногда называют «тестами ЕГЭ», не является ни тестами, ни педагогическими измерениями¹⁵.

Н.Ф. Ефремова считает, что тесты являются наиболее эффективной и объективной формой *оценивания* знаний¹⁶. Она относит тест не к формам педагогического измере-

ния, а к формам оценивания, таким, например, как экзамен, зачёты, КИМы ЕГЭ и т.п. С этим трудно согласиться. Отнесение разных объектов к одному классу может способствовать лишь риску потери уникальных свойств теста, в отличие от экзаменов и пр.

Тесты обычно являются средством итогового, по содержанию, внешнего контроля. В России методом внешнего, относительно школ, контроля стал единый государственный экзамен. Но, по сути, это не тест и не экзамен в обычном понимании смысла этих педагогических форм¹⁷.

Измерения можно также определить как процесс перехода от изучения интересующего свойства к созданию переменной величины, являющейся операциональным выражением данного свойства. Педагогические измерения имеют дело с латентными педагогическими свойствами личности, формируемыми в процессе образовательной деятельности и воспитания.

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 надёжности и валидности тестовых результатов с методами шкалирования.

Результат измерения интересующего свойства выражается точкой на прямой линии, представляющей меру интересующего свойства. Отсюда и возникло одно из

¹⁴ Каргин Ю.Н. Идея педагогических измерений в шкале отношений // Педагогические Измерения № 2, 2012 г.

¹⁵ Аванесов В.С. Являются ли КИМы ЕГЭ методом педагогических измерений? // Педагогические Измерения № 1, 2009. С. 3–26.

¹⁶ Ефремова Н.Ф. Тестовый контроль в образовании. Учебное пособие. М.: Университетская книга, Логос; Москва; 2007.

¹⁷ Аванесов В.С. И не тест, и не экзамен... Ж. «Университет и школа» № 1–2, 2002. С. 26–27.

самых коротких определений педагогического измерения¹⁸.

Цели и задачи педагогических измерений

Ещё одна проблема, которая решалась в последние годы — это формулирование целей и задач.

Сейчас цели тестирования и педагогических измерений лучше разделять. Цель тестирования — получение тестовых результатов уровня подготовленности испытуемых, цель же педагогических измерений в основном — получение шкалированных значений уровня подготовленности испытуемых и уровня трудности заданий.

Цели педагогических измерений можно разделить на метрические, педагогические, психологические и социально-политические.

Пример метрической цели — построить шкалу уровня подготовленности испытуемых или шкалу уровня трудности заданий. В рамках метрической цели нередко используется сравнительно узкое, фокусированное определение цели педагогического измерения — как позиционирование испытуемого на прямой линии, представляющей латентную переменную величину¹⁹.

Для позиционирования испытуемых на шкале их надо тестировать посредством заданий, определяющих переменную величину, а затем посредством шкалирования убедиться в том, что полученные тестовые

результаты помогают определить место испытуемого на латентной переменной величине.

Пример педагогической цели — провести классификацию испытуемых по уровню их подготовленности, для комплектования уровневых групп учащихся. Психологи используют тестовые результаты в целях выявления психологических затруднений в учебном процессе.

И наконец, пример социально-политической цели — применение тестов для обеспечения равного доступа учащихся к качественному образованию. Для достижения последней цели вместо педагогических измерений в России стали использоваться некачественные КИМы ЕГЭ.

Основную задачу педагогических измерений Ю.Н. Каргин видит в поиске оптимальных оценок показателей подготовленности испытуемых и трудности заданий по наблюдаемым результатам тестирования²⁰.

С момента возникновения психометрики считается, что латентная переменная величина образует так называемый континуум измеряемого свойства²¹. Это означает непрерывность значений интересующего свойства на шкале, отсутствие каких-либо ограничений на принимаемые значения слева и справа от среднего арифметического результата.

Контрольно-измерительные материалы обернулись метафорой, не содержащей в себе ничего метрического. В реальности был только госконтроль посредством преи-

¹⁸ A measure is a location on a line. Measurement is the process of constructing lines and locating individuals on lines. In: *Wright, D. N. and M. H. Stone. Best Test Design. Chicago, 1979.*

¹⁹ *Wright B.D., Stone M.H. Best Test Design. P.3.*

²⁰ *Каргин Ю.Н. Педагогические измерения в шкале отношений // Педагогические Измерения, № 2, 2012. С.14.*

²¹ *Guilford J.P. Psychometric Methods. McGraw-Hill, 1954.*

мущественно нетестовых заданий. Противоречия между теорией педагогических измерений и практикой разработки КИМов ЕГЭ скрыть не удавалось.

Образовательно-политический фактор

*В российской системе образования одновременно сосуществуют три процесса, три состояния: реформа, застой и откат назад*²².

Действительно, в России много лет мирно уживаются три названных *процесса, три состояния: реформа, застой и откат назад* или, иначе, деградация образования. Это вызвано образовательной политикой, действующими силами и традициями. Развитие образования востребовано родителями и общественностью, обеспечивается педагогами-новаторами снизу, разрушительные же реформы навязываются бюрократией.

А потому в России всё может превратиться в политический фактор, даже ЕГЭ и теория педагогических измерений. Высокие значения погрешностей получаемых оценок ЕГЭ оказались неприемлемыми для экзаменуемых и их родителей. На дефекты ЕГЭ в Правительстве РФ не обращали должного внимания²³. Вопрос качества ЕГЭ власть не считает важным. А потому этот вопрос с момента возникновения некачественного ЕГЭ стал постепенно превращаться в фактор ошибочной образовательной политики. Наверное, поэтому перед началом ежегодного государственного экзамена выс-

шие руководители страны обещают его улучшить.

В.В. Путин признал, что существует *очень много проблем с Единым государственным экзаменом, которые надо как-то решать*²⁴. Но вслед за этой сентенцией ничего путного не последовало. Д.А. Медведев неоднократно заявлял о необходимости разобраться в сути проблем, возникающих с проведением ЕГЭ²⁵ и о корректировке²⁶ этого экзамена. Была даже создана Кремлёвская комиссия по ЕГЭ, которая тоже не смогла в нём что-то улучшить.

Никто так и не объяснил гражданам, что конкретно нужно изменить в этом ошибочно скроенном экзамене, у которого по сей день нет ни автора, ни проекта, ни теоретического обоснования, ни научно проверенных результатов, но на который вхолостую тратится много государственных денег — по некоторой информации, порядка 33 миллионов рублей в год.

Похоже, что В.В. Путин и Д.А. Медведев теперь решили дистанцироваться от этой политически проигрышной темы, поскольку процент населения, поддерживающий идею проведения ЕГЭ, с каждым годом снижается.

В 2012 году очередные обещания улучшить ЕГЭ было поручено дать главе админи-

²² Основные положения концепции очередного этапа реформирования системы образования // Народное образование № 8 1997. С. 5.

²³ Впрочем, обращают внимание, но реагируют весьма специфично.

²⁴ Юлия Котова. Президенту доложили об успехе ЕГЭ. <http://www.monich1.ucoz.ru/publ/1-1-0-78FirefoxHTML%5CShell%5COpen%5CCommand>

²⁵ Дмитрий Медведев поддерживает идею о президентской комиссии по ЕГЭ. http://www.rian.ru/edu_egrus/20090810/180375075.html

²⁶ А. Черных, Х. Дачаева, М. Юришина, А. Саргин. Президент поставил оценку системе ЕГЭ. 31.08.2009. <http://gzt.ru/topnews/business/258180.html>

страции Президента РФ Сергею Иванову. В его речи прозвучали элементы критики. Ещё критичнее выступил новый министр образования и науки, который в довольно жёстком тоне заявил, что «российские научные и образовательные организации за последние 20 лет оказались полностью неконкурентоспособными. А потому нуждаются в радикальной реформе»²⁷.

Ввязавшись в проведение ЕГЭ без должной теоретической проработки проблемы, бывшее Министерство образования допустило стратегическую ошибку. Были келейные обсуждения, оппонентов не слышали, да и слушать не хотели. В общем, всё было так же, как и сейчас. Всё это — последствия своеобразного стиля работы. А потому ошибки в управлении образованием продолжают множиться.

О необходимости изменений немало говорилось и бывшими министрами. Например, в 2006 году А. Фурсенко признал, что, например, качество содержания российского образования совершенно не отвечает современным требованиям, и что 80% учебников допускаться к обучению не должны. После чего он допускал эти учебники в школу и продолжал управлять сферой образования ещё шесть лет, ничего не сделав для исправления ситуации. Итог таков — учебники в России становятся всё хуже и хуже, образование же продолжает деградировать²⁸.

Не было ясности в ключевых целях и задачах проводимых реформ, в методах

и формах образовательной деятельности, в критериях достижения обтекаемо сформулированных целей. В общем, во всех российских образовательных реформах последних 20 лет не было заметной опоры на существующие теории и методологии. Это были бюрократические реформы.

С годами ЕГЭ всё больше удалялся от требований, предъявляемых к педагогическим тестам, постепенно превращаясь, таким образом, в отчасти технологизированный массовый государственный экзамен. Это и есть его настоящее название. Используемые в нём так называемые контрольно-измерительные материалы не имеют признаков пригодности для проведения педагогических измерений. И это уже доказанные факты, признавать которые министерство не спешит.

С точки зрения теории педагогических измерений *не бывает* тестов в бланковом формате, одинаково точно измеряющих испытуемых разного уровня подготовленности. Это аксиома, которую никому ещё не удалось опровергнуть. Практика проведения некачественного ЕГЭ лишь подтверждает высокую погрешность получаемых оценок. От такого метода больше вреда, чем пользы. Такая погрешность оценок унижает для страны. Дети имеют право на объективное измерение уровня их подготовленности. К сожалению, ЕГЭ это право нарушает. Кодекс профессиональной чести тестологов на Западе не разрешает использовать методы измерения с неизвестными показателями точности. В России это разрешает делать государство.

Сложившееся отношение к педагогическим измерениям абсурдно. Альбер Камю

²⁷ Дмитрий Ливанов. РИА Новости. http://ria.ru/edu_news/20120621/678237385.html

²⁸ Петренко В. Учебники в России всё хуже и хуже. http://www.gazeta.ru/education/2006/09/20_a_837211.shtml

был прав, утверждая, что чувство абсурдности поджидает нас на каждом шагу. Мы уже много лет живём в антиобразовательном абсурде. И не видно пока просвета.

Основные результаты развития педагогической теории измерений

Как уже отмечалось выше, в работах последних лет был сформулирован понятийный аппарат педагогической теории, даны определения основным понятиям²⁹. В качестве таковых выделены: *задание в тестовой форме, тестовые задания, педагогический тест, испытуемые, тестовый процесс и содержание теста*.

Предметом специального анализа стали также два других важных понятия педагогической теории измерений. Первое понятие — *форма тестовых заданий*. Это способ связи, упорядочения элементов задания. Тысячи примеров заданий четырёх основных форм по многим учебным дисциплинам читатель найдёт в книгах автора^{30, 31}.

Показано, что логическое преимущество задания в тестовой форме, по сравнению с привычными вопросами экзаменов и госэкзамена, заключается в возможности превращения заданий в тестовой форме, после ответа испытуемых, в форму высказываний,

²⁹ 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>

³⁰ Аванесов В.С. Форма тестовых заданий // Труды Иссл. центра. М. 1991. 33 с. Аванесов В.С. Форма тестовых заданий. М.: Центр тестирования, 2005 г.

³¹ Аванесов В.С. Композиция тестовых заданий. М.: Аванесов В.С. Педагогическое измерение латентных качеств // Педагогическая диагностика, №4, 2003 г.;

истинных или ложных. Что создаёт надёжную основу для безошибочного дифференцирования испытуемых по уровню их подготовленности. Задания в тестовой форме выгодно отличаются от обычных вопросов свойствами краткости, лучшей понимаемости смысла заданий, быстроты ответа учащихся, технологичности и эффективности.

В современном тестировании часто применяется дифференцирование испытуемых, по их ответам на каждое задание, не только на два, но и на большее число уровней. Это означает, что оценки даются за ответы правильные, неправильные, а также за частично правильные, с разными уровнями градации правильности. Такое оценивание ответов испытуемых называется по-английски Partial Credit Model. Оно имеет преимущества перед традиционным дихотомическим оцениванием посредством единицы и нуля. Главное преимущество — повышение надёжности тестовых результатов.

В качестве главной *проблемы* теории педагогических измерений названа разработка показателей качества и эффективности образовательного процесса.

Предмет педагогических измерений — уровень подготовленности испытуемых, который представляет собой единство знаний, умений, навыков, представлений. К понятию «подготовленность» можно отнести и часто используемое сейчас понятие «компетентность», представляющую собой прагматическую, или утилитарную, проекцию общего содержания образования.

При педагогических измерениях одно и то же свойство личности проявляется у разных испытуемых в разной степени. Обычно большему проявлению интересую-

шего свойства ставится в соответствие и большее число. Реже поступают наоборот. Это бывает в случаях измерения общественно неодобряемых психологических свойств личности.

Кроме упомянутых выше результатов, в рамках развития педагогической теории измерений исследованы два традиционных критерия качества³² — надёжности и валидности тестовых результатов, принятых в статистической теории измерений, и два относительно малоизвестных критерия — объективности³³ и эффективности³⁴ тестовых результатов, используемых преимущественно в РМ и в адаптивном тестировании.

Сформулированы также принципы отображения содержания тестовых заданий³⁵, концепция пяти этапов тестового педагогического процесса³⁶, дана обновлённая классификация видов проверяемых знаний школьников и студентов³⁷. Самое большое внимание в рамках разработки педагогической теории измерений было уделено вопросам формы³⁸

и принципам композиции тестовых заданий³⁹.

Три главных свойства заданий подлежат обязательному шкалированию и доскональному теоретическому и эмпирическому педагогическому исследованию — это уровень понимаемости смысла содержания заданий, мера их трудности и дифференцирующей способности. В России система такого рода метрической работы по изучению главных свойств педагогических заданий пока ещё не сложилась.

Не были обойдены вниманием и вопросы становления Национальной системы педагогических измерений⁴⁰. Эту проблему Правительство РФ пыталось решить по-своему, и безуспешно, посредством ОСОКО (Общероссийской системы оценки качества образования), в центре которой стоял тот же ЕГЭ. Новые варианты ОСОКО не лучше прежней системы.

Определены пять основных направлений развития теории и практики педагогических измерений:

1. Демаркация (отграничение) педагогических измерений от контрольных материалов государственных экзаменов и от иных псевдометрических форм и методов.

2. Разработка новых форм тестовых заданий.

³² *Аванесов В.С.* Проблема качества педагогических измерений // Педагогические Измерения, № 2 2004.

³³ *Аванесов В.С.* Проблема объективности педагогических измерений // Педагогические Измерения № 3, 2008. С. 3–40.

³⁴ *Аванесов В.С.* Проблема эффективности педагогических измерений // Педагогические Измерения № 4, 2008. С. 3–24.

³⁵ *Аванесов В.С.* От заданий в тестовой форме — к тестовым заданиям // Педагогические измерения № 4, 2010 г. С. 105–110.

³⁶ *Аванесов В.С.* Пять этапов педагогических измерений. <http://testolog.narod.ru/Theory58.html>

³⁷ *Аванесов В.С.* Знания как предмет педагогического измерения. // Педагогические измерения, № 3, 2005.

³⁸ *Аванесов В.С.* Форма тестовых заданий. М.: Центр тестирования, 2006. 137 стр.

³⁹ *Аванесов В.С.* Теоретические основы разработки заданий в тестовой форме. Уч. пособие для профессорско-преподавательского состава высшей школы. М.: МГТА, 1995. 95 с. *Аванесов В.С.* Композиция тестовых заданий. М.: Центр тестирования. 2003. 217 с.

⁴⁰ *Аванесов В.С.* Проблема становления Национальной системы педагогических измерений. Харьков, 13–14 декабря, 2007 г. <http://testolog.narod.ru/Education19.html>

3. Существенное обновление содержания тестов и тестовых заданий.

4. Централизованная разработка показателей качества и эффективности педагогических измерений и образовательной деятельности.

5. Улучшение языка педагогических измерений.

Выводы и предложения

1. История тестов и практический современный опыт стран с развивающимся образованием позволяют сделать вывод о необходимости скорейшей трансформации государственной формы контроля знаний учащихся и студентов в общественно-профессиональные формы.

2. История педагогических измерений подтверждает важную роль науки в организации систем контроля качества образования. Как показывает опыт внедрения российского ЕГЭ, без развития педагогических измерений существующие формы государственного экзамена вырождаются в примитивные бюрократические и асоциальные формы контроля.

3. Практика игнорирования педагогической теории измерений уже нанесла большой вред образовательной деятельности в России, и может нанести ещё больший

вред, если образовательная политика не начнёт меняться в лучшую сторону.

4. Для каждого уровня обучения и для каждого учебного предмета предлагается создать тысячу тщательно отобранных по содержанию и технологичных по форме заданий возрастающей трудности, доведённых до каждого педагога и учащегося. Эта тысяча способна радикально изменить качество массового образования к лучшему. Потому что появляется принципиально новая и эффективная единица учёта качества собственной учебной работы школьников и студентов.

5. Наличие двух и более тысяч таких заданий, умелое их применение для самостоятельной работы каждого учащегося позволят совершить рывок в сфере образования. Потому что такие технологичные задания начинают выполнять главную функцию развития учащихся и студентов. Всё остальное лишь способствует этой функции.

Это предложение можно рассматривать как одну из современных электронных и метрических версий известного в прежней педагогике задачного подхода в обучении. Разработка систем заданий возрастающего уровня трудности по всем изучаемым темам — проблема государственной важности.