
ЖУРНАЛ НАЧИНАЕТ ДИСКУССИЮ НА ТЕМУ РЕЙТИНГОВ

Вопрос об объективных методах сравнения и измерения качества образовательных институтов и систем сравнительно нов. Для постсоветского образования он революционен. Ведь от идеи обязательного соответствия эталону мы приходим к картине существенно более многообразной. Но это многообразие нуждается в упорядочивании, выделении внутренних структур и иерархий. Именно в этом (а вовсе не в конкуренции) состоит интеллектуальный пафос всевозможных рейтингов. Именно к этой проблеме — проблеме сравнения сложных социальных систем — и обращается наш журнал в рубрике «Дискуссии». Мы планируем публиковать как собственно рейтинги, так и обсуждение различных методологических подходов к их составлению. Новая под рубрика открывается полемической статьей, содержащей критический анализ первых результатов эксперимента по формированию системы рейтингования регионов по показателям развития образования, и ответом на нее разработчиков этого эксперимента.

Е.Е. Чепурных, А.Н. Майоров

РЕЙТИНГ ЗАБЛУЖДЕНИЯ ИЛИ ЗАБЛУЖДЕНИЯ РЕЙТИНГА

Статья поступила
в редакцию
в декабре 2006 г.

Критический анализ результатов
научного исследования

Статья содержит критический анализ работы по построению рейтинга региональных образовательных систем. Показано, что рейтингование объектов в сфере образования имеет ряд существенных ограничений и что этот специфический инструмент адекватно может быть использован в довольно узком сегменте. Рассмотрены вопросы качества исходных данных, процедуры работы с ними, необходимые для подготовки высококачественной управленческой информации. Обсуждается также вопрос об адекватности выводов и оценок на основе анализа количественных данных. Даны конкретные предложения по использованию рейтингования в сфере образования и по направлениям перспективных научных исследо-



ваний, которые позволят в значительной мере решить проблему информационного обеспечения управленческих процессов в этой сфере.

Введение

В тематическом приложении № 1 за 2006 г. сборника приказов и инструкций Министерства образования и науки Российской Федерации «Вестник образования» был опубликован «Рейтинг субъектов Российской Федерации по показателям развития образования».

К сожалению, какого-то комментария от Минобрнауки России в сборнике нет, как нет и комментария от редакции сборника, поэтому читателям приходится только догадываться об отношении Министерства к этим материалам и об их статусе. Из авторских комментариев и оговорок следует, что этот материал представляет собой результаты научной работы на первом промежуточном этапе ее реализации. В таком случае правомерен вопрос о смысле опубликования этой работой под грифом «приказы и инструкции»?

Учитывая серьезность поднятой темы и почти официальный характер публикации, возникает необходимость обсудить поднятие авторами проблемы, рассматривая приведенную публикацию отчасти как предмет для анализа и, учитывая проделанную авторами работу, в незначительной мере как и предмет для критики. Мы считаем возможным обсудить проблему рейтингов не только применительно к данному случаю, но шире, в общем контексте информационного обеспечения управления в сфере образования и использования в нем рейтингов.

В статьях, предваряющих рейтинги, совершенно справедливо отмечено, что сегодня отсутствует «информация о результативности реализации государственной политики в области образования, что в распоряжении образовательного сообщества нет данных о состоянии образования, как в целом, так и по наиболее важным проблемам, сформулированным в качестве главных задач модернизации, таким, как обеспечение доступности и повышение качества образования». Однако в полной мере отсутствие этой информации характерно для системы управления образованием на федеральном уровне, о чем свидетельствует вступительная статья председателя Комитета образования Новгородской области В.Н. Аверкина. Сегодня практически в каждом субъекте Российской Федерации созданы информационно-диагностические службы, которые в меру собственных сил и понимания решают проблему информационного обеспечения управления необходимой информацией. К сожалению, несмотря на давно ведущиеся разговоры и несколько неудачных попыток, на федеральном уровне не создано эффективного центра, который бы мог координировать работу региональных информационных служб, унифицировать инструментарий, проводить обработку информации для получения сравнительных данных по регионам Российской Федерации.

В этом отношении факт публикации рейтингов выглядит как надежда авторов на некую панацею — вот будет рейтингование и



будет получен «надежный инструмент образовательной политики», будет решена проблема «ответственности за результаты проводимой образовательной политики» и «неэффективного использования бюджетных средств» и т.д. Это несколько наивная позиция, не имеющая, как мы увидим позже, никаких оснований для реализации через предложенный инструмент. Вместе с тем авторы очень четко и однозначно определяют, чего же все-таки не хватает в информационном обеспечении управления образованием на федеральном уровне.

Нельзя сказать, что на федеральном уровне не предпринимается никаких усилий в этом направлении: есть работы по представлению данных российской образовательной статистики по международным индикаторам, разработаны макеты публичных докладов о деятельности региональных и муниципальных органов управления образованием, проведена унификация показателей деятельности образовательных систем, выполнен целый ряд других достаточно полезных и значимых работ.

Однако заметных системных изменений, к сожалению, не происходит. Давно ведутся обсуждения и финансируются разработки по модернизации системы государственной статистики. Работы над изменением статистики идут, по крайней мере, последние 10 лет, но до сих пор существенных изменений не произошло. Работу по рейтингованию регионов можно было бы рассмотреть в ряду этих усилий.

Мы попытаемся проанализировать рассматриваемую работу с двух точек зрения: во-первых, рейтинг как метод оценки в образовании, и во-вторых, рейтинговая оценка регионов в поле информационного обеспечения управленческих решений.

Анализ имеющихся определений и практики дает три основных понятия для рейтинга социальных систем. Первое — числовой показатель, характеризующий систему и привязанный к ранжированному ряду. Второе — ранжированный ряд, который может быть представлен как с данными, на основании которого он был получен, так и без него. Третье — это, как правило, буквенно-числовой индекс, относящий ту или иную систему к группе однотипных по определенному показателю.

С этой точки зрения представленные в рассматриваемой работе данные имеют некоторую лукавость. Структура таблицы представленных данных выглядит следующим образом:

№п/п	Субъект Российской Федерации	Рейтинг	% к первому месту
1	2	3	4

В данном случае первый столбец и представляет собой никак не номер по порядку, а собственно рейтинг, а третий столбец — это численное значение показателя, на основании которого построен рейтинг. И это не только проблема измерения и правильной терминологии, это проблема управленческая. Первое, на что смот-



рят управленцы, да и все, кто интересуется данными, — какое место в списке занимает тот или иной объект. И дальше уже мало кто разбирается, насколько значимы различия между регионами, почему относительные данные просчитаны с ошибками и т.п.

К сожалению, в рассматриваемой нами работе был принят самый примитивный способ построения рейтингов, при котором не учитывались ни особенности, ни условия развития образовательных систем, агрегировались случайные наборы показателей, не группировались объекты рейтингования, а выстраивался только ранжированный ряд, не было и какого-то описания полученных данных, раскрывающих их педагогический или управленческий смысл.

Таким образом, в неудобном положении оказываются руководители региональных органов управления образованием, которые вынуждены искать объяснения, искать ошибки и т.д. Да и с точки зрения развития рейтингования как инструмента оценивания такая работа наносит вред — ну кто в следующий раз будет серьезно относиться к таким оценкам?

Рейтинги в образовании

В сфере управления образованием существует сложившийся сегмент информационного обеспечения, где рейтинги используются достаточно давно и успешно — это рейтинги вузов и рейтинги учащихся. Что касается рейтингов вузов, то достаточно часто они печатаются без приведения числовых данных: например, журнал «Studentspiegel» составил по 50 параметрам рейтинг немецких университетов. Кроме общего рейтинга вузы оценивались также по ряду предметов, таких, как экономика, биология, химия, электротехника, германистика, информатика, право, математика, медицина, физика, политология, психология, социология.

В Америке признанный профессиональным сообществом рейтинг по 50 ведущим вузам США начал публиковать в 1983 г. журнал «US News and World Report». В его основу легли два показателя: репутация вуза и карьерный рост выпускников. Сегодня ряд известных изданий — такие, как «Times», «Newsweek», «Guardian» и др., — публикуют рейтинги вузов на основе их вклада в общественное развитие.

Среди главных критериев, которые используются в западных рейтингах вузов: 1) оценка вуза академическим сообществом; 2) оценка работодателями; 3) какой конкурс был в этот вуз; 4) количество выпускников, избравших научную деятельность; 5) количество членов национальной академии; 6) количество преподавателей, имеющих докторскую степень. Всегда важно и количество иностранных профессоров с мировым именем.

Рейтинги готовятся как отдельно по разным показателям, так и в обобщенном виде.

Есть своя история и у рейтингования вузов в России. Официальный рейтинг вузов Минобрнауки России размещен на сайте, опубликована методика его определения. В нем рассматриваются формальные показатели ресурсов вузов: книги, площади, ученые



степени преподавателей, наличие доступа в Интернет, компьютеров и т.д. Достаточно хорошо известны проблемы рейтингования вузов в России. Это разрешение ситуации, когда и при хороших ресурсных показателях вуза выпускник не пользуется спросом на рынке труда, это проблемы рейтингования негосударственных вузов и отдельного рейтингования филиалов и др. Для решения этих проблем необходимы внешние социологические, статистические и экономические исследования, например, карьерных достижений выпускников вузов, уровень их зарплаты и т.д.

Рейтинги в образовательных учреждениях различных стран мира широко применяются для оценки:

конкурса в вузы — количество абитуриентов на имеющиеся места;

студентов — назначение стипендий, отбор для программ обучения;

выпускников учреждений профессионального образования — для первичного отбора на работу.

Вводимый в нашей стране ЕГЭ по своей сути (при некоторых оговорках) представляет собой систему рейтингования выпускников общеобразовательных учреждений.

Рейтинги учащихся достаточно активно использовались в дореволюционной России, где оценки гимназистам давались по тому месту, которое они занимали в итоговой сводной ведомости успеваемости по классу, параллели или всему образовательному учреждению. Кстати, заметим, что это давало возможность решить проблему учета достижений ученика, его продвижения: если в предыдущем году он был 62-м из 115, а в этом стал 42-м, то это и было отражением его продвижения, роста. Вместе с тем такая система оценивания имела и целый ряд существенных недостатков, что привело к их вымыванию из системы общего образования, но об этом мы поговорим немного позже.

Таким образом, и в системе образования есть достаточно широкий сегмент, в котором использование рейтингов целесообразно, оправдано и имеет достаточно большую историю.

А теперь давайте посмотрим, что общего у объектов рейтингования, какие у них особенности, что дает основание применять к ним такую форму оценивания.

Общим для них является то, что они находятся в одинаковых условиях относительно пользователя информации рейтинга. Например, ученики, студенты обучались по одной и той же программе или в рамках одного стандарта, а система оценивания соответствует ему, так же как рейтинг спортсменов, рассчитывается при условии, что они имеют одинаковые возможности для участия в соревнованиях.

Специфика оценивания такова, что состояние объекта и условия его деятельности не важны с точки зрения пользователя. Например, абитуриента при выборе вуза не интересуют сложности его существования, он выбирает абсолютно лучшее для себя, то



же можно сказать и про инвестора, который выбирает соответствующее кредитное учреждение.

Такие разные
объекты
для сравнения

Таким образом, можно говорить о первой из выделенных нами проблем — это проблема идентичности объектов измерения относительно использования результатов рейтинга и, как следствие, идентичности условий их функционирования и развития. Однако закономерны вопросы: насколько внешние условия, в которых существуют эти объекты, оказывают влияние на измеряемый в рейтинге показатель, насколько эти объекты должны быть одинаковы относительно измеряемых свойств и насколько это важно для потребителей информации в сфере образования?

Рассмотрим эти вопросы на примерах. Для рейтингов учащихся в условиях школы, класса или вуза, в общем-то, сохраняется идентичность условий. Как правило, учащиеся занимаются в более-менее одинаковых условиях. Вместе с тем хорошо известно, что существуют девять взаимосвязанных факторов, оказывающих влияние на обучение детей (первые пять имеют наибольшее влияние): биологический (физиологический), семья, класс, школа, соседи, город, регион, нация, глобальные влияния.

Таким образом, рейтингуя учеников, мы должны понимать, что кроме собственно вклада, который дает результат систематического обучения, он оказывается нагруженным и другими факторами, на которые школа может оказать влияние (например, улучшить качество преподавания) или может оказать влияние при определенных условиях (например, получить право проведения отбора) или не может повлиять никак (например, изменить инфраструктуру социокультурного окружения школы). Именно здесь кроются причины проблем системы оценивания учеников вообще и рейтингования в частности. Выход из создавшейся ситуации многие видят в том, что основой рейтинга должна служить не сама величина показателя в когнитивной области, а ее изменение для отдельного ученика. Это представляется совершенно верным, но, во-первых, это значительно усложняет и удорожает процедуру оценивания, во-вторых, существует целый ряд практических задач, при которых абсолютных значений вполне достаточно, а в-третьих — и в этом заключается существенный позитивный потенциал рейтинговой системы оценки, — при проведении повторных, и достаточно частых измерений изменение места в рейтинге будет нагружено нужными нам факторами, поскольку остальные будут рандомизированы. Но здесь мы сталкиваемся с изменением смысла измерения — по своей сути это уже не рейтингование, а система мониторинга.

Переместимся на более высокий уровень и посмотрим, как обстоят дела при попытках построить рейтинг класса или образовательного учреждения. Достаточно очевидно, что, выбирая в качестве единицы измерения класс в каком-то образовательном учреждении, следует учитывать его социально-демографические характеристики. В настоящее время никому не придет в голову срав-



нивать результаты образовательной деятельности городской гимназии с конкурсом в несколько человек на место и школы, в которую принимают всех учащихся микрорайона, расположенного на окраине города.

Посмотрим, какой опыт накоплен для этого случая. Сразу скажем, что отечественная практика стратификации образовательных учреждений отличается от западной, где территориальная стратификация районов застройки определяет и качество жизни, и качество образовательных систем: социально благополучный район — благополучные родители — большие налоги — благополучная школа. (Заметим, что и здесь бывают определенные сложности: в наиболее благополучных районах часть родителей предпочитает обучать детей в частных школах, и в связи с этим сравнительный социальный статус и сравнительное благополучие школ могут быть нарушены.)

Западные специалисты в качестве единицы измерения охотно используют образовательное учреждение, считая, что внутри одного микрорайона социальные различия достаточно гомогенизированы. Именно поэтому они постоянно отмечают некорректность сравнения образовательных учреждений, расположенных в разных районах.

Еще один прием, который достаточно широко используется в западных странах, — это использование индикаторов, отражающих содержательные особенности образовательных учреждений, связанных с контингентом, финансированием, отбором и т.д. Таким индикатором в ряде стран является количество бесплатных завтраков, которые выделяются ученикам той или иной школы. Это, как правило, хороший индикатор, поскольку завтраки назначаются социальными службами, независимыми от системы образования, потому что бесплатные завтраки назначаются с учетом дохода и статуса семьи. Таким образом, формируются группы образовательных учреждений, внутри которых сравнения можно считать корректными.

У нас стратификация проходит иначе — не по районам застройки, а по отдельным образовательным учреждениям и классам, а количество бесплатных обедов не является индикатором, характеризующим статус образовательного учреждения. Здесь в качестве фактора, смещающего результаты измерения, могут оказаться как территориальные, так и социально-демографические особенности отдельных классов. Причем это смещение может быть вызвано не только официально сложившейся системой отбора в образовательные учреждения, но и скрытым отбором. Как мы уже говорили, сравнивать престижную гимназию и обычную школу в окраинном микрорайоне некорректно, однако вопрос о том, как сделать такое сравнение корректным, остается открытым.

Если говорить о регионах, то ситуация является еще более неопределенной, нежели в случае с образовательными учреждениями. Тот факт, что регионы Российской Федерации являются

разными, признается практически всеми. Однако вопрос, в чем заключается эта разница и насколько она велика, тоже остается во многом открытым. Используемые достаточно объективные показатели о донорстве или дотационности или показатель уровня бюджетной обеспеченности являются скорее экономическими и в слабой степени отражают специфику образования. Да и они используются в основном как дихотомические, поскольку характеризуют сам факт дотационности и при этом никак не характеризуют уровень дотационности. Можно предположить, что за счет хорошо развитой системы трансфертов собственные доходы регионов оказывают не столь большое влияние на развитие системы образования, но опять же это только предположения. Очевидно, что данный вопрос нуждается в серьезном научном изучении.

Как нам представляется, существуют достаточные основания зафиксировать недостаточность наших знаний о разнице в условиях деятельности региональных образовательных систем для построения объективной системы рейтингов. Достаточно ясно и то, что необходимо проведение в этом направлении серьезных научных исследований, с тем чтобы выяснить, какие факторы социально-экономического развития оказывают существенное влияние на развитие системы образования, каковы механизмы этого воздействия, как можно учесть их влияние при построении системы рейтингов региональных систем образования.

Здесь можно предложить и другой путь — поиск и обоснование системы показателей, не зависящих от условий, в которых развивается та или иная региональная система образования. Речь в данном случае должна идти именно об обосновании и доказательствах, понятных управленцам системы образования.

Решения по результатам рейтингования

Напомним, что изначально наши рассуждения исходили из анализа особенностей объектов, для которых практически используется рейтинговая система оценивания. Второй аспект из рассмотрения реально действующих различных систем рейтингования заключается в том, чтобы посмотреть, какие решения принимаются на основе рейтингов.

Оказывается, спектр возможных решений очень ограничен. Первое и, по нашему мнению, основное — решение о выборе лучшего объекта (или поддержка такого решения) — ученика, школы, вуза, инвестиционной или кредитной кампании и т.д. Второе, это скорее латентная цель, — привлечение внимания общественности, PR-акции, собственное позиционирование объектов рейтингования и т.д. Мы говорим даже в большей степени о поддержке решения, ведь для реального выбора необходим учет еще целого ряда дополнительных факторов: доступности, рисков и т.п.

В работе, которая послужила отправной точкой для нашего исследования, есть не совсем внятная фраза о том, что «анализ ситуации в региональном разрезе станет отправной точкой для принятия адекватных управленческих решений в каждом субъекте Рос-



сийской Федерации». Указывается только одно реальное решение, которое было проведено на основе рейтингования — это отбор школ для получения грантов, да и то не в самой работе, а во вступительной статье. Таким образом, авторами предложена инновационная идея принятия управленческих решений на основании рейтинга, не реализованная ни в одной из областей применения рейтингования. Эта идея нуждается в соответствующем обеспечении. Было бы полезно, если бы авторы указали даже не то, какие решения могут быть приняты на основании рассчитанных ими рейтингов (представляется, что это слишком сложная задача), но хотя бы группы возможных решений или механизмы или направления анализа рейтингов для использования в практике управленческой деятельности.

Признанным методическим подходом для изучения и измерения какого-то сложного феномена являются: анализ, декомпозиция понятий на составляющие, построение модели, подбор и описание показателей для измерения составляющих данного понятия, подбор или создание методик, их методическое оснащение и т.д. Мы намеренно упростили процедуру — декомпозиция может быть проведена несколько раз, модель критически осмыслена и соответствующим образом оценена и описана, при измерениях могут быть подобраны адекватные индикаторы и т.д.

При этом каждый из показателей должен быть достаточно точно и надежно измеряем, а главное, он должен быть валиден относительно того феномена, часть которого он описывает. Количество показателей, которое необходимо для адекватного описания, зависит от сложности объекта изучения, но в любом случае показатели должны покрывать максимально возможное поле составляющих исходного понятия. В этом случае становится понятно, что, собственно говоря, измеряется, каковы возможные ошибки, огрубления модели и т.д. К сожалению, такой работы авторы предложенной методики рейтингования не провели, в чем заключается их методологический просчет, вызвавший целый ряд ошибок и недочетов.

Посмотрим на примере показателей, которые входят в расчет рейтинга «Доступность получения общего образования». Все сложное понятие «доступность образования» свелась к двум показателям — удельному весу детей, не обучающихся в образовательных учреждениях и индексу потребительских цен на услуги образования. Если смысл первого показателя достаточно очевиден, то валидность второго показателя относительно доступности образования очень сомнительна. Как они связаны? Насколько индекс цен детерминирован системой образования. Как известно уровень цен зависит помимо прочего от платежеспособного спроса, наличия конкурентной среды и т.д. А вот его влияние или возможность характеризовать доступность образования нуждается в серьезном обосновании.

Точность,
надежность
и валидность

Доступность для общего образования представлена вообще в одном показателе — удельном весе учащихся в общей численности детей в возрасте 7—15 лет.

Если говорить о повышении доступности образования в концепции модернизации образования, то это целый ряд направлений деятельности. Это и использование новых технологий, и реструктуризация сетей образовательных учреждений, и новое оборудование для образовательных учреждений, и повышение квалификации, и выравнивания стартовых возможностей детей из разных социальных групп и слоев населения, и введение профильного обучения. В итоге ни одно из мероприятий, на которые были направлены усилия как федеральных органов исполнительной власти, так и региональных органов управления образованием, не нашли отражения в представленном рейтинге. Поле понятия доступности образования не раскрыто совсем, даже с точки зрения модели, представленной в концепции модернизации образования.

Но ведь понимание доступности, данное через мероприятия, — это примитивное понимание, не отражающее сути проблемы доступности. Соответствующая работа была выполнена под руководством Д.Л. Константиновского, в ней проблема доступности рассматривалась через понимание барьеров как совокупности факторов, затрудняющих получение образования в школе данного конкретного кластера ребенком, не обладающем тем или иным ресурсом. Ключевая характеристика всякого барьера — его высота, т.е. степень выраженности. Высота определяется по вероятности успешного поступления (барьер на «входе») и завершения обучения («внутренний» барьер) ребенком, не обладающим ресурсом для преодоления этого барьера. Например, если в школе кластера n у 100% учащихся доход на одного человека в семье существенно превышает средние показатели, то шансы ребенка из «небогатой» семьи получить образование в этой школе стремятся к нулю. Авторами была построена модель барьеров, дано их описание:

- социокультурный барьер: социальный и культурный капитал семьи;
- территориальный барьер: перемещения и ресурсы мобильности;
- экономический барьер: явное и латентное неравенство;
- институциональный барьер: структурные ограничения доступности;
- специфические барьеры: информация и мотивация.

Наверное, этот подход может служить примером адекватного подхода к оценке деятельности как федеральной, так региональных образовательных систем в сфере повышения доступности образования, по крайней мере на уровне исходной модели.

Коль речь зашла об этой работе под руководством Д.Л. Константиновского, нужно отметить и еще один важный вывод, к которому пришли ее авторы, описывая понятие качественного образования, — это использование кластерной модели образова-



тельных учреждений при описании региональных образовательных систем.

Авторы этой работы предложили выделить несколько наиболее распространенных (типичных) разновидностей учебных заведений, обладающих определенными сочетаниями признаков (качеств), выраженных в высокой степени. Это будут учебные заведения, дающие качественное образование в том или ином аспекте, т.е. относительно того или иного признака (качества) или сочетаний качеств. Такие группы учебных заведений они и предложили назвать *кластерами*. Соответственно для анализа различных «типов качества» общего образования была использована *кластерная модель*. К пониманию и использованию кластерной модели пришли и разработчики методики оценки уровня информатизации образовательных учреждений. Нам кажется, что для случая рейтингования образовательных систем это наиболее адекватная и перспективная модель, которую нужно попытаться использовать.

Сама ситуация неадекватного оценивания совсем не так безобидна, как это может показаться. Ведь эффект обратной связи от оценочной деятельности хорошо известен, равно как и смещение оценок при нагрузке показателя значимыми факторами, — то, что западные специалисты называют «высокими ставками».

Хорошо знакомы примеры смещения оценочных индикаторов, такие, как количество регистрируемых преступлений в органах внутренних дел, количество решенных вопросов в системе социальных служб. Известен также пример со средней оценкой аттестата, которая в 1970—1980-х гг. принималась как одна из оценок при поступлении в вуз. Смещение достигло такой величины, что от самой оценки пришлось отказаться. Аналогичный процесс наблюдается в настоящее время с медалистами общеобразовательных учреждений. Такая же участь угрожает и другим достаточно информативным индикаторам и показателям, если они известны испытуемым и используются долгое время.

Несомненно, что для региональных руководителей опубликование рейтинга весьма значимое событие. В том случае если публикации таких рейтингов станут регулярными, то и ответная реакция системы не заставит себя ждать.

Еще одной существенной проблемой использования рейтингов как системы оценивания является огрубление шкалы.

После того как проведено измерение и получены числовые значения (естественно, что данные могут быть представлены в абсолютных величинах или приведены к относительному значению) с точки зрения процедуры измерения, несущественно, каким образом будут представлены результаты. Большинство измерений предусматривает дальнейшие манипуляции с данными — использование различных шкал, перевод данных из одних единиц измерения в другие и т.д. Цель этих манипуляций — более удобное или более наглядное представление данных. В подавляющем большинстве случаев такие преобразования носят линейный характер,



т.е. если у одного объекта какое-то свойство выражено сильнее, чем у другого, то в любых линейных преобразованиях это соотношение сохраняется. Например, для ЕГЭ данные об абсолютном количестве выполненных заданий в 10- или 100-балльной шкале будут, по сути, одинаковыми. Если первый ученик получил больше баллов, чем второй, а тот в свою очередь — больше, чем третий, то такое же соотношение останется во всех реально используемых шкалах. Может меняться разница между их оценками: по 100-балльной шкале она может составить 20 единиц, по 10-балльной — 2 единицы, а по 5-балльной вообще нивелироваться до одинаковой оценки. Здесь мы столкнулись с проблемой, которую условно можно назвать проблемой огрубления шкалы. Точно такая же проблема существует и при составлении рейтинга. Разрыв между исходными данными нивелируется (или расширяется) до одного — шага шкалы рейтинга.

Данные и данные

Посмотрим, какие требования должны предъявляться к исходным данным, которые могут входить в рейтинг. Наверное, таких требований может быть не одно. Помимо *объективности*, необходимость которой достаточно очевидна, нужно еще несколько.

Каждый показатель, который входит в состав интегрального, должен обладать собственной *содержательной значимостью*, т.е. он должен отражать оценку какой-то части того, что собственно оценивается, и каждый показатель должен отражать свою часть. С точки зрения измерения использование в составе обобщенного (интегрального) показателя двух исходных показателей с высокой степенью корреляции равнозначно удвоению одного показателя, присвоению ему весового коэффициента 2.

Приведем пример из рассматриваемой нами системы построения рейтинга регионов. Коэффициент корреляции показателей 2.1.10 «Удельный вес населения в возрасте 25—34 лет не ниже среднего специального» и 2.1.11 «Удельный вес занятых с образованием не ниже среднего специального» практически равен единице, то же можно сказать о показателях 6.2.1 «Численность учащихся учреждений НПО» и 6.2.5 «Охват молодежи программами НПО» и о ряде других показателей, т.е. в данном случае произошло удвоение одной и той же содержательной единицы.

Есть среди показателей и пары, которые находятся в обратной пропорции: коэффициент корреляции для них приближается к -1 , например, для показателей 3.4.7 «Удельный вес ДОУ, требующих капремонта» и 3.4.8 «Удельный вес ДОУ имеющих все виды благоустройства (без учета мегаполисов)». В этом случае ситуация еще более странная, поскольку данные, по сути, взаимно уничтожаются.

Второе требование к показателям — это их *приемлемая дифференцирующая способность*. Они должны разделять регионы, выходя за рамки ошибки измерения. Создается впечатление, что авторы не очень использовали свои знания и понятия об ошибках измерения, доверительном интервале и т.д.



Если брать международный опыт, то решение проблемы важных показателей, не обладающих достаточной дифференцирующей способностью, — объединение регионов в группы, как это сделано, например, в презентации результатов международных сравнительных исследований, где страны отнесены к нескольким группам, которые лежат в определенном интервале, значительно отличающемся от других. С этой точки зрения очень сомнительно выглядит показатель 2.1.13 «Удельный вес детей и подростков, не обучающихся», вся разница между регионами заключается даже не в десятых, а в сотых долях процента, а для некоторых значений показателя, как мы увидим позже, и в тысячных долях. Несомненно, что включения такого рода данных в состав рейтингов должно иметь серьезное статистическое обоснование.

Третье требование к исходным показателям — это их *надежность* в понимании помехоустойчивости. В этом отношении просто необходимо знание того, насколько показатели носят случайный характер, насколько они дают стабильные значения от года к году, насколько они подвержены флуктуационным изменениям и т.д. Сомнение в надежности есть, например, относительно показателей 4.5.7 и 4.5.8 удельного веса детей, выбывших из учреждений общего образования.

Весьма примечательным является рейтинг «Здоровье детей и подростков» для общего образования. В него входят три показателя.

Первый — «Число заболеваний с впервые установленным диагнозом у подростков в возрасте 15—17 лет». Если предположить, что ВСЕ причины заболеваемости детей сосредоточены в образовательных учреждениях, то, наверное, этот показатель может быть использован. На самом деле это совсем не так. Социальная ситуация в регионе, экологическая обстановка, наследственность, деятельность системы здравоохранения, эпидемиологическая обстановка и целый ряд других факторов являются определяющими для числа заболеваний. Естественно, какой-то вклад — не очень пока понятно, позитивный или негативный, — вносит и образование, но он настолько плотно экранируется другими, что выделить его при всем желании невозможно. Просматривая сами данные и отмечая их многократный разброс, можно предположить, что, скорее всего, этот показатель отражает уровень объективности медицинских служб по регистрации обращений подростков, нежели несет в себе хоть какое-то содержательное наполнение. Таким образом, делать ответственным за заболевания детей систему образования, по крайней мере, несправедливо.

Второй показатель — «Удельный вес государственных и муниципальных ОУ, имеющих физкультурный зал». Хорошо известно, что влияние на состояние здоровья оказывает двигательная активность, а она зависит от того, как организована деятельность учащихся, и в первую очередь внеклассная и внешкольная. Скорее всего, его можно было использовать как характеристику инфра-



структуры образования. В данном контексте совершенно не понятно, как факт наличия спортивного зала влияет на состояние здоровья учеников.

Третий показатель — «Удельный вес учащихся, пользующихся горячим питанием». Скорее всего, какое-то влияние хорошая организация горячего питания может оказать на состояние здоровья учащихся. Однако о чем, собственно, идет речь? Об организации питания на бюджетные средства, поэтому остающимся в поле здравого смысла этот показатель может говорить об уровне социальной защищенности учащихся. Статистический факт питания или непитания за средства бюджета мало о чем говорит в контексте здоровья учеников.

Рейтинг
и здоровый
смысл:
проблема
каузального
вывода

Как отдельную проблему, характерную и для информационного обеспечения управления образованием, и для исследования в социальной области, необходимо рассмотреть то, что можно назвать «отсутствие валидизации здравым смыслом» (термин введен Кэмпбеллом). Суть этой проблемы можно пояснить на примере естественнонаучного эксперимента. Если экспериментатор, проводя измерения, например, силы тока, неожиданно получает отрицательные показания на приборе вместо предполагавшихся положительных, то он, в первую очередь, начинает искать то место в схеме, где неверно соединены провода, и, исправив схему, продолжает измерение. В социальных науках исследователь, получив «ошеломляющие» результаты, в первую очередь придумывает им «разумное» объяснение и сразу же начинает оповещать всех о совершенном им открытии.

Приведем несколько примеров из анализируемой нами работы.

Авторы используют в расчетах коэффициент охвата начальным и основным образованием 83,36% для России в целом. Это означает, что один ребенок из каждых шести не получает обязательно в России 9-летнего образования, а в Новгородской области, если верить расчетам авторов, не учится в 9-летней школе один ребенок из четырех — вот они ошеломляющие результаты, опубликованные в официальном издании. На самом деле авторы просто не учли переход с 3- на 4-летнее обучение в начальной школе.

Охват молодежи программами ВПО для Москвы — почти 90%, а в Московской области — 14, в Ленинградской области — 6, в Санкт-Петербурге — 63%. Но если в Москве 89% населения соответствующего возраста получают высшее образование, 17 — среднее и 16 — начальное профессиональное, то в сумме получается 122%! И в этом случае ситуация разрешается с помощью простого здравого смысла: конечно же, в Москве и Санкт-Петербурге учатся дети практически из всех регионов России и относить их к населению только мегаполисов мягко говоря некорректно.

Судя по опубликованным данным, в вузах Ханты-Мансийского автономного округа и ряда других регионов не оказалось ни одно-



го персонального компьютера, но даже это не насторожило авторов.

Все эти «открытия» не выдерживают критики ни с точки зрения статистики, ни с точки зрения элементарного здравого смысла.

Итак, речь идет том, что полученные результаты, особенно выраженные в цифрах, должны быть проанализированы с точки зрения здравого смысла и, что очень важно, в анализе должно присутствовать максимально возможное количество альтернативных гипотез объяснения полученных результатов. Нужно отметить, что проблема причинно-следственных связей, верного каузального вывода, свойственна в той или иной мере всей гамме измерений и анализа в социальной области, и данная работа просто лишнее тому подтверждение.

Приведем пример из другой работы. В достаточно хорошей и грамотной работе, посвященной анализу состояния образовательной системы Саратовской области, можно прочесть, что при переходе от 1-й ко 2-й ступени обучения качество образования падает на 30%. Причинами называются отсутствие преемственности в обучении, неадекватность методов обучения в начальной школе. Однако явные альтернативные гипотезы о том, что в школе 1-й и 2-й ступени различаются требования и подходы к оцениванию, что при этом переходе дети испытывают дискомфорт и стресс, почему-то во внимание приняты не были.

Дело в том, что одни и те же эффекты могут объясняться различными причинами. Например, низкий охват дошкольным образованием может объясняться как нехваткой мест в ДДО, так и хорошей социальной защитой, которая заключается в том, что пособия на воспитание детей являются достаточными для того, чтобы родители оставались с детьми дома до достижения ими школьного возраста. Однако в рассматриваемой работе низкий охват дошкольным образованием рассматривается как исключительно негативная характеристика системы образования.

Можно предположить, что и сами исходные данные не совсем надежны. Если проанализировать данные по отсеву из образовательных учреждений, в разных возрастных группах, то становится понятно, что они зависят от уровня урбанизации той или иной территории, т.е. чем больше городского населения, тем меньше цифра отсева. В такой ситуации вполне резонно предположить, что в сельской местности отсев из образовательных учреждений значительно выше. Однако это не совсем согласуется со здравым смыслом и реальным положением дел в сельских школах. Дело здесь, скорее всего, в том, что в сельской местности сохранилась достаточно эффективная системы всеобщего обучения, которая включает строгий учет учащихся, проведение подворовых обходов и т.д. В городах, особенно крупных, эта система давно развалилась, а какая-то адекватная до сих пор не появилась. Может оказаться так, что там, где работа по учету детей поставлена лучше, будет негативная оценка в рейтинге.



Незначитель-
ные ошибки —
маленькие
ошибки —
ошибки —
большие
ошибки

Рассматривая более подробно приведенные данные, мы можем заметить, что в работе есть элементарные ошибки в расчетах, которые содержатся в таблицах. Например, в таблице 1 «Уровень образования населения» данные рейтинга Смоленской и Волгоградской областей одинаковы — 3,025, но результаты расчета процентов к первому месту почему-то разные — 87,33 для Смоленской области и 87,35 для Волгоградской. То же и для Красноярского края и Кемеровской области, где разрыв еще больше и составляет три пункта. Такой разброс теоретически может быть объяснен округлением, однако тогда нужно говорить о том, что уровень дискриминации в данном случае катастрофически низок, учитываются уже четвертый и пятый знаки после запятой, что не выдерживает никакой критики с точки зрения точности измерения, впрочем, как и с точки зрения представления полученных результатов.

Из математической статистики хорошо известно, что в случае порядковых переменных недопустимы математические преобразования, в том числе и по вычислению среднего. Мало того, проведенное Тейлором исследование по применению более сложных преобразований (вычисление средней дисперсии, стандартного квадратичного отклонения и др.) доказало, что для порядковых переменных эти вычисления могут быть корректными, но при условии, что порядковые переменные не носят характер дихотомии или трихотомии. Таким образом, дихотомические переменные использованы в рейтинге некорректно, как с точки зрения вычислений, так и с точки зрения включения их в рейтинг с другими видами измерений.

Есть в рассматриваемой нами работе и просто непонятные места. Например, в методике расчета рейтингов авторы предлагают различные варианты, при этом они предлагают использовать или среднероссийское значение показателя, или среднее арифметическое значение. Авторы выбрали среднероссийское в качестве основания для формирования, оставив читателя в недоумении относительно того, чем оно отличается от среднеарифметического.

Достаточно показательна ошибка с использованием данных ЕГЭ. Рособрнадзор неоднократно предупреждал о невозможности использования данных ЕГЭ для сравнительных оценочных суждений. Логика в этом совершенно простая. Часть регионов делают экзамен по тому или иному предмету обязательным для выпускников, а в другой части регионов экзамены по тому же предмету сдаются по выбору, да еще и количество выбравших этот экзамен оказывается незначительным. Вполне естественно, что выбор в пользу сдачи экзамена по предметам делают те ученики, которые лучше к ним подготовлены. Поэтому ситуацию, когда результаты сотни лучших учеников одного региона сравниваются с результатами всех выпускников другого, нормальной назвать нельзя даже с точки зрения элементарного здравого смысла.

Достаточно показательной кажется оценка по рейтингу «Региональная образовательная политика». Для получения места в высшей части рейтинга нужно:



а) чтобы дети сдавали экзамен по математике и русскому языку по выбору, при этом чем меньше детей сдавали ЕГЭ, тем выше место в рейтинге;

б) зафиксировать факт создания региональной службы оценки качества, вернее, сказать, что она создана;

в) сделать то же в отношении общественного совета по проблемам образования при руководителе субъекта федерации.

И это все! Дальше каждый может самостоятельно судить о том, какое отношение этот рейтинг имеет к образовательной политике и ее реализации.

Можно сколько угодно говорить о методологии и объективности, о научной обоснованности и грандиозном опыте, но все это ни на шаг не приблизит нас к решению реальных задач оценки образовательных систем.

Краткое рассмотрение проблематики использования рейтингов в различных областях управления говорит о том, что это специфический инструмент, который адекватно может быть использован в весьма узком сегменте управления, там, где нужно провести отбор лучшего объекта, при этом не просто лучшего, а абсолютно лучшего, не учитывая условий деятельности объекта рейтингования.

Рейтингование как процедура и рейтинг как оценка в представленном виде не могут быть признаны адекватными для такой сложной социальной системы, как региональная система образования, и особенно с точки зрения оценки деятельности региональной системы управления.

Работы по поиску адекватной системы оценки деятельности и состояния региональных систем образования необходимо не только продолжить, но и значительно активизировать. При этом не следует ограничиваться только рейтинговой системой оценивания в том понимании, в котором она представлена в анализируемой нами работе.

Очевидно, что есть необходимость продолжить научные разработки в направлении сравнительных исследований региональных образовательных систем. При этом представляется наиболее перспективным использование кластерной модели сети образовательных систем регионального уровня.

Важно вернуться к целесообразности использования рейтингов для региональных систем образования после решения вопроса об учете условий их деятельности.

Казалось бы, банально говорить о том, что при предложениях каких-то оценочных схем, методик технологий необходимо использовать элементарный инструмент математической статистики, обоснование показателей, какое-то моделирование и т.д. Однако исследованная нами работа делает такой вывод необходимым.

Необходима проработка вопроса о тех решениях, которые могут быть приняты на основе рейтингов или других оценок.

Заключение



С организационной точки зрения представляется очень важным решение вопроса о федеральном методическом и координационном центре, который занялся бы решением обозначенного комплекса проблем. Давно назрела необходимость создания реально действующего федерального центра по вопросам мониторинга, статистики и социологии системы образования, оценки региональных образовательных систем, который вел бы систематическую и планомерную работу, неся соответствующую ответственность за ее результаты.

Наиболее адекватной может быть признана модель мониторинга образования, построенная на основе распределенной системы региональных центров с сильной, работоспособной и эффективной федеральной структурой во главе.