

Развитие стимулирующей среды для оценки достижений учащихся на базе школы в штате Квинсленд, Австралия

Рег Аллен



**Развитие стимулирующей
среды для оценки
достижений учащихся на
базе школы в штате
Квинсленд, Австралия**

Рег Аллен

© 2012 Международный банк реконструкции и развития / Всемирный банк
1818 H Street NW
Washington DC 20433
Телефон: 202-473-1000
Интернет: www.worldbank.org

1 2 3 4 15 14 13 12

Настоящий документ является плодом совместных усилий сотрудников Всемирного банка и коллег из иных организаций. Сведения, толкования и выводы, изложенные в данной работе, могут не совпадать с взглядами Всемирного банка, его Совета директоров или правительств стран, которые они представляют.

Всемирный банк не гарантирует точности данных, приводимых в данной работе. Границы, цветовые обозначения и иная информация, представленные на содержащихся в данной работе картах, не являются отражением каких-либо суждений Всемирного банка в отношении правового статуса каких-либо территорий, равно как и подтверждением или принятием с его стороны таких границ.

Права и разрешения

Материалы, представленные в данной работе, подпадают под действие законов об авторских правах. Поскольку Всемирный банк поощряет распространение своих знаний, он разрешает частично или полностью воспроизводить данный документ для некоммерческих целей при условии обязательного указания полной информации о правообладателе.

Любые запросы, касающиеся прав и лицензий, в том числе вторичных издательских прав, просим направлять издателю по адресу: The World Bank, 1818 H Street NW, Washington, DC 20433, USA, факсу: 202-522-2422, или по электронной почте: pubrights@worldbank.org.

Оформление обложки: Патриция Хорд. «Graphik Design», г. Александрия, шт. Вирджиния

Перевод: Александр Резанов

Содержание

Сокращения	v
Информация о данной серии публикаций	vii
Об авторе.....	ix
Слова благодарности от автора	xi
Обзор	xiii
Введение	1
Главные особенности оценки на базе школы в штате Квинсленд.....	5
Стимулирующая среда.....	18
Факторы перемен.....	20
Извлеченные уроки	23
Библиография	29

Сокращения

ВВП	Валовый внутренний продукт
МСКО	Международная стандартная классификация образования
NAPLAN	Национальная программа оценки — грамотность и умение считать
ОП	Общая позиция
PIRLS	Международное исследование по оценке качества чтения и понимания текста
PISA	Международное исследование по оценке достижений учащихся
QCE	Сертификат об образовании шт. Квинсленд
QCS	Ключевые навыки (по классификации шт. Квинсленд)
QSA	Орган управления системой образования шт. Квинсленд
READ TF	Трастовый фонд Российской программы содействия образованию в целях развития
SAI	Показатели (индикаторы) достижений в изучении предмета
SABER	Системный подход к улучшению результатов образования
TIMSS	Международное исследование по оценке качества математического и естественнонаучного образования

Информация о данной серии публикаций

Построение эффективных образовательных систем, обеспечивающих качественное обучение, имеет исключительно важное значение для развития и экономического роста. В последние несколько лет развивающимся странам удалось открыть новые школы и принять на учебу миллионы новых учащихся. На этом фоне образовательное сообщество начало активно внедрять идею использования измеримых показателей для оценки результативности обучения всех школьников. Однако получение знаний зависит не только от вкладываемых в школьную систему ресурсов, но и от качества методик и обеспечивающих их использование организаций, а также от того, насколько хорошо эти совокупности методов и приемов внедряются на практике.

В 2011 году Группа организаций Всемирного банка начала реализацию «Стратегии в сфере образования на период до 2020 года: обучение для всех», в которой содержится план достижения цели «Обучение для всех» в развивающемся мире в течение предстоящего десятилетия. Для поддержания процесса реализации стратегии Всемирный банк начал многогодичную программу, в рамках которой странам оказывается содействие в систематическом анализе и повышении результативности работы их систем образования. В рамках данной научно обоснованной инициативы, названной сокращенно SABER (от англ. Systems Approach for Better Education Results — системный подход к улучшению результатов образования), создается набор диагностических инструментов для исследования образовательных систем и направлений политики в сфере образования на основе глобальных стандартов и передовых практических методов, а также проведения сравнительного анализа их с методиками и практическими подходами стран в различных частях земного шара. За счет привлечения этого глобального ресурса научных знаний программа SABER ликвидирует дефицит данных о наиболее важных факторах, влияющих на повышение качества образования и достижение лучших результатов, и научного обоснования этих факторов.

SABER — оценка достижений учащихся — одна из составляющих программы, содержит разработанные инструменты

для анализа методик и систем оценки учащихся во всем мире и установки соответствующих контрольных показателей с целью укрепления оценочных систем, способствующих повышению качества образования и обеспечивающих достижение цели «обучение для всех». Для повышения качества изучения научных знаний в данной сфере группа, занимающаяся оценкой достижений учащихся в рамках SABER, привлекла ведущих ученых, специалистов по оценке и практиков из развивающихся и индустриально развитых стран для обсуждения проблемных вопросов оценки, важных для повышения качества образования и учебных результатов. Публикуемые в данной серии документы и конкретные примеры оценки достижений учащихся явились результатом такой дискуссии и исследовательской работы. До публикации все документы были подвергнуты самому тщательному анализу, в ходе которого были учтены замечания сотрудников Всемирного банка, ученых, практиков в области развития и экспертов по страновым оценкам.

Подготовка всех документов по оценке достижений учащихся по системе «SABER» в данной серии публикаций стала возможна благодаря поддержке со стороны трастового фонда Российской программы содействия образованию в целях развития (READ TF). READ TF является совместным специализированным фондом Российской Федерации и Всемирного банка, предназначенным для оказания содействия в улучшении результатов обучения в странах с низкими уровнями доходов за счет создания надежных систем оценки достижений учащихся.

Серия рабочих документов в рамках программы SABER была создана под общим руководством директора по вопросам образования Элизабет Кинг и руководителя по направлению «Образование» Сети по развитию человеческого потенциала Всемирного банка Хэрри Энтони Патриноса. Документы по оценке достижений учащихся в данной серии были подготовлены под техническим руководством Маргарит Кларк, старшего специалиста по образованию и координатора группы по оценкам достижений учащихся в рамках SABER Сети по развитию человеческого потенциала Всемирного банка. В документах данной серии представлены независимые взгляды авторов.

Об авторе

Д-р Рег Аллен является руководителем Квалификационного управления Тасмании (Австралия). Он имеет 30-летний опыт работы учителем, старшим преподавателем по предмету, членом совета, представителем коллектива учителей, консультантом, заместителем директора и директором в различных образовательных учреждениях и органах Австралии, Англии и США. Он работал в качестве эксперта-консультанта по вопросам образования в Австралии и США, а также ведущим сотрудником исследовательско-консультационной группы по программе New Basics штата Квинсленд с 2001 по 2004 г. В 2012 году Национальный центр по совершенствованию оценочной деятельности (шт. Нью-Гемпшир, США) пригласил его на один из крупнейших коллоквиумов в г. Боулдере (шт. Колорадо), чтобы он поделился международным опытом в области проработки ключевых технических вопросов в использовании оценки результативности при приеме в университеты. Ранее он занимал пост заместителя директора Совета по учебному процессу на старшей ступени среднего образования шт. Квинсленд. Работая на этой должности, он играл ключевую роль в составлении и реализации Комплекса образовательных параметров учащихся шт. Квинсленд для учащихся старшей ступени средней школы, а также в интеграции профессионально-технического образования в систему образования на этапе после получения обязательного всеобщего образования. Он является автором более 40 публикаций, посвященных различным аспектам образования, в том числе национальным оценочным исследованиям и разработке политики Органов Океании по оценке учебных программ и сертификации.

Слова благодарности от автора

Автор благодарит: Питера Лакстона, и.о. директора Управления образования шт. Квинсленд; Д-ра Габриэль Мэттерс, бывшего заместителя директора Совета по учебному процессу на старшей ступени среднего образования шт. Квинсленд; Джона А. Питмана, бывшего директора Совета по учебному процессу на старшей ступени среднего образования шт. Квинсленд.

Обзор

Австралийский штат Квинсленд демонстрирует уникальный пример того, как программа оценки может воплотить в себе принципы оценки обучения, оценки как обучения и оценки для обучения. Программа оценки штата Квинсленд может служить примером для других стран, которые намерены найти более эффективные пути измерения учебных достижений учащихся средней школы. Этот вопрос приобретает особую актуальность на фоне глобальной дискуссии о путях создания систем оценивания, которые были бы более достоверны, требовательны и не ограничены форматом письменного теста, основанного на вопросах с несколькими вариантами ответов. Актуальность этого вопроса также подтверждается идущей в настоящее время дискуссией о том, как повысить качество педагогических кадров и обеспечить постоянное повышение квалификации учителей.

В начале 1970-х годов в штате Квинсленд произошла замена программы внешних экзаменов на систему оценки с координацией извне на базе школ. Источником информации для оценки на базе школы по каждому учебному предмету являются достижения учащихся на верхней ступени среднего образования (11-й и 12-й классы). Для того чтобы убедиться в том, что учителя применяют единые нормы для вынесения суждений относительно достижений учащихся и выставления оценок, внешние группы специалистов анализируют выборку работ учащихся.

За проведение программы оценки на уровне школ отвечают одновременно орган управления штата и школы. Первый является независимым правительственным органом, отвечающим за полноту и достоверность результатов, которые он заверяет. Этот орган отвечает за то, чтобы внешняя координационная комиссия работала должным образом, и присваивает учащимся штата рейтинг, который используется в процессе отбора потенциальных студентов высшей школы. Школы отвечают за сбор доказательств работы учащихся, оценивая эту работу на основании стандартов, и передают выборки таких работ во внешнюю координационную комиссию. Учителя выполняют эти связанные с оценкой функции как часть своих основных должностных обязанностей.

Каковы же факторы, позволившие укорениться этой уникальной оценочной программе? В 1960-е годы в Квинсленде, как и везде, среди сотрудников системы образования нарастали сомнения в ценности для системы образования централизованно проводимых экзаменов, особенно в условиях роста числа учащихся, продолжающих обучение в 11-м и 12-м классах. Социально-политические условия того времени позволили этим сомнениям педагогов сформироваться в радикально иной подход к оценкам с высокими ставками. Мысль о том, что именно учитель является тем человеком, который наилучшим образом способен оценить достижения своего ученика, утвердилась среди основных заинтересованных сторон. Низкие прямые затраты, обусловленные тем, что учителя осуществляют оценочную деятельность как одну из своих прямых должностных обязанностей, также способствовали институционализации программы.

Другие страны, намеренные реформировать свои оценочные системы путем введения в них элементов оценки на уровне школ, могут воспользоваться опытом штата Квинсленд. Подход, использованный в Квинсленде, требует согласия всех заинтересованных сторон со следующими убеждениями и основополагающими принципами: самыми квалифицированными экспертами, способными наилучшим образом оценить достижения учащихся, являются их учителя; оценочную деятельность никогда не следует отделять от учебной программы и преподавания, и структура оценки и последующая достоверность ее результатов (и воздействие на результаты обучения) должны преобладать над узкой сосредоточенностью на психометрических сомнениях в отношении надежности, равенства и ценности стандартизированного тестирования.

Развитие стимулирующей среды для оценки достижений учащихся на базе школы в штате Квинсленд, Австралия

Рез Аллен

Введение

В ходе всемирных дебатов о путях повышения качества образования традиционные системы оценки часто подвергаются критике на том основании, что они обычно не вносят вклада в улучшение преподавания и приобретения знаний, или что в рамках этих систем не происходит измерения самых насущных навыков (таких как решение задач, критическое мышление, творческое мышление, коммуникативность и способность работать в коллективе), необходимых для достижения учащимися успеха в жизни. В австралийском штате Квинсленд была введена в действие программа оценки достижений учащихся, которая в значительной степени устраняет эти недостатки. Данная программа отражает прогрессивные идеи в области образования и оценки приобретения знаний с целью повышения эффективности обучения. Страны, желающие реформировать свои системы оценки, могут извлечь большую пользу из опыта этого штата, который уже прошел путь к внедрению этой инновационной системы оценки.

Рассмотрение данного конкретного примера преследует три цели: (1) охарактеризовать основные особенности координируемой извне программы штата Квинсленд по оценке на базе школы на старшей ступени среднего образования¹; (2) проанализировать

¹ В Квинсленде имеются другие программы оценки, которые не описываются в данной работе. Например, Квинсленд регулярно участвует в общенациональном

стимулирующую среду и движущие силы перемен, которые позволили внедрить эту программу и поддерживать ее² и (3) поделиться извлеченными уроками с другими странами, намеревающимися провести реформу своих систем оценки.

Для понимания основных особенностей программы штата Квинсленд важно понимать обстановку, в которой происходило ее внедрение. Квинсленд является одним из шести штатов, составляющих Австралийскую Федерацию³. Это богатый штат богатой страны: экономика штата находится на третьем месте в стране, а ВВП на душу населения в штате составляет 60 тыс. долларов США. Международные оценки достижений учащихся свидетельствуют о том, что Австралия находится среди самых успешных в этой области стран, а штат Квинсленд имеет показатели эффективности, подобные таким показателям других пяти штатов Австралии (Австралийское управление по программе школьного обучения, оценкам и отчетности, 2011 г.). По состоянию на 2011 год в 1237 государственных школах (начальных и средних) штата Квинсленд, в которых трудятся более 36 тыс. учителей, обучалось почти полмиллиона учащихся (70 процентов учащихся в штате). Остальные 30 процентов учащихся штата посещали частные (независимые или католические) школы. Около 500 средних школ, из которых в приблизительно 400 имеются учащиеся 11-го и 12-го года обучения, участвуют в координируемой извне программой оценки на базе школы шт. Квинсленд.

Зарождение, развитие и реализацию системы оценки штата Квинсленд следует рассматривать в географическом и демографическом контекстах. Квинсленд — крупный по территории

(NAPLAN) и международных (PIRLS, PISA, TIMSS) широкомасштабных оценочных исследованиях, проводимых на начальной и младшей ступенях среднего образования для мониторинга качества образования в сравнении с его качеством в других штатах и государствах. С обзором оценочных программ можно ознакомиться на интернет-портале Управления образования штата Квинсленд по адресу <http://www.qsa.qld.edu.au/3111.html>.

² Для получения общей информации о стимулирующей среде для оценки достижений учащихся см. Кларк (2012 г.).

³ По Конституции Австралии руководство системой образования является прерогативой властей штатов, однако в течение последних 50 лет правительство Содружества действует все более активно в этой области.

штат (около 1,8 млн. кв. км), находящийся в северо-восточной части Австралии. Хотя большая часть населения штата проживает в его юго-восточной части, штат в целом относительно децентрализован. Столица, в которой расположены органы управления системой образования штата, расположена неподалеку от южной границы. Это означает, что много небольших и несколько крупных школ расположены на удалении 1-2 тысячи километров от «штаб-квартиры». В условиях системы Квинсленда чрезвычайно важно, по организационным и политическим причинам, поддерживать эффективную связь с этими удаленными школами и кластерами школ в центрах провинций, находящихся на большом расстоянии от столицы.

В 1972 году в Квинсленде были отменены внешние экзамены для учащихся 12-го класса школ, а на смену этим экзаменам пришла координируемая извне программа оценки на базе школы («система Квинсленда») для всех школ (государственных и частных), в которых имеются учащиеся, желающие сертифицировать свои результаты учебы по окончании 12-го класса. Оценка используется для получения информации, необходимой для улучшения преподавания и получения знаний, сертификации⁴ достижений и принятия решений о поступлении в высшие учебные заведения. В школах Квинсленда нет внешних экзаменов. Вместо этого решения с высокими ставками в отношении отдельных учащихся, завершающих программы обучения на уровне 3A⁵ по классификации ISCED, принимаются на основании их достижений, продемонстрированных в течение последних двух лет обучения в средней школе (старшая ступень среднего образования — 11-й и 12-й классы). Учителя выносят суждения о работе своих учащихся и выставляют отметки, пользуясь централизованно установленными программными нормативами и методическими указаниями по

⁴ Термин «сертифицировать» используется здесь для описания процесса письменного представления результатов в официальном сертификате — юридически зарегистрированном документе, принимаемом как доказательство изложенного в нем содержания.

⁵ ISCED — международная стандартная классификация в образовании. Программы на уровне ISCED 3A ориентированы на прямой доступ к уровню ISCED 5A (университет).

оценке. Для подтверждения того, что все учителя выставляют оценки учащимся, пользуясь одинаковыми нормативами, образцы работ учащихся рассматриваются внешними координационными⁶ комиссиями. Затем результаты учащихся проходят официальную сертификацию как соответствующие выпускным требованиям средней школы и используются как основа для принятия решений о приеме в высшие учебные заведения.

В течение 11-го и 12-го годов обучения учащиеся штата Квинсленд изучают в школе предметные курсы (предметы с такими названиями, как «физика», «современная история», «японский язык»). Их успехи в изучении данных предметов оцениваются школой по мере прохождения ими предметного курса. По окончании 12-го года они получают официальный сертификат с результатами по данным предметам, выдаваемый центральным органом штата⁷ (законодательно созданный государственный комитет), который подтверждает соответствие их достижений определенному набору нормативов, распространяющихся на всех учащихся, изучающих предмет во всех школах штата.

Хотя за оценку достижений учащихся отвечает школа, центральный орган должен убедиться в том, что он может сертифицировать результаты, будучи уверенным в их

⁶ Координация означает выполнение процедур, необходимых для приведения оценочных решений одной школы в соответствие с такими решениями другой школы на этапе до окончательного вынесения этих решений. В некоторых системах этот процесс принимает форму «статистической координации», для которой используется статистическая методика. В других системах процесс принимает форму «совещаний для выработки консенсуса» или собраний учителей, на которых они согласовывают методику применения стандартов. В некоторых других системах процесс имеет форму «внешнего анализа», проводимого группой экспертов. Термин «социальная координация» иногда используется для того, чтобы отличить две последних формы процесса от статистической координации.

⁷ За последние 40 лет этот центральный орган несколько раз менял свое название. В настоящее время он называется Управление образования штата Квинсленд (QSA). Для целей данного примера очень важно подчеркнуть, что именно это управление является *центральным* органом, независимым от Департамента просвещения штата. Использование таких органов являлось и является обычной практикой в штатах Австралии. Поэтому не возникло никаких трудностей при создании данного органа, предназначенного для управления системой оценки в новом, 2012 году: нужно было лишь изменить функции ранее существовавшего органа.

сопоставимости в границах всего штата — обязанность органа состоит в том, чтобы гарантировать, что, например, одинаковые результаты по физике двух учащихся из двух разных школ, находящихся на расстоянии нескольких тысяч километров одна от другой, фактически соответствуют одному и тому же стандарту.

В последние несколько лет в Квинсленде результаты программы оценки на базе школы также зачитываются при определении соответствия требованиям на получение сертификата, подтверждающего завершение обучения на старшей ступени средней школы. Начиная с 2009 года, Сертификат об образовании штата Квинсленд (QCE) выдается на основании факта прохождения достаточного курса обучения с удовлетворительными результатами. Краеугольным камнем основания для выдачи сертификата являются результаты по учебным предметам, оцененные в ходе проведения координируемой внешним органом оценки на базе школы.

В остальной части данной работы автор дает подробное описание основных особенностей данной программы оценки на уровне средней школы, указывает ключевые факторы, которые способствовали введению и реализации такой программы, а также рассматривает извлеченные из опыта внедрения программы уроки, которые будут полезны для других стран, намеренных реформировать свои системы оценки.

Главные особенности оценки на базе школы в штате Квинсленд

В Квинсленде координируемая внешним органом оценка на базе школы используется для нескольких целей: оказать поддержку процессам преподавания и получения знаний, официально сертифицировать достижения учащихся старшей ступени средней школы в изучении ими предметов школьного цикла и предоставить результаты, которые будут использоваться в процессе отбора студентов в высшие учебные заведения.

Оценка с высокими ставками, которая является информационной основой непрерывного процесса преподавания и

обучения, является одной из главных особенностей программы оценки на базе школы в штате Квинсленд. Учителя регулярно оценивают своих учащихся в ходе обучения по предметным курсам, согласованным со школьной программой штата. Как и во многих других странах, учителя используют оценку на уровне класса для мониторинга процесса преподавания, для предоставления учащимся информации о том, как они усваивают необходимые знания и для выставления оценок учащимся за их работу. Однако в отличие от других стран, в Квинсленде эти решения учителей в отношении достигнутых их учениками стандартов имеют прямое влияние на окончательные, официально сертифицируемые результаты, которые не подвергаются пересмотру на основе комбинирования решения учителя с результатами внешнего экзамена или на основе шкалирования.

Для обеспечения достоверности, целостности и качества официальных сертифицированных табелей, содержащих достижения учащихся, действует жесткая система контроля качества (внешняя координация), которая гарантирует, что все учащиеся, получившие одинаковую оценку⁸ по одному и тому же предмету, фактически выполнили требования сопоставимых⁹ стандартов.

Существует четкое разграничение полномочий в рамках оценки на базе школы. Школы отвечают за обеспечение того, что достижения учащихся оцениваются в соответствии с официально согласованными стандартами и требованиями, предъявляемыми к содержанию и оценке. Учитель¹⁰, преподающий конкретный

⁸ В течение первых 10 лет существования системы учащиеся получали результат, оцениваемый по шкале от 1 (низший балл) до 7 (самый высокий балл). С 1980-х годов результат учащихся описывается при помощи одной из пяти формулировок — от «очень ограниченные достижения» до «очень высокие достижения». Каждая из пяти формулировок имеет определение, в котором четко прописаны стандарты по каждому предмету.

⁹ Садлер (1995 г.) отмечает, что «сопоставимость — это обеспечение эквивалентного качества показателей эффективности всех учащихся, получивших определенную оценку по предмету (в диапазоне показателей эффективности, отвечающем определенной оценочной категории) вне зависимости от того, какой орган проводил оценку, и в каком году проводилась оценка».

¹⁰ Ссылки на «учителя» не следует понимать так, что каждый учитель действует независимо от других учителей, преподающих тот же самый предмет в той же самой школе. Именно школа отвечает за качество решений, принимаемых в отношении

предмет в школе, должен принимать решения в отношении конкретных компонентов оценки как части преподаваемого им предметного курса. Это означает, что учитель отвечает за инструменты оценки, за график проведения оценочных мероприятий, за критерии отметок, за выставление отметок и за то, чтобы показанные результаты повлияли на окончательное решение школы о достижениях учащегося в изучении данного предмета.

В Квинсленде учитель является частью формализованной официальной итоговой оценочной системы и отвечает перед школой за предоставление достоверных фактов, на основании которых школа принимает решения о достижениях учащихся. Это создает проблему: учитель должен быть на стороне учащегося и в то же время защищать интересы системы. Временами эта напряженность находит выражение в гневных заявлениях о том, что внешняя экспертная комиссия «отвергла» выводы школы (читай: учителя), или в полных озабоченности заявлениях о том, что учителя других школ манипулируют системой.

За официально упорядоченный процесс сертификации несут солидарную ответственность школы и Управление образования (QSA), орган на уровне штата. Тогда как школы отвечают за оценку достижений учащегося в изучении предметных курсов, QSA несет ответственность за документальное подтверждение того, что результаты оценки по каждому предмету сопоставимы с такими результатами во всех школах на территории штата. Именно QSA должно гарантировать, что, например, два учащихся с одинаковыми результатами по физике из двух разных школ, находящихся на расстоянии тысяч километров друг от друга, фактически выполнили требования одних и тех же стандартов. Эта задача решается Управлением посредством отлаженной системы внешней координации.

Эта система внешней координации требует наличия партнерских отношений между центральным органом (QSA) и школой. Центральный орган:

оценивания учащихся. Учитель или учителя, преподающие конкретный предмет в данной школе, являются средством, при помощи которого школа выполняет эту свою обязанность.

- создан на основании законодательных актов,
- не зависит от правительства в своих действиях,
- финансируется правительством,
- занимается сертификацией¹¹ достижений учащихся,
- создает школьную программу («учебный план») по каждому предмету, опираясь на которую, школы составляют собственные учебные курсы,
- устанавливает порядок и обеспечивает исполнение процедур, необходимых для гарантирования достаточного уровня сопоставимости результатов по предметам во всем штате,
- предоставляет высшим учебным заведениям (преимущественно, университетам) рейтинговый список учащихся, место в котором определяется, исходя из уровня общих академических достижений¹², выведенного на основании результатов, показанных учащимся по разным предметам,
- формирует, разрабатывает и проводит тест общих навыков (Ключевые навыки по классификации шт. Квинсленд — тест QCS), преследуя, прежде всего, цель сбора информации о группах учащихся (а не отдельных учащихся).

По каждому из предметов с высокими ставками (обычно, это предметы, актуальные для поступления в университет):

- центральный орган создает основу программы по каждому предмету («учебный план»),

¹¹ Термин «сертификация» выражает идею оформления официального документа, который широко принимается в качестве доказательства достижений. Например, университет, требующий от абитуриента наличия определенных минимальных достижений в математике («обязательное предварительное условие»), продемонстрированных ранее, может убедиться в выполнении данного требования, если абитуриент предъявит (приемлемый) официальный сертификат, подтверждающий уровень его достижений по предмету в 12 классе.

¹² Термин «общие академические достижения» используется в данной работе в том смысле, что комбинация (совокупность или средняя сумма) результатов по различным предметам представляет собой меру достижений учащегося в целом. Средний академический балл является одним из примеров меры общего уровня достижений, основанный на предположении, что оценки по предметам сопоставимы без какого-либо шкалирования.

- школа определяет программу изучения предмета в деталях, включая в нее желательную программу оценки («программа работы»),
- центральный орган утверждает программу работы, как отвечающую требованиям учебного плана, включая оценку, которая будет использоваться для определения окончательного результата («выпускной уровень достижения») согласно стандартам, установленным в учебном плане,
- школа выполняет программу работы,
- школа передает в центральный орган *образцы* принятых ей решений в отношении уровней достижений по каждому из небольшой группы учащихся в двух случаях в течение курса (один раз в 11 классе и один раз в 12 классе), предоставляя, если требуется, дополнительную информацию по окончании 12-го года обучения,
- используя свои экспертные группы, работающие на уровне района и штата, центральный орган проверяет правильность принятия решений школой в отношении уровней достижений учащихся в трех случаях (однократно в 11 классе и дважды в 12 классе). В результате такого анализа центральный орган может выдать рекомендации школе о внесении изменений в ее решения,
- центральный орган сертифицирует достижения учащихся по предмету в том случае, если он удовлетворен тем, как школа применила требуемые учебным планом стандарты к работе учащихся по данному предмету.

В течение 11-го и 12-го года обучения могут также заниматься программами по предметам с более низкими ставками (оценки по которым не используются для составления рейтингового списка для поступления в вузы). В отношении этих предметов применяется менее тщательно проработанная система обеспечения качества.

В течение первых десяти лет действия программы оценки на базе школы центральный орган по каждому предмету с высокими ставками организовывал систему координационных совещаний по выработке консенсуса, которые проводились под руководством

координаторов от района и штата¹³. На этих совещаниях учителя обсуждали образцы работы, инструменты оценки и стандарты¹⁴. В целом в те времена центральный орган ожидал, что распределение отметок¹⁵ по всему штату будет соответствовать «нормальному» распределению, а, следовательно, будет приблизительно одинаковым по каждому предмету. На практике были предметы, по которым распределение высочайших результатов по всему штату было постоянно *меньше*, чем требовалось. То есть, несмотря на надежды официального органа на то, что процентная доля высших оценок по каждому предмету будет одинакова, в отношении некоторых предметов учителя применяли одинаково воспринимаемое ими понятие стандартов и принимали решение, что у них недостаточно учащихся, которые достигали бы, с точки зрения учителей, стандартных показателей для высшего результата. Произошло явление, противоположное инфляции оценок, которую можно было ожидать.

Ключевое направление политики, сохранившееся до сегодняшнего дня, присутствовало и в самом начале. Каждый из образцов работы, представляемых на координационном совещании, являл собой представительную подборку работы *отдельного* учащегося, а не выборку ответов учащихся на вопросы конкретных инструментов оценки. Это позволяет экспертам, проводящим анализ сосредоточиться на общем стандарте,

¹³ Географическое положение штата Квинсленд, количество школ, политический контекст, в котором школа принимает решения — все эти факторы подталкивают к использованию местной районной модели. При этом штат осуществляет надзор за вопросами сопоставимости районов. Этот подход, естественно, требует денежных затрат, но использование централизованной модели устранило бы для учителей необходимость брать на себя ответственность за толкование стандартов.

¹⁴ Стандарт можно определить (что и делается на сегодняшний день) словами учебного плана. Стандарт это еще и набор субъективных ожиданий учителя, применяемый им к работе учащихся. На координационных совещаниях учителя анализировали работу учащихся, исходя из собственного понимания стандартов; учебные планы тогда не включали в себя четких определений стандартов. Важно признать тот факт, что определения сами по себе не устанавливают стандартов — стандарты определяются и укрепляются в ходе практической оценки и принятия решений.

¹⁵ В то время результаты оценивались по шкале от 1 (низший балл) до 7 (высший балл).

продемонстрированном в работе конкретного учащегося, а не на отметках учителя, выставленных за конкретный тест или задание¹⁶.

Участие в координационных совещаниях для выработки консенсуса или, на сегодняшний день, в районных встречах экспертно-аналитических групп, является стержневым направлением деятельности «учителей-профессионалов». На этих мероприятиях они рассматривают доказательства результативности работы учащихся, оценивают эти доказательства с точки зрения выполнения стандартов учебного плана и дают соответствующие советы школам по выставлению отметок. Более того, получают большее профессиональное признание (и профессионально совершенствуются), принимая участие в работе экспертных координационных групп.

В течение первых лет реализации программы сертификации средних школ в Квинсленде имели место ощутимые трения. Учителя сомневались в правильности процессов координации, другие были обеспокоены тем, что учителя по-прежнему сохраняют верность традиционному пониманию того, как должна выглядеть оценка.

После целого ряда опросов и отчетов в методику и процедуры программы были внесены изменения. Эти изменения включали в себя:

- открытое признание целей оценки, которые заключались в оценивании обучения, оценивании как обучения и оценивании для обучения¹⁷,

¹⁶ Люди, не знакомые с сопоставлением массы работ с набором стандартов, часто обращаются непосредственно к переоценке отдельных пунктов оценки или к критике отдельных тестов вместо того, чтобы сопоставлять весь массив работ как единое целое с набором стандартов. Точно так же люди, неопытные в реализации подхода, основанного на применении критериев и стандартов системы шт. Квинсленд, не в состоянии вынести суждение о стандартах, если учащиеся не выполняли один и тот же набор тестов.

¹⁷ Оценка обучения: итоговая, официальная сертификация; оценка как обучение: учащиеся анализируют результативность своей работы, используя критерии эффективных практических методов в изучении предмета; оценка для обучения: формирующая, сбор информации о том, какие знания усваиваются учащимися, и использование этой информации как основания для последующей работы.

- учебные планы, в которых приводилось описание процессов более высокого порядка, а также содержания и навыков, актуальных для конкретных учебных предметов,
- экспертно-аналитические группы на уровнях штата и районов, которые проводили аккредитацию рабочих программ и утверждали стандарты, отраженные в решениях школ о результатах учащихся.

В начале 1990-х годов департамент образования штата изучил, насколько эффективны его процедуры в достижении сопоставимости в рамках учебных предметов¹⁸. Департамент проводил (и проводит) эту работу, поручая районным экспертным группам провести анализ произвольной выборки работ учащихся и определить, насколько верно этой работе была дана оценка в соответствии со стандартами штата. Недавние отчеты показывают, что уровень совпадения выводов школы и районной экспертной группы достигает приблизительно 84 процентов. Серьезные разногласия¹⁹ имеют место в приблизительно 7 процентах случаев из всех рассмотренных выборов.

В штате Квинсленд также практикуется уникальный подход к отбору учащихся для продолжения образования в вузах. Тогда как большинство стран при отборе студентов опираются на стандартизированные экзамены, в Квинсленде вместо экзаменов используется описанная выше программа оценки на базе школ. Иными словами, отбор полностью основан на работе учащегося в течение учебного курса последних двух лет средней школы.

С 1974 года учащиеся подают заявления о поступлении в вузы, используя свой действующий на всей территории штата учебный

Программа обеспечения качества соединяет эти функции, а не рассматривает итоговую оценку как нечто «единственно важное».

¹⁸ В данном контексте понятие «сопоставимость» отражает мысль о том, что учащиеся, демонстрирующие конкретный результат по конкретному предмету, провели работу фактически равного качества, даже несмотря на то, что они выполняли разные оценочные задания и что оценки им выставляли разные учителя в разных школах.

¹⁹ Существует 5 (пять) уровней достижений. Каждый уровень подразделяется на 10 подуровней, что в итоге дает 50-балльную шкалу. Серьезным разногласием считается расхождение на один или более уровней достижений и как минимум на 8 (восемь) баллов по 50-балльной шкале.

рейтинг, основанный на их достижениях, показанных в средней школе. Рейтинг, с 1992 года называемый «Общая позиция» (ОП), отражает общие академические достижения учащихся, складывающиеся из их достижений в изучении различных предметов.

Это присвоение рейтинга на основе общих достижений отражает один из принципов политики (не пользующийся широкой поддержкой населения как таковой²⁰), введенный через два года после начала программы. Принцип заключается в построении системы, в которой нет излишнего давления ни на один из ее компонентов. Излишнее давление возникает, когда результаты по одному компоненту — экзамену, оценке с использованием координационной системы — становятся единственной основой для принятия решений с высокими ставками. Последствия такого излишнего давления могут проявляться по-разному: в случае с внешними тестами оно может привести к преобладанию надежности над достоверностью; в случае же использования координационных систем оно может привести к различного рода злоупотреблениям.

С 1974 года в конце августа — начале сентября центральный орган проводит тест (с 1992 года называемый «Ключевые навыки по классификации шт. Квинсленд» — тест QCS) на общие академические достижения для всех (подлежащих тестированию) учащихся 12 классов. Главная цель теста — собрать групповую (по школе и по предметному курсу) информацию (параметры основной тенденции и распространения), которая позволит создать действующий во всем штате рейтинг учащихся для поступления в вузы (ту самую ОП).

Тест QCS не является внешним экзаменом в том смысле, что он не используется для индивидуальной сертификации с высокими ставками: учащиеся, действительно, получают индивидуальные результаты, но, в отличие от случая с внешним экзаменом, эти результаты не используются для принятия каких-либо решений с высокими ставками в отношении отдельного учащегося: два учащихся с идентичными результатами по предмету и одинаковым

²⁰ Признание этого содержится в отчете об анализе системы 1990 года.

рейтингом для поступления в вуз могут иметь очень сильно различающиеся результаты теста QCS. Рейтинги учащихся или баллы по тесту QCS не соединяются с результатами по предмету.

До 1989 года этот тест представлял собой подборку из 100 вопросов с несколькими вариантами ответов, содержался в секрете и по техническим характеристикам должен был отвечать требованию проверить вербальное и количественное мышление. Ограниченная природа теста (несколько вариантов ответа) и его секретность (которая вызывала подозрения в его сути и достоверности) создавали напряжение, которое привело к следующим значительным изменениям:

- включение в тест письменного задания с 1989 года,
- с 1992 года начал действовать реструктурированный тест по 49 общим элементам программы для старших классов средней школы шт. Квинсленд,
- с 1992 года в тест введен элемент с самостоятельным кратким ответом,
- опубликование всех протоколов теста, начиная с 1992 года.

Эти изменения существенно улучшили репутацию теста, его структуру и повысили вытекающую из нее достоверность (за счет снижения надежности — альфа Кронбаха обычно составляет примерно 0,86 — 0,89) и, что естественно, затраты на его проведение: новый тест каждый год плюс проверка и оценка проверяющим письменного задания и задания, требующего краткого ответа. В то же время тест оказал положительное воздействие на практическую работу учителей по оценке и заставил школы сосредоточить внимание на проверяемых навыках (общих элементах программы).

Результаты тестирования выражаются в индивидуальной оценке (от А до Е, имеется также оценка А⁺), которая указывается в сертификате учащегося и в ведомости баллов²¹. Школы присваивают каждому учащемуся по каждому предмету номер по интервальной

²¹ Хотя неофициально эти баллы рассматривались как информация об эффективности работы школы, официально они не являются показателем качества образования.

шкале²², проводя сравнение достижений учащихся в данной школе по данному предмету. В настоящее время этот метод известен как Индикатор достижений по предмету (SAI). В интервальных шкалах по предметам не используются абсолютные величины — они отражают относительные достижения в рамках группы учащихся. Различия между школами в рамках *предметов* не имеют значения при условии, что каждая школа проводит справедливое сравнение успехов собственных учащихся.

Информация о таких группах в каждой школе по каждому предмету (основанная на полученных учащимися баллами за тест) используется для наложения каждой из этих интервальных шкал на общую шкалу (общих достижений, а не достижений по конкретному предмету). В результате объединения этих шкалированных баллов по предметам создается интервальная шкала по всей школе, а перешкалирование этих интервальных шкал школьного уровня создает общую базу по всему штату. Интервальная шкала в масштабах штата затем используется для создания рейтингового порядка по всему штату. Целью шкалирования является восстановление справедливости за счет *удаления* любого влияния разницы в уровнях академических²³ характеристик по школам и предметам.²⁴

²² Интервальные шкалы показывают порядок и пробелы. Ноль в этих шкалах отсутствует. С 1992 года тот же самый диапазон значений используется по всем предметам во всех школах: наименее успешному учащемуся в каком-либо предмете в какой-либо школе присваивается балл 200, а самый успешный учащийся по данному предмету в данной школе получает 400 баллов. До 1991 года школы использовали шкалу от 1 до 99 баллов, что создавало впечатление, что успех оценивается в процентах.

²³ Конструкт, лежащий в основе рейтингового порядка в масштабах штата, это «общие академические достижения», а не достижения по конкретным предметам. Конструкт, лежащий в основе теста QCS, — это, следовательно, общие академические достижения. Лучше всего представлять себе конструкт теста как набор из 49 элементов *общей* учебной программы. Когда групповые результаты теста QCS используются для сравнения академических успехов группы учащихся в школе, изучающих физику с успехами группы в этой же школе, изучающей французский, опорным конструктом в этой ситуации будет не физика и не французский, а общие академические достижения.

²⁴ Примите во внимание, что место каждого отдельного учащегося в рейтинговом списке на уровне штата определяется, исходя из достижений по предмету, а не по результату теста QCS: два учащихся, имеющие одинаковый результат по одному и

Процедура установления рейтинга для поступления в вузы была введена в действие после двухлетнего опыта использования механизма отбора, основанного на средних оценках, с целью удовлетворения потребности университетов в справедливом и разумно жестком подходе к выбору абитуриентов по «академическим заслугам». При использовании средних баллов по предметам возникает очевидная проблема, связанная с тем, что при данном подходе баллы по всем предметам рассматриваются как эквивалентные по значимости²⁵, в связи с чем возникает как реальное, так и воспринимаемое давление на сопоставимость баллов по отдельным предметам.

На уровне штата суммы баллов по тестированию ключевых навыков и отдельные расчетные показатели по общим достижениям, выведенные на основании оценки, проведенной учителями, достаточно хорошо коррелируются (~0,7–0,75), если тест использовать как тест для шкалирования: на практике возникают проблемы с достоверностью в случае его использования для малых групп и для всех групп учащихся — в целом удовлетворительное общее соотношение может скрывать аномалии.

Главные допущения в отношении предоставления школами интервальной шкалы по каждому предмету, изучаемому учащимися в данной школе (SAI), сводятся к следующему:

- учителя-предметники могут справедливо сравнивать достижения учащихся, которых они учили и оценивали тщательнейшим образом²⁶

тому же предмету в одной школе, но разные результаты по тесту QCS, получают одно и то же место в рейтинговом списке на уровне штата.

²⁵ Как бы ни определялись оценки, нет оснований предполагать, что это правильное допущение. Если оценки определяются «нормативно», они сопоставимы лишь в рамках предметов и только если изучающие предмет когорты аналогичны, а это, как известно, не так. Если оценки определяются по критериям и стандартам, они относятся к конкретному предмету — оценка по французскому языку отражает достижения в изучении французского, а оценка по физике не отражает достижений в изучении французского языка.

²⁶ Суть внешней координации на уровне штата заключается в том, что за счет ее обеспечивается разумный уровень сопоставимости оценок по предмету, если существует ряд из пяти-семи четко разграниченных оценок. Она не может обеспечить разумный уровень сопоставимости на шкале от 50 до 100 пунктов, и, что более важно,

- отсутствует межпредметная корреляция ошибок в данном процессе — каждый учитель по предмету определяет результаты независимо от учителей по другим предметам.

В течение нескольких лет процедуры работы по программе были более тщательно проработаны. Этому способствовала необходимость иметь возможность внесения коррективов в результаты учащихся в тех случаях, когда имеют место существенные отклонения от указанных допущений.

Поскольку первостепенным назначением²⁷ теста является его использование при шкалировании, значение сопоставимости оценок по годам незначительно. Это в настоящее время делается за счет уравнивания на основе информации по проверке вопросов теста²⁸.

Индивидуальные оценки по тесту, тем не менее, в некоторой степени используются в альтернативных схемах приема студентов на определенные курсы в университетах. Данная тенденция, по-видимому, нарастает. Доля учащихся, попадающих в стандартный рейтинговый список штата, снижается (например, в 1987 году — 84 процента от общего числа учащихся 12 классов, в 2001 году — 72 процента). Это порождает особое внимание к достоверности процедур, используемых для шкалирования оценки по предметам²⁹. Оба эти фактора создают фон, требующий перемен.

для участников очевидна эта ее неспособность. У экзаменов имеется похожая проблема, грубо отслеживаемая на стандартной погрешности измерения или на стандартной погрешности оценки, но это не очевидно для учителей, учащихся и родителей.

²⁷ Учащиеся также могут использовать свои результаты теста QCS, чтобы показать, что они отвечают требованиям нового сертификата (QCE) по стандартам грамотности и навыкам счета.

²⁸ Вопросы теста подвергаются проверке, обычно, в различных формах, на учащихся за пределами штата Квинсленд. Анализ результатов таких проверок (проводящихся в закрытом режиме) используется для определения целесообразности использования этих вопросов в тесте. Сам тест публикуется ежегодно.

²⁹ Например, уменьшение доли учащихся, попадающих в рейтинг, означает большее количество малых групп (предметные классы, в которых занимается менее 15 учащихся); большее количество случаев, когда группа учащихся в школе находится в ограниченном диапазоне достижений.

Стимулирующая среда

Какие компоненты необходимы для обеспечения работы столь комплексной и сложной программы оценки достижений учащихся? В данном разделе представлено описание политик, учреждений, кадровых и финансовых ресурсов и более широкого социального контекста. Именно сочетание всех этих компонентов позволило внедрить в штате Квинсленд оценочную программу на базе школ с внешней координацией.

Управление образования штата Квинсленд (QSA) осуществляет надзор за качеством программы на уровне штата. Этот орган был создан на основании соответствующего законодательного акта в 2002 году, но и до этого года он существовал под другими названиями и выполняя другой набор функций. Это в высокой степени стабильное и организационно оформленное учреждение, финансируемое штатом, организует свою деятельность независимо от правительства.

QSA имеет устойчивый и квалифицированный штат сотрудников. В 2011 году в Управлении работали 248 сотрудников на условиях занятости, эквивалентной постоянной. Большинство сотрудников работало в центральном офисе, и 13 человек — в районных филиалах в различных частях штата. Замещение должностей осуществляется на основе комбинации постоянных и временных назначений, а также откомандирований.

Подсчитать общие затраты на осуществление программы оценки на базе школ в Квинсленде не просто. С одной стороны, полный расчет затрат на программу должен состоять из затрат на оплату рабочего времени учителей, затрачиваемого на подготовку программ работы, оценку достижений учащихся и внесение вклада в процессы, обеспечивающие сопоставимость результатов в масштабах штата. С другой стороны, практически все эти виды деятельности входят в круг обычных профессиональных обязанностей учителя. Эти затраты не включены в расходы центрального органа и не фигурируют в смете затрат, приведенной ниже.

В приведенных ниже расчетах затрат нашли отражение все прямые издержки на сертификацию результатов по предметам в

старших классах средней школы и определение рейтинга для его дальнейшего использования при отборе студентов вузов. Имеющие место в настоящее время ежегодные расходы QSA на оценку на базе школы с внешней координацией составляют 19 млн. долларов США, из которых около 5,3 млн. долларов США расходуется на внешнюю оценку для шкалирования (тест основных навыков по классификации шт. Квинсленд), и ориентировочно 10,7 млн. долларов расходуется на программу внешней координации.³⁰

В течение нескольких лет работы программы приблизительно половина общей суммы расходов центрального органа направлялась на заработную плату сотрудников, из которых большая часть трудится в главном офисе, и несколько человек работают в районных центрах. Согласно расчетам, затраты этого органа штата с 1987 года возросли более чем вдвое, тогда как количество учащихся за этот период выросло на 50% (Совет по делам старшей ступени средней школы, 1987; и Управление образования штата Квинсленд, 2009 г.).

Судя по всему, данный подход к сертификации учащихся средних школ обходится дешевле, чем традиционные программы внешних экзаменов, используемой для тех же целей. На самом деле, текущие прямые затраты на оценку достижений одного учащегося (один результат по одному предмету одного учащегося) составляют, по-видимому, 70 долларов США. Затраты на учащегося в рамках программы оценки на базе школ в Квинсленде составляют менее 60 процентов таких затрат в рамках любой из сопоставимых программ внешних экзаменов, действующих в Австралии.

³⁰ Это меньше половины текущих общих расходов центрального органа, который в настоящее время имеет расширенный круг обязанностей и полномочий, включающий в себя проведение тестирования учащихся 3, 5, 7, и 9 классов во всем штате, составление и организацию Задач сопоставимой оценки шт. Квинсленд (QCAT) в 4, 6, и 9 классах, регистрацию школ как учреждений профессионально-технического образования и составление учебных программ для всех уровней школьного образования в штате Квинсленд.

Факторы перемен

Программа оценки достижений учащихся на базе школ в Квинсленде впервые проинформировала учащихся 12-го года обучения о результатах их тестирования в 1972 году, вслед за принятием правительством штата рекомендаций из доклада педагогической комиссии. Эта комиссия была создана в ответ на имевшиеся у населения сомнения в отношении достоверности внешних экзаменов как измерителя достижений учащихся³¹, озабоченность строгостью выставления оценок по результатам экзамена³², нарастающим разнообразием когорты учащихся старшей ступени, и необходимостью подготовки молодежи к выходу на рынок труда.

Последовавшие за опубликованием доклада публичные дебаты имели предсказуемые элементы, но реформа была поддержана профсоюзом учителей (который отметил, что его более молодые члены, с большей степенью вероятности ее поддержат), так как она вела к повышению профессионального уровня; а ведущий университет Квинсленда не выступил против нее. Ряд совещаний, прошедших по всему штату, с участием населения, представителей бизнеса и групп родителей, по-видимому, помог погасить потенциальную волну озабоченности. Потенциал, навыки, личная преданность делу и стаж работы сотрудников департамента образования, участвовавших в этих совещаниях, способствовали «подключению» людей к процессу. Большие расстояния усложняли

³¹ Очевидно, что педагоги того периода времени нашли бы отражение своих философских мыслей в таких идеях, как «оценка обучения, оценка как обучение и оценка ради обучения»; важность оценки крупных и важных аспектов обучения, не ограничиваясь оценкой лишь тех аспектов, которые проще всего измерить при помощи стандартизированных тестов; и необходимость повышения результативности обучения путем создания прочной связи между программой обучения, преподаванием и оценкой.

³² В частности, письменный экзамен по физике 1967 г. Экзаменационные задания составлялись преподавателями университетов. Данный экзамен по физике был составлен таким образом, что большинство учащихся его провалило; бессмысленный результат с точки зрения учителей и общества, которые понимали, что экзамен сдавали очень способные учащиеся. После массового возмущения, проходной балл был снижен, а результаты пересмотрены.

процесс достижения общественного согласия, однако те факты, что население Квинсленда невелико, а влиятельные люди имеют тесные связи друг с другом, помогли в решении данной задачи.

Политический аспект сыграл серьезную роль во введении новой программы оценки на базе школы. С 1968 по 1988 г. в правительстве Квинсленда преобладающим влиянием пользовалась политическая партия, электорат которой состоял из населения небольших городов и сельской местности. Очень небольшая часть взрослого электората имела более, чем 8-летнее образование. Иметь больше 8 классов образования было необычным явлением и для большинства министров. Политику внедрения программы оценки в штате Квинсленд необходимо рассматривать в контексте:

- сомнений негородского электората правительства в возможности удовлетворения образовательных потребностей их детей и их подозрений, что будет иметь место излишний диктат со стороны столицы;
- ограниченной осведомленности среди министров правительства о более поздних ступенях образования с вытекающей из этого тенденцией следовать совету экспертов, когда нет конфликта с другими политическими вопросами;
- наличия министра образования, желающего и способного действовать, опираясь на указания самого высокопоставленного чиновника, который взял на себя роль лидера в сфере образования и, свою очередь, в своих действиях опирается на консультативную помощь выдающихся специалистов в области образования.

Введение в действие программы оценки на базе школы опиралось на убежденность в том, что именно учителя обладают наивысшей квалификацией, необходимой для вынесения суждений об эффективности работы учащихся. Заинтересованные стороны не доверяли внешним оценкам. Заинтересованные стороны пришли к консенсусу в отношении важности наделения учителей новыми правами и возможностями в рамках новой программы оценки.

Внедрение программы оценки на базе школ требует времени. Учителям необходимо научиться интегрировать оценочные мероприятия в профессиональную практику. Необходимо

постоянно освежать знания учителей, укреплять их убежденность в необходимости оценки и поддерживать практическое применение ими методов оценки.

Создание рейтинга учащихся по всему штату на основании оценки на базе школы. Хотя идея кажется достаточно простой и интуитивно понятной, обеспечение достоверности и справедливости рейтинга всех учащихся требует очень тонкого подхода. Данный подход основан на ключевом допущении, что учителя обладают наивысшей квалификацией, необходимой для анализа достижений учащихся, которых они учили.

Некоторые сомневаются в том, что программа штата Квинсленд является эффективной и правильной и в том, что она когда-нибудь станет таковой. Некоторые считают, что стандартизированный анонимно выполненный тест — это единственный способ обеспечить достаточную степень надежности, последовательности и справедливости. Сторонники анонимных стандартизированных тестов по-прежнему придерживаются своего мнения, несмотря на хорошо известный эффект сужения преподаваемого и усваиваемого программного материала (а, следовательно, и снижения достоверности, когда планируемая программа содержит учебные результаты, которым в стандартизированном тесте не уделяется внимания). С точки зрения сторонников этих тестов, программа штата Квинсленд *априори* неработоспособна, недостаточно надежна и строга. Данные об эффективности координационного процесса — ежегодного изучения произвольной выборки — с этой точки зрения не могут быть достаточным доказательством противоположного взгляда.

Однако программа штата Квинсленд, тем не менее, работает и работает уже в течение длительного времени. Ее эффективность зависит не только от процесса и процедур программы, но и уверенности общества (характеристика культуры и обстановки) в том, что учителя могут действовать справедливо и разумно при вынесении суждений о работе учащихся. Для построения такого доверия необходимы время, тщательно проработанные процедуры и наличие должного лидерства.

Извлеченные уроки

В той или иной форме Квинсленд уже в течение 40 лет располагает действующей программой оценки на базе школы с внешней координацией. В настоящем разделе изложены уроки, которые могут извлечь для себя другие страны из опыта штата Квинсленд.

Введение в действие и поддержание работоспособности программы оценки на базе школы требуют времени и усилий

Руководители системы образования в Квинсленде воспользовались социально-культурной и политической обстановкой, климатом, способствовавшим изменениям, и конкретными событиями для реализации предложений, которые на сегодняшний день могут показаться радикальными. Реформы касались самой сущности оценки как культурной практики, практики, оказывающей глубочайшее воздействие на усваиваемые знания. Провести сегодня подобные радикальные изменения было бы трудно по ряду причин. Им могли бы помешать, например, современные ожидания быстрого эффекта от реформ в образовании, природа публичной дискуссии об образовании в круглосуточных СМИ, прямая вовлеченность политических руководителей в принятие решений в сфере образования.

Проведение такой реформы — особенный вызов. Если люди не считают, что изменения действительно необходимы, нужного для реализации изменений давления не будет создано. С другой стороны, если люди убеждены в насущной необходимости изменений, возникнет соответствующее давление и напор, требующие безотлагательного и результативного проведения таких изменений.

Продолжающемуся существованию программы Квинсленда способствуют следующие факторы:

- низкие прямые затраты правительства штата;
- отношение органа управления образования штата к школам как к партнерам;
- долгосрочное бесменное руководство в центральном органе;

- небольшое количество школ в Квинсленде, позволяющее поддерживать тесные связи между руководителями в центре (центральный орган) и периферией (школами);
- орган управления образованием штата независим, а, следовательно, может действовать в партнерстве со всеми государственными и частными школами.

Начать с малого

Координируемая извне оценка на базе школы является сертификацией с высокими ставками, которая была введена для всех школ (частных и государственных) одновременно. Воздействие оценки на практическую работу учителей приобретает размах и вес только тогда, когда оценки имеют реальное значение, то есть, когда результаты оценки становятся поворотным пунктом с высокими ставками для учащихся. Может ли такая программа внедряться поэтапно?

Одним из предварительных шагов может стать начало наращивания опыта и специальных знаний у учителей путем реализации оценочной программы на базе школы, которая не имеет целей с высокими ставками, а, следовательно, не вызовет беспокойства у населения. Затем можно будет поступательно наращивать программу, пользуясь извлеченными из реализации программы в Квинсленде в 1980-е годы уроками, касающимися важности поэтапного введения существенных изменений в оценку. Сложности будут связаны с необходимостью убедить оставшихся в предыдущей программе в том, что они не оказались в невыгодном положении, а участников новой программы в том, что они не подвергаются эксперименту.

Разумно на начальном этапе создать программу простейшей структуры, заложив в нее возможность дальнейшего усложнения по мере необходимости. Изначально сложную по структуре программу трудно вводить в действие и невозможно развивать в будущем.

Поддержание добропорядочности

Допущение о том, что учителя и школы могут действовать и действуют профессионально (при поддержке со стороны системы сдержек и противовесов, обеспечиваемой эффективной внешней

координацией), включает в себя вотум доверия в отношении добропорядочности. Иными словами, ожидается, что учителя и школы противостоят попыткам лиц извне оказывать нежелательное влияние на процесс принятия решений. В основе программы штата Квинсленд лежит предположение, что попытки оказывать неправомерное давление будут ограничиваться внутренне присущей программе легитимностью, а также системой сдержек и противовесов, обеспечивающей прозрачность решений.

Принятие программой решений школы, основанное на том, что эти решения строятся на применении к работе учащихся четких стандартов, на предъявляемых по требованию доказательствах и являются результатом процесса, который демонстрирует прозрачность принимаемых решений. Такой механизм эффективен, когда он используется в условиях, когда попытки оказать неправомерное давление редки и обычно воспринимаются как неприемлемое, даже коррупционное, явление. В конце 1980-х годов существенные коррупционные явления, вскрытые в то время в других областях работы правительства штата, по-видимому, не распространились на область действия программы оценки достижений учащихся верхней ступени среднего образования. Этому, возможно, способствовали установленная законодательно независимость центрального органа и его преданность идее добропорядочности. Также возможно и то, что коррумпированные чиновники не увидели в результатах учебы учащихся 12-го класса источника доходов или влияния.

Особо подчеркивать профессионализм учителей как ключевой компонент

С самого начала в рамках программы штата Квинсленд была определена не только ответственность за оценочные решения с высокими ставками на уровне школы, но и ответственность за детальную проработку курсов и программ оценки, в которых участвуют учащиеся конкретных школ. Ответственность на местах основывалась на профессионализме учителей и зависела от него. Кроме того, программа основывалась на понимании того, что наилучшие результаты учебы достигаются, когда учитель разрабатывает и реализует учебную программу, определяет, как

будет собираться информация о достижениях учащихся, собирает и использует эту информацию.

Позиционирование учителя как справедливого и разумного арбитра, применяющего стандарты к работе учащихся, накладывает на учителя дополнительную ответственность. Это, кроме того, не позволяет учителю играть традиционную роль в рамках внешней экзаменационной программы, то есть принимать сторону учащегося и пытаться выиграть соревнование с экзаменатором.

Имеется некоторая обеспокоенность «излишней нагрузкой» на учителей. Ведь в конечном итоге, круг их обязанностей включает в себя формирование целей обучения, принятие решения, как им предстоит собирать информацию о достижениях их учащихся на пути к этим целям, и принимать участие в процессе внешней координации с тем, чтобы оценочные решения были частью сертификационной программы в масштабах всего штата. Рабочая нагрузка всегда является деликатным вопросом для профсоюзов учителей. Они должны на должном уровне выражать заботы членов союза, и одновременно способствовать общественному восприятию учителей как профессионалов³³.

Можно допустить, что теория и практика этих профессиональных обязанностей уже давно должна была бы быть обычной частью курса педагогического образования учителей. Приблизительно до начала последнего десятилетия это было, скорее, исключением, чем правилом.

Что касается образования без отрыва от работы, один из главных уроков, извлеченных из реализации программы оценки штата Квинсленд, заключается в том, что наиболее успешное профессиональное обучение имеет место как побочный эффект участия учителей в оценках с высокими ставками либо на уровне школы, либо на уровне внешних аналитических экспертных групп.

Значение того, что программа оценки основана на идеале высокого профессионализма учителей, не следует недооценивать.

³³ Тенденции на государственном уровне в 2005 году, имевшие целью переход на систему внешних экзаменов, в Квинсленде натолкнулись на сильное противостояние, выразившееся в заявлениях против таких изменений, прозвучавших из уст политических руководителей, руководителей сектора образования, профсоюзов учителей и родительских групп.

Учителя-профессионалы гарантируют, что практическая работа в классе способствует глубокому усвоению учебного материала, что чрезвычайно важно для будущего учащихся. Программы, поддерживающие на словах одно отношение к учителям, а на деле подразумевающее другое, или системы, действующие так, как будто большинство учителей не может быть профессионалами, столкнутся в итоге с тем, что многие учителя начнут отвечать именно таким ожиданиям, хотя и будут иметь место заслуживающие уважения исключения. Программы же, построенные на ожидании профессионального поведения, с течением времени столкнутся с ситуацией, когда многие (хотя и не все) учителя будут отвечать этому ожиданию.

Вывести воздействие оценки на преподавание и усвоение учебного материала в число приоритетов

Введение в действие и непрерывная реализация программы оценки в 12 классах основаны на допущении, что эффективность программы оценки и сертификации зиждется, прежде всего, на ее способности вводить стремление к высочайшему качеству в обучении учащихся, а не на точности измерения. Несомненно, что программы сертификации должны стремиться к тому, чтобы сертифицируемые ими результаты имели уровень достоверности и надежности, ожидаемый лицами, использующими эти результаты для принятия решений.

Из всех факторов, на которые система образования хотела бы оказывать прямое влияние, фактор практической учительской работы, вероятно, является наиважнейшим. Именно то, что происходит в классах (практические методы и используемые стандарты), непосредственно воздействует на усвоение учебного материала. В управлении практическими методами преподавания и стандартами в классе невозможно переоценить роль учителя. При разработке эффективной программы оценки и сертификации необходимо учитывать исключительную важность того, что делают учителя, и как программа может вносить вклад в совершенствование их педагогических методов.

То, что учителя делают в классах, формируется их участием в «сообществах практиков», то есть в сетях профессионалов, где

происходит формирование, распространение, оформление и изменение на базе опыта идей, понимания, знаний, методов и подходов к обучению. Программа штата Квинсленд создает и поддерживает функционирование высоко эффективных сообществ практиков, участвуя в которых учителя формируют, совершенствуют и углубляют методы своей работы и методы познания того, что, как и насколько хорошо узнают в процессе обучения их учащиеся.

Суждения о ценности программы штата Квинсленд зависят от того, рассматривают ли ее значение, прежде всего, через призму психометрического или стандартизированного тестирования, или через призму «учения о действии». Учение о действии сосредоточивает внимание не на точности и последовательности измерения как такового, а на воздействии программы оценки на то, что преподается и усваивается учащимися.

Возможно, самым важным уроком, который следует извлечь из опыта работы программы оценки на базе школ в штате Квинсленд, является понимание возможности интеграции учебной программы, преподавания, практических методов работы учителей в классе и системы оценочных мероприятий с высокими ставками по всему штату. Чтобы такая система эффективно работала, необходимо сделать приоритетом воздействие программ оценки на практику преподавания и усвоения учебного материала. Это означает, что необходимо сделать приоритетом достоверность (как на уровне дизайна, так и на второстепенных уровнях), а не показатели технической надежности и составление уравнений. Взяв на вооружение такой подход, страны окажутся в более выгодном положении для создания сильной системы оценки, которая сможет вносить реальный вклад в повышение качества образования.

Библиография

- Allen, J. R. 1988. *ASAT and TE Scores: A Focus on Gender Differences*. Brisbane, Queensland: Board of Secondary School Studies.
- — —. 2002. "Tertiary Entrance." In Pitman, *The Senior Certificate: A New Deal*. Brisbane, Queensland: Department of Education.
- Australian Curriculum, Assessment, and Reporting Authority (ACARA). 2011. *NAPLAN Achievement in Reading, Persuasive Writing, Language Conventions, and Numeracy: National Report for 2011*. Sydney: ACARA.
- Board of Secondary School Studies. 1987. *Annual Report 1987*. Brisbane, Queensland: Board of Secondary School Studies.
- Butler, J., L. Bartlett, W. Beasley, N. Hardy, and L. Logan. 1984. *Overview of Rosba Consultancy Activities*. Brisbane, Queensland: University of Queensland.
- Clarke, E. 1987. *Assessment in Queensland Secondary Schools: Two Decades of Change 1964–1983*. Brisbane, Queensland: Department of Education.
- — —. 1990. *Assessment in Queensland Secondary Schools: 1983–1990*. Brisbane, Queensland: Department of Education.
- Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." *SABER–Student Assessment Working Paper No. 1*. World Bank, Washington, DC.
- Lingard, R. 1990. "Accountability and Control: A Sociological Account of Secondary School Assessment in Queensland." *British Journal of Sociology of Education* 11(2): 171–88.
- Luxton, P., and R. Dudley. 2008. "The Development of P-12 Assessment Policy in Queensland, Australia." Paper given at the 34th International Educational Assessment Annual Conference 2008. Cambridge, England.
- Matters, G., J. Pitman, and J. O'Brien. 1998. "Validity and Reliability in Educational Assessment and Testing: A Matter of Judgement." *Queensland Journal of Education Research* 14.
- McCollow, J. 2008. "Square Pegs, Round Holes: Defending School-Based Assessment." *Queensland Teachers Union Professional Magazine*, October.

- McGaw, B. 1989. "Comments on Tertiary Education in Queensland: A Review." *Queensland Journal of Education Research* 5.
- Organisation for Economic Co-operation and Development (OECD). 2010. *PISA 2009 Results: What Students Know and Can Do—Student Performance in Reading, Mathematics and Science (Volume I)*. Paris: OECD.
- Pitman, J. P. (chair). 1987. *Tertiary Entrance in Queensland: A Review*. Brisbane, Queensland: Board of Secondary School Studies.
- Pitman, J. P. 2002. *The Senior Certificate: A New Deal*. Brisbane, Queensland: Department of Education.
- Queensland Department of Education. 1985. *Education 2000: Issues and Options for the Future of Education in Queensland*. Brisbane, Queensland: Department of Education.
- Queensland Studies Authority (QSA). 2009. *P-12 Assessment Policy*. Brisbane, Queensland: QSA.
- . 2010a. *2009 Data Summary: Year 12 Enrolment and Certification*. Brisbane, Queensland: QSA.
- . 2010b. *School-based Assessment: The Queensland System*. Brisbane, Queensland: QSA.
- . 2011. *Annual Report 2010–11*. Brisbane, Queensland: QSA.
- . n.d. Web site information about Year 11 and 12 Assessment. Brisbane, Queensland: QSA. Available at: www.qsa.qld.edu.au.
- Radford, W. C. (chair). 1970. *Public Examinations for Queensland Secondary School Students*. Brisbane, Queensland: Department of Education.
- Sadler, D. Royce. 1995. "Comparability of Assessments, Grades, and Qualifications." Address to 1995 AARE Conference. Hobart.
- Scott, E. (chair). 1976. *A Review of School-Based Assessment in Queensland Secondary Schools (1978)*. Brisbane, Queensland: Board of Secondary School Studies.
- Viviani, N. 1990. *The Review of Tertiary Entrance in Queensland 1990*. Brisbane, Queensland: Department of Education.

Список работ, ранее опубликованных в данной серии публикаций

1. Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper" (*также в переводе на русский: Кларк, М. 2012. «Что является наиболее важным в системах оценки достижений учащихся: основные ориентиры» и португальский языки*).
2. Ramirez, M. 2012. "Developing the Enabling Context for Student Assessment in Chile" (*также в переводе на русский: Рамирес, М. 2012. «Развитие стимулирующей среды для оценки достижений учащихся в Чили»*).
3. Ramirez, M. 2012. "Disseminating and Using Student Assessment Information in Chile" (*также в переводе на русский язык: Рамирес, М. 2012. «Распространение и использование информации о результатах оценки достижений учащихся в Чили»*).
4. Liberman, J. and M. Clarke. 2012. "Review of World Bank Support of Student Assessment Activities in Client Countries, 1998–2009" (*также в переводе на русский язык: Либерман Дж. и М. Кларк. 2012. «Обзор работы Всемирного банка по поддержке деятельности в области оценки достижений учащихся в государствах-клиентах в период с 1998 по 2009 г.г.»*).
5. Flockton, L. 2012. "The Development of the Student Assessment System in New Zealand" (*также в переводе на русский язык: Флоктон, Л. 2013 «Развитие системы оценки достижений учащихся в Новой Зеландии»*).
6. Allen, R. 2012. "Developing the Enabling Context for School-Based Assessment in Queensland, Australia" (*также в переводе на русский язык: Аллен, Р. 2013. «Развитие стимулирующей среды для оценки достижений учащихся на базе школы в штате Квинсленд, Австралия»*).
7. Castro, M. 2012. "Developing the Enabling Context for Student Assessment in Brazil" (*также в переводе на русский язык: Кастро, М. 2013. «Развитие стимулирующей среды для оценки достижений учащихся в Бразилии»*).

Готовящиеся к публикации материалы

8. Kanjee, A., and S. Acana. 2013. "Developing the Enabling Context for Student Assessment in Uganda."

Австралийский штат Квинсленд дает нам уникальный пример того, как программа оценки может воплотить в себе принципы оценки учебы, оценки как учебы и оценки для учебы. Программа оценки штата Квинсленд может служить примером для других стран, желающих найти более эффективные пути измерения учебных достижений учащихся средней школы. Этот вопрос приобретает особую актуальность на фоне глобальной дискуссии о путях создания систем оценивания, которые были бы более достоверны, требовательны и не ограничены форматом письменного теста, основанного на вопросах с несколькими вариантами ответов. Актуальность этого вопроса также подтверждается идущей в настоящее время дискуссией о том, как повысить качество педагогических кадров и обеспечить непрерывное повышение квалификации учителей.

Рег Аллен, *руководитель Квалификационного управления Тасмании (Австралия)*