

Национальная оценка учебных достижений

Во всем мире эффективная система образования играет ключевую роль в формировании политики, направленной на оптимизацию развития человеческого капитала. Серия «Национальная оценка учебных достижений» знакомит читателя с основными моментами, касающимися оценки образовательных достижений учащихся на разных уровнях обучения. В пяти книгах рассматривается широкий круг вопросов, начиная с политических аспектов и заканчивая моментами, которые касаются проектирования и проведения оценки, такими как разработка тестов и дизайн анкет, формирование выборки, организация сбора данных, метод статистического анализа, чистка данных, написание отчета и использование результатов национальной оценки для повышения качества образования.

Что изучают школьники и студенты? В настоящее время, работая над повышением качества образовательных услуг, во всем мире правительства все чаще обращаются к национальной оценке, которая предоставляет информацию о достижениях учащихся по ключевым областям учебного плана. В последние годы важность национальной оценки существенно возросла, однако использование ее результатов пока не получило широкого распространения. Эта книга призвана повысить ценность результатов национальной оценки в глазах общественности и помочь странам, принимающим участие в оценочных процедурах, в их распространении.

В книге «Использование результатов национальной оценки учебных достижений» освещены основные факторы, которые определяют практическое применение результатов оценочных процедур. Это политический контекст, вид оценки (определяется исходя из того, основывается ли она на генеральной совокупности или выборке), характер основного отчета и качество оценочного инструментария. Приводятся данные о том, какие сведения должен содержать отчет о результатах национальной оценки и каковы возможности другого инструментария. Выявлены особенности перевода полученных данных национального оценивания в практические действия и политические решения, а также описаны конкретные процедуры использования результатов для формирования государственной политики в области образования, особенностей обучения и общественного восприятия.

Книга будет полезна политикам, исследователям и практическим деятелям системы образования, решающим задачи его развития.



THE WORLD BANK

CICEP



THE WORLD BANK

CICEP

Национальная оценка учебных достижений

Томас Келлаган, Винсент Грини, Т. Скотт Мюррей

Использование результатов национальной оценки учебных достижений

КНИГА 5

Использование
результатов
национальной оценки
учебных достижений

Using the Results
of a National
Assessment of
Educational
Achievement

National Assessments of Educational Achievement

VOLUME 5

Using the Results of a National Assessment of Educational Achievement

**Thomas Kellaghan
Vincent Greaney
T. Scott Murray**



THE WORLD BANK
Washington, DC

Национальная оценка учебных достижений

Томас Келлаган, Винсент Грини, Т. Скотт Мюррей

Использование результатов национальной оценки учебных достижений

КНИГА 5



Москва
Логос
2014

УДК 378
ББК 74.58



*Издание подготовлено при поддержке
Международного банка реконструкции и развития*

К34



*Перевод подготовлен по инициативе
Центра международного сотрудничества
по развитию образования Российской
академии народного хозяйства
и государственной службы при
Президенте Российской Федерации*

Келлаган, Томас

К34 Использование результатов национальной оценки учебных достижений / Томас Келлаган, Винсент Грини, Т. Скотт Мюррей; пер. с англ. С.В. Фирсовой; под. ред. В.И. Звонникова. – М.: Логос, 2014. – 248 с.: ил. – (Национальная оценка учебных достижений)

ISBN 978-5-98704-575-6

Рассмотрены факторы, влияющие на использование результатов национальной оценки учебных достижений, к которым прежде всего относятся политический контекст, подотчетность, качество оценочного инструментария и вид оценки. Раскрыты подходы к сообщению результатов национальной оценки и переводу полученных данных в конкретные решения и действия. Проанализирован инструментарий для представления результатов оценки, показано их влияние на политику и менеджмент в образовании и на систему обучения в целом. Затронуты ключевые аспекты использования результатов национальной оценки для определения политики в сфере образования, уточнения его стандартов, обеспечения учебных заведений ресурсами, пересмотра учебных планов и программ, профессионального развития преподавателей. Показаны особенности общественного восприятия национальной оценки. Предлагаются пути оптимизации использования результатов национальной оценки и повышения ее роли в дальнейшем развитии образования.

Для лиц, чья деятельность связана с принятием решений в области образования, руководителей учебных заведений и органов управления ими. Будет полезна специалистам, участвующим в программах оценки качества образования. Представляет интерес для широкого круга читателей, учителей, преподавателей и методистов учебных заведений всех видов и уровней.

УДК 378
ББК 74.58

ISBN 978-5-98704-575-6

© Международный банк
реконструкции и развития/
Всемирный банк, 2008, 2012
© Логос, 2014

1818 H Street, NW
Washington, DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org
E-mail: feedback@worldbank.org

Все права защищены
1 2 3 4 12 11 10 09

Эта книга подготовлена под эгидой Международного банка реконструкции и развития. Представленные в ней материалы не всегда отражают точку зрения исполнительных директоров Всемирного банка или правительств государств, которые они представляют.

Всемирный банк не гарантирует точности приведенных данных. Границы, цвета, географические наименования и иная информация, приведенная на картах, не отражает позицию Всемирного банка относительно правового статуса территорий.

Авторское право и разрешение на воспроизведение

Данная работа защищена законодательством об авторских правах. Всемирный банк приветствует распространение знаний, которыми он обладает, поэтому данная работа может быть воспроизведена, частично или полностью, для использования в некоммерческих целях при условии указания ссылки на источник.

Для получения разрешения на воспроизведение любой части данной публикации в коммерческих целях необходимо направить запрос с предоставлением полной информации в Центр по проверке авторских прав (The Copyright Clearance Center Inc.) по адресу: 222 Rosewood Drive, Danvers, MA 01923, USA; телефон: 978-750-8400; факс: 978-750-4470; Интернет: www.copyright.com.

Все остальные запросы, связанные с получением прав и лицензий, в том числе субсидиарных прав, следует направлять в Отдел публикаций Всемирного банка (The Office of the Publisher, The World Bank) по адресу: 1818 H Street NW, Washington, DC 20433, USA;

факс: 202-522-2422; адрес электронной почты: pubrights@worldbank.org.

Дизайн обложки: Naylor Design, Washington, DC
ISBN: 978-0-8213-7929-5
eISBN: 978-0-8213-7966-0
DOI: 10.1596/978-0-8213-7929-5

Выражение признательности

Научные консультанты, переводчики, редакторы Издательской группы «Логос» выражают признательность Московскому представительству Международного банка реконструкции и развития, особенно Исаку Фрумину и Тиграну Шмису, за помощь и поддержку в подготовке и выпуске этой книги и всей серии книг «Национальная оценка учебных достижений». Мы благодарим В.А. Болотова, вице-президента Российской академии образования, и В.А. Мау, ректора Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, которые направляли наши усилия на достижение целей, имеющих исключительное значение для реформирования российской системы образования. Высокой оценки заслуживает вклад Центра международного сотрудничества по развитию образования в лице научного руководителя А.Е. Волкова, который выступил инициатором перевода книг по оценке учебных достижений на русский язык.



СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ	15
ПРЕДИСЛОВИЕ	23
БЛАГОДАРНОСТИ	25
СПИСОК СОКРАЩЕНИЙ	27
ГЛАВА 1 ФАКТОРЫ, ВЛИЯЮЩИЕ НА ВОЗМОЖНОСТЬ ИСПОЛЬЗОВАНИЯ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ	29
Политический контекст национальной оценки	32
Подотчетность.....	34
Качество инструментария для оценки	48
Вид оценки	52
Неполное использование результатов национальной оценки	57
Выводы.....	63
ГЛАВА 2 СООБЩЕНИЕ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ: ОСНОВНОЙ ОТЧЕТ	69
Контекст национальной оценки.....	72
Цели национальной оценки	72

Структура для национальной оценки.....	73
Процедуры проведения национальной оценки	74
Описание достижений в национальной оценке.....	74
Корреляты достижений	80
Изменение достижений в течение длительного периода времени	88
Выводы.....	91

ГЛАВА 3 ОТЧЕТ О НАЦИОНАЛЬНОЙ ОЦЕНКЕ: ДРУГОЙ ИНСТРУМЕНТАРИЙ ДЛЯ СООБЩЕНИЯ

РЕЗУЛЬТАТОВ	93
Лист спецификации продуктов	94
Брифинг министров и старшего персонала в политике	95
Публикация итоговых отчетов	97
Публикация технических отчетов.....	102
Публикация тематических отчетов.....	106
Обеспечение сообщений для СМИ	108
Издание пресс-релизов.....	109
Проведение пресс-конференций.....	111
Проведение отдельных брифингов	115
Регистрация сообщений на сайте.....	116
Создание доступных данных оценки	119
Другой инструментарий для распространения данных	120
Выводы.....	121

ГЛАВА 4 ПЕРЕВОД РЕЗУЛЬТАТОВ ОЦЕНКИ В ПОЛИТИКУ И ДЕЙСТВИЯ

Сформированная способность к восприятию и использованию информации.....	125
Состоятельность и релевантность информации, обеспечиваемой оценкой	126
Состоятельность и релевантность информации, обеспечиваемой оценкой	127

Процедуры для идентификации соответствующей политики или действий по результатам оценки	128
Определение широкомасштабного системного или целевого вмешательства	133
Сложность формирования политики и принятия решения.....	134
Выводы.....	137
ГЛАВА 5 РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ, ПОЛИТИКА И УПРАВЛЕНИЕ В ОБРАЗОВАНИИ	139
Описание достижений	140
Описание ресурсов	141
Мониторинг достижений	144
Анализ системы образования.....	146
Выработка общей политики и помощь в принятии решений.....	146
Определение стандартов	152
Обеспечение ресурсов для школ	153
Поддержка пересмотра учебного плана	157
Пересмотр учебников.....	160
Выводы.....	161
ГЛАВА 6 РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ОБУЧЕНИЕ	163
Профессиональное развитие учителей.....	164
Основное внимание школам и обучению в классе.....	178
Выводы.....	196
ГЛАВА 7 РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ОСВЕДОМЛЕННОСТЬ ОБЩЕСТВЕННОСТИ.....	201
Примеры использования результатов оценки для сообщения общественности.....	204
Роль средств массовой информации.....	206

Развитие стратегии коммуникации.....	208
Выводы.....	211
ГЛАВА 8 ОПТИМИЗАЦИЯ ИСПОЛЬЗОВАНИЯ ПРОГРАММ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ИХ ЗНАЧЕНИЕ	215
Оптимизация использования результатов национальной оценки	217
Развитие национальной оценки для повышения его ценности	221
Выводы.....	230
СПИСОК ЛИТЕРАТУРЫ	231
ДОПОЛНИТЕЛЬНЫЕ МАТЕРИАЛЫ	
1.1. Проблемы, на которые национальная оценка может пролить свет.....	31
1.2. Оценка достижений по ограниченному кругу свидетельств.....	39
1.3. Отрицательный эффект использования результатов тестов для принятия ответственных решений	42
1.4. Установление подотчетности за достижения учащихся	46
1.5. Переложение вины за неудачи школ	47
1.6. Два примера использования данных международной оценки	57
1.7. Степень использования результатов национальной оценки: Гондурас	59
3.1. Итоговый отчет штата Иллинойс, уровни достижений учащихся 4-го класса по чтению в 2007 году в программе NAEP: Соединенные Штаты Америки.....	98
3.2. Предложения для применения подхода программы PISA к обучению и изучению математики: Ирландия.....	101

3.3. Выдержка из пресс-релиза, программа NAEP: США.....	112
3.4. Пример пресс-релиза: США	113
3.5. Сайт. Достижения по математике в начальных школах: Ирландия	117
3.6. Карта сайта, программа NAEP Web Site: США	118
5.1. Использование результатов оценки для продвижения реформ: Доминиканская Республика	149
5.2. Мифы об образовании в США	151
6.1. Задание на установление соответствия между числом и его словесным эквивалентом: Пакистан	170
6.2. Использование результатов национальной оценки для повышения уровня подготовки учителей: Минас-Жерайс, Бразилия	175
6.3. Основные особенности программы для работающих учителей, основанной на результатах национальной оценки: Уругвай.....	176
6.4. Отрывок постера, опубликованного после национальной оценки: Уганда	179
6.5. Задание по математике.....	180
6.6. Рекомендации по результатам национальной оценки по математике, 5-й класс: Ирландия	186
6.7. Темы для обсуждения: результаты национальной оценки и школьные переменные для анализа эффективности деятельности школ	188
6.8. Темы для обсуждения: результаты национальной оценки и переменные для анализа эффективности деятельности учителей.....	188
6.9. Темы для обсуждения: результаты национальной оценки и индивидуальные факторы учащегося и факторы семьи, влияющие на учебную деятельность учащегося	190

7.1. Законы о свободе информации	203
7.2. Соглашения по релизу результатов национальной оценки: Уругвай	204
7.3. Меры для придания результатам национальной оценки максимального эффекта воздействия на образование.....	208
7.4. Обложка брошюры для национальной оценки: Эфиопия.....	212

РИСУНКИ

2.1. Умения по чтению, стратегии и оценки пороговых баллов, полученные для эталонных точек в объединенной шкале грамотности чтения для 4-го класса, PIRLS, 2001	78
2.2. Учащиеся мужского и женского пола, достигшие уровня мастерства в национальной оценке, 4-й класс, предметы учебного плана: Шри-Ланка, %	82
2.3. Учащиеся, смотрящие телевизор различные промежутки времени, в сочетании с гендерными признаками и средней оценкой по чтению: Кувейт, %	83
2.4. Региональные различия в достижениях, 4-й класс: Эфиопия	85
2.5. Средние тестовые оценки по чтению учащихся по отношению к числу книг у них дома, 1-й и 5-й классы: Ирландия.....	87
2.6. Тенденции в шкале средних оценок по чтению для учащихся 9, 13 и 17 лет NAEP, 1971–2004 годы: Соединенные Штаты Америки	89
2.7. Оценки среднего процента правильных ответов по математике по содержательным разделам в национальной оценке, 5-й класс, 1999 и 2004 годы: Ирландия	90
2.8. Оценки среднего процента правильных ответов для умений по математике в национальной оценке 5-й класс, 1999 и 2004 годы: Ирландия.....	91

3.1. Средние баллы по чтению по расовым и этническим группам, NAEP, 4-й класс, 2005 год: США	107
6.1. Распределение оценок по чтению учащихся и учителей: Вьетнам.....	169
6.2. Программа национальной оценки по математике во Вьетнаме: корреляция между средними баллами учителей и учащихся в провинциях	173
6.3. Факторы, которые влияют на достижения: Уганда.....	197

ТАБЛИЦЫ

1.1. Причины неполного использования результатов национальной оценки, действия для коррекции неполного использования и структуры, ответственные за действия.....	60
2.1. Средние баллы (и стандартные ошибки) для мальчиков и девочек в национальной оценке по языку и математике.....	70
2.2. Средние баллы (и стандартные ошибки) и оценки различных процентильных рангов в национальной оценке по естественным наукам в провинциях	71
2.3. Средние баллы (и стандартные ошибки) в национальной оценке по предъявлению тестов в два момента времени.....	71
2.4. Корреляция между средними баллами учебных достижений по чтению в школе и школьными факторами в программе национальной оценки для 5-го класса	72
2.5. Учащиеся, чьи оценки расположены на минимальном и желательном уровнях мастерства в тестах по грамотности, действиям с числами и жизненным навыкам: Маврикий, %.....	75
2.6. Уровни достижений по математике в программе NAEP, 4-й класс: Соединенные Штаты Америки	76

2.7. Уровни умений по чтению в 5-м классе в национальной оценке: Вьетнам	79
2.8. Средние оценки по математике, 8-й класс, в соответствии с расовыми группами: Южная Африка	82
2.9. Средние оценки в национальной оценке по непальскому языку, 5-й класс, по регионам: Непал	84
3.1. Технический отчет: предлагаемое содержание	103
5.1. Школы, обладающие основными материально- техническими средствами: Кения, %	142
5.2. Школы, обладающие различным оборудованием, 1990–2002 годы: Малави, %	143
5.3. Примеры использования результатов национальной оценки в анализе систем образования различных стран	147
5.4. Ранги учащихся 4-го класса, оцененных на высоком уровне или выше по тестам штата и в национальной оценке, 2005 год: США	154
6.1. Число учащихся, учителя которых сообщили о готовности преподавать математику, данные TIMSS, 8-й класс, %	168
6.2. Учителя, оцененные на каждом уровне профессионализма по математике: Мозамбик, %	171
6.3. Доля учащихся, ответивших на задания теста по математике правильно: Британская Колумбия, Канада, %	181



ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Выход в свет в переводе на русский язык серии из пяти книг «Национальная оценка учебных достижений», подготовленной под эгидой Центра международного сотрудничества по развитию образования и Международного банка реконструкции и развития, представляет собой важную веху в реформировании системы образования Российской Федерации. Благодаря этому уникальному изданию учителя и преподаватели, методисты, руководители учебных заведений, специалисты органов управления ими, правительственные учреждения, да и все наше общество в целом получили надежный, объективный и, что весьма существенно, признанный в мире инструмент оценки результатов обучения всех категорий учащихся с первого класса общеобразовательных школ до выпускного курса вузов.

Уместно предположить, что издание серии книг «Национальная оценка учебных достижений» на русском языке будет полезно для педагогических кадров и сотрудников органов управления систем образования стран – участниц СНГ. Тем более что во многих странах Содружества решаются те же проблемы, что и в России, а в некоторых из них широко используется русский язык, в том числе в сфере образования.

Известно, что оценка учебных достижений – неотъемлемый элемент учебного процесса в любом учебном заведении. Она всегда служила стимулом учебного труда, показателем его результатов и регулятором взаимоотношений тех, кто учится, и тех, кто учит. И тем не менее с древности до наших дней оценка учебных достижений носит преимущественно субъективный и относительный характер.

Как правило, педагог ставит оценки на основе суммы своих впечатлений об успеваемости обучаемого и результатов выполнения им конкретного задания, определяя оценочный балл с учетом сложившегося мнения об успеваемости других учащихся. Поэтому шкала оценок всегда носила и продолжает носить расплывчатый и почти недифференцированный характер. У нас в России это школьные оценки «5», «4», «3», а также «двойка» и «единица», которые в действительности находятся уже за шкалой оценки успеваемости. Во многих странах используется 10-балльная шкала оценок, а баллы выставляются в обратной последовательности, но суть повсюду одна и та же.

До тех пор, пока учебные достижения обучаемых не имели большого общественного значения, а оставались внутренним делом академических коллективов, устоявшиеся подходы к оценке и учету успеваемости воспринимались как нечто само собой разумеющееся. Но в XX веке ситуация стала меняться. В условиях все более широкого использования достижений научно-технического прогресса резко возросли требования к компетентности большинства категорий занятых, да и всего населения обществ, насыщенных знаниями и техникой. Успеваемость учащихся превратилась во все более влиятельный фактор развития цивилизации, экономического роста и личного успеха представителей вступающих в жизнь поколений.

Разумеется, это оказало влияние и на сферу образования, где встали задачи повышения объективности оценки учебных достижений обучаемых и обеспечения их соответствия потребностям общества и экономики. В ответ на эти новые задачи

в ряде стран – в Соединенных Штатах Америки, Великобритании, Франции и других вводится в практику тестирование учащихся школ и поступающих в высшие учебные заведения. Для разработки тестов и оценки их выполнения широко используются методы математической статистики и теории вероятностей. Создается стройная научная теория тестов, а в 40-х годах XX века складывается современная теория тестов, позволяющая получить более надежные и содержательные в информационном отношении данные об учебных достижениях учащихся.

Этот опыт приобретает институциональный характер, воспринимается другими странами и международными организациями. К настоящему времени во многих государствах, в том числе и странах СНГ, создаются и совершенствуются национальные системы оценки учебных достижений обучаемых, которые используются для управления развитием образования в этих странах – от реформирования учебных планов и программ до повышения квалификации преподавателей и создания отдельным категориям учащихся, показавших пробелы в знаниях, более благоприятных условий для занятий.

В советской России в 20-е годы – первой половине 30-х годов прошлого века тестирование также получило распространение и применялось не только для оценки учебных достижений учащихся, но и для выявления наиболее способных представителей подрастающего поколения и развития их творческих качеств. Но в конце 1930-х годов работы в этом направлении были свернуты, а их научная база подверглась критике, прежде всего по идеологическим мотивам.

Более чем сдержанное отношение к строгим методам оценки учебных достижений не пошло на пользу советской системе образования. Медалисты-отличники – выпускники школ с самым различным уровнем преподавания – по результатам всего лишь одного экзамена принимались в лучшие высшие учебные заведения Москвы, Ленинграда, Новосибирска

и других признанных центров высшего образования. Принимались, а потом многие из них не могли сдать первую сессию. Невозможно было даже поставить рядом выпускников разных школ, особенно центральных и удаленных, либо молодых специалистов, получивших образование по одной и той же специальности в разных вузах. В силу этих (и многих иных!) причин Советский Союз стал все заметнее отставать в темпах научно-технического развития и в производительности труда от других стран северного полушария. Так одна, казалось бы, неприметная и неочевидная проблема образования приобрела общенациональное значение и привела к растрате значительных человеческих и материальных ресурсов.

Смена государственно-политического и социально-экономического устройства России в начале 90-х годов XX века создала возможности и обусловила потребности в реформировании сложившейся в прошлом системы образования. В ходе реформ развернулся поиск новых подходов к оценке учебных достижений учащихся и студентов. Под влиянием сотрудничества с учебными заведениями зарубежных стран стал осваиваться международный опыт в этой области. Российские школы на регулярной основе вошли в число участников различных программ международной оценки учебных достижений учащихся (PISA, PIRLS, TIMSS и др.).

Влиятельным фактором, способствующим более энергичному освоению вузами современных методов оценки учебных достижений студентов, явилось присоединение России к Болонскому процессу. Введение единого приложения к диплому о высшем образовании и основанной на общеевропейских принципах системы зачетных единиц, реализация компетентностного подхода к подготовке кадров, как и некоторые другие меры, реализуемые в странах – участницах Болонского процесса, в конечном счете опираются именно на достоверные и объективные методы оценки результатов обучения студентов.

В инициативном порядке и при поддержке органов управления образованием в стране стали создаваться научные учреждения, занявшиеся исследованиями проблематики качества обучения и подготовки кадров, педагогических измерений, психодиагностики и других аспектов определения и оценки учебных достижений. В частности, в системе Российской академии образования был организован Центр оценки качества образования, а в системе Министерства образования и науки – Федеральный институт педагогических измерений.

Но по-настоящему переломные последствия имел масштабный социально-экономический эксперимент по проведению единого государственного экзамена, объединившего выпускные экзамены в общеобразовательных учреждениях и вступительные экзамены в вузах. Он стимулировал дискуссию, охватившую не только педагогические круги, но и широкую общественность. Несмотря на выявленное в ходе этой дискуссии многообразие точек зрения, она подтвердила единство общественного мнения в том, что методы и результаты оценки учебных достижений должны носить точный, объективный и сопоставимый характер.

Создание в России системы национальной оценки учебных достижений – следующий этап реформирования образования в стране. Мы видим, что отправные предпосылки для этого уже заложены. Более того, в некоторых авторитетных вузах – Национальном исследовательском университете «Высшая школа экономики», Московской высшей школе социальных и экономических наук, Российском государственном педагогическом университете им. А.И. Герцена и других в магистратуре по профилю «Менеджмент образования» развернута подготовка кадров, способных выступать в качестве организаторов, ключевых специалистов и экспертов этой системы.

Незаменимым пособием для проведения всей этой работы призваны стать книги серии «Национальная оценка учебных достижений», которые будут особенно полезны для реализации

программ подготовки, переподготовки и повышения квалификации педагогических кадров и специалистов органов управления образованием по тематике оценки знаний и менеджмента образования. В книгах серии со всей тщательностью обобщен национальный и международный опыт в этой области. По своей тематике они образуют логическую последовательность, одна за другой раскрывая основные стадии строительства национальной системы оценки учебных достижений, освещая ее принципы, методы, процедуры. В них приводится множество конкретных примеров успешного функционирования этой системы, а также анализируются типичные трудности и ошибки. Еще одна привлекательная сторона этого издания состоит в том, что в нем предлагается не какой-либо единый шаблон, а представлено широкое многообразие моделей, которые могут использоваться в зависимости от конкретной ситуации в стране, объемов финансирования, выделяемых на национальную оценку, способов использования ее результатов.

Все книги, вошедшие в серию, написаны понятным, хорошо воспринимаемым языком. Переводчикам и научным редакторам удалось сохранить это ценное качество английского оригинала в русской версии издания, которая практически полностью идентична своему прототипу. Думаем, что читатели будут благодарны всем, кто участвовал в подготовке перевода на русский язык, за хорошую работу, выполненную не только со знанием дела, но и со вкусом, творчески.

Что вправе ожидать от создания в России, да и в любой другой стране, национальной системы оценки учебных достижений?

Во-первых, появления нового источника достоверной информации о результатах обучения в отдельных учебных заведениях и их типах, предназначенной как для лиц, принимающих решения в сфере образования, так и для заинтересованных кругов общественности.

Во-вторых, создания условий для более целенаправленного проведения реформ в сфере образования в интересах обеспечения его более высокого качества и доступности на основе использования данных об учебных достижениях учащихся для улучшения работы образовательных учреждений и повышения эффективности учебного процесса.

В-третьих, совершенствования системы и практики управления образованием на всех уровнях и во всех звеньях, приближения его результатов к потребностям общества, личности и государства, нашедшим отражение в образовательных стандартах.

Книги серии «Национальная оценка учебных достижений» начали свой путь к читателю. Хотелось бы пожелать, чтобы они были благосклонно приняты своей читательской аудиторией и повлияли на практическую деятельность систем образования России и стран СНГ, а главное – способствовали бы повышению уровня учебных достижений учащихся наших государств.

В.А. Болотов,
вице-президент Российской
академии образования

И.Д. Фрумин,
руководитель международных
образовательных программ
Международного банка
реконструкции и развития России



ПРЕДИСЛОВИЕ

В гораздо большей степени, чем число лет, потраченных на обучение, на рост производительности труда и доходов индивидуумов влияют результаты учебной деятельности или приобретенные познавательные умения. От них также зависит совокупный экономический доход государства. Например, в результате осуществления программ международной оценки по грамотности и компетентности по математике было доказано, что повышение оценок учащихся на одно стандартное отклонение и двухпроцентное увеличение ежегодных темпов роста ВВП на душу населения – два неразрывно связанных между собой факта.

Измерение результатов образовательной деятельности с каждым днем набирает обороты. Признается его необходимость не только для мониторинга успехов школьной системы, но и для повышения качества образования в целом. Информация об учебных достижениях может использоваться для формирования политики в образовании множеством разнообразных способов, включая проектирование и реализацию программ для совершенствования деятельности учащихся и преподавателей, выявления отстающих учащихся с тем, чтобы они могли получить необходимую поддержку, предоставление соответствующей

щей методической помощи; обучение преподавателей школ, учащиеся которых получили низкие баллы.

Использование результатов оценки для совершенствования учебной деятельности учащихся является важнейшей проблемой для развивающихся стран. Далекое не в каждой из них уровни учебной деятельности измеряются на регулярной основе, и даже там, где это все-таки имеет место, полученные данные не всегда позволяют в полной мере информировать политиков, ответственных за развитие образования. Тем не менее грамотная образовательная политика должна обладать потенциалом для обеспечения значимого прироста познавательных умений. Если развивающиеся страны действительно хотят достичь поставленной цели – «образования для всех», то они будут нуждаться в помощи как при измерении уровней учебной деятельности учащихся, так и в поиске способов перевода этой информации в сферу формирования эффективной образовательной политики и методов ее осуществления.

Книга, которую вы держите в руках, может оказать в этом отношении весьма ощутимую поддержку. Она вобрала в себя большую часть существующих на сегодняшний день исследований относительно актуальных мер по переводу данных, полученных в ходе национальной оценки учебных достижений, в полезную для практической работы информацию. Кроме того, она включает сведения о том, как найти компромисс между заинтересованными лицами в обсуждении этой информации способами, которые повысят качество образования и результаты учебной деятельности учащихся. Книга стремится максимизировать значение национальных данных об учебной деятельности учащихся и помочь странам в их стремлении использовать информацию, к которой приводят их программы оценки, полностью.

Элизабет Кинг,
директор сети по развитию образования,
Международный банк реконструкции и развития



БЛАГОДАРНОСТИ

Работу над серией книг под названием «Национальная оценка учебных достижений» осуществляла команда во главе с Винсентом Грини (консультант Сети развития человеческого потенциала, группа по образованию, Всемирный банк) и Томасом Келлаганом (Центр исследований в образовании, колледж Святого Патрика, Дублин). Также большой вклад внесли Сильвия Акана (Национальная экзаменационная комиссия, Уганда), Пру Андерсон (Австралийский совет по научным исследованиям в образовании), Фернандо Картрайт (Канадское национальное статистическое агентство), Жан Дюмей (Канадское национальное статистическое агентство), Крис Фримен (Австралийский совет по научным исследованиям в образовании), Хью Гоф (Канадское национальное статистическое агентство), Сара Хауи (Университет Претории), Джордж Морган (Австралийский совет по научным исследованиям в образовании), Т. Скотт Мюррей (Институт статистики, ЮНЕСКО) и Джерри Шил (Центр исследований в образовании, колледж Святого Патрика, Дублин).

Общее руководство осуществляли Руфь Кагиа, директор сектора образования, которая впоследствии передала свои полномочия Элизабет Кинг, и Робин Хорн, директор по вопросам образования, Всемирный банк. Инициатором проекта был

Роберт Праути, он же руководил им до августа 2007 года. На его более поздних этапах руководство перешло к Маргерит Кларк (подготовка обзоров и публикаций). Мы выражаем благодарность внешнему эксперту Олу Битону (Бостонский колледж). Большую пользу принесли также советы Патрисии Аррегуи, Луиса Беневисте, Маргарит Кларк, Шабаны Сосале и Эмилианы Вегас. Ценную поддержку мы получили от Дэвида Хардинга, Эйдана Малкина и Мириам Уэйзер. Особой благодарности заслуживают Хилари Уолш из Научно-исследовательского центра по образованию, которая напечатала рукопись, и Мэри Рохан, которая помогала этой работе.

Мы благодарим за разрешение использовать в издании свои материалы: Сильвию Акану, Мартина Экли, Абдулу Кхану аль-Базасу, Аише аль-Рудхан, Патрисию Аррегуи, Сервисный центр образования и развития (Катманду), Центр исследований в образовании (Дублин), Агентство по экзаменам и проверке качества образования Эфиопии, Экзаменационный совет Лесото, Люсьен Финетт, Зеуду Джеберкидану, Лауру Грегори, Синтию Гатмен, Сару Хауи, Международную ассоциацию по эвалюации образовательных достижений, Терт Кханию, Экзаменационный синдикат Маврикия, Клаудию МакЛошлен, Отдел образования штата Мичиган, Национальный центр статистики в образовании Департамента по образованию США, Национальный центр исследований и эвалюации в образовании (Коломбо, Шри-Ланка), Боба Праути, ЮНЕСКО (Общий отчет по мониторингу), Мацеко Рамокоена, Ребекку Росс, Морин Шейфер, Берта Стонеберга, Садию Тайаб и Ганса Уоджемейкера.

Работы по оформлению книг и производственный процесс были организованы Джанет Сэссер и Паоле Скалабрин, Издательский офис Международного банка реконструкции и развития. Серьезную поддержку оказали также Ирландский образовательный доверительный фонд, Банк Партнерских программ Нидерландов, Центр исследований в образовании Дублина и Австралийский совет по исследованиям в образовании.



СПИСОК СОКРАЩЕНИЙ

МБРР	Международный банк реконструкции и развития (Всемирный банк)
ООН	Организация Объединенных Наций
ОЭСР	Организация экономического сотрудничества и развития
ЮНЕСКО	Организация Объединенных Наций по вопросам образования, науки и культуры
PIRLS	Progress in International Reading Literacy Study: Международное исследование прогресса в грамотности чтения
PISA	Program for International Student Assessment: Программа международного оценивания учащихся
TIMSS	Trends in International Mathematics and Science Study: Международное исследование тенденций в математическом и естественнонаучном образовании

ФАКТОРЫ, ВЛИЯЮЩИЕ НА ВОЗМОЖНОСТЬ ИСПОЛЬЗОВАНИЯ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ

Как было изложено в первой книге данной серии, национальная оценка учебных достижений предназначена, главным образом, чтобы определить:

- каковы успехи учащихся в системе образования (по отношению к общим ожиданиям, целям учебного плана и подготовке к дальнейшему обучению и к жизни);
- есть ли свидетельства отдельных сильных и слабых сторон в знаниях и умениях учащихся;
- есть ли отдельные подгруппы среди участников оценки, которые демонстрируют низкие результаты;
- какие факторы связаны с достижениями учащихся;
- соблюдаются ли требования государственных стандартов;
- изменяются ли учебные достижения учащихся с течением времени (Greaney and Kellaghan, 2008).

Именно для получения ответов на эти вопросы с помощью специальных процедур, установленных в социальных науках, и собираются данные по учащимся и другим заинтересованным лицам в системе образования. Это необходимо для придания результатам управления и практики в образовании более прозрачной формы и в конечном счете для обеспечения предста-

вителей образовательных учреждений информацией с целью улучшения их практической деятельности (Fetter, 2006).

Достижение целей национальной оценки имеет немало важное значение для определения таких важных аспектов деятельности системы образования, как обеспечение доступности, качества, эффективности и активности (Braun and others, 2006) (см. дополнительный материал 1.1). Оценка будет более правдоподобной и адекватной при условии взаимосвязанности указанных аспектов. Во многих системах образования прослеживается следующая закономерность: в школах, учащиеся которых демонстрируют низкие результаты, обучается немало представителей малообеспеченных слоев населения и национальных меньшинств. Зачастую там наблюдается весьма низкий уровень ресурсного обеспечения (учебники присылают слишком поздно, если присылают вообще). Кроме того, нередко такие школы испытывают трудности в привлечении учителей – из-за своего отдаленного местоположения либо по этническим или языковым причинам. Ясно, что любая информация по данным проблемам, которую можно собрать в ходе национальной оценки, будет представлять интерес для широкого круга заинтересованных лиц: политических деятелей, управленцев сферы образования, педагогов, учителей, разработчиков учебного плана, родителей, учителей и широкой общественности.

В предыдущих книгах серии был всесторонне освещен процесс получения информации в ходе национальной оценки. В них рассматривались следующие вопросы: как проектируется инструментарий для сбора данных о достижениях учащихся и связанных с ними переменных; как формируется выборка учащихся, чтобы представить достижения системы образования в целом (или ясно определенной части, например учащихся 4-х классов или одиннадцатилетних обучаемых); какие процедуры должны осуществляться при сборе и обработке данных; какие методы могут использоваться для анализа данных. Книга, которую вы держите в руках, посвящена сообщению

Дополнительный материал 1.1**Проблемы, на которые национальная оценка может пролить свет**

Доступность. Препятствия для посещения школы, такие как ограниченная пригодность мест расположения школы или значительное расстояние от домов учащихся до школы.

Качество. Качество входных характеристик школы для получения результатов обучения, таких как ресурсы и средства обслуживания, доступных для поддержки учебного процесса (чувствительность учебных планов к потребностям учащихся, компетентность учителей, качество учебников); учебных методов; взаимодействия учащихся с учителями и учения.

Эффективность. Оптимальное использование человеческих и финансовых средств, отраженных во взаимодействии учащихся с учителями, способности учащихся к восприятию нового материала и норм повторного обучения в одном и том же классе.

Справедливость. Условия для реализации образовательных возможностей учащихся и достижения паритета в успехах учащихся независимо от их характеристик, таких как пол, язык или этническая принадлежность к определенным группам и географическое место проживания.

Источник: разработки авторов; основано на работе: Braun and others, 2006.

и использованию данных, полученных в ходе национальной оценки с конечной целью, направленной на улучшение качества учебной деятельности учащихся. Она предназначена для двух категорий начинающих читателей: тех, кто несет ответственность за подготовку отчетов о результатах оценки и за их сообщение и распространение, и тех, кто эти результаты использует.

В данной вводной главе рассмотрены пять тем. Во-первых, это политический контекст национальной оценки и его значение для использования его результатов. Во-вторых, тема подотчетности, которая является центральной для большинства органов власти и во многом определяющей при плани-

ровании мероприятий по проведению национальной оценки. В-третьих, качество инструментария, используемого в ходе национальной оценки для получения информации относительно результатов обучения (знания, умения, отношения и привычки, которые учащиеся приобрели в результате обучения в школе), что имеет немаловажное значение для использования данных оценки с целью улучшения учебного процесса. В-четвертых, как характеристики национальной оценки (основанной на генеральной совокупности обучаемых или их выборке) или международной оценки влияют на использование результатов. Наконец, в главе рассматриваются возможные причины затруднений при использовании результатов национальной оценки.

ПОЛИТИЧЕСКИЙ КОНТЕКСТ НАЦИОНАЛЬНОЙ ОЦЕНКИ

Хотя одна национальная оценка во многих отношениях может очень походить на другую, в действительности между ними всегда будут различия, которые повлияют на использование их результатов. Эти различия объясняются тем, что оценка является политическим явлением в той же степени, что и методическим. Она отражает потребности дня, напряженность отношений, установленные нормы и характер расстановки сил между представителями власти. Идентификация политического контекста, в котором проводилась оценка, может помочь в объяснении различий между странами в их стратегиях оценки (Benveniste, 2002). Даже внутри США системы подотчетности в разных штатах различаются, что отражает административные решения и традиции, которые развивались в течение долгого времени (Linn, 2005b).

Роль процесса оценки в осуществлении контроля сферы образования и руководства многогранна и многоаспектна. Во-первых, оценка проходит в политическом процессе, часто вдохновляемом и направляемом политическими побуждениями. Во-вторых, форма оценки, как правило, является

результатом конкурентной борьбы между представителями различных социальных групп за влияние на нормы и ценности, которым будет отдаваться предпочтение в государстве. В-третьих, оценка может затронуть социальные связи между, например, управленцами в образовании и учителями или между учителями и родителями. В-четвертых, контроль размещения и интерпретации результатов оценки демонстрирует способность влияния на политику, распределение ресурсов и общественное восприятие результатов образования. Наконец, оценка в скрытой или в явной форме может включать механизмы регулирования деятельности представителей различных социальных групп, ответственных за результаты образования (Benveniste, 2002).

Социальные группы, обладающие способностью влиять на характер процесса оценки и использование ее результатов, довольно многочисленны. То, как образовательная политика будет развиваться в стране, зависит от множества факторов, к числу которых можно отнести:

- степень, в которой решения относительно образовательных условий (например, учебных планов, финансирования учебного процесса) являются функциями централизованного или децентрализованного управления;
- существование и сфера влияния неофициальных учреждений, сетей и специальных групп как внутри, так и вне правительства;
- влияние союзов учителей, которые могут играть ключевую роль если не в формировании политики, то в ее реализации;
- роль внешних (многосторонних и двусторонних) агентств, зависящих от административного влияния в улучшении условий обучения в своей системе образования и в обеспечении или поддержке развития способности по взаимодействию с ними образовательных учреждений.

Значение политической позиции для национальной оценки может быть проиллюстрировано на двух примерах (Benveniste,

2002). В Чили акцент смещается на отчетность перед общественностью, способствуя таким образом развитию конкуренции на рынке, вызванной публикацией результатов оценки для отдельных школ. Противоположная позиция представлена в Уругвае, где государство принимает на себя ответственность за достижения учащихся и за обеспечение ресурсов, необходимых для поддержки процесса обучения учащихся, особенно тех, которые относятся к неимущим слоям населения.

Следующий аспект политического контекста, значимый для использования результатов оценки, – это степень «открытости» или «закрытости» системы образования. Некоторые системы были описаны как «взаимоисключающие». В таких системах доступ к некоторым важным сведениям, включая информацию о результатах исследования, политической элите или руководителях, принимающих решения, ограничен. Последним не разрешается участвовать в общественном распространении этих данных. В более открытых системах, напротив, предпринимаются все усилия для привлечения интереса СМИ, мобилизации политических сил и проведения дебатов по проблемам образования (Reimers, 2003). Возможна также «промежуточная» ситуация, предполагающая умеренное использование данных национальной оценки. В этом случае обращение к информации о системе образования, включающей данные о достижениях учащихся, является ограниченным, но не исключается полностью. Например, в Уругвае данные о достижениях учащихся предназначены прежде всего для использования внутри образовательного сообщества (Benveniste, 2002).

ПОДОТЧЕТНОСТЬ

В последние десятилетия в правительственных административных кругах многих стран набирает обороты движение за рост подотчетности. Это происходит в ответ на политическое,

социальное и экономическое давление. В данном разделе подотчетность рассматривается в контексте образования. Особое внимание уделяется тому, каким образом интерпретация понятий влияет на использование данных национальной оценки. При этом необходимо учитывать, что большая часть рассуждений строится исходя из опыта США и касается, в основном, подотчетности школ (McDonnell, 2005).

Развитие движения за рост подотчетности в общественном секторе (включая образование) объясняется множеством факторов, которые не являются взаимоисключающими. Они перечислены ниже.

- Необходимость управления ограниченными (а в некоторых случаях и сокращающимися) ресурсами и повышения результативности для данного количества входных данных.
- Использование планирования и идей менеджмента, которые заимствованы из мира бизнеса, особенно в той части, которая касается проверки качества, удовлетворенности потребителя и непрерывного совершенствования (это основные черты Движения нового государственного управления и корпоративного подхода к администрированию). Такой подход может, в свою очередь, включать: определение деятельности в терминах результатов, постановку целей деятельности, использование показателей деятельности для определения степени, в которой цели достигнуты, осуществление стратегического и текущего планирования и обоснование распределения ресурсов деятельности.
- Введение рыночных механизмов распределения и контроля, включая системы стимулирования и конкуренции, заключение контрактов, аудит и перевод властных отношений в механизмы самоконтроля, чтобы минимизировать потребности во внешнем наблюдении и заставить людей усвоить нормы, ценности и ожидания ключевых заинтересованных лиц и менталитет, требуемый для самоуправления.

- Движение к расширению практики, основанной на фактических данных. Такое движение требует получения данных для поддержки требований, которые индивидуумы или учреждения должны выполнять профессионально и эффективно, а также данных, на которых базируются решения относительно распределения ресурсов (см. Blalock, 1999; Clegg and Clarke, 2001; Davies, 1999; Normann and Brinek, 2007; Kellaghan and Madaus, 2000).

Национальная оценка хорошо согласуется со многими из этих факторов, обеспечивая относительно простую статистическую информацию (доказательства) о системе образования на своевременной основе. Кроме того, с ее помощью можно выявить подгруппы или единицы обследуемой совокупности, которые по своим достижениям отвечают определенным стандартам, и подгруппы, чьи достижения этим стандартам не соответствуют. Такая информация может быть полезна для планирования и управления, в частности для выбора действий, необходимых для улучшения качества или повышения эффективности образования. Также она может использоваться для поддержки социальных групп, неявно или явно ответственных за результаты образования – тем самым на них возлагается ответственность и отчетность за изменение или регулирование качества образования.

Центр внимания в этой подотчетности варьируется по странам, раскрываясь с различной скоростью и испытывая разную интенсивность воздействия (Normann and Brinek, 2007). Таким образом, неудивительно, что оперативные и стратегические цели многих систем национальной оценки, особенно в развивающихся странах, или пути согласования оценки с системой подотчетности, могут быть не вполне ясны. В ситуации, когда политика подотчетности не развита в достаточной степени, результаты национальной оценки вряд ли будут достаточно эффективны (Normann and Brinek, 2007).

Однако по крайней мере неявное признание подотчетности представляется совершенно необходимым в тех случаях, когда по результатам оценки должны быть сделаны выводы. В противном случае как могут быть получены решения относительно действий, которые необходимо предпринять после оценки, и относительно людей или учреждений, которые должны выполнять эти действия?

Установить подотчетность многих ключевых заинтересованных лиц, вовлеченных в такую сложную систему, как образование, задача не тривиальная. В ходе ее решения нужно рассмотреть шесть вопросов, ответы на которые могут помочь при разъяснении существующих проблем, особенно в тех случаях, когда результаты национальной оценки используются для обеспечения подотчетности школы и учителей.

Должна ли система подотчетности сосредоточиться на результатах?

Сместив внимание на образовательные результаты учащихся, в частности на процесс обучения, можно получить формальные доказательства того, что многие дети проводят в школе много времени, но при этом не приобретают никаких полезных знаний и умений. Необходимость получения гарантий того, что в результате образовательной практики дети фактически усваивают учебный материал, была выдвинута на первый план на Всемирной конференции по образованию для всех, проведенной в Джамтjene (Таиланд) в 1990 году (ЮНЕСКО, 1990) и повторно в Дакаре в рамках Дакарского плана действий (ЮНЕСКО, 2000).

Если использовать данные о результатах образования в качестве единственного основания для вынесения решения о подотчетности за них, то можно упустить из вида тот факт, что различные условия, в которых работают школы (школьные

здания, учебные планы, образовательные материалы, методы обучения, используемые учителями, и деятельность по подготовке детей к школе) также относятся к оценке качества. Даже если нет других оснований для их учета, эти факторы важны по той причине, что от них зависит обучение. «Не следует ожидать, что учащиеся достигнут должного уровня мастерства в результатах обучения до тех пор, пока содержание обучения и учебный процесс в классе не будут в состоянии хорошо подготовить их к этому» (Haertel and Herman, 2005: 21).

Должна ли система подотчетности сосредоточиться на познавательных результатах?

Большинство людей, вероятно, согласятся с тем, что школьное обучение преследует множество целей – как личных (познавательное, моральное и социальное развитие учащихся), так и социальных (развитие социальной сплоченности или национального достояния). Трудно поспорить и с тем, что познавательные результаты являются первостепенно важными. Кроме того, грамотность и начальные математические знания, измеряемые во всех системах национальной оценки, необходимы как база для последующего продвижения учащихся в образовании. Однако едва ли можно счесть удовлетворительным, если полная уверенность в соответствии принимаемых мер целям образования будет достигнута в условиях пренебрежения другими ценными результатами обучения в школе, связанными с отношением учащихся к обучению, сформированными у них ценностями, мотивацией, стремлениями, способностью к самопознанию, способностью работать в группах, умениями устной презентации и социализацией. Работодатели и экономисты обычно рассматривают многие из этих результатов (часто описываемые как «мягкие» навыки) как очень важные аргументы при устройстве на работу (Cheng and Yip, 2006).

Должна ли система подотчетности основываться на единственном измерении достижений учащихся?

В большинстве систем национальной оценки проводится единственное тестирование, хотя при оценке компетентности учащихся по различным предметам учебного плана (по математике, чтению или естественным наукам) для него может использоваться несколько вариантов – форм теста. Таким образом, возникает вопрос: даже если познавательные достижения учащихся принимаются как установленный критерий качества обучения, действительно ли разумно основывать оценку этого качества (и, возможно, присваивать определенный уровень подотчетности) на единственном измерении выполнения учащимися теста в одном или двух классах?

Ответ, казалось бы, очевиден – нет. Тест может обеспечить только ограниченное количество информации о достижениях учащихся (см. дополнительный материал 1.2). Точная картина обученности учащихся, вне зависимости от того, оценивается ли она на национальном уровне или на уровне отдельной школы, требует многочисленных измерений (Guilfoyle, 2006).

Если тест будет ограничен использованием заданий с множественным выбором, то, вероятно, возникнут дополнительные проблемы, потому что чрезвычайно трудно использовать эту форму заданий для измерения познавательных умений высокого уровня.

Дополнительный материал 1.2

Оценка достижений по ограниченному кругу свидетельств

«Любая система, которая основывает оценку результатов всей школы на одной средней тестовой оценке единственной группы учащихся одного класса, не может надеяться на точные оценки этой школы».

Источник: Guilfoyle, 2006: 13.

Должны ли санкции быть основаны на результатах тестирования в национальной оценке?

Ключевой вопрос, на который следует ответить при использовании результатов национальной оценки, – должны ли санкции быть связаны с результатами выполнения теста учащимися. В любом случае некоторая степень подотчетности, пусть она даже не будет признана открыто, будет сопровождать процесс оценки, но не обязательно следует, что будут применены санкции. Однако в некоторых системах национальной оценки санкции все же применяются, как правило, по отношению к школам, преподавателям и даже к учащимся. Это происходит, например, в ходе оценки национального учебного плана в Англии, которая первоначально была введена как инструмент для формирования подотчетности. То же самое можно сказать о нескольких системах оценки на государственном уровне в США. На основе результатов выполнения тестов учащимися решаются вопросы, связанные с наградами и наказаниями. Так, школы или учителя могут получить денежно-кредитные премии, учителя – уволены, а сами учащиеся – получить отказ в поощрении или выдаче сертификата об окончании учебного заведения.

Существует ряд аргументов в пользу ответственных управленческих решений по результатам выполнения теста учащимися. Во-первых, этот подход поощряет индивидуумов (в частности, преподавателей) усваивать нормы, ценности и ожидания, которые возлагаются на результаты образования заинтересованными лицами (в частности, представителями министерства образования), и нести ответственность за их соблюдение. Во-вторых, он поддерживает действие рыночных механизмов в системе образования, привлекая элементы конкуренции, договорных отношений и аудита. В-третьих, он

способствует сосредоточению усилий учителей и учащихся на целях обучения и выполнения стандартов ожидаемых достижений, к которым учащиеся и преподаватели должны стремиться, создавая таким образом систему обучения, управляемую на основе измерений. В этой ситуации разумно ожидать, что результаты выполнения тестов улучшатся, если обучение будет вестись в тесной связи с инструментарием для оценки. Рост результатов, однако, может не быть очевидным, если достижения учащихся будут оцениваться с помощью другого инструментария. Так, например, в соответствии с программой NAEP в США, которая предоставляет результаты тестирования для принятия ответственных решений, достижения учащихся с течением времени растут. Сравнение же их с результатами оценки в государствах, где тестирование для принятия ответственных управленческих решений не используется, не всегда будет корректным (Amrein and Berliner, 2002; Braun, 2004).

Аргументы против того, чтобы результаты выполнения теста стали основой для принятия ответственных решений относительно учащихся, базируются больше на наблюдении и исследовании результатов государственных экзаменов, нежели на данных программ национальной оценки за длительный период времени (Kellaghan and Greaney, 1992; Madaus and Kellaghan, 1992; Madaus, Russell and Higgins, 2009). Подобные заключения касаются эффектов тестирования, связанного с законом «Ни одного отстающего ребенка в соответствии с законами США» (оригинальное название – No Child Left Behind Act (NCLB)). Закон был подписан 8 января 2002 года президентом Джорджем Бушем – младшим. Он представляет собой результат реформ в образовании и плана президента и отражает наиболее радикальные изменения в начальном и среднем образовании. См. подробнее: The Elementary and Secondary Education Act)

(Guilfoyle, 2006). Имеющиеся доказательства указывают, что когда санкции связываются с результатами выполнения тестов учащимися, то наблюдаются следующие отрицательные последствия.

Преподаватели будут уделять внимание тем знаниям и умениям, которые будут оцениваться в тесте («обучение под тест»), пренебрегая при этом предметами учебного плана (такими как искусство, общественные науки, физкультура), которые с помощью теста не оцениваются.

Обучение будет иметь преимущественную ориентацию на механическое запоминание, выполнение типовых заданий и накопление фактологических знаний, что приведет к пассивному подходу к усвоению учебного материала и исключит акценты на обучение общим навыкам рассуждения на более высоком уровне и умение решать проблемы.

Учителя, вероятно, затратят значительное время на обучение учащихся стратегиям выполнения тестов (в частности, как отвечать на задания с множественным выбором) и могут даже использовать форму с множественным выбором в обучении (см. дополнительный материал 1.3).

Дополнительный материал 1.3

Отрицательный эффект использования результатов тестов для принятия ответственных решений

«Системы оценки, которые являются полезными источниками информации, теряют большую часть своей значимости и доверия при использовании их результатов в случаях, когда на них возлагается высокая ответственность. Непреднамеренные отрицательные эффекты при использовании результатов оценки для ответственных управленческих решений часто перевешивают ожидаемые положительные эффекты».

Источник: Linn, 2000: 14.

Должны ли публиковаться ранговые таблицы по результатам национальной оценки?

Специфический пример использования результатов национальной оценки в целях принятия особо ответственных решений – публикация результатов в форме ранговых таблиц (рейтингов), в которых школы ранжируются в порядке результатов выполнения тестов учащимися. Есть мнение, что данный подход вызовет соревнование среди школ и в свою очередь повысит достижения учащихся (Reimers, 2003). Информацию о результатах можно сообщить родителям и обществу, а родители в некоторых ситуациях могут использовать ее в процессе выбора школы для своих детей. Даже когда такая проблема выбора не стоит вовсе или когда родители не используют результаты оценки учебных достижений в процессе выбора школы (Vegas and Petrow, 2008), простая публикация информации о работе школ может оказать воздействие на них с целью улучшения их работы.

В дополнение к неблагоприятным факторам, влияющим на обучение и усвоение учащимися материала, которые уже были приведены выше, можно перечислить и ряд других проблем, возникающих, когда результаты просчитаны и опубликованы для отдельных школ (Clotfelter and Ladd, 1996; Kane and Staiger, 2002; Kellaghan and Greaney, 2001; Linn, 2000).

Во-первых, работа школ (и, таким образом, их положение в ранговой таблице) может меняться в зависимости от итоговых результатов, которые оцениваются (например, достижения по чтению или математике).

Во-вторых, даже ранжирование, которое основано на одной и той же мере, может меняться в зависимости от используемого критерия успешности (например, средний балл или доля учащихся, которые получили высокие оценки).

В-третьих, недостаточная точность в процедурах оценки означает, что обязательно будут наблюдаться маленькие разли-

чия в результатах между школами (которые способны оказать большое влияние на их ранг).

В-четвертых, оценки достижений из года в год могут меняться из-за факторов, которые находятся вне контроля школы (например, различия в выборках (учащихся)). В связи с этой проблемой особенно уязвимы маленькие школы.

В-пятых, достижения учащихся в школе отражают не только результаты усилий учителей. Как показывают факты, ранжирование школ, основанное на достижениях и социально-экономических данных, почти идентично (Vegas and Petrow, 2008).

В-шестых, чтобы принять во внимание факторы, которые находятся вне контроля школы (например, способности учащихся, домашние условия), в качестве показателя работы школы может быть использовано среднее значение прироста тестовых оценок учащихся в течение года. Однако эта мера, как правило, свидетельствует об очень небольшом различии между школами, и поэтому она признана неудовлетворительной. Кроме того, не принимается во внимание тот факт, что скорость прироста достижений учащихся связана с их начальными достижениями. Более сложные статистические подходы, которые учитывают диапазон факторов, не находящихся под контролем школ, могут использоваться при вычислении прироста школьных оценок (модели с добавленной величиной). Проблемы, которые возникают в рамках этих подходов, связаны со сложностью процедур предъявления тестов, необходимых для сбора данных, уровнем статистической экспертизы, требуемой для анализа, трудностями в выборе переменных, которые будут включены в статистические модели, и тем фактом, что учет начальных достижений может привести к более низким ожиданиям относительно результатов наиболее слабо подготовленных учащихся.

Наконец, ранговые таблицы допускают привлечение коррумпированных методов, таких как исключение слабо подготовленных учащихся из процедуры оценки или упор на результаты сильных учащихся, чтобы повысить среднюю оценку школы.

Может быть предоставлена ложная информация относительно условий в школах (как это случилось в Чили), чтобы манипулировать социально-экономической классификацией школы, если отнесение школ к низкой категории может привлечь дополнительные вложения средств.

Что должно быть оценено при анализе подотчетности?

Главный аргумент против привлечения результатов оценки для принятия ответственных решений сводится к тому, что сами по себе оценки не выявляют те аспекты достижений, которые касаются работы школ или учителей. Факторов, которые способны повлиять на итоги тестирования, огромное множество. Даже если мы приведем не все, а только некоторые из них, это все равно заставит задуматься и взять паузу перед определением подотчетности. К таким факторам относятся: характеристики учащихся, включая их более ранние достижения; условия, в которых живут учащиеся, включая семейные и общественные ресурсы, поддержку родственников; образовательная политика, учебные планы и подготовка учителей, которая обеспечивается соответствующими органами государственной власти; школьные условия и ресурсы, в том числе особенности управления и компетентность педагогического состава (Kellaghan and Greaney, 2001).

Кажется вполне естественно ожидать, что личности или учреждения, которые так или иначе могут повлиять на эти факторы, должны считаться подотчетными и отвечать только за те действия, которые они могут контролировать. Таким образом, подотчетность за результаты обучения разделена между: учащимися; учителями; школами; политиками, администраторами и управленцами школьной системы (на национальном, государственном, региональном или муниципальном уровнях, в зависимости от того, как система образования организована); сотрудниками службы поддержки (разработчиками учебных

планов, педагогами, обучающими учителей, и издателями учебной литературы); родителями; другими лицами (включая политических деятелей, налогоплательщиков и общество в целом). По факту распределить подотчетность среди этого множества заинтересованных лиц чрезвычайно трудно (см. дополнительный материал 1.4). Однако отказ от признания этой проблемы может привести к неправильному приписыванию подотчетности, что в свою очередь повлечет неадекватные действия (см. дополнительный материал 1.5).

Дополнительный материал 1.4

Установление подотчетности за достижения учащихся

Установить подотчетность за достижения учащихся достаточно трудно.

- Хотя *учителя* до некоторой степени подотчетны за достижения своих учащихся, означает ли это, что они должны целиком нести ответственность, если школа, в которой они преподают, испытывает недостаток в основных средствах обучения? Если они были неадекватно подготовлены? Если учащиеся отсутствуют на занятиях в течение долгих периодов? Если учебный план, по которому они обязаны преподавать, является неподходящим для их учащихся?
- Можно ли предполагать, что *учащиеся* являются полностью подотчетными за результаты обучения, если дома не ценят их образование или не обеспечивают средства или возможности для того, чтобы учиться? Если учебный план является неподходящим? Если их ранний образовательный опыт был испорчен некомпетентным обучением?
- Можно ли возлагать полную подотчетность на *родителей*, если они не посылают своих детей в школу регулярно, потому что не могут позволить себе соответствующие затраты или потому что финансовое положение семьи вынуждает детей работать?
- Допустимо ли считать *управленцев в образовании* полностью подотчетными за результаты, если они не обеспечивают фонды, способные удовлетворить потребности школ?
- Должны ли *политические деятели* считаться полностью подотчетными за результаты, если средства, которые правительство в состоянии предоставить, неадекватны потребностям системы образования?

Источник: разработки авторов.

Многие системы национальной оценки, по крайней мере по умолчанию, признают влияние внешкольных факторов на достижения обучаемых. Даже в ходе масштабных оценочных процедур для принятия наиболее ответственных решений данные приводятся не по школам, а в зависимости от социально-экономического статуса учащихся, которые в них учатся. Анализ результатов позволяет также учесть такие характеристики, как предшествующие достижения учащихся или социально-экономический статус их семей. Кроме того, даже если оценка проводится с целью принятия наиболее ответственных решений, школам, испытывающим трудности, обычно выделяют дополнительные средства. Это своеобразное признание того непреложного факта, что учителям, учащиеся которых продемонстрировали невысокие результаты, возможно, необходима дополнительная помощь и поддержка в профессиональном развитии (Linn, 2000).

Анализ деятельности ключевых лиц, которые могут повлиять на результаты обучения, позволяет сделать вывод, что вопрос о подотчетности является весьма сложным и что отвечать на него, основываясь на ограниченной статистике, которую обеспечивает национальная оценка, недопустимо. В случае возложения подотчетности на учителей ее оценка подотчетности требует развернутых суждений. При этом необходимо принимать во внимание широкий диапазон факторов,

Дополнительный материал 1.5

Переложение вины за неудачи школ

«Проверка подотчетности может мягко и незаметно переложить вину за неудачу школы на неадекватные школьные ресурсы, слабую подготовку учителя или внешкольные факторы, а также на учителей и учащихся, посчитав, что они просто недостаточно упорно трудятся, и таким образом отвлечь внимание от более дорогостоящих и необходимых реформ».

Источник: Haertel and Herman, 2005: 3.

и в первую очередь условия, в которых они преподают. Желательно, чтобы суждения выносили профессионалы (директор, инспектор или наблюдатель). Вопросы об подотчетности других ключевых лиц являются в равной степени сложными. Так, в любой демократической системе политические деятели подотчетны электорату (что выражается в процедуре всенародных выборов). Но отсюда вовсе не следует, что граждане, делая выбор в пользу того или иного кандидата, голосуют за образование и за достижения учащихся (последнее и вовсе маловероятно). Управленцы в образовании подотчетны перед своими руководителями и политическими деятелями за выполнение своих обязанностей, но и в данном случае совсем не очевидно, что в сферу этой подотчетности должны входить достижения учащихся. Дистанцированность управленцев от практической деятельности школы, в отличие от учителей, вероятно, должна гарантировать, что на оценку их деятельности достижения учащихся не оказывают никакого влияния.

Большая ясность и прозрачность по вопросам подотчетности и обоснованности решений отдельных личностей и учреждений, которые вносят свой вклад в систему образования (включая усвоение учебного материала учащимся), должны устранить ряд двусмысленностей, которые существуют в функционирующих в настоящее время системах подотчетности. Кроме того, использование системы подотчетности, которая включает всех индивидуумов, учреждения и агентства, осуществляющие контроль за ресурсами и деятельностью школ, должно способствовать сплочению всех ответственных участников системы образования (Clegg and Clarke, 2001).

КАЧЕСТВО ИНСТРУМЕНТАРИЯ ДЛЯ ПРОВЕДЕНИЯ ОЦЕНКИ

Термин «качество» относится к практике образования, которая включает в том числе исследование окружающей среды с точки зрения того, насколько она безопасна и адекватна

по своим ресурсам, и учебных планов, которые должны отвечать потребностям учащихся, методов обучения, компетентности учителей, участвующих в активном обучении, и усвоения учащимися учебного материала (см., например, Schubert, 2005; ЮНЕСКО, 2000; ЮНИСЕФ, 2000). Однако в национальных исследованиях по оценке, как мы видели, основное внимание при рассмотрении качества сосредоточено на познавательных результатах образовательного процесса, т.е. на том, какой материал освоили учащиеся, с позиций развития стратегий для улучшения результатов обучения. Этот акцент сделан в соответствии с шестой целью, провозглашенной в Дакарском плане действий, которая выдвигает на первый план улучшение качества образования «так, чтобы признанные и измеримые результаты обучения были достигнуты всеми, особенно в области грамотности, действий с числами и основных жизненных навыков» (ЮНЕСКО, 2000: 8).

С учетом центральной роли, которая отводится результатам обучения в национальной оценке, в этом разделе рассмотрены четыре условия, которые нужно соблюдать для получения гарантий того, что используемый тест направлен на проверку именно тех достижений, которые школы стремятся развивать у учащихся, и что полученная информация удовлетворяет потребностям пользователей (Beaton and Johnson, 1992).

Во-первых, поскольку тест может измерить только часть знаний и умений, определенных в учебном плане или с помощью конструктора (например, чтение), то важно точно знать, что он обеспечивает адекватное отображение наиболее значимых знаний и умений (см. Haertel and Herman, 2005; Linn and Baker, 1996; Messick, 1989). Кроме того, тестовые задания должны отображать самые важные аспекты учебного плана, познавательную сложность, лингвистическую уместность и значимость для учащихся. Следовательно,

если целью системы образования является развитие познавательных умений и навыков высокого уровня (способности рассуждать, идентифицировать и решать проблемы и творческие задачи), тест нельзя ограничивать измерением изолированных компонентов умений или заданиями, которые требуют от учащихся только воспроизведения фактов или информации (особенность многих программ национальной оценки). Разработчики тестов должны иметь в виду, что желательно конструировать инструментарий, который обеспечит информацию для выбора политики и решений, приводящих с определенной долей вероятности к изменениям в учебных планах и системе обучения, что в свою очередь будет способствовать развитию оцениваемых знаний и умений (см. Frederiksen and Collins, 1989).

Чтобы обеспечить верное представление содержательной области или конструкта, целей тестирования или подобластей содержания (например, содержательных разделов или умений в математике), тест должен включать адекватное число заданий. Если заданий будет мало, это может вызвать вопросы относительно адекватности теста в содержательных аспектах. Например, число заданий тестов в программах оценивания в Латинской Америке (20–40 заданий во всех странах, кроме Бразилии) свидетельствует о том, что охват содержания был недостаточным. Кроме того, трудно отстоять точку зрения, согласно которой мастерство в достижении определенной цели можно определить только с помощью трех или четырех заданий (González, 2002). Подобные случаи характерны не только для программ национальной оценки в Латинской Америке.

Во-вторых, тест должен оценивать знания и умения на уровне, подходящем для учащихся, которые его выполняют. Проблема возникнет в том случае, если тест основывается исключительно на учебном плане, содержащем нереалистичные ожидания относительно достижений

учащихся. Подобная ситуация является вполне типичной для развивающихся стран. Тест окажется слишком трудным для учащихся с низким уровнем достижений и не сможет отразить их результаты обучения. Для решения этой проблемы при разработке теста необходимо принять во внимание не только стандарты намеченного учебного плана, но также и ту информацию, которая известна о фактических достижениях учащихся в школах. На практике очень небольшая доля учащихся выполняет все задания абсолютно правильно или, напротив, неверно. Лучшие результаты при выборе адекватной трудности достигаются благодаря привлечению к разработке и отбору заданий теста учителей-практиков и путем проведения перед основной процедурой оценки тщательной апробации заданий на выборке учащихся, которая охватывает все вариации достижений учащихся в генеральной совокупности школ.

В-третьих, если необходимо убедиться в том, что тест обеспечивает валидную информацию о знаниях и умениях учащихся в определенной области учебного плана, то необходимо, чтобы результаты не определялись их компетентностью в иных содержательных областях на планируемых к включению в тест (Messick, 1989). Например, тест, разработанный для оценки достижений в естественных науках или математике, не должен содержать большое количество текстового материала. Ведь в таком случае результаты будут зависеть от различий в способностях учащихся в чтении, а не в естественных науках или математике. Такая проблема возникает тогда, когда выполняющие тест учащиеся не обладают одинаковым уровнем умений в чтении, скажем, в том случае, если для некоторых учащихся язык теста не является родным, используемым в общении.

Наконец, если результаты оценки предназначаются для использования в мониторинге изменений достижений в течение длительного времени, то оценочный инструментарий

должен обеспечивать сопоставимость результатов. Чтобы достичь этого, можно использовать один и тот же тест, который необходимо хранить в условиях информационной безопасности в периоды между предъявлениями его учащимся. Если используются различные тесты, то шкалирование результатов с помощью современной теории тестов позволяет представить их на одних и тех же шкалах мастерства (см. четвертую книгу данной серии).

Передовой опыт основывается на переносе субтеста заданий из теста в тест, чтобы обеспечить эффективное средство для связи результатов разных тестов. Также важно, чтобы выборки учащихся и процедуры, которые соблюдаются при предъявлении теста, были эквивалентными. Если критерии исключения учащихся из выборки (например, учащихся с трудностями в усвоении материала) изменяются от одного случая проведения оценки к другому или если условия, не контролируемые при предъявлении теста (например, скорость ответов учащихся), различаются, то такие факторы обязательно должны быть приняты во внимание, когда проводится сравнение между достижениями учащихся в различные моменты времени.

ВИД ОЦЕНКИ

Потенциальные возможности использования информации, полученной по результатам оценки, зависят от характеристик оценок. Использование результатов изменяется в зависимости от вида оценки, который может быть основан на оценке генеральной совокупности, в которой участвуют все (или большинство) школ и учащихся целевой группы (как, например, в Бразилии, Чили и Англии); оценке выборки, в которую включают совокупность учащихся или школ, ото-

бренных для обеспечения репрезентативности всей группы (как и делают на практике в большинстве стран); международной оценке, в ходе которого ряд стран следуют подобным процедурам для получения информации об усвоении учащимися учебного материала.

Оценка, основанная на генеральной совокупности

Национальная оценка, в которой участвуют все (или почти все) школы и учащиеся, обычно определенного класса или возраста участников, называют основанными на генеральной совокупности. Эта оценка имеет большой потенциал с точки зрения предоставления информации о достижениях учащихся для: системы образования в целом, сектора системы, школ, учителей или классов, отдельных учащихся и по факторам, связанным с достижениями. Поскольку в этом случае доступна информация обо всех школах, то могут быть выявлены школы с плохими результатами выполнения тестов и могут быть приняты решения о возможных мерах вмешательства, таких как предоставление возможности для обучения на курсах повышения квалификации учителей, дополнительные услуги или дополнительные ресурсы. Оценка будет иметь высокий статус, влечь за собой подотчетность, если по ее результатам будут применены санкции к работе школ или если будет опубликована информация о работе отдельных школ.

Оценка, основанная на выборке

Поскольку вся целевая группа не принимает участия в оценке, охватывающей выборочную совокупность учащихся, она может обеспечить информацию относительно достижений учащихся только для системы образования в целом, одного сектора системы и по конкретным факторам, связанным с до-

стижениями. Хотя подобная ориентация на выборку ограничивает использование результатов оценки, у такого подхода есть много преимуществ. Во-первых, оценка, охватывающая выборочную совокупность учащихся, требует значительно меньше расходов при предъявлении тестов по сравнению с оценкой всех учащихся. Во-вторых, не нужно оценивать всех учащихся для достижения основной цели национальной оценки, которая должна обеспечить валидную, надежную и своевременную информацию о деятельности системы образования, и в частности о качестве усвоения учебного материала учащимися. В-третьих, поскольку участие школ не является обязательным, оценка, основанная на выборке, не оказывает негативного влияния на школы. Исследование результатов выборочной оценки не влияет на принятие санкций по отношению к школам или учителям или по отношению и к тем и другим. Наконец, выборочную оценку можно проводить чаще, что позволяет последовательно выявлять появляющиеся в образовании проблемы и думать над их решением. Некоторые программы национальной оценки постоянно проводятся на смещающихся выборках учащихся, тем самым предоставляя педагогам доступ к данным оценки на непрерывной основе.

Международная оценка

Другое отличие, которое относится к рассматриваемым вопросам использования данных оценки, состоит в том, является ли оценка автономной процедурой или выполняется в контексте международных исследований. Международные исследования имеют обязательства по предоставлению информации, которая в национальной оценке недоступна. Они помогают в определении того, что достижимо (какой

материал могут освоить учащиеся и в каком возрасте), в наблюдении деятельности определенных областей систем образования; они позволяют исследователям увидеть и охарактеризовать последствия применения различных методов и политики в образовании; они проливают свет на некоторые концепции, чтобы понять, какие упущения были в системе образования страны; они помогают установить спорные вопросы и предположения, которые могут считаться само собой разумеющимся (Chabbott and Elliott, 2003). Кроме того, международные исследования, как правило, достигают более высоких методических стандартов, чем программы национальной оценки. Также они позволяют участникам разделить затраты на разработку и проведение оценки, которые могли бы в противном случае сделать эти методы недостижимыми для многих систем образования. Результаты программ международной оценки обычно привлекают значительное внимание СМИ и становятся материалом для дебатов об адекватности образовательных условий и достижений учащихся и для предложений по изменениям в учебных планах, особенно в математике и естественных науках (Robitaille, Beaton and Plomp, 2000).

Программы международной оценки на самом высоком уровне могут обеспечить те сравнительные данные, которые национальная оценка никогда не предоставит. Однако когда дело доходит до использования результатов международной оценки во внутренней образовательной политике, необходимо принять во внимание одно обстоятельство. Одна из потенциальных опасностей в использовании международных данных связана с необходимостью предъявления тестов в нескольких странах, поэтому их содержание, возможно, не будет адекватно представлять учебный план каждой отдельно взятой страны. Также общепризнано, что в ходе между-

народных исследований не уделяется достаточно внимания контекстам, внутри которых функционируют системы образования. Действительно, маловероятно, что технология, которую используют международные исследования, может представить специфику отдельных систем образования или обеспечить фундаментальное понимание процесса усвоения учебного материала и то, как на процесс усвоения влияют местные культурные и контекстные факторы (Porter and Gamoran, 2002; Watson, 1999). Если это действительно так, то нельзя утверждать, что подходы, принятые в международных исследованиях, эффективные в одних системах образования, будут столь же эффективны и в других. Предположение об одинаковой эффективности может не только привести к неграмотному принятию решений в политике и их проведению в жизнь, но и нанести фактический вред (Robertson, 2005).

Данные, полученные в международной оценке, могут использоваться отдельной страной не только для сравнения ее системы образования с системами образования иных стран, но и для всестороннего изучения своей собственной системы (основанном на исследованиях внутри страны) в той части, что в действительности относится к национальной оценке (Министерство образования Кувейта, 2008; Postlethwaite and Kellaghan, 2008; см. дополнительный материал 1.6). Действительно, одна из целей Международной ассоциации эвалюации образовательных достижений в 1990–1991 годах в программе по исследованию грамотности чтения, состояла в том, чтобы обеспечить основные данные на национальном уровне по грамотности чтения у учащихся 9 и 14 лет для мониторинга изменений в течение долгого времени (Elley, 1992).

Дополнительный материал 1.6**Два примера использования данных международной оценки**

Различие между использованием данных, собранных в результате международной оценки для кросс-национальных сравнений, и данных, собранных для национального анализа, может быть выявлено путем сравнения примеров использования данных, собранных Международной ассоциацией по эвалюации образовательных достижений при исследованиях в Венгрии и Финляндии. В Венгрии основное внимание уделялось сравнению между системой образования в стране и системами образования в других странах. Сравнение показало, что достижения в математике и естественных науках удовлетворительные; достижения в понимании прочитанного в целом ниже, чем в других странах; наблюдается большой разброс в результатах оценки достижений по школам по сравнению с другими странами; семейные фоновые характеристики оказывают большее влияние на достижения в чтении учащихся в Венгрии (Báthory, 1992). В Финляндии, напротив, данные международной оценки, полученные в исследованиях Международной ассоциации по эвалюации образовательных достижений, использовались для анализа качества математического и естественнонаучного образования внутри страны. Данные, которые предоставляли только эмпирические доказательства, доступные в это время о достижениях учащихся, использовались в анализе, который был направлен на специфические потребности, и предоставлялись специалистам по планированию, политикам и специальным национальным рабочим группам.

Источник: Howie and Plomp, 2005; Leimu, 1992.

НЕПОЛНОЕ ИСПОЛЬЗОВАНИЕ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ

При рассмотрении использования (и отказа от использования) результатов национальной оценки вначале следует признать, что информации по этой теме очень мало. Еще меньше известно об оптимальном использовании результатов и об эффектах политических решений, принятых на их основе. Возможно, этот недостаток данных не является истинным отражением

фактического использования результатов оценки, потому что информация, связанная с использованием результатов правительственными органами, не может быть представлена публично.

Доступные доказательства указывают, что использование результатов национальной оценки распространено не очень широко, несмотря на потенциал для развертывания реформ, который имеет информация, полученная из программ оценки, и на расходы, необходимые для ее получения. Это наблюдение было сделано, например, в контексте определения специфической политики и решений (см. Arregui and McLauchlan, 2005; Himmel, 1996; Olivares, 1996; Rojas and Esquivel, 1998). Оно предполагает, что использование данных национальной оценки во многом схоже с использованием результатов другого исследования, связанного с политикой (см. главу 4). Описание опыта Гондураса, вероятно, типично и совпадает с опытом, полученным в другой стране (дополнительный материал 1.7). Однако, хотя установление специфического использования данных оценки, скорее всего, в этом случае невозможно, тот факт, что данные влияли на общественное мнение и повышали его осведомленность, сам по себе является значимым.

Можно привести множество причин неполного использования результатов национальной оценки (табл. 1.1).

Во-первых, результаты, вероятно, будут использованы не в полной мере в том случае, если национальная оценка рассматривается как автономная деятельность, слабо связанная с другой образовательной деятельностью, либо не связанная с ней вовсе. Эта ситуация чаще всего возникает в тех случаях, когда национальная оценка осуществляется впервые, или когда она выполняется внешними агентами, или по требованию спонсоров. Так, Rust (1999) указала, что в районе Сахары политические документы часто воспринимаются местными бюрократами как принадлежащие агентству спонсоров и не имеющие отношения к местной политике.

Дополнительный материал 1.7**Степень использования результатов национальной оценки:
Гондурас**

Результатом выполнения этих проектов было большое количество информации, описывающей образовательные достижения учащихся. Полученные данные являются результатами исследований, которые оказывают значительное влияние на общественное мнение, но не сделали образовательную систему более эффективной и эффективной.

Результаты исследований используются весьма слабо, они не закладывают базу для реального механизма контроля и установления подотчетности, и вплоть до настоящего времени полученные сведения не имели никаких значимых последствий, кроме выполнения чисто информационных функций, более глубокого восприятия проблем образования в общественном сознании.

Источник: Moncada and others, 2003: 73, цит. по: Arregui and McLauchlan, 2005: 6.

Во-вторых, неполное использование результатов национальной оценки характерно для ситуации, когда разработчики, политики, управленцы в сфере образования и другие заинтересованные лица, которые имеют возможность влиять на результаты, имеют ограниченное отношение к разработке проекта оценки и к его выполнению (если имеют вообще).

В-третьих, учитывая, что оценка несет важную информацию и что первый шаг состоит в ее передаче соответствующим участникам, таким как разработчики политики, руководство педагогических учебных заведений и спонсоры, удивительно, что оно не всегда носит законченный характер. Это явно ограничивает потенциал использования его результатов. Данная проблема может быть вызвана недостаточным финансированием этапа распространения результатов. В ситуации, когда для разработки и предъявления инструментария и анализа данных требуется большая часть доступного планируемого времени и ресурсов, вполне возможно, что на производство

ТАБЛИЦА 1.1

Причины неполного использования результатов национальной оценки, действия для коррекции неполного использования и структуры, ответственные за эти действия

Причина	Действие	Структура
Деятельность по национальной оценке рассматривается как автономная, слабо связанная с другой деятельностью в образовании	Усилить интеграцию деятельности по оценке в существующие структуры, политику и процессы принятия решения	Министерство образования
Неадекватное привлечение ключевых лиц для проектирования и проведения оценки	Привлечь релевантных заинтересованных лиц для проектирования и проведения оценки	Национальное агентство по оценке, министерство образования
Отказ от сообщения результатов всем лицам, имеющим возможность действовать	Обеспечить условия при планировании бюджета для распространения, планирования действий и подготовки множества отчетов, приспособленных к потребностям пользователей	Национальное агентство по оценке
Недостаточная конфиденциальность результатов национальной оценки	Обеспечить гарантии того, что команда по оценке имеет необходимую методическую компетентность и что релевантные заинтересованные лица привлечены к работе с самого начала	Министерство образования
Политическая чувствительность к приданию результатам оценки публичного характера	Повысить вероятность получения результатов, которые подойдут для разглашения путем проведения их регулярного обсуждения заинтересованными лицами	Министерство образования

Неудачный выбор соответствующих действий после оценки на уровне общей политики в сфере образования	Ввести деятельность по национальной оценке в политику и управленческие действия, и рассмотреть результаты их определения, приложенной и стратегий	Министерство образования
Неудачный выбор соответствующих действий после оценки на школьном уровне	Обеспечить гарантии адекватной связи при передаче результатов школам, подготовить обзор результатов и предложить стратегии улучшения достижений учащихся и обеспечить непрерывную поддержку их реализации	Национальное агентство по оценке, министерство образования, школы и учителя; обучающий состав для учителей, ответственные за учебные планы; поставщики учебников

Источник: разработки авторов.

и распространение информационных продуктов и услуг ничего не останется.

В-четвертых, многие программы оценки испытывают дефицит в инструментарии, представительных выборках и методах научного анализа. Это может вызвать вопросы о валидности данных, которые они обеспечивают, заставляя потенциальных пользователей сделать паузу, прежде чем повлиять на результаты или вовсе отклонить их.

В-пятых, если в ходе национальной оценки выявляется социально-экономическое и образовательное неравенство, связанное с принадлежностью к этнической, расовой или религиозной группе, то у политических деятелей могут возникнуть затруднения, которые возможно заставят их прибегнуть к попытке скрыть результаты от общественности.

В-шестых, национальная оценка вряд ли приведет к разработке соответствующей политики и управленческих решений, если не будут точно расписаны процедуры и механизмы для рассмотрения результатов в контексте другой политической или управленческой деятельности и для определения действий на основе результатов оценки.

Наконец, результаты национальной оценки, вероятно, будут использоваться не полностью до тех пор, пока заинтересованные лица, которые могут влиять на результаты, не будут адекватно о них информированы; не оценят значимость результатов для своей работы; не разработают стратегии, направленные на то, чтобы учащиеся лучше усваивали учебный материал. Так, если не будут предприняты усилия по представлению результатов национальной оценки школам и учителям тем способом, который адекватен их потребностям, и заранее не будут обеспечены фонды для создания механизмов, на основе которых учителя могут использовать информацию, полученную благодаря оценке, для проведения реформ, то в лучшем случае школьный персонал может проигнорировать результаты национальной оценки, а в худшем – просто сорвать их использование.

Эти соображения должны послужить в качестве предостережения всем тем, у кого есть нереалистичные ожидания быстрых изменений политики, вызванных оценкой. Тем не менее в данной книге мы делаем попытку доказать, что данные оценки могут помочь политикам и управленцам в разработке действий для коррекции неполного использования данных оценки (см. табл. 1.1). Чтобы пробудить интерес общественности, а также политиков и управленцев, следует привести примеры фактического использования результатов оценки, собранные по итогам исследования в большом количестве стран. Гораздо меньше свидетельств существует по поводу критикуемой в большей степени и более сложной области использования результатов национальной оценки в классах для улучшения усвоения учебного материала учащимися.

ВЫВОДЫ

Использование результатов национальной оценки зависит от ряда факторов. Политический контекст, в котором проводится оценка, предполагает значительное внимание к проблемам использования ее результатов. Признание того факта, что саму оценку можно считать актом политическим, отражающим характер власти, идеологию и интересы социальных групп, влияющих на систему образования, может сделать саму процедуру оценки и решения, основанные на нем, более прозрачными.

Поскольку краеугольным камнем системы национальной оценки является инструментарий, используемый для измерения достижений учащихся, его качество влияет на использование результатов. Для оптимального использования тестовый инструментарий должен предоставлять информацию, которая отвечает следующим критериям:

- является точной и всесторонней;
- является мерой диапазона достижений;
- обеспечивает направления для корректирующих действий;

- чувствительна к изменениям в обучении.

Тесты, используемые во многих программах национальной оценки, данным условиям не отвечают. Зачастую они ограничиваются измерениями знаний и умений более низкого уровня, чем это необходимо, не содержат достаточное число заданий и могут быть слишком трудными, что в итоге приводит к тому, что у потенциальных пользователей не будет надежных оснований для выбора политики и принятия решений. Ценность национальной оценки для потенциальных пользователей повысится, если фоновые данные собираются по практической деятельности учащихся и процедуры, которые используются для анализа данных, обращаются к факторам, которые влияют на результаты учебной деятельности учащихся и могут быть изменены при выборе политики.

Ключевое решение для политиков и управленцев сферы образования, рассматривающих национальную оценку с точки зрения результатов, предназначенных для использования в будущем, состоит в том, должна ли процедура основываться на выборке или генеральной совокупности. Оценка, основанная на выборке, предоставит информацию и основания для действий на различных уровнях системы образования, тогда как оценка, основанная на генеральной совокупности, кроме того, обеспечит информацию об учащихся, школах и необходимых действиях в них. Решение о проведении оценки, основанной на выборке или на генеральной совокупности учащихся, должно быть принято с учетом как информационных потребностей политиков и управленцев, так и предполагаемой стоимости.

Оценка, основанная на генеральной совокупности учащихся, позволяет возложить на школы ответственность за результаты их учебной деятельности. Прежде чем принять решение об использовании результатов оценки с этой целью, политики должны уделить серьезное внимание: проблеме ограниченности информации, которую национальная оценка может обеспечить о качестве образования, предоставляемого школами;

совокупности людей, учреждений и условий, влияющих на результаты учебной деятельности учащихся; негативным (а может и преднамеренным) последствиям принятия ответственных административно-управленческих решений по результатам выполнения тестов учащимися. Хотя оценка, используемая как механизм власти, в ближайшей перспективе может быть скорректирована. В дальнейшем бюрократический императив, связанный с ней, может ввести коррупцию в систему, которая была предназначена для коррекции или улучшений. (Madaus and Kellaghan, 1992).

Если оценка не имеет значимых прямых следствий (а это характерно для большинства программ национальной оценки), ее результаты нельзя использовать для ответственных административных решений. Однако она может стать прекрасным инструментом для планирования и управления (McDonnell, 2005). Информация, полученная в ходе оценки, рассматривается для политических деятелей, управленцев, педагогов, родителей и общественности как достаточный стимул к действию. Хотя государство, возможно, и не будет отвечать за фактические достижения учащихся, нельзя исключать его подотчетность за создание адекватных условий для системы государственного образования и снижения различий в его качестве, планируемого и достигнутого учащимися – представителями разных национальностей или социальных классов (Reimers, 2003).

Если государство занимает такую позицию, для описания достижений учащихся, идентификации школ и методик учителей, которые направлены на повышение этих достижений, потребуется подробный анализ результатов тестирования. После этого необходимо широко распространить результаты, а также помочь школам выявить проблемы, которые они испытывают (обеспечить их соответствующими ресурсами и методической помощью). Кроме того, школам должна быть оказана непрерывная поддержка в процессе совершенствования.

Настоящая серия книг была создана прежде всего в помощь лицам, проводящим национальную оценку, основанную на выборках. Однако другие книги, за исключением частей, касающихся формирования выборки и статистического анализа, могут быть полезны и при выполнении оценки, основанной на генеральной совокупности учащихся. Данная книга также в основном релевантна последнему подходу, хотя многие затронутые здесь проблемы не связаны с генеральной совокупностью (например, проблема выявления школ, которые нуждаются в информационной поддержке на основе выборочной оценки). Предпосылки для эффективного использования результатов национальной оценки, которые будут рассматриваться далее, относятся к самой оценке, основанной как на выборке, так и на генеральной совокупности. Среди них можно выделить следующие.

- Привлечение политиков и лиц, принимающих решения, к разработке проекта оценки, чтобы обратиться к проблемам, которые они должны идентифицировать как представляющие неотложный интерес.
- Сообщение результатов своевременно и в форме, которая понятна ключевым пользователям.
- Ввод оценочной информации в существующие бюрократические структуры и перевод ее в политику, стратегии и инструменты для политики (например, мандаты, стратегии роста потенциала, стимулы и повышение квалификации для мотивации действий).
- Обеспечение гарантий того, что результаты оценки оказывают влияние на практическую работу учителей в классах с целью улучшения результатов учебной деятельности.
- Обеспечение непрерывной политической поддержки для использования результатов с целью внесения изменений и разработки механизмов для поддержки их применения в реформах на уровне работы в классах.

В книге подробно описаны действия, которые может повлечь национальная оценка. Мы обращаемся к оценке, основанной на генеральной совокупности, и к международным исследованиям в тех случаях, когда они обеспечивают углубленное понимание при использовании результатов или когда они описывают методы, которые относятся к оценке, основанной на выборках учащихся.

В главах 2 и 3 описаны типы отчетов, которые необходимы для сообщения пользователям результатов оценки. В главе 4 в общих чертах представлены проблемы, которые заслуживают внимания при переводе результатов оценки в политику и действия. Далее следует описание специфического использования данных национальной оценки для политики и управления в образовании (глава 5), для обучения (глава 6) и повышения осведомленности общественности (глава 7). В главе 8 определены условия, которые должны оптимизировать использование результатов национальной оценки. В ней также предлагаются конкретные способы, которые могут преобразовать и расширить деятельность по национальной оценке в целях увеличения ее ценности для пользователей.

СООБЩЕНИЕ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ: ОСНОВНОЙ ОТЧЕТ

В этой главе в общих чертах охарактеризованы наиболее важные компоненты инструментария для сообщения результатов национальной оценки – отчета. Эти компоненты должны включать не только результаты, но и основные процедуры, благодаря чему читатели смогут вынести свои суждения относительно их адекватности и уместности. В отчете также содержатся основания для выбора вспомогательных средств при сообщении результатов (например, замечания для брифинга, пресс-релизы, отчет для школ – см. главу 3).

В основном отчете по национальной оценке в обязательном порядке должно присутствовать описание следующих компонентов: контекст оценки; цель оценки; структура, которой руководствуются в проекте оценки; сопровождающие процедуры; достижения в национальной оценке; корреляции между достижениями; изменения в достижениях в течение длительного времени (если соответствующие данные доступны). Число деталей, представленных в основном отчете, зависит от того, будет ли подготовлен отдельный технический отчет. Большинство читателей обладают ограниченными методическими знаниями, и для них представляет интерес лишь то, в какой мере отчет применим к их работе. Большая часть

методических деталей может быть помещена в приложениях к основному отчету.

Вначале члены команды по национальной оценке и ключевые заинтересованные лица должны договориться о том, как проектировать основной отчет, как собрать данные и сообщить о результатах. Прийти к согласию относительно сообщения результатов поможет набор незаполненных или предполагаемых таблиц и обсуждение точных переменных и данных, связанных с каждой таблицей. Так, табл. 2.1 (незаполненная) демонстрирует, как можно представить данные национальной оценки на уровне учащихся в сочетании с предметами учебного плана и гендерными признаками. В табл. 2.2 показано, как могут быть представлены результаты на уровне провинций, чтобы политики могли сравнить уровни достижений среди учащихся с наиболее низкими (не выше 5-го перцентиля) и наиболее высокими (не ниже 95-го перцентиля) оценками в каждой провинции. В табл. 2.3 сравниваются уровни достижений учащихся в двух точках оси времени. Табл. 2.4 предназначена для установления связей между достижениями учащихся и рядом переменных, представляющих интерес для разработки новой политики.

ТАБЛИЦА 2.1

Средние баллы (и стандартные ошибки) для мальчиков и девочек в национальной оценке по языку и математике

Дисциплина	Язык		Математика	
	Мальчики	Девочки	Мальчики	Девочки
Учащиеся по половому признаку				
Средние баллы				
Стандартные ошибки				

Источник: разработки авторов.

ТАБЛИЦА 2.2

Средние баллы (и стандартные ошибки) и оценки различных процентильных рангов в национальной оценке по естественным наукам в провинциях

Провинции	Средние баллы	Стандартные ошибки	Процентильные ранги			
			5-й	25-й	75-й	95-й
1						
2						
3						
4						
5						
6						
Национальные данные						

Примечание: эти данные могут использоваться для подготовки таблиц и частей графиков.

Источник: разработки авторов.

ТАБЛИЦА 2.3

Средние баллы (и стандартные ошибки) в национальной оценке по предъявлению тестов в два момента времени

Первое предъявление		Второе предъявление		Является ли величина статистически значимой?
Средний балл	Стандартная ошибка	Средний балл	Стандартная ошибка	

Примечание: необходимо учитывать, что оба средних значения являются выборочными, основанными на вычислении значимых отличий между ними.

Источник: разработки авторов.

ТАБЛИЦА 2.4

Корреляция между средними баллами учебных достижений по чтению в школе и школьными факторами в программе национальной оценки для 5-го класса

Переменные	r
Средний размер класса, 5-й класс	
Среднее число учебников на учащегося, 5-й класс	
Процент учащихся, которые могут выполнить письменное задание, 5-й класс	
Число учителей в школе	
Число зарегистрированных учащихся в школе	
Опыт преподавания, в годах (учителя в 5-м классе)	
Уровень квалификации учителя (учителя в 5-м классе)	
Количество средств обучения в классе	

Источник: разработки авторов.

КОНТЕКСТ НАЦИОНАЛЬНОЙ ОЦЕНКИ

В описании контекста можно обосновать важность получения информации об учебной деятельности учащихся как основании для выбора политики и принятия управленческих решений. Вполне уместно будет также рассмотреть доказательства, полученные в более ранних исследованиях достижений учащихся (если они доступны).

ЦЕЛИ НАЦИОНАЛЬНОЙ ОЦЕНКИ

Необходимо провозгласить основную цель оценки: скажем, «Оценка проводится, чтобы выявить результаты учебной деятельности учащегося в системе образования». Также могут быть заявлены и более специфические цели: например, оценка

для установления текущих стандартов чтения у учащихся 4-го класса; оценка для сравнения достижений учащихся в частных школах и государственных школах; оценка для мониторинга тенденций в результатах учебной деятельности учащегося в течение длительного времени; оценка для описания школьных ресурсов, проверки школы, семейного фона и факторов учащегося, которые могут быть связаны с достижениями по чтению и обеспечению оснований для будущей оценки.

СТРУКТУРА НАЦИОНАЛЬНОЙ ОЦЕНКИ

Структура – это общий план, который представляет то, что оценивается, в терминах знаний, умений и других признаков. Также в ней содержится информация о том, как эти признаки оцениваются. Структура направляет процесс оценки, делает его прозрачным, во-первых, для тех, кто конструирует оценочный инструментарий, и, во-вторых, для более широкой аудитории, которая анализирует ответы учащихся при оценке. В главе 2 второй книги данной серии подробно рассматривается, как разработать структуру оценки (Anderson and Morgan, 2008).

Исследование PIRLS содержит пример описания конструкта, оцениваемого в исследовании достижений по чтению учащихся в возрасте 9 лет (Mullis and others, 2006; см. также приложение В2 в первой книге из этой серии, Greaney and Kellaghan, 2008). Чтение представлено в терминах достижения двух целей (чтение для приобретения литературного опыта и чтение для получения и использования информации) и четырех процессов (ориентация на цель и восстановление явно заданной информации, получение прямых выводов, интерпретация и объединение идей и информации, анализ и оценка содержания).

Структура также описывает инструментарий, используемый для оценки достижений. Наличие в структуре примеров заданий, используемых в ходе оценки, полезно для читателей. Они получают представление о природе включаемых в тест задач. Конечно, заданий, которые запланированы для использования в будущих процедурах оценки, там быть не должно.

ПРОЦЕДУРЫ ПРОВЕДЕНИЯ НАЦИОНАЛЬНОЙ ОЦЕНКИ

Необходимо привести описание того, как и когда были собраны данные национальной оценки. Это описание будет включать определение целевой группы учащихся, на которой основывается оценка, выбор школ или учащихся и данных по исключенным и не участвующим школам либо учащимся.

ОПИСАНИЕ ДОСТИЖЕНИЙ В НАЦИОНАЛЬНОЙ ОЦЕНКЕ

Принимая решение о том, как представить результаты национальной оценки, необходимо помнить, что эта информация должна быть релевантна потребностям политиков и лиц, принимающих решения, а также быть полезной при обращении к политическим проблемам. Вполне уместно решить вопрос о выборе единственного показателя достижений учащихся (например, общего балла по математике) или множества показателей (например, отдельных баллов для вычислений и решения проблемных заданий). Хотя политики могут вообще предпочесть итоговую статистику. Сообщение единственного показателя достижений, скорее всего, приведет к потере важной информации, что ограничит действия, следующие за оценкой (Kupermintz and others, 1995).

Выполнение тестов чаще всего описывается в терминах уровней мастерства. Описание включает якорное шкалиро-

вание, у которого есть два компонента. Первый компонент – статистический, выявляющий задания, дифференцирующие учащихся между последовательными точками шкалы мастерства с помощью специальных характеристик (например, доли правильных ответов на задания на различных уровнях шкалы оценок). Второй – компонент согласованности, в котором определенные задания используются специалистами по учебному плану для интерпретации материала, который учащиеся, чьи оценки находятся в определенных точках шкалы или рядом с ними, хорошо усвоили и могут выполнить соответствующие задания теста, направленные на его проверку (Beaton and Allen, 1992). Уровням можно присвоить названия (например, удовлетворительный / неудовлетворительный; минимальный / желательный; базовый / профессиональный / продвинутый) и установить долю учащихся, достигших каждого из них. В табл. 2.5 представлены данные, полученные в ходе национальной оценки в Маврикии.

ТАБЛИЦА 2.5

Учащиеся, чьи оценки расположены на минимальном и желательном уровнях достижений в тестах по грамотности, действиям с числами и жизненным навыкам: Маврикий, %

Предмет	%	
	На минимальном уровне или выше	На желаемом уровне или выше
Грамотность	77,6	35,4
Действия с числами	70,3	26,4
Жизненные навыки	71,6	32,4

Источник: Синдикат экзаменов Маврикия, 2003. Воспроизводится с разрешения правообладателя.

В табл. 2.6 описаны уровни достижений по математике в американской программе NAEP. Данные в ней приведены не так, в табл. 2.5. Они должны обеспечить описание результатов выполнения тестов для различных уровней мастерства. Доля учащихся (в государственных школах) колебалась от 44% на базовом уровне до 30% на профессиональном уровне и до 5% на продвинутом уровне. Таким образом, 79% учащихся выполнили тест базовом уровне или выше него (Perie, Grigg and Dion, 2005).

ТАБЛИЦА 2.6

**Уровни достижений по математике в программе NAEP, 4-й класс:
Соединенные Штаты Америки**

Уровень	Ожидаемые достижения для 4-го класса
Базовый	Учащиеся должны продемонстрировать свое понимание математических понятий и процедур в пяти содержательных областях программы NAEP, оценить и использовать основные факты, чтобы выполнить простые вычисления с целыми числами, показать определенное понимание простых и десятичных дробей и уметь решать простые проблемы реального мира во всех содержательных областях программы NAEP. Они должны быть в состоянии выполнять, хотя не всегда точно, четыре действия при вычислениях, знать правила и геометрические фигуры. Их письменные ответы могут быть краткими и не сопровождаться поясняющей информацией
Профессиональный	Учащиеся должны последовательно применять интегрированные процедурные знания и концептуальное понимание при решении проблем в пяти содержательных областях программы NAEP. Они должны быть в состоянии использовать целые числа, чтобы оценить, вычислить и определить, разумны ли результаты. Они должны обладать пониманием сущности простых и десятичных дробей, быть в состоянии решать проблемы реального мира во всех содержательных областях NAEP и использовать четыре действия при вычислениях, знать правила и геометрические формы. Они должны применять стратегии по решению проблем, такие как идентификация и использование

	соответствующей информации. Их письменные решения должны быть организованы и представлены совместно с поясняющей информацией и объяснениями того, как получено решение
Продвинутый	Учащиеся должны уметь применять интегрированные процедурные знания и демонстрировать концептуальное понимание для решения сложных и нетрадиционных проблем в пяти содержательных областях программы NAEP. Они должны быть в состоянии решать сложные и необычные проблемы реального мира в пяти содержательных областях программы NAEP. Они должны показать мастерство в использовании четырех действий при вычислениях, в знании правил и геометрических форм. Ожидается, что они будут в состоянии получить логические выводы и обосновать ответы и процесс решения, объясняя, почему и каким способом решения были получены. Они должны выйти за пределы очевидного в своих интерпретациях и быть в состоянии изложить свои мысли ясно и кратко.

Источник: Американский национальный центр статистики образования, 2006а.

Подход к установлению уровней мастерства отличается от того, который использовался в 2001 году в исследовании PIRLS. Пороговые баллы были определены вначале с помощью установления процента учащихся для каждой эталонной категории и затем на основе проверки умений по чтению и стратегий, связанных с каждым уровнем мастерства (рис 2.1).

Вьетнамские политики, работающие с другими заинтересованными сторонами, такими как разработчики учебного плана, установили шесть уровней достижений по чтению для учащихся 5-х классов, используя статистическую информацию и суждения экспертов (табл. 2.7). Политики использовали данные для проведения сравнений достижений по всей стране в провинциях и т.д.

Во многих программах национальной оценки дисперсию достижений учащихся разделяли на межшкольные и внутри-

Рис. 2.1

Умения по чтению, стратегии и оценки пороговых баллов, полученные для эталонных точек в объединенной шкале грамотности чтения для 4-го класса, PIRLS, 2001



школьные компоненты. Этот процесс требует вычисления коэффициента внутриклассовой корреляции (ρ – коэффициент корреляции между членами совокупности), который

ТАБЛИЦА 2.7

**Уровни умений по чтению в 5-м классе в национальной оценке:
Вьетнам**

Уровень умений	Достижения	% учащихся	Стандартная ошибка
1	Учащийся устанавливает соответствие тексту отдельных слов или на уровне предложений с помощью рисунков. Умение ограничено диапазоном словарного запаса, связанного с рисунками	4,6	0,17
2	Учащийся определяет местонахождение разделов текста, выраженных в кратких сходных по смыслу предложениях и может иметь дело с текстом, не сопровождаемым рисунками. Вид текста ограничен краткими предложениями и фразами с повторяющимися примерами	14,4	0,28
3	Учащийся читает и понимает более длинные связанные по смыслу разделы текста. Учащийся может найти нижележащую или вышележащую по тексту информацию и понимает перефразировки. Более обширный словарный запас позволяет ему понять предложения с несколько более сложной структурой	23,1	0,34
4	Учащийся связывает информацию из различных частей текста. Учащийся выбирает и соединяет части текста, чтобы получить и вывести его различные возможные смыслы	20,2	0,27
5	Учащийся выводит связи и устанавливает намерения автора из информации, представленной различными способами в текстах разных видов и в документах, где нужное сообщение не содержится в явном виде	24,5	0,39
6	Учащийся комбинирует текст с внешними знаниями для вывода различных значений, включая скрытый смысл. Учащийся устанавливает цели автора, его отношения, ценности, верования, мотивы, неустановленные предположения и аргументы	13,1	0,41

Источник: Всемирный банк, 2004. Книга 2, табл. 2.1.

является мерой гомогенности достижений учащихся внутри школ. Он показывает, какова вариация достижений учащихся внутри каждой школы (внутри кластеров) и какова вариация достижений между школами (между кластерами).

Низкие значения коэффициента внутриклассовой корреляции означают, что школы работают на сопоставимых уровнях, в то время как рост величин коэффициента указывает на увеличивающуюся дисперсию учебных достижений по школам (Postlethwaite, 1995). Результаты международных исследований (например, PIRLS или PISA) показывают, что значительные различия существуют между системами образования согласно величинам внутриклассовой корреляции. Кроме того, системы, в которых уровень национальных достижений низок, имеют тенденцию показывать большие различия между школами в их уровнях достижения.

КОРРЕЛЯТЫ ДОСТИЖЕНИЙ

Программы национальной оценки обычно собирают информацию относительно демографических и других фоновых факторов, для того чтобы можно было сравнить достижения подгрупп в составе совокупности учащихся. В тех случаях, когда эта информация связывается в статистическом анализе с достижениями учащихся, она помогает ответить на вопросы, которые являются центральными применительно к роли национальной оценки. В число этих вопросов входят следующие.

- Имеет ли место ситуация, когда система образования не обеспечивает необходимого качества предоставляемых услуг для конкретных групп учащихся?
- Являются ли различия между группами по результатам выполнения тестов статистически значимыми, чтобы гарантировать правильный выбор корректирующих воздействий?

- Какие факторы связаны с низкими достижениями?

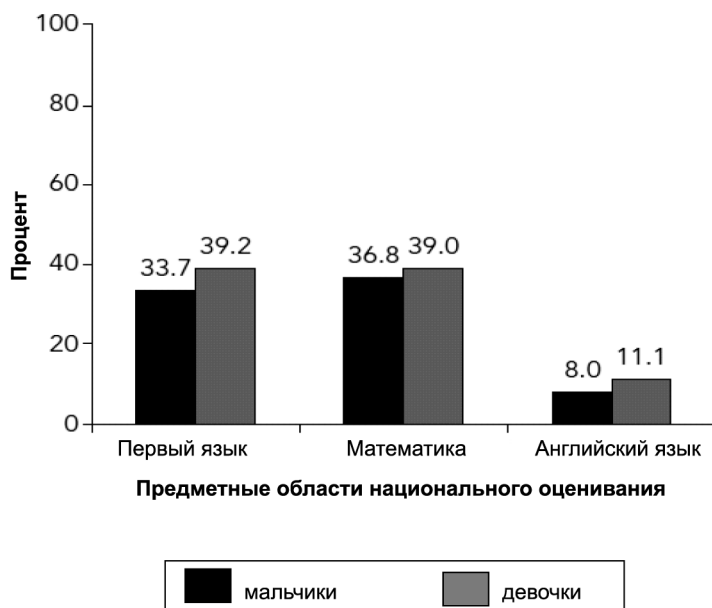
Таким образом, связывая результаты образования с входными характеристиками и процессами обучения, национальная оценка показывает, «что есть». Однако показав, что некоторые секторы системы образования имеют более высокие достижения, можно увидеть, «что могло бы быть» и попытаться идентифицировать факторы, связанные с этими относительными успехами.

Если выборка учащихся в национальной оценке достаточно большая, то ответы на эти вопросы можно получить с помощью учета ряда характеристик: гендерных и региональных признаков (например, по провинциям), характеристик местоположения школы (городская или сельская); принадлежности к этническим или языковым группам и вида образовательного учреждения (государственное или частное). Отчасти в соответствии с актуальным акцентом на гендерном равенстве в программах «Образование для всех» и «Кратчайшие пути введения инициатив» в ходе национальной оценки обычно исследуются различия между достижениями мальчиков и девочек. На рис. 2.2 суммированы результаты национальной оценки в Шри-Ланке, приведены доли учащихся мужского и женского пола, достигших «мастерства» в родном языке, математике и английском языке.

Национальная оценка в Кувейте также выявляет гендерные различия в достижениях по чтению (рис. 2.3). В этом случае гендерные различия были связаны с длительностью периода времени, которое учащиеся разных групп смотрели телевизор. Это не удивительно, поскольку вполне возможно, что время, истраченное на длительный период просмотра телевизора, было связано с более низкими оценками достижений по чтению.

Рис. 2.2

Учащиеся мужского и женского пола, достигшие уровня мастерства в национальной оценке, 4-й класс, предметы учебного плана: Шри-Ланка, %



Источник: Шри-Ланка, Центр национальных исследований в образовании и эвалуации, 2004: рис. 4.44. Воспроизводится с разрешения правообладателя.

В табл. 2.8 приведена информация о достижениях по математике учащихся разных расовых групп в Южной Африке. Она основана на данных, собранных в программе TIMSS.

ТАБЛИЦА 2.8

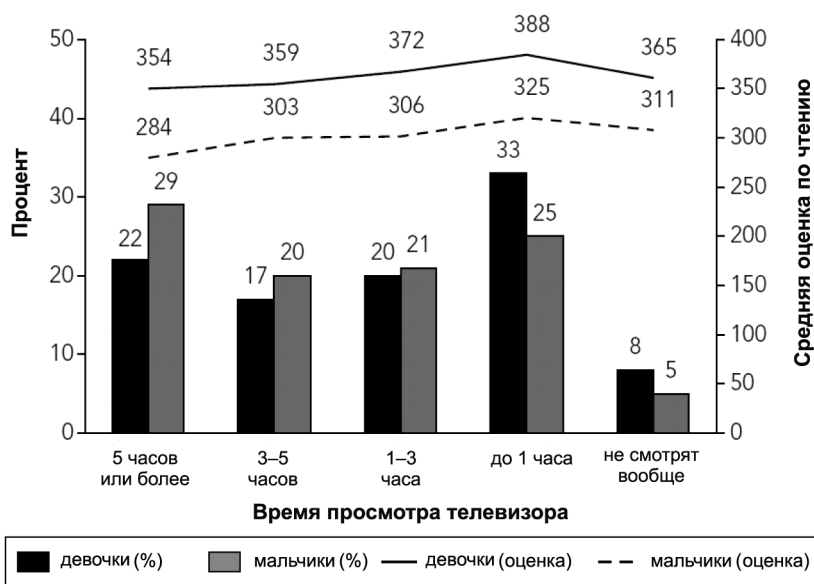
Средние баллы по математике, 8-й класс, в соответствии с расовыми группами: Южная Африка

Расовая группа	Число учащихся	Средний балл	Стандартная ошибка	Минимум	Максимум
Африканец	5412	254	1,2	5	647
Выходец из Азии	76	269	13,8	7	589

Выходец из Африки	1 172	339	2,9	34	608
Индус	199	341	8,6	12	612
Белый	831	373	4,9	18	699
Общее или полное среднее	8147	275	6,89		
Международное среднее		487	0,7		

Рис. 2.3

Учащиеся, смотрящие телевизор различные промежутки времени, в сочетании с гендерными признаками и средним баллом по чтению: Кувейт, %



Источник: Кувейт, Министерство образования, 2008. Воспроизводится с разрешения правообладателя.

В табл. 2.9 представлены сведения о региональных различиях в программе национальной оценки в Непале.

ТАБЛИЦА 2.9

Средние баллы в национальной оценке по непальскому языку, 5-й класс, по регионам: Непал

Регион	Число	Средние баллы	Стандартное отклонение
Восточный	802	51,32	16,7
Центральный	932	50,91	19,5
Западный	1 018	52,89	13,2
Средний Запад	465	50,78	12,7
Дальний Запад	293	49,71	13,2

Источник: Непал, Центр образования и развития, 1999. Воспроизводится с разрешения правообладателя.

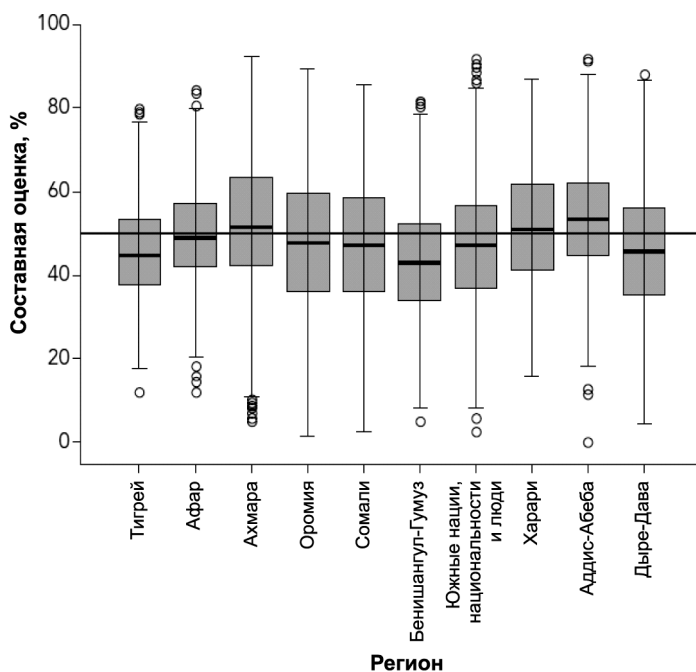
Программа национальной оценки Эфиопии также дает информацию о распределении достижений по регионам (рис. 2.4).

Из результатов оценки следует по крайней мере два возможных варианта политики. С одной стороны, необходимость поддержки регионов с наибольшим числом учащихся с низкими оценками (Оромия, Сомали, Дыре-Дауа) может быть обоснована показателем в виде меры, составленной из оценок в тестах по чтению, математике, естественным наукам и английскому языку. С другой стороны, решение об оказании помощи могло бы быть основано на самых низких средних баллах по тестам (Тигрей, Бенишангул-Гумуз).

Часто (но не всегда) в ходе национальной оценки с помощью анкет собирается информация о различных аспектах образовательной деятельности учащихся, условиях у них дома и атмосфере в коллективе. Это помогает определить школьные и дополнительные факторы, влияющие на формальные оценки

Рис. 2.4

Региональные различия в достижениях, 4-й класс: Эфиопия



Источник: Джебрекидан, 2006. Воспроизводится с разрешения Агентства по обеспечению гарантий качества и экзаменов в Эфиопии.

Примечание: Среднее значение представлено с помощью жирной черной линии; боксы простираются от -1 стандартного отклонения (SD) до $+1$ SD, и конец линий от $-1,96$ SD до $+1,96$ SD. Символы вне линий в их конце представляют выбросы или экстремальные величины.

прогресса учащихся. Во Вьетнаме, например, систематическое исследование домашних факторов всегда ассоциируется с более высокими достижениями, даже в тех случаях, когда проводится отдельная, более глобальная оценка домашнего фона. Школы и учащихся с высокими и низкими достижениями также можно дифференцировать на основе других факторов, например, регулярности питания учащихся и числа пропущенных занятий в школе (Всемирный банк, 2004).

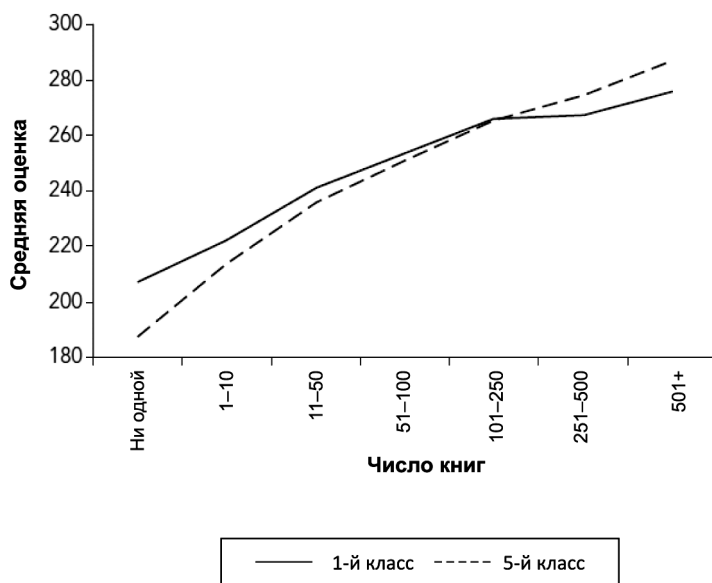
Интерпретировать результаты исследований, в которых достижения учащихся связываются с другими переменными, необходимо с осторожностью. Вывод, что факторы, связанные с достижениями учащихся, можно рассматривать в контексте влияния как «причину» достижений, не может быть гарантирован по ряду причин, которые перечислены ниже.

- Причины связей, устанавливаемых на основе взаимодействующих данных, можно получить только в случае, если они подкреплены другими свидетельствами.
- Число школ или учащихся в некоторых категориях может быть слишком маленьким, чтобы получить надежные выводы.
- Методы статистического анализа могут оказаться неподходящими. Анализ, который исследует связи между двумя переменными (например, между размером класса и достижениями) и который не в состоянии принять во внимание сложные взаимодействия данных (например, между квалификацией учителя и местоположением школы), может привести к ошибочным интерпретациям.
- Требуются сложные исследования, чтобы учесть влияние на результаты образования совокупности взаимодействующих факторов, работающих на уровне учащихся, школы или класса. По этой причине необходимо применять многоуровневые, многомерные статистические методы. Они изолируют эффекты переменной, которые очищаются от других путем систематического удаления или регулирования эффекта кластеров переменных. Это выявляет существование (или отсутствие) значимых различий среди учащихся и школ после такого регулирования. Например, заключение о том, что частные школы превосходят государственные, основанное на данных, показывающих, что у учащихся частных школ более высокий уровень достижений, чем у тех, кто посещает государственные школы, может быть неправильным в тех случаях, когда оценки учащихся не выравниваются с учетом их социально-экономического фона.

Выявить причины изменения достижений совсем не просто. Об этом свидетельствует и тот факт, что зачастую фоновые переменные, по которым собираются данные в национальной оценке, являются всего лишь отображением тех факторов, которые влияют на учебную деятельность учащихся на более глубинном уровне и разнообразными способами. В этом случае необходимо рассмотреть другие данные исследования. И даже сложный статистический анализ может быть не адекватен его задачам. Например, хотя анализ мог бы выявить положительную корреляцию между учебной деятельностью учащихся и числом книг у них дома (см. рис. 2.5), даже в тех случаях, когда учитываются другие переменные, нельзя утверждать, что достижения учащихся имеют причинную

Рис. 2.5

Средние тестовые баллы по чтению учащихся по отношению к числу книг у них дома, 1-й и 5-й классы: Ирландия



Источник: Eivers and others, 2005a: рис. 4.2.

зависимость от от числа имеющихся у них книг. Хотя доступность книг очень важна, результаты улучшения учебной деятельности учащихся вряд ли являются ее непосредственным следствием. Они также меняются под влиянием характеристик окружающей среды, если в ней демонстрируется уважение к книгам, родители придают высокую ценность формальным оценкам достижений учащихся, обеспечивают академическое руководство и поддержку детей, стимулируют своих детей к исследованию и обсуждению идей и событий и в ней установлены высокие нормы и прогнозы школьных достижений (Kellaghan and others, 1993).

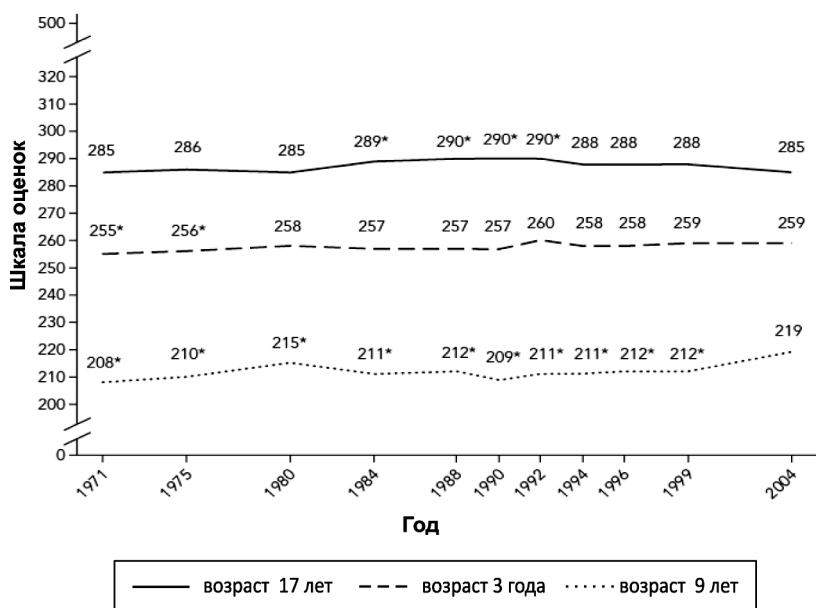
ИЗМЕНЕНИЕ ДОСТИЖЕНИЙ В ТЕЧЕНИЕ ДЛИТЕЛЬНОГО ПЕРИОДА ВРЕМЕНИ

Национальная оценка позволяет судить об изменениях достижений учащихся в течение длительного времени в том случае, если инструментарий связан должным образом. В тех случаях, когда такая возможность есть, результаты могут быть представлены так же, как они приводились в американской программе национальной оценки NAEP на протяжении ряда лет – с 1971 по 2004 год (рис. 2.6). График средних оценок в шкале чтения показывает, что для девятилетних учащихся средний балл по чтению была выше в 2004 году, чем в любом предыдущем году. Для учащихся тринадцати лет средняя оценка в 2004 году была выше, чем в 1971 и 1975 годах, но не отличалась от средней оценки в другие годы. Для учащихся семнадцати лет с 1992 года средняя оценка снижается. В первой книге данной серии (Greaney and Kellaghan, 2008) приводится информация об изменении оценок грамотности в течение длительного периода времени в ряде африканских стран.

Если оценочный инструментарий обеспечивает адекватное представление подобластей учебного плана, то национальная оценка может выявить изменения не только в грубых мерах

Рис. 2.6

Тенденции в шкале средних баллов по чтению для учащихся 9, 13 и 17 лет NAEP, 1971–2004 годы: Соединенные Штаты Америки



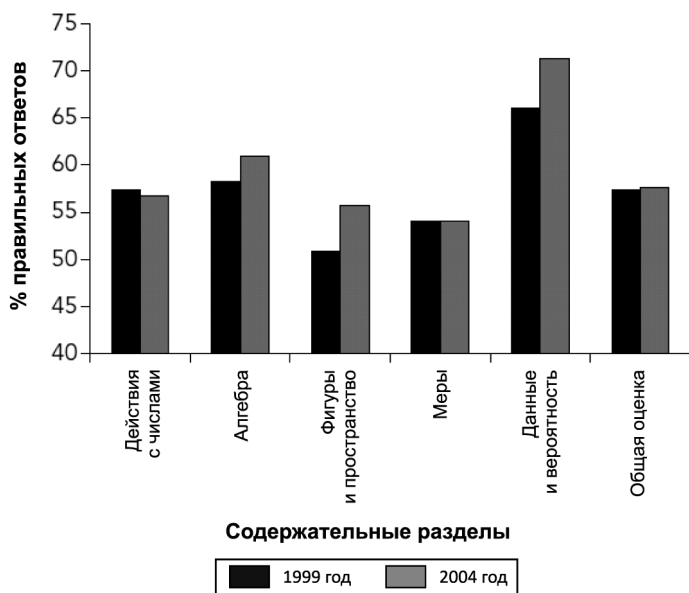
Примечание: *указывает на то, что оценка значительно отличается от оценки 2004 года.

Источник: Национальный центр образовательной статистики США, 2005.

достижений учащихся, но также и в предметных подобластях учебного плана. Например, тесты, использованные в национальной оценке учащихся 5-х классов в Ирландии, позволили оценить результаты выполнения тестов учащимися для ряда содержательных разделов и умений в математике. Результаты оценки, проведенной в 1999 и 2004 годах для пяти содержательных разделов, представлены на рис. 2.7. Различия между годами являются статистически значимыми только для разделов «формы и пространство» и «данные и вероятность». В обоих случаях в промежутке между 1999 и 2004 годами было отмечено повышение результатов.

Рис. 2.7

Оценки среднего процента правильных ответов по математике по содержательным разделам в национальной оценке, 5-й класс, 1999 и 2004 годы: Ирландия



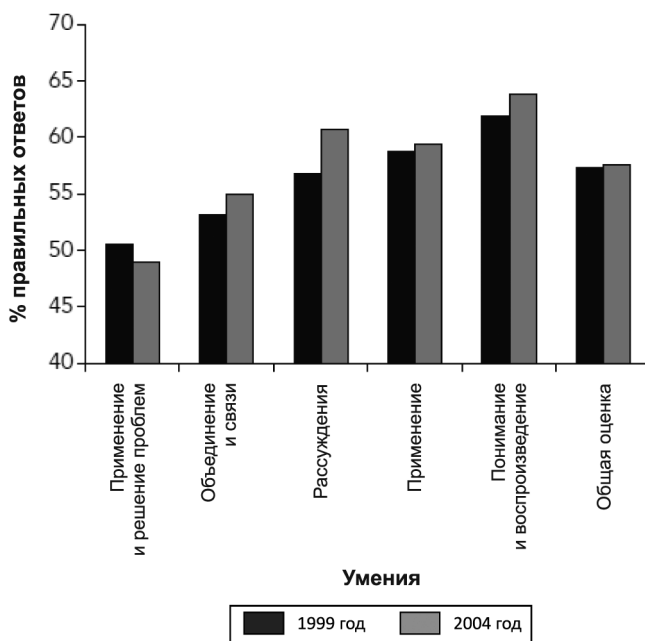
Источник: Surgenor and others, 2006, рис. 3.1.

Данные той же самой национальной оценки для математических умений по пяти областям приводятся на рис. 2.8. В этом случае в промежутке между 1999 и 2004 годами значимым оказалось только увеличение результатов по разделу «рассуждения».

Во многих странах очевидна необходимость проведения мониторинга прогресса для достижения «Целей развития тысячелетия всеобщего начального образования до 2015 года». Это укрепляет интерес к оценке достижений в течение длительного периода времени. Однако усилия по повышению качества образования, вероятно, могут встретить ряд проблем, порожденных ростом затрат и уменьшением бюджетных ассигнований. В Малави, например, была проведена кампания по введению

Рис. 2.8

Оценки среднего процента правильных ответов для умений по математике в национальной оценке 5-й класс, 1999 и 2004 годы: Ирландия



Источник: Surgenor and others, 2006, рис. 3.2.

всеобщего обучения. Однако необходимые ресурсы для отмены платы за обучение в школе и увеличения числа учащихся выделены не были. Как следствие, результаты выполнения тестов резко ухудшились (Altinok, 2008).

ВЫВОДЫ

Основной отчет о результатах национальной оценки должен представлять собой достаточно подробное исследование. В таком случае он станет первичным источником информации для оценки. В то же время отчет не должен перегружать читателей беспорядочным количеством методической информации.

Некоторые отчеты действительно готовятся в соответствии с этими правилами. Иногда допускаются неточности, связанные с предоставлением неадекватных сведений относительно методических аспектов исследования, содержания тестов достижений, используемых методов или ошибок и смещений в оценках.

Другая ошибка состоит в представлении данных исключительно в табличной форме. Немногие потенциальные пользователи умеют делать выводы из статистических таблиц. Поэтому лучше всего иллюстрировать ключевые результаты с помощью диаграмм или графиков. Кроме того, в табличном отчете может не уделяться должное внимание результатам, которые помогают объяснить связи, выявленные в ходе оценки, например, почему переменная (такая, как домашний фон) является важной и что известно о ее влиянии на достижения учащихся. Однако эта информация становится очень значимой для пользователей, если от них требуется оценить возможность импортирования результатов. Она позволит им судить о том, какие действия могут потребоваться для установления недостатков.

Относительно включения рекомендаций в отчет о национальной оценке мнения о форме и методах представления расходятся. Некоторые национальные (и международные) отчеты содержат рекомендации, другие нет. Команды по национальной оценке в самом начале должны разъяснить, должен ли отчет, с точки зрения представителей министерства образования, содержать какие-либо рекомендации. Если должен, то необходимо, чтобы команды по национальной оценке установили, какие именно – только те, которые следуют непосредственно из исследования по национальной оценке, или более широкие, требующие релевантных исследований результатов.

ГЛАВА

3

ОТЧЕТ О НАЦИОНАЛЬНОЙ ОЦЕНКЕ: ДРУГОЙ ИНСТРУМЕНТАРИЙ ДЛЯ СООБЩЕНИЯ РЕЗУЛЬТАТОВ

Довольно часто единственный отчет, который составляется по итогам национальной оценки, готовят по правилам, описанным в предыдущей главе. Но потенциальных пользователей данных, полученных по результатам оценки довольно много, и требования различных участников образовательного процесса сильно отличаются. Поэтому в данной главе мы рассмотрим дополнительные способы сообщения результатов, которые адаптированы применительно к множеству потребностей. Эти методы охватывают широкий диапазон: от публикации длинных детальных отчетов до издания коротких пресс-релизов и проведения индивидуальных брифингов, которые подводят итоги оценки. Хотя разработчики учебного плана и авторы учебников будут требовать подробной информации относительно результатов учебной деятельности (возможно, более детальной, чем та, что доступна в основном отчете), ряд суммарных результатов должен быть представлен гражданам, которые имеют ограниченную статистическую подготовку.

Определим следующие общие принципы сообщения результатов исследований, которые могут быть рассмотрены и применительно к результатам национальной оценки (Pérez, 2006).

- Используйте простой язык, которому отдается предпочтение в наиболее распространенных СМИ.
- Четко выявляйте ключевые заинтересованные лица и адаптируйте сообщения о событиях и продуктах применительно к их потребностям.
- Рассматривайте общественных волонтеров и вероятных лидеров как защитников.
- Распространяйте информацию средствами масс-медиа. События нужно хорошо рекламировать.
- Используйте лозунги и простые сообщения, которые будут ясны и понятны. Например, утверждение о том, что «восьмилетний ребенок должен быть в состоянии прочесть историю с 60 словами за одну минуту и ответить на три вопроса о ее содержании», четко приводит требования стандарта.
- Неоднократно обращайтесь ко всем радиопередачам или материалам, распространяемым для большой аудитории (включая презентации PowerPoint) с сопровождающей методической информацией.

В этой главе описаны процедуры и инструментарий, которые команда по национальной оценке может использовать в дополнение к основному отчету. Они включают: брифинг министерства и руководящего персонала в политике; подготовку примечаний к брифингу; публикацию итоговых отчетов, технических отчетов и тематических сообщений; обеспечение сообщений для СМИ; издание пресс-релизов; проведение пресс-конференций; проведение отдельных брифингов; регистрацию сообщений на сайте и создание доступных данных по оценке.

ЛИСТ СПЕЦИФИКАЦИИ ПРОДУКТОВ

Разработка листа спецификации продуктов, с точки зрения любого менеджера по национальной оценке, – очень полезная и важная процедура в процессе подготовки отчетов. Лист должен включать следующие элементы:

- название итогового информационного продукта, который должен быть подготовлен по результатам оценки (например, основной отчет);
- суммарное описание совокупности информационных продуктов, освещающих материалы оценки;
- предполагаемые пользователи;
- приоритеты информационного продукта по итогам оценки;
- детальное описание информационного продукта по итогам оценки;
- ключевые действия по их получению, обязанности сотрудников и расписание работ;
- стоимость получения информационного продукта по итогам оценки;
- планируемая дата получения (и место его нахождения, если это возможно);
- цена информационного продукта по материалам оценки;
- зависимость от других информационных продуктов по итогам оценки или входных данных.

Поскольку потенциальное число конечных результатов оценки и услуг почти неизбежно будет превышать технические и финансовые ресурсы, доступные для команды по национальной оценке, продукты необходимо ранжировать в соответствии с приоритетами. Те, которые обладают самым большим потенциалом для проведения реформ, должны быть получены в первую очередь. От команды по национальной оценке требуется запланировать и определить стоимость полного спектра конечных информационных продуктов по материалам оценки и услуг. Однако только в том случае, если будет доступно дополнительное финансирование, команда сможет выполнять свою работу быстрее.

БРИФИНГ МИНИСТРОВ И СТАРШЕГО ПЕРСОНАЛА В ПОЛИТИКЕ

Команды по национальной оценке должны готовить материалы к брифингам для министра и руководителей отделов министерства образования. Им понадобится информация об

основных результатах оценки в краткой форме и о возможном использовании этих результатов (Beaton and Johnson, 1992). Поскольку министры ежедневно прочитывают очень большое количество документов, заметки для брифинга должны быть очень краткими и затрагивать только суть.

Министрам и руководителям вряд ли нужно знать полный текст отчета. Но они обязательно должны ознакомиться с ключевыми результатами и соответственно подготовиться к решению проблем, которые могут поднять СМИ, парламент или заинтересованные лица в системе образования (например, союзы учителей) после публикации отчетов об оценке. Они нуждаются в этой информации, даже если новости о результатах оценки «плохи». Некоторые участники образовательного процесса не всегда приветствуют «хорошие» новости, потому что они могут привести к выводу, что дополнительные вложения в образование не нужны.

Особое внимание необходимо обратить на то, как сообщаются и интерпретируются отличающиеся результаты отдельных групп учащихся. Если определенные группы учащихся показывают плохие результаты, то это может стать основанием для обвинения некоторых политических деятелей. Ведь такие результаты можно интерпретировать как свидетельство пренебрежения этими слоями учащихся. Еще одним следствием может стать отказ от поддержки национальной оценки в целом или решение о неразглашении ее результатов.

У многих министерств имеются стандартные формы для брифингов. Их следует использовать, только если они адекватны задачам представления результатов национальной оценки. Примерный список эффективных заметок к брифингу по национальной оценке представлен ниже.

- Кратко сформулированная цель оценки.
- Одно или два предложения о том, почему для министра важно знать результаты оценки.

- Краткое описание фона оценки (например, кто ее проводил и какая целевая группа учащихся было оценена).
- Ключевые результаты, особенно те, которые могли бы иметь политическое значение (уровни достижений учащихся, региональные и гендерные различия, распределение ресурсов по школам).
- Возможные действия, которые необходимо предпринять в будущем (должен ли министр делать публичное заявление о результатах, если да, то нужно ли ему рекомендовать их исследование на основе учебного плана, должны ли его проводить образовательные учреждения в системе педагогического образования, предложить ли проведение дальнейшей национальной оценки в различных областях учебного плана).
- Рекомендации министру для подготовки к ответам на возможные вопросы.
- Информация о наличии приложений – пресс-релиза, резюме по национальной оценке или полного отчета.
- Имя участника команды по национальной оценке, с которым можно связаться в том случае, если министр запросит дополнительную информацию.



ПУБЛИКАЦИЯ ИТОГОВЫХ ОТЧЕТОВ

Для методически неподготовленных читателей в дополнение к основному отчету часто публикуется итоговый отчет. Он может быть очень кратким. Так, например, в США по итогам NAEP на сайте программы публикуются резюме или краткие версии итогового отчета. В дополнительном материале 3.1 представлен итоговый отчет по результатам выполнения тестов учащимися 4-го класса в Иллинойсе в 2007 году при оценке чтения.

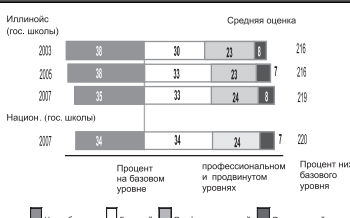
Некоторые итоговые сообщения могут быть более длинными. Например, после национальной оценки по географии учащихся 4-х, 8-х и 12-х классов Национальный центр образо-

Дополнительный материал 3.1

Итоговый отчет штата Иллинойс, уровни достижений учащихся 4-го класса по чтению в 2007 году в программе NAEP: Соединенные Штаты Америки



Чтение-2007
 Снимок государственного сообщения
 Иллинойс
 4-й класс
 Государственные школы

Программа «Национальная оценка образовательного прогресса» (NAEP) оценивает чтение в двух содержательных областях в 4-м классе: чтение для приобретения литературного опыта и чтение для получения информации. Шкала по чтению NAEP охватывает диапазон баллов от 0 до 500.

Результаты по чтению для Иллинойса в целом	Процент учащихся в NAEP для уровней достижений и средняя оценка																								
<ul style="list-style-type: none"> В 2007 году средний балл на шкале для учащихся 4-го класса в Иллинойсе составил 219. Она мало отличается от среднего балла в 2005 году (216) и в 2003 году (216). Средний балл (219) в Иллинойсе в 2007 году незначительно отличался от полученного в национальной оценке в государственных школах (220). В 52 штатах и других юрисдикциях, которые участвовали в программе 2007 года по оценке учащихся 4-го класса, средний балл по шкале для учащихся Иллинойса был выше, чем в 12 юрисдикциях, но незначительно отличался от тех, что были в 14 юрисдикциях, и был ниже, чем в 25 юрисдикциях. Доля учащихся в Иллинойсе, чьи результаты были на высоком уровне или выше него в программе NAEP составила в 2007 году 32%. Этот показатель незначительно отличался от того, который был в 2005 году (29%) и от того, который был в 2003 году (31%). Доля учащихся в Иллинойсе, чьи результаты выполнения тестов были на базовом уровне или выше него, составила 65% в 2007 году. Этот показатель незначительно отличался от того, который был в 2003 году (61%). 	 <p>Иллинойс (гос. школы) Средняя оценка</p> <table border="1"> <tr> <td>2003</td> <td>38</td> <td>30</td> <td>22</td> <td>6</td> <td>216</td> </tr> <tr> <td>2005</td> <td>38</td> <td>33</td> <td>23</td> <td>7</td> <td>216</td> </tr> <tr> <td>2007</td> <td>35</td> <td>33</td> <td>24</td> <td>8</td> <td>219</td> </tr> </table> <p>Национал. (гос. школы)</p> <table border="1"> <tr> <td>2007</td> <td>34</td> <td>34</td> <td>24</td> <td>7</td> <td>220</td> </tr> </table> <p> ■ Ниже базового ■ базовый ■ Профессиональный ■ Продвинутый </p> <p><i>Примечание:</i> в программе NAEP для оценки учащихся по чтению в 4-м классе уровни достижений соответствуют следующим точкам шкалы: ниже базового: 207; базовый: 208–237; высокий: 238–257; продвинутый: 268 или выше</p>	2003	38	30	22	6	216	2005	38	33	23	7	216	2007	35	33	24	8	219	2007	34	34	24	7	220
2003	38	30	22	6	216																				
2005	38	33	23	7	216																				
2007	35	33	24	8	219																				
2007	34	34	24	7	220																				

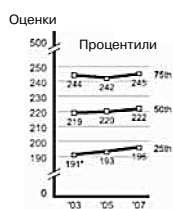
Результаты выполнения тестов в NAEP. Отчет по группам в Иллинойсе за 2007 год						
Группы для отчета	Процент учащихся	Средний балл	Процент учащихся ниже базового уровня	Процент учащихся на базовом уровне или выше		Процент учащихся на продвинутом уровне
Мальчики	51	217	37	63	30	7
Девочки	49	222	33	67	35	9
Белые	55	230	23	77	42	12
Афроамериканцы	20	201	56	44	14	2
Латинская Америка	20	205	50	50	IS	3
Азия/Острова Тихого океана	3	240	13	87	54	16
Американские индейцы/Коренные жители Аляски	#	↓	↓	↓	↓	↓
Имеет право на бесплатный завтрак в Национальной школьной программе	44	204 ↑	52 ↓	48 ↑	16	2
Не имеет право на бесплатный завтрак в Национальной школьной программе	56	232	21	79	45	12

Дополнительный материал 3.1

Различия в средних баллах между отобранными группами

- В 2007 году средний балл учащихся мужского пола в Иллинойсе был ниже, чем средний балл учащихся женского пола на 5 пунктов шкалы.
- В 2007 году средний балл афроамериканских учащихся был ниже, чем средний балл белых учащихся, на 29 пунктов шкалы.
- В 2003 году афроамериканские учащиеся имели средний балл, ниже чем у белых учащихся, на 34 пункта шкалы.
- В 2007 году испаноговорящие учащиеся имели средний балл, который был ниже, чем средний балл белых учащихся, на 234 точки шкалы.
- В 2003 году испаноговорящие учащиеся имели средний балл, ниже чем у белых учащихся, на 31 пункт шкалы.
- В 2007 году учащиеся, у которых было право на бесплатный завтрак, имели средний балл ниже, чем учащиеся, у которых этого права не было, на 28 пунктов шкалы. Этот разброс результатов выполнения был меньше, чем тот, который был в 2003 году (35 пунктов).
- В 2007 году разброс баллов между учащимися с результатами на 75-м процентиле и учащимися с результатами на 25-м процентиле был 49 пунктов шкалы. В 2003 году разброс оценок между учащимися с результатами на 75-м процентиле и на 25-м процентиле был 53 пункта шкалы

Оценки по чтению в отобранных процентилях



Примечание: баллы для отобранных перцентилей на шкале оценок по чтению в программе NAEP указывают, как хорошо учащиеся выполнили тест на нижнем, среднем и более высоком уровнях.

Источник: США, Национальный центр образовательной статистики, 2007.

вательной статистики готовит резюме под названием «География: что знают учащиеся и что они могут выполнить?» Его объем занимает приблизительно шесть страниц. В отчете рассматривается, с какими заданиями справились учащиеся с результатами на 25-м, 50-м и 90-м процентилях в каждом классе (Vanneman, 1996). Другие отчеты могут быть более длинными и включать краткое описание всех компонентов основного отчета.

В Интернете доступны основные и итоговые отчеты по национальной оценке чтения на английском языке в Ирландии (Eivers and others, 2005a, 2005b). К обеим формам публикации можно получить доступ на сайте Научно-исследовательского центра в образовании (<http://www.erc.ie>). Основной отчет США по программе NAEP для результатов по истории (для учащихся 4-х, 8-х и 12-х классов) и итоговый отчет могут быть загружены с сайта <http://nces.ed.gov/nationsreportcard/ushistory/>.

Чаще всего с итоговыми отчетами первыми знакомятся учителя. В этом случае итоговые отчеты должны включать

рекомендации по обучению, которые являются результатом оценки (см. главу 6). Например, руководство для учителей, подготовленное на основе оценки достижений по математике пятнадцатилетних ирландских учащихся, было издано после предъявления тестов по программе PISA в Ирландии в 2003 году (Shiel and others, 2007). Наряду с информацией относительно выполнения теста учащимися и степени, в которой преподавание математики соответствует требованиям PISA, отчет содержит примеры тестовых заданий. В дополнительном материале 3.2 приведены три из множества рекомендаций для учителей, которые содержатся в отчете.

Целевой аудиторией при подготовке итоговых отчетов могут быть и другие заинтересованные группы пользователей, перечисленные ниже.

- Союзы учителей, которые представляют их общие интересы и могут выступать в качестве сильных противников или сторонников изменений в образовании. Часто у них есть большая личная заинтересованность в использовании информации, полученной в ходе оценки для поддержки своих позиций.
- Местные руководители, в том числе местные политические деятели, которые нуждаются в информации для определения того, совпадают ли результаты функционирования системы образования с тем, что общество должно достигнуть на социальном, культурном и экономическом уровнях.
- Работодатели и бизнес-лидеры, которые нуждаются в объективной информации о достижениях в учебной деятельности учащихся, рассматриваемых как показатель качества подготовки будущих служащих.
- Граждане, которые могут судить о том, отвечает ли система образования критериям доступности, качества, эффективности и равноправия, и, благодаря своему влиянию и авторитету, оказывать давление с целью улучшения условий образования.

Дополнительный материал 3.2**Предложения для применения подхода программы PISA к обучению и изучению математики: Ирландия**

Следующие рекомендации взяты из руководства для учителя, подготовленного Научно-исследовательским центром в образовании и изданного Ирландским департаментом образования и науки после завершения национальной оценки.

Применяйте чаще диалоговый подход в обучении математике, подразумевающий, что учащиеся привлекаются к обсуждению проблем как до, так и после их решения. Обсуждение должно быть сосредоточено на идентификации математического аппарата, необходимого для решения проблем и рассуждения учащихся при их решении.

В процессе обучения делайте акцент на полном диапазоне когнитивных компетенций. Если на занятиях и на экзаменах уделять чрезмерное внимание процессу воспроизведения, то это приведет к тому, что многие учащиеся не научатся применять компетенции высокого уровня, например, установление связей и рефлексию. Вероятно, что применение именно этих компетенций поможет учащимся с различными способностями достичь более высокого уровня концептуального понимания и самостоятельности в решении задач.

Постарайтесь достичь равновесия между контекстно-свободными ситуациями и вопросами, которые представлены в контексте реальных жизненных ситуаций. В существующих учебниках и экзаменационных материалах многие из заданий являются контекстно-свободными. Хотя они играют важную роль в развитии основных умений по математике, не менее важно предоставить учащимся возможность иметь дело с проблемами реального мира. Такое требование поможет сделать математику более релевантной для учащихся и предоставит им возможности для развития широкого диапазона математических компетенций.

Источник: Shiel and others, 2007, 48.

- Спонсорские агентства, которые ищут объективную информацию для оценки эффективности программ или для обоснования необходимости оказать помощь сектору образования, специфической подгруппе учащихся или географическому региону.

ПУБЛИКАЦИЯ ТЕХНИЧЕСКИХ ОТЧЕТОВ

Ключевым элементом программ национальной оценки является технический отчет. Он обеспечивает членам исследовательских и научных сообществ подробную информацию об оценке, которая позволяет им осуществить анализ программ оценки и критически осмыслить их. Технические отчеты также предоставляют запись последовательных действий, включенных в оценку, которая необходима для осуществления будущих циклов оценочных процессов.

Некоторые программы национальной оценки публикуют только один отчет, который служит и общим, и техническим, например австралийский отчет о достижениях учащихся от 6 до 10 лет в информационных и коммуникационных сферах. Он содержит технические детали относительно уровней мастерства и процедур формирования выборок (Ainley, Fraillon and Freeman, 2007).

Другие сообщения сосредотачиваются в большей степени на технических аспектах формирования выборки, анализе заданий, методиках оценки результатов и критериях правильности выполнения заданий, шкалировании, статистическом анализе и контроле качества информации. Примеры такого подхода могут быть найдены в техническом отчете TIMSS 2003 года (Martin, Mullis and Chrostowski, 2004) и в программе «Исследование грамотности чтения в Соединенных Штатах Америки»: технический отчет США, Международная ассоциация по эвалюации образовательных достижений, раздел «Исследование грамотности чтения» (Binkley and Rust, 1994).

В табл. 3.1 представлены виды деятельности, которую должен охватывать технический отчет. В частности, внимание должно быть уделено разработке инструментария, описанию генеральной совокупности учащихся или выборки, которая оценивается, шкалированию результатов по заданиям и статистическому анализу.

ТАБЛИЦА 3.1

Технический отчет: предлагаемое содержание

Раздел	Описание деятельности	Примеры или комментарии
Цель	Опишите контекст и главные цели национальной оценки	Проводите мониторинг изменений по уровням достижений, начиная с момента последней национальной оценки. Сообщите о региональных различиях в достижениях учащихся или о том и о другом
Определение предмета оценки	Определите оцениваемый предмет, список аспектов оцениваемого предмета (например, словарный запас, понимание, поведение)	Грамотность чтения определяется как «способность понимать и использовать те формы письменного языка, которые требуются обществом и/или ценятся индивидуумом» (Campbell and others, 2001, 3)
Детали того, что предназначается к измерению	Опишите содержательные области и познавательные уровни, которые будут оценены для каждого предмета и класса. Включите детали о заданиях	Включите проект или таблицу спецификаций. Укажите число заданий с множественным выбором, закрытых заданий с конструируемым ответом и заданий с развернутым ответом
Разработка инструментария	Сообщите подробности по конструированию апробационных тестов, анкет и руководства по предьявлению тестов, включая их коррекцию	Включите в качестве руководства суммарный учебный план и результаты экспертиз заданий учителей для анализа их пригодности. Если тесты или анкеты были переведены с другого языка, то опишите, как была проверена точность перевода

Оцениваемая генеральная совокупность или выборка	Для выборки укажите ее размер и критерии для исключения из нее учащихся, группировки школ и замены школ	Подготовьте отчет относительно желаемой, определенной и исключенной выборки (например, возраст или класс; государственная или частная школа); нормах участия; стратификации выборки; вида выборки (например, кластерная, число стадий отбора; метод определения размера выборки и метод вычисления ошибки)
Действия	Опишите выбор мер проверки качества информации и подготовки администраторов тестирования	Опишите процедуры для гарантии информационной безопасности при поставке, хранении и возвращении всего инвентаря оценки
Оценивание	Опишите оценочные процедуры и меры контроля качества	Укажите процент тестовых заданий различных видов, которые были предметом независимой экспертизы и переоценки
Ввод и чистка данных	Опишите процедуры и меры контроля качества	Объясните, как определенные ошибки в тестовых бланках, заполненных учащимися, были идентифицированы и скорректированы
Анализ заданий	Интегрируйте информацию о трудности заданий и их уровнях дискриминативности	Укажите, были ли задания методически адекватны особенностям регионов и лингвистических групп, если они применяются в них. Приведите причины для любого удаления заданий

<p>Шкалирование результатов по заданиям</p>	<p>Если используется современная теория тестов (Item Response Theory, см. Крокер Линда. Введение в классическую и современную теорию тестов / Линда Крокер, Джеймс Алгина. М.: Логос, 2010. – <i>Примеч. ред.</i>), объясните, как будут вычисляться шкалированные оценки и определяться диапазоны уровней мастерства</p>	<p>Опишите роль специалистов по предметам в определении диапазонов уровней мастерства</p>
<p>Анализ данных оценки</p>	<p>Представьте итоговые статистические результаты, включая стандартные ошибки. Сравните результаты с более ранними результатами национальной оценки, если это возможно. Проанализируйте данные по проблемам, предложенным комитетом по регулированию</p>	<p>Составьте точный список статистических процедур; установите используемое программное обеспечение; опишите методы вычисления стандартных ошибок, уровни статистической значимости и способы, которыми пропущенные данные были обработаны. Объясните, как показатели (например, школьные ресурсы, родительский интерес к образованию), если они использовались, были вычислены. Сравните региональные различия в достижениях. Установите связь между достижениями и характеристиками учащихся (возраст, пол, отношение к предмету); школы (квалификация учителей, школьные ресурсы или домашние фоновые факторы (состав семьи, образование родителей)</p>
<p>Выводы</p>	<p>Составьте резюме о главных полученных данных. Предложите ограничения по применению результатов</p>	<p>Разработайте допустимые рекомендации, основанные на результатах, если это запланировано</p>

ПУБЛИКАЦИЯ ТЕМАТИЧЕСКИХ ОТЧЕТОВ

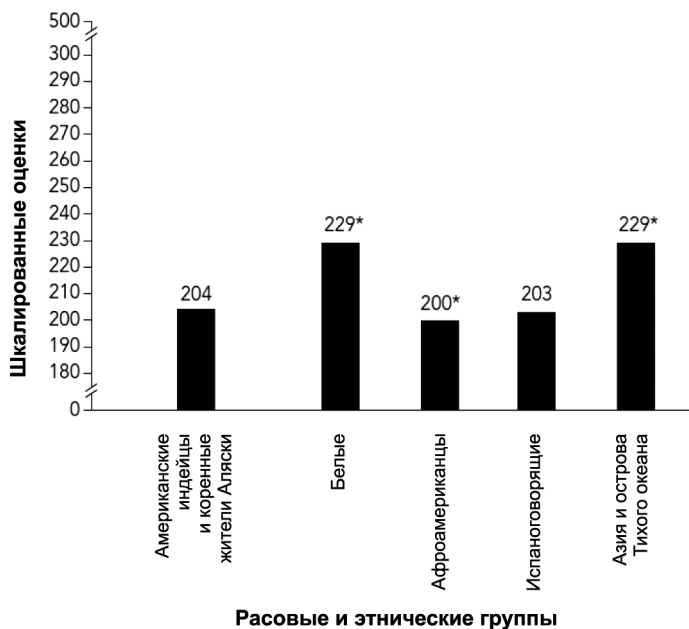
Тематические отчеты исследуют аспекты результатов оценки, связанные с определенной темой, которая не рассмотрена подробно в основном отчете. Тематический отчет может содержать анализ ошибочных профилей в ответах учащихся по отдельным аспектам учебного плана или по ряду заданий в тесте учебных достижений. Такой анализ поможет идентифицировать, в каких переделках нуждается учебный план или по каким темам обучение должно быть усилено. Например, для учащихся в Катаре проводился анализ ошибочных профилей ответов на задания в тесте по естественным наукам в программе PISA 2006 года (DataAngel Policy Research, 2007). Несмотря на потенциал такого анализа, он осуществлялся в очень незначительных масштабах.

Тематический отчет может затрагивать в основном субпопуляцию учащихся, которая представляет интерес для специфической аудитории или имеет отношение к определенным политическим проблемам (например, мальчики или девочки, учащиеся, принадлежащие к разным расовым или этническим группам, учащиеся в невыгодных фоновых условиях, сельские учащиеся). На рис. 3.1 сравниваются достижения по чтению учащихся, классифицированных по принадлежности к расовым или этническим группам на основе данных национальной оценки в США.

Примеры тематических отчетов, которые включают продвинутые статистические методы и существующие доступные результаты могут быть найдены в исследованиях, использующих данные Южно- и Восточноафриканского консорциума по мониторингу качества образования (SAC-MEQ) и Программы анализа систем образования (d'Analyse des Systèmes Educatifs de la CONFEMEN, или PASEC), по рентабельности входных характеристик школ (Michaelowa and Wechtler, 2006) и различий в грамотности между учащимися

Рис. 3.1

Средние баллы по чтению по расовым и этническим группам, NAEP, 4-й класс, 2005 год: США



Примечание: результаты основаны на национальной выборке программы NAEP.

Афроамериканцы: учащиеся – выходцы из Африки; *испаноговорящие:* учащиеся из стран Латинской Америки; *Азия и Острова Тихого океана:* учащиеся, чьей родиной являются Гавайские острова.

Расовые группы исключают учащихся из Испании.

* указывает, что оценка значительно отличается от той, которую получили учащиеся индейцы и уроженцы Аляски.

Источник: США, Национальный центр образовательной статистики, 2006b. Рис. 2.17.

сельских и городских школ (Zhang, 2006) и исследования, использующие данные программы PISA по обязательным результатам чтения для пятнадцатилетних учащихся (Kirsch and others, 2002).

ОБЕСПЕЧЕНИЕ СООБЩЕНИЙ ДЛЯ СМИ

Производство отчетов в текстовом виде является довольно дорогим. Кроме того, этот вид публикации, возможно, не будет подходящим для многих людей, которые могут проявить интерес к результатам национальной оценки. Напротив, печатные СМИ, радио, телевидение или видео могут обеспечить недорогой способ распространения основных сообщений о национальной оценке для людей, которые, вероятно, иначе не узнали бы о нем. С одной стороны, отказ от освещения в печати может означать, что оценка останется в значительной степени незамеченным. С другой стороны, когда СМИ критически настроены по отношению к оценке либо публикуют сенсационные или неточные сообщения, то это тоже может представлять угрозу для системы национальной оценки.

Радио охватывает значительное число слушателей и является особенно подходящим для людей с низкими навыками грамотности. Телевидение также достигает широкого охвата аудитории и играет важную роль в пробуждающемся общественном интересе к проблемам национальной оценки. Однако радио- и телевизионные передачи обычно слишком поверхностны и коротки и потому не обеспечивают реальное понимание результатов исследований или их значения для образования.

Наиболее всестороннее использование телевидения для сообщения общественности результатов оценки было реализовано в Латинской Америке. В процессе доведения до всеобщего сведения результатов международной оценки, проведенной Латиноамериканской лабораторией по оценке качества образования (см. первую книгу данной серии, приложение С.3), был подготовлен специальный видеофильм

с материалами, который жители всего континента позже увидели по телевидению (Ferrer and Arregui, 2003). Опыт использования видеоресурсов для сообщения результатов национального или международного оценивания получил более масштабное распространение в Перу. Там обнаружили, что видеофильмы оказались намного более эффективными, чем лекции или презентации в Microsoft PowerPoint в режиме диалога с заинтересованными лицами в образовательной политике (Pérez, 2006).

ИЗДАНИЕ ПРЕСС-РЕЛИЗОВ

Пресс-релиз – короткий письменный обзор, подготовленный для СМИ. Его формат и содержание зависят от конкретной выпускающей организации. В пресс-релизе министерства образования на первый план выдвигаются положительные моменты, тогда как в пресс-релизе, выпущенном какой-либо исследовательской структурой, будет отражена более нейтральная позиция.

Вначале специалист, ответственный за составление пресс-релиза, должен выяснить, для каких читателей он предназначен – широкой общественности, правительственных чиновников или экспертов. Знание целевой аудитории поможет определить необходимое количество методической информации и характер пресс-релиза.

Пресс-релиз помогает уменьшить тенденцию к упрощению результатов оценки, но, к сожалению, не устраняет ее полностью. Некоторые репортеры вместо нейтрального сбалансированного освещения результатов склонны выдвигать на первый план результаты, имеющие политически разрушительные последствия.

Пресс-релиз необходимо начинать с указания даты выпуска, а также наименования и адреса ответственной организации

(министерство образования, научно-исследовательский институт или другое агентство). Если это возможно, то в заголовке помещают ее логотип. Заголовок нужно набирать жирным шрифтом, он должен быть кратким и интересным, чтобы привлечь внимание читателей. Так, «Сообщение о национальной оценке» – просто короткий заголовок, а «Новый отчет, освещающий успехи в секторе образования» – заголовок еще и привлекательный. В сообщении обязательно должна быть начальная или ведущая фраза, направленная на пробуждение интереса читателя и дополненная одним или двумя высказываниями, которые уточняют ведущую идею.

Основной раздел пресс-релиза должен содержать факты и краткие ответы на следующие вопросы.

- Кто осуществлял национальную оценку?
- Почему она проводилась?
- Когда она была проведена?
- Как она выполнялась?
- Каковы были основные результаты?
- Почему они важны?

Разработчикам пресс-релиза обычно дают следующие рекомендации.

- Составьте ясное представление о том, какие выводы ожидают читатели от пресс-релиза.
- Ограничьте материал непосредственными фактами и не приукрашиваете результаты.
- Избегайте длинных предложений, методических терминов и специальной статистической лексики.
- Используйте глаголы действительного залога.
- Используйте фразы, которые похожи на те, что могли бы появиться в газетах.
- Ограничьте релиз одной или максимум двумя страницами с текстом (используйте двойной пробел).
- Проверьте текст на предмет правильности фактов и наличие грамматических ошибок.

- Обязательно укажите контактные данные: телефонный номер, почтовый адрес и адрес электронной почты.
- Чтобы обозначить конец пресс-релиза, используйте знак ### под последней строчкой текста.

Если используются диаграммы, они должны быть ясными и легко читаемыми. В них не должно быть слишком много деталей или сносок. Диаграммы лучше представлять в форме столбцов, а не линий, это считается более эффективным (Hambleton and Slater, 1997).

В дополнительном материале 3.3 представлены выдержки из пресс-релиза, занимающего две страницы и подготовленного в 2007 году по материалам национальной оценки в США. Другой пример пресс-релиза приведен в дополнительном материале 3.4.

ПРОВЕДЕНИЕ ПРЕСС-КОНФЕРЕНЦИЙ

Для привлечения широкого диапазона СМИ, включая представителей прессы, радио и телевидения, проводится пресс-конференция, в ходе которой руководство национального комитета по управлению и национальный координатор исследования представляют результаты оценки. Сама презентация должна занимать приблизительно 20 минут, еще 10 минут следует оставить для вопросов. Опыт показывает, что зачастую докладчики отступают от согласованных ранее текстов сообщений, поэтому им лучше предварительно попрактиковаться и заранее продумать ответы на возможные вопросы.

Участники должны получить копию пресс-релиза, итоговый отчет о результатах национальной оценки, резюме с биографическими примечаниями о выступающих и фоновую информацию по агентству, которое проводило оценку (если это была структура, отличающаяся от министерства образования).

Дополнительный материал 3.3**Выдержка из пресс-релиза, программа NAEP: США****ПРЕСС-РЕЛИЗ**

**Внимание: информация сохраняет свою актуальность до вторника,
25 сентября, 10:00 утра по восточному времени**

КОНТАКТЫ: Мэтт Маурер, (202) 955-9450, доб. 322,
mmaurer@communicationworks.com

**США, Прогресс учащихся по математике и чтению,
в соответствии с Национальным отчетом о достижениях™
Группа меньшинства показала наивысший прирост результатов**

ВАШИНГТОН (25 сентября 2007 года). Согласно результатам Национального отчета по успеваемости™ 2007 года, уровень достижений учащихся по математике и чтению в США в целом возрастает. Часть наиболее высоких достижений принадлежит представителям группы меньшинства выборки учащихся.

Два отчета, выпущенные в настоящий момент: Национальный отчет по успеваемости™: Математика-2007 и Национальный отчет по успеваемости™: Чтение-2007 – детализируют достижения учащихся 4-х и 8-х классов в программе NAEP, которую ранее в этом году осуществил Департамент образования США. В отчетах национальные данные и данные штата за 2007 год сравниваются с результатами тестов предшествующих периодов, начиная с 1990 года по математике и с 1992 года по чтению. Основываясь на национальных средних баллах, можно сделать вывод, что знания по математике учащихся 4-х и 8-х классов с 1990 года продолжают повышаться. Кроме того, доля учащихся, показавших результаты, соответствующие базовому и профессиональному уровню или выше, за прошедшие годы заметно увеличилась. Начиная с 2003 года прирост достижений является статистически значимым, хотя не столь большим, как за несколько более ранних периодов. Тем временем, средний балл по чтению для учащихся 4-х классов была самым высоким за 15 лет и увеличился с 2003 года, хотя полный прирост достижений с 1992 года был более скромным, чем отмеченный прирост по математике. Средний балл для учащихся 8-го класса по чтению с 2005 года

Дополнительный материал 3.3

увеличился незначительно, но остается ниже уровня достижений, показанного в 2002 году, и является примерно таким же, как средний балл в 1998 году.

...

Копии Национального отчета успеваемости™: Математика-2007 и Национального отчета успеваемости™: Чтение-2007 и обширная информация по 2007 году по математике и чтению программы оценки NAEP будут доступны на сайте <<http://nationsreportcard.gov>> 25 сентября в 10:00 утра по восточному времени.

#

Национальный отчет об успеваемости является единственным представительным на национальном уровне материалом, продолжающим оценку условий образования в США. Он служит национальным критерием достижений учащихся с 1969 года. Посредством программы NAEP Национальный отчет об успеваемости сообщает общественности о том, что учащиеся Америки знают и могут выполнить в различных предметных областях, а также сравнивают данные достижения между штатами и различными демографическими группами учащихся.

Источник: http://www.nationsreportcard.gov/math_2007/media/pdf/newsrelease.pdf.

Дополнительный материал 3.4

Пример пресс-релиза: США

Этот пример пресс-релиза был издан министерством образования штата Мичиган в материале «Опустите карандаши. Руководство для использования и сообщения результатов тестирования».

ДЛЯ НЕМЕДЛЕННОГО ВЫПУСКА:

ДАТА:

КОНТАКТЫ:

Связи с общественным офисом

Мэри Грин

555-1313

Дополнительный материал 3.4

Школы округа Мэсвилл приводят несоответствие программ обучения и оценки как возможную причину маленького прироста в тестовых баллах.

Учащиеся 4-го, 7-го и 10-го классов школ округа Мэсвилл показали только небольшое повышение результатов в некоторых областях учебного плана, проверенных по Программе оценки образования в штате Мичиган (MEAP) с помощью тестов, предъявленных учащимся в конце сентября.

В сообщении для министерства образования штата, где расположен округ Мэсвилл, руководитель Фред Смарт отметил, что, «хотя мы рады видеть некоторое повышение результатов, необходимо проделать значительную работу, особенно в области математики». Несмотря на то что результаты тестирования учащихся округа Мэсвилл оказались выше среднего в штате Мичиган по нескольким предметам, Фред Смарт сказал: «Мы рассматриваем нашу математическую программу в свете того факта, что многие из протестированных понятий не преподаются нашим учащимся до даты испытания».

___ % четвероклассников школ округа Мэсвилл, ___ % семиклассников и ___ % десятиклассников школ, достигших освоения трех четвертей учебных целей на каждом уровне.

Смарт указывает, что математика для 7-го класса должна быть подвергнута значительному исследованию районной комиссией по изучению математики. Комиссия по изучению математики составлена из учителей с 1-го по 9-й класс, консультанта по математике школьного района и координатора школ.

___ % четвероклассников школ округа Мэсвилл, достигших освоения трех четвертей учебных целей по чтению по сравнению с ___ % учащихся, полученным для всего штата.

Баллы семиклассников были выше, чем среднее по штату по ___ целям из 25 проверенных целей по чтению, среди них ___ % учащихся овладели по крайней мере тремя четвертями целей по сравнению с ___ % учащихся в среднем по штату. ___ % десятиклассников овладели тремя четвертями целей по чтению по сравнению с ___ % учащихся, полученным для всего штата.

Тесты по программе MEAP выдаются каждому 4-му, 7-му и 10-му классу в государственных школах штата Мичиган каждую осень. Это тестирование требуется в соответствии с законом Мичигана с 1969 года. Тест используется учителями для проектирования программ, с тем чтобы они отвечали специальным потребностям учебной деятельности учащихся.

Дополнительный материал 3.4

Программа MEAP помогла определить, овладели ли учащиеся специальными умениями по чтению и математике. Тестовые вопросы берутся из банка вопросов, предназначенных для отражения базовых умений и рассматриваемых как обязательные учителями Мичигана.

Копии результатов оценки по районам доступны для общественности района в административных офисах, 242, 13-я улица.

Источник: Gucwa and Mastie, 1989. Воспроизводится с разрешения министерства образования штата Мичиган.

Иногда на подобных пресс-конференциях присутствуют элементы риска. У некоторых журналистов и заинтересованных пользователей может возникнуть желание вызвать сенсацию, что приведет к искажению ключевых сообщений. Представители СМИ часто задают вопросы, которые касаются именно плохих результатов. У популярных СМИ во многих странах существует тенденция к упрощенному объяснению сложных проблем, таких как причины снижения или роста стандартов достижений учащихся. Важно, чтобы пресс-конференция была использована для исправления непредусмотренных выводов, таких как переложение вины за плохие результаты на единственную группу участников (например, на учителей или министерство образования) или выделение единственной причины неудач (например, плохие учебники).

ПРОВЕДЕНИЕ ОТДЕЛЬНЫХ БРИФИНГОВ

Эффективности основного сообщения по итогам национальной оценки зачастую больше способствует даже не пресс-конференция, а проведение встреч с отдельными людьми. Ин-

дивидуальное отражение событий по результатам программ национальной оценки помогает докладчику установить контакты с журналистами, с радио- или телерепортерами. Такое представление событий также позволяет сократить время на разъяснение возможных неправильных представлений о результатах оценки.

РЕГИСТРАЦИЯ СООБЩЕНИЙ НА САЙТЕ

Рост числа посетителей сайта в будущем может гарантировать, что больше людей получают доступ к результатам национального оценивания. В Чили детализированные результаты по программам национальной оценки представляются на сайте (<http://www.simce.cl>). Министерство образования и науки Ирландии приводит резюме по результатам национальной оценки на своем официальном сайте (http://www.education.ie/servlet/blobServlet/des_cos_preface.htm). На странице сайта с главным меню для достижений по математике просто перечислены заголовки глав отчета, каждую из которых можно загрузить путем щелчка по соответствующему названию (дополнительный материал 3.5).

Ниже мы приводим страницу американского сайта для программы NAEP. Он дает заинтересованным педагогам и представителям общественности ответы на определенные вопросы с помощью проведения сравнительного анализа данных оценки (дополнительный материал 3.6).

Приступая к работе с сайтом <http://nces.ed.gov/nationsreportcard>, малокомпетентные или абсолютно некомпетентные в статистике пользователи могут получить ответы на следующие вопросы, связанные с результатами выполнения тестов в их собственном штате.

Дополнительный материал 3.5**Сайт. Достижения по математике в начальных школах:
Ирландия**

На сайте Ирландского департамента образования приведено содержание национального отчета об оценке достижений по математике. Чтобы посмотреть соответствующие разделы отчета, пользователи могут просто кликнуть на следующие названия.

Предисловие

Глава 1 – Контекст оценки

Глава 2 – Инструментарий и процедуры оценки

Глава 3 – Достижения учащихся по математике в 4-м классе

Глава 4 – Характеристики учащихся и достижения по математике

Глава 5 – Домашние фоновые факторы и достижения по математике

Глава 6 – Окружающая среда в классе и достижения по математике

Глава 7 – Характеристики школы и достижения по математике

Глава 8 – Поддержка изучения математики

Глава 9 – Взгляды инспекторов на обучение и изучение математики

Глава 10 – Рекомендации

Приложения

Глоссарий

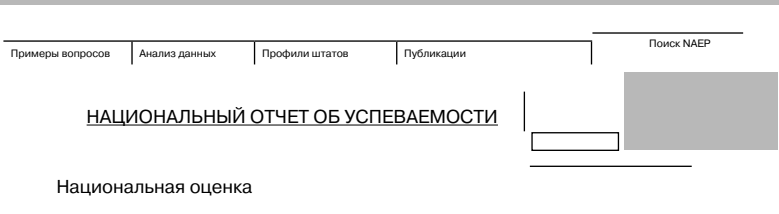
Источник: http://www.education.ie/servlet/blobServlet/des_cos_preface.htm.

- Как изменились стандарты выполнения тестов (повысились или понизились) начиная с момента выполнения предыдущего исследования по программе NAEP?
- Есть ли гендерные отличия в достижениях по чтению?
- Как результаты учащихся из групп меньшинства сравниваются с общим национальным средним баллом по чтению?

- Как штат сравнивает свои баллы с такими же результатами в соседних штатах или с общим национальным средним баллом по естественным наукам?

Дополнительный материал 3.6

Карта сайта, программа NAEP Web Site: США



О программе NAEP	Предметы NAEP	Специальные правила
Обзор FAQs Деятельность NAEP Информация для отобранных школ Информация для родителей Включение SD/ELL учащихся Национальная программа NAEP Штаты NAEP Городские районы NAEP (TUDA) Длительные тенденции NAEP Исследования по расшифровке средних школ Исследования по обучению индейцев Специальные исследования Партнеры Правительственный комитет по национальной оценке	Искусство Граждановедение Экономика Иностранный язык География Математика Чтение Естественные науки История США Письмо Мировая история	Вопросы по инструментарию NAEP Исследование данных NAEP Сравнения по штату Профили штатов <hr/> <p style="text-align: center;">Дополнительные ресурсы</p> Контакты Публикации Глоссарий терминов Расписание оценивания Новый сервис Ни одного отстающего ребенка Техническая документация Исследовательский центр Фоновые вопросы Пример буклета вопросов Архив

Источник: США, Национальный центр образовательной статистики, 2008.

СОЗДАНИЕ ДОСТУПНЫХ ДАННЫХ ОЦЕНКИ

Фактические данные оценки со временем часто превращаются в заброшенный актив. Однако существуют различные организации и частные лица, которые заинтересованы в проведении вторичных исследований. Это представители министерств, отличных от министерства образования (например, здравоохранения, социальных работ или финансов), ученые-исследователи, представители факультетов образовательных учреждений и спонсорских агентств, находящихся в поиске справедливых аргументов для поддержки национальных или целевых вложений в образование. При этом существуют барьеры, ограничивающие привлечение специалистов или агентств, отличных от команд по национальной оценке. Они касаются не только методики использования, но и затрагивают проблемы частной жизни и конфиденциальности.

Традиционный метод удовлетворения потребностей вторичных пользователей должен быть приспособлен для получения и публикации большого числа объемных статистических таблиц. Эти публикации, однако, являются трудными для использования, а таблицы, которые они содержат, часто не удовлетворяют точные потребности пользователей. Коммерческие продукты, такие как iVision, позволяют получить легкий доступ к большим объемам предварительно запланированных таблиц.

Многие системы национальной оценки заинтересованы в разработке для пользователей файлов с микроданными и связанной с ними методической документации, которая высвечивается на экране, чтобы защитить конфиденциальность результатов отдельных учащихся, учителей и школ. В отличие от обычных агрегированных данных, разрешающих дальнейшую компьютерную статистическую обработку, микроданные доступны только для специализированных обзоров с индиви-

дуальными сценариями доступа. Они включают специальные переменные, которые позволяют идентифицировать пользователя (например, идентификационный номер служащего, код доступа и др.). Некоторые микроданные менее чувствительны, чем другие, относительно режима конфиденциальности при раскрытии. Обычно социальные обзоры содержат меньше степеней защиты, чем деловые обзоры, которые защищены намного больше.

Там, где пользователям требуется доступ к данным по учащимся, команды по национальной оценке могут настроить обслуживание отдаленного доступа, по которому внешние пользователи для проведения анализа вводят свой пароль, управляющий доступом к базе конфиденциальных данных. Результаты тогда исследуются с соблюдением конфиденциальности. Хотя это несколько тяжело, тем не менее такой метод позволяет командам по национальному исследованию защитить конфиденциальность, важность которой не подвергается никакому сомнению, поскольку участие стран часто зависит от гарантий того, что результаты отдельных школ, учителей или учащихся не будут обнародованы.

Команды по национальным исследованиям, возможно, должны убедить политических деятелей и официальных представителей министерств в том, что создание доступных данных национальной оценки входит в сферу их долгосрочных интересов, даже если это может вызвать некоторую краткосрочные трудности из-за попадания данных в руки потенциальных критиков.

ДРУГОЙ ИНСТРУМЕНТАРИЙ ДЛЯ РАСПРОСТРАНЕНИЯ ДАННЫХ

Другой инструментарий, используемый для распространения результатов национальной оценки, включает конференции, платные курсы, семинары, информационные бюллетени и брошюры, ориентированные на специфические группы за-

интересованных лиц (например, представителей министерства или учителей).

Конференции, платные курсы и семинары обеспечивают механизмы для рекламы соответствия результатов оценки потребностям ключевых заинтересованных лиц множеством способов. Они обеспечивают возможность достижения согласия по значению наиболее важных результатов и по шагам, которые должны быть сделаны, чтобы исправить любые выявленные проблемы.

Эти методы являются наиболее подходящими для информирования следующих групп:

- Разработчики учебного плана, которые могут найти информацию в результатах национальной оценки, выявляющую недостатки в достижениях учащихся и имеющих значение для пересмотра учебного плана.
- Авторы учебников, которые, как в случае разработчиков учебного плана, возможно, должны пересмотреть учебники и вспомогательные материалы (например, руководства для учителей) в свете свидетельств достижений учащихся, показанных в национальной оценке.
- Педагоги, обучающие учителей школ (обучение до начала работы и в системе повышения квалификации), которые могут использовать результаты национальной оценки для идентификации недостатков в подготовке учителей или в их педагогической практике.

ВЫВОДЫ

Многие страны принимают на вооружение подходы к распространению результатов национальной оценки. Например, в Колумбии результаты были представлены с помощью средств массовой информации. Кроме того, был организован цикл национальных и местных семинаров для обсуждения результатов оценки и их значения. На национальном семинаре также обсуж-

дались стратегии улучшения качества образования. Местным администраторам и отдельным учителям отправили информацию относительно национальных и региональных результатов в виде информационных бюллетеней и брошюр. В Ирландии широкое освещение результатов оценки осуществлялось с помощью основного отчета, итогового отчета (нацеленного прежде всего на потребности учителей) и пресс-релизов. О результатах также сообщали по радио и телевидению и в газетах. Совместно с министерством образования и агентством, ответственным за осуществление национальной оценки (Образовательный научно-исследовательский центр) был организован национальный семинар, в котором приняли участие администраторы, учителя и ученые.

Многие руководящие органы по национальной оценке сообщают результаты не вполне адекватно по отношению к тем разнообразным группам индивидуумов и структур, которые имеют вполне легитимный интерес к информации об учебной деятельности учащихся в школах. Этот недостаток неизбежно накладывает серьезные ограничения на использование результатов. При анализе этой ситуации могут приниматься во внимание различные причины, такие как отсутствие интереса со стороны части руководителей министерства; недостаточное развитие инфраструктуры, сопутствующей деятельности по национальной оценке; недооценка роли, которую множество заинтересованных лиц могут играть в реагировании на результаты национальной оценки; нехватка ресурсов для подготовки нескольких отчетов и управления множеством информационных каналов, связанных с потребностями потенциальных пользователей.

Как только системы национальной оценки достигнут зрелости, будет выдвинуто множество требований к ограниченным ресурсам, доступным для оценки. Это требования, связанные с частотой оценки, повышением качества методического обеспечения систем оценки, совершенствованием навыков

персонала, который выполняет оценку и повышением способности политиков и управленцев в образовании к осмыслению и использованию информации, полученной в результате оценки. В этой ситуации трудно сделать правильный выбор относительно оптимального распределения ресурсов. При осуществлении такого выбора нельзя оставить без внимания важность процедур для оптимизации использования оценки с помощью сообщения результатов в форме, адекватной запросам заинтересованных сторон.

ПЕРЕВОД РЕЗУЛЬТАТОВ ОЦЕНКИ В ПОЛИТИКУ И ДЕЙСТВИЯ

Первичная цель системы национальной оценки – дать общее представление о результатах учебной деятельности учащихся. Однако ее роль этим не ограничивается. Чтобы оправдать потраченные усилия и понесенные расходы, полученная в ходе оценки информация о сильных и слабых аспектах подготовки испытуемых и о распределении достижений в выборке учащихся (в частности, по гендерному признаку или местоположению школы), должна быть полезной для разработки новой политики и принятия важных решений, касающихся, например, ресурсов. Также необходимо тщательно продумать вопрос о том, как обеспечить информацией разработчиков учебного плана, авторов учебников, преподавателей педагогических вузов и общественность. С этой точки зрения оценка должна не только предоставить необходимые сведения, но и гарантировать их последующее распространение. Благодаря этому она станет главной движущей силой реформ в образовании.

В данной главе раскрыты пять важных вопросов, связанных с тем, как перевести информацию о результатах оценивания в сферу политики и конкретных действий: установление способности к восприятию и использованию информации;

надежность и релевантность информации, предоставленной оценкой; процедуры для идентификации соответствующей политики или действий после оценки; определение системного или целенаправленного вмешательства; сложность формирования политики и принятия решений. Основное внимание сосредоточено на том, какие меры требуется принять в системе образования. Последующие главы обращены к более специфическому использованию национальной оценки для политики и управления в сфере образования (глава 5) и обучения (глава 6).

СФОРМИРОВАННАЯ СПОСОБНОСТЬ К ВОСПРИЯТИЮ И ИСПОЛЬЗОВАНИЮ ИНФОРМАЦИИ

Чтобы достичь оптимального использования результатов национальной оценки, необходимо соблюдать ряд условий, связанных с формированием способности к восприятию (Kellaghan and Greaney, 2004; Postlethwaite, 1987; Vegas and Petrow, 2008). Во-первых, со стороны разработчиков политики очень важно проявить политическое желание, открытость к новой информации и готовность к изменениям и проведению реформ. Во-вторых, политики и лица, принимающие решения в министерстве образования, должны обладать способностью (знаниями и умениями) к интерпретации и использованию информации о результатах национальной оценки. Для развития этой способности могут потребоваться значительные вложения, особенно в тех странах, где национальная оценка проводится впервые. В-третьих, нельзя допустить, чтобы национальная оценка превращалась в изолированную деятельность. Ее необходимо интегрировать в существующие структуры. Она должна осуществляться параллельно с выработкой политики, принятием решений и распределением ресурсов.

В-четвертых, команда по национальной оценке должна гарантировать выбор таких систем и стратегий, которые обеспе-

чат сообщение результатов оценки учреждениям и агентам, играющим важную роль в осуществлении политики (например, представителям местных органов власти; наблюдателям, инспекторам и советникам; руководителям процесса развития учебного плана; руководителям школ и учителям). В-пятых, после того, как была сформирована способность к восприятию информации, необходимо обеспечить непрерывность участия сотрудников в организации оценки и в интерпретации ее результатов. Систематический обмен официальными и методическими экспертами, который является общим явлением во многих развивающихся странах, окажет негативное воздействие на способность принимать информацию. Наконец, со стороны всех заинтересованных лиц требуются поддержка и выполнение обязательств. Следовательно, цель, результаты и значение национальной оценки должны быть ясно сообщены именно тем заинтересованным лицам, которые чувствуют свою зависимость от политики или политических решений.

СОСТОЯТЕЛЬНОСТЬ И РЕЛЕВАНТНОСТЬ ИНФОРМАЦИИ, ОБЕСПЕЧИВАЕМОЙ ОЦЕНКОЙ

Прежде чем рассмотреть, какие решения можно принимать исходя из результатов национальной оценки, специалисты, принимающие решения, должны убедиться в том, что информация, предоставленная системой оценки, заслуживает доверия и является реалистичной. Это можно гарантировать, если: команда по национальной оценке будет методически компетентна; знания и умения учащихся, для измерения которых планировался тест, адекватно представлены в оценочном инструментарии; процессы сбора и анализа данных сопровождались надлежащими процедурами; учебная деятельность учащихся описана в соответствующих деталях и способом, который чувствителен к потребностям и методической экспертизе пользователей (глава 1; см. также Postlethwaite, 2004a, глава 5).

Политики и управленцы запрашивают информацию, связанную с касающимися их проблемами и обеспечивающую основания для формирования политики, планирования и принятия решений. Существует достаточно много примеров систем национальной оценки, которые не в состоянии удовлетворить эти информационные потребности. Например, если тест, используемый в процессе оценки, будет слишком трудным и не позволит дифференцировать учащихся с более низкими уровнями достижений, то оценка не будет обеспечивать информацию, на основании которой старшие чиновники министерства могут принять решение о распределении ресурсов для учащихся с низким уровнем достижений. Результаты оценки, которые обеспечивают только средние баллы или ранжируют географические регионы на их основе без анализа достижений, которые эти баллы представляют, не будут обеспечивать достаточно подробную информацию для разработчиков учебных планов и учителей педагогических вузов. Наконец, отчеты о программах оценки, которые длятся три или четыре года до момента завершения, вряд ли будут очень интересны для разработчиков политики, потому что результаты, возможно, устареют и не будут более иметь отношение к проблемам или персоналу, проводившему оценку.

ПРОЦЕДУРЫ ДЛЯ ИДЕНТИФИКАЦИИ СООТВЕТСТВУЮЩЕЙ ПОЛИТИКИ ИЛИ ДЕЙСТВИЙ ПО РЕЗУЛЬТАТАМ ОЦЕНКИ

В некоторых случаях результаты национальной оценки диктуют выбор непосредственных действий, которые необходимо осуществить для решения выявленной проблемы. Но чаще требуемые действия не являются столь очевидными. Оценка помогает идентифицировать проблемы, но не приводит к принятию конкретного решения или к выбору между альтернативными направлениями действий. Например,

оценка может выявить, что учащиеся плохо выполнили тест в некоторых областях, установить связь между достижениями и фоновыми факторами, указывающую, что баллы у девочек ниже, чем у мальчиков, или что учащиеся маленьких сельских школ выполняют тест хуже, чем учащиеся городских школ. Однако сама идентификация фактов, связанных с достижениями учащихся, не дает объяснения, почему гендерные признаки или принадлежность учащихся к школам с различным местоположением приводят к появлению различий в их достижениях. Многие системы национальной оценки частично обращаются к этой проблеме в процессе сбора дополнительных сведений об условиях, в которых проходит учебная деятельность. Такой анализ, хотя и не приводит к конкретным действиям, способствует лучшему пониманию факторов, влияющих на результаты и обеспечивающих базис для формирования политики и решений (Blalock, 1999).

Помимо этого, политики и лица, принимающие решения, должны быть осведомлены о выполнимости и рентабельности последующих инициатив или реформ. Ясно, что приоритет отдается определению входных условий, которые имеют значение для учебной деятельности. Таким образом, некоторые даже очень привлекательные входные условия (например, школьное оборудование) могли бы быть очень хорошими, но они не могут быть более эффективными, чем более скромные входные условия обучения. Также вполне уместно учесть влияние стоимости обучения на бюджет образования. В малочисленных классах результаты обучения учащихся могли бы быть выше, однако стоимость перехода к таким классам в развивающихся странах может быть слишком высока, поэтому отнести такую стратегию к числу жизнеспособных нельзя (Lockheed and Verspoor, 1991).

В этой главе установлены четыре (не взаимоисключающие) процедуры, в которых результаты оценки могут играть роль при выборе соответствующей политики или действий после оценки.

Обсуждение результатов

Наиболее важно после национальной оценки привлечь заинтересованных лиц к обсуждению и интерпретации его результатов. Это поможет извлечь нужную информацию, выдвинуть предположения о причинных связях и предложить подходы для решения выявленных проблем. В интерпретации результатов и поиске решений значительную роль играют источники информации, экспертизы и многие заинтересованные лица. Отдельно следует выделить вклад людей, деятельность которых так или иначе связана со школой (учителя, инспекторы, наблюдатели или советники).

Во многих странах после выполнения программ национальной оценки проводятся семинары и платные занятия. Благодаря этому становится возможным услышать множество точек зрения. Также используются другие, более формальные подходы к рассмотрению результатов – например, обсуждение их национальной комиссией или при выборе национальной стратегии сектора образования (см. главу 5).

Предоставление возможности «прояснить» процесс выработки политики с помощью результатов

Использование результатов некоторого ориентированного на политику исследования в образовании предполагает, что влияние результатов национальной оценки должно быть скорее концептуальным, нежели прямым. Таким образом, результаты оценки не используются на арене политики путем прямого применения, а выполняют обобщающую, ориентирующую и общую руководящую роль, благодаря чему они «проясняют» процесс

выработки политики, формируя способы осмысления проблем и информирования при проведении общих дебатов и обсуждений (Husyn, 1984; Weiss, 1979). При такой точке зрения использование результатов исследования является постепенным и распределенным процессом, обеспечивающим концептуальные и теоретические перспективы, которые проникают в процесс выработки политики, воздействуя на понимание образовательных проблем и предлагая их соответствующие решения.

Такой процесс «прояснения» не должен быть ограничен сферами политики и принятия решений. Предоставление результатов национальной оценки общественности может способствовать их более глубокому осмыслению не только политиками и управленцами в образовании, но также всеми членами общества вообще. Рост осведомленности общественности, в свою очередь, поможет подчеркнуть важную роль образования, которую оно играет на стыке национальных целей и общественного мнения, чтобы поддержать усилия общества по улучшению условий образования.

Хотя преимущества такого прояснения довольно очевидны, подобное использование результатов исследования может иметь недостатки. Во-первых, этот процесс открыт для всяческих упрощений и искажений. Во-вторых, плохое исследование может столь же сильно привлечь внимание общественности, как и хорошее. В-третьих, некоторые важные результаты могут не достичь сфер политики или принятия решений. Использование результатов национальной оценки для прояснения заинтересованным лицам (включая общественность) различных аспектов системы образования в государстве должно всячески поощряться, но необходима и другая дополнительная работа. Министерство образования и иные учреждения и ключевые заинтересованные лица должны тщательно изучить результаты национальной оценки, чтобы разработать определенную политику и действия для обращения к выявленным недостаткам.

Реакция на определенные проблемы, выявленные национальной оценкой

В некоторых системах национальной оценки действия, которые нужно предпринять, вполне очевидны. Так, если будет установлено, что у школ нет ресурсов, регламентированных в инструктивных материалах министерств (например, требуемого числа учебников), необходимо выполнить те шаги, которые эти ресурсы обеспечат. Если учителя недостаточно знают свой предмет, то самым эффективным способом исправления подобной ситуации являются курсы повышения квалификации, нацеленные на выявленные недостатки в подготовке учителей.

Соотнесение с другими результатами исследования

Рассмотрев другие результаты исследования, мы можем сделать выводы относительно причинных связей, с помощью которых национальная оценка обеспечит формирование политики на более безопасной основе. Эти результаты являются следствием множества исследований, включая исследования эффективности деятельности учителей и школы; исследования относительного влияния школьных и домашних факторов на учебную деятельность учащихся и исследования процессов, протекающих в классе, его размеров и эффектов повторного обучения в том же классе. В качестве предложения о том, как можно использовать результаты национальной оценки для улучшения обучения в классе и учебной деятельности учащихся, в главе 6 приведен литературный обзор в качестве руководства, в частности, при исследовании эффективности деятельности школ и учителей.

Использовать результаты исследования для подкрепления выводов, полученных по данным национальной оценки, необходимо с осторожностью. Возможно, исследования не будут ме-

тодически адекватными или их результаты не будут относиться к тому контексту, в котором проводилась национальная оценка. Потенциальных проблем станет больше в том случае, если исследование было выполнено в других странах. В частности, если это экономически высокоразвитые страны, то результаты исследования вряд ли будут релевантны проблемам развивающейся страны. Например, заключение о том, что очень высокие достижения учащихся связаны с малочисленностью классов, как правило, основывается на исследованиях, проводившихся в США, где в классе было 20 или меньше учащихся (см., например, Finn and Achilles, 1990). В развивающихся странах классы такой численности встречаются редко.

ОПРЕДЕЛЕНИЕ ШИРОКОМАСШТАБНОГО СИСТЕМНОГО ИЛИ ЦЕЛЕВОГО ВМЕШАТЕЛЬСТВА

При определении конкретного действия важно решить, будет ли оно носить широкомасштабный системный характер или предназначается для целенаправленного воздействия на отдельные категории учащихся или секторы системы образования. Широкомасштабное системное действие направлено на улучшение результатов деятельности всех учащихся. Оно включает улучшение методов преподавания, реформирование учебного плана и учебников. Инициативы по обучению педагогов могут быть как системными, так и целенаправленными. Иные целевые вмешательства включают множество условий. Во-первых, вмешательство может предоставить дополнительные ресурсы учащимся со специальными потребностями в осуществлении учебной деятельности (например, ранние программы предотвращения отставания в обучении или коррекционные программы по чтению). Во-вторых, вмешательство может включать различные действия и политические решения, такие как устранение или снижение влияния трудностей в педагогической деятельности (например, нехватка учебных материалов или плохо

выполняющих свои обязанности учителей) и экономических барьеров (например, плата за обучение в школе или затраты на приглашение дополнительных учителей). Эти меры могут непосредственно влиять на результаты деятельности некоторых категорий учащихся, например учащихся из семей с более низким социально-экономическим статусом; детей, переданных в другие семьи из-за нарушений общественного порядка; сирот; детей из «лингвистических» меньшинств, учащихся маленьких сельских школ; девочек (в некоторых обществах). В-третьих, вмешательство может быть запланировано таким образом, чтобы косвенно воздействовать на познавательную деятельность учащихся, например, программы привлечения родителей к обучению или школьные программы льготного питания (см. Willms, 2006; Всемирный банк, 2004). Целевое вмешательство часто требует, чтобы школы разработали и представили план, который описывает, как они предлагают использовать дополнительные ресурсы для улучшения учебной деятельности учащихся.

СЛОЖНОСТЬ ФОРМИРОВАНИЯ ПОЛИТИКИ И ПРИНЯТИЯ РЕШЕНИЯ

В главе 1 было отмечено, что диапазон политических факторов может играть важную роль в определении формы национальной оценки. Теперь мы подробно рассмотрим вопрос о сложности формирования политики и принятия решения после оценки в окружающей политической среде.

Выработка политики – не простой линейный процесс, в ходе которого выявляется проблема, рассматриваются возможные решения и разрабатываются стратегии их реализации. В большей степени он включает сложную политическую деятельность, в том числе проведение переговоров, поиск компромисса и ответы на давление и лоббирование, следующие из широкого диапазона источников, которые соотносятся с существующими образцами веры и системами ценностей (см.

Blalock, 1999). Принимая решение, политик может руководствоваться результатами оценки, но он или она также должны учитывать многочисленные политические и другие аспекты и на национальном, и на локальном уровнях. Это взгляды политических деятелей (идеологического характера или положения, касающиеся ответов избирателям и электорату), доступность бюджетных средств, представления и интересы заинтересованных лиц и заинтересованных групп, традиционные представления и современные увлечения.

Связи между национальной оценкой и политическими факторами имеют некоторые плюсы. Например, политические силы иногда могут применяться для осуществления поддержки при использовании результатов оценки. В частности, поддержка для формирования политики и принятия решений, основанных на результатах оценки, возрастет, если ключевые заинтересованные лица были активно вовлечены в процесс оценки на ранней стадии его проектирования. Конечно, на всех этапах заинтересованные лица должны иметь отношение к соответствующим аспектам формирования политики и принятия решения, чтобы гарантировать, что они понимают потребности и аргументы для проведения реформ в образовании. Привести к завершению этого процесса поможет представление интересов заинтересованных лиц руководящему комитету, который контролирует выполнение программы национальной оценки (см. Greaney and Kellaghan, 2008 – первую книгу в этой серии). Политики в процессе поиска поддержки для проведения реформ также нуждаются в обращении к принятым в обществе принципам, таким как обеспечение равных возможностей для всех учащихся, гарантия того, что качество результатов учебной деятельности учащихся оправдывает понесенные расходы, и потребность в обеспечении устойчивого базиса для развития как учащихся, так и национальной экономики.

Тесная связь между системой национальной оценки и политической ареной, хотя и содержит в себе потенциальные возможности для улучшения процессов формирования политики и принятия решений, включает в себя элементы опасности. Если оценка будет вовлечена в политический конфликт, она не сможет играть ведущую роль в повышении результатов учебной деятельности учащихся. В Аргентине, вследствие напряженных отношений между центральными и провинциальными властями по вопросам сфер их влияния, информация, полученная в ходе оценки, использовалась прежде всего для узаконивания национальной политики и реформ, а также для регулирования статуса провинциальных властей, а вовсе не для проектирования политики компенсаторного обучения в образовании (Benveniste, 2002; Gvirtz and Larriga, 2004). Схожая ситуация возникла в Египте, где местные органы управления образованием отказались от сотрудничества в проектировании программы национальной оценки, потому что чувствовали, что она приведет в скрытой форме к увеличению контроля их деятельности со стороны центрального правительства (Carrol, 1996).

В Уругвае также имел место конфликт между руководящими заинтересованными лицами, однако ситуация там была несколько иной. Опасаясь роста собственной ответственности в результате оценки, союзы учителей долгое время оставались в оппозиции, пока не было достигнуто соглашение по ряду важных пунктов. Так, было решено, что отчеты о результатах деятельности школ публиковать не нужно; что необходимо уделить должное внимание влиянию фоновых факторов на достижения учащихся; что учителя не должны нести полную ответственность за результаты учащихся (Benveniste, 2002; Ravela, 2005). Конфликт был разрешен путем переговоров. Результаты оценки не привели к возложению

ответственности на учителей, и в настоящее время Уругвай является одним из наиболее интересных примеров использования результатов национальной оценки для улучшения результатов учебной деятельности учащихся (см. главу 6).

ВЫВОДЫ

В данной главе мы рассмотрели, как влияют результаты национальной оценки на формирование политики государства, осуществление конкретных действий после нее. Это помогло выявить две главные проблемы. Первая связана со сложностью принятия решений, развитием способности к восприятию и использованию информации и необходимостью учета интересов множества лиц. Вторая касается интерпретации результатов оценки и выбора способов улучшения результатов учебной деятельности учащихся.

Можно утверждать, что базис для принятия решения усилится, если для стратегий по улучшению результатов учебной деятельности учащихся, предложенных на основе результатов национальной оценки, будут проведены в жизнь программы оценки в экспериментальном или квазиэкспериментальном исследованиях (см. главу 8). Однако по ряду причин, таких как временные ограничения, стоимость и отсутствие грамотного персонала с необходимыми методическими навыками, это маловероятно. В большинстве случаев политики и лица, принимающие решения, будут возлагать надежды на то, что распространение результатов оценки приведет к лучшему осмыслению и пониманию проблем образования, обсуждению результатов заинтересованными лицами и проведению соответствующего исследования, пусть даже и не в той стране, в которой была проведена оценка.

РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ, ПОЛИТИКА И УПРАВЛЕНИЕ В ОБРАЗОВАНИИ

Результатами национальной оценки в первую очередь пользуются разработчики политики – политические деятели, администраторы в образовании и управленцы, в частности руководители структурных подразделений министерства образования. Даже в тех случаях, когда результаты имеют значение для работы других заинтересованных лиц (разработчиков учебного плана, преподавателей педагогических вузов или школьных учителей), наиболее важную роль в формировании политики, анализе информации, издании директив и обеспечении ресурсов будут играть именно структурные подразделения министерства. В данной главе рассказывается о том, как можно теоретически использовать данные, полученные в ходе оценки, в решении вопросов политики и управления в образовании и как это происходит на практике.

Можно выделить четыре основных направления использования результатов оценки, связанных с предоставлением информации о состоянии образования в государстве, особенно – о достижениях учащихся. К ним относятся: описание достижений; описание ресурсов; мониторинг достижений; анализ состояния системы образования. Следующие пять направлений касаются использования информации, связанной

с выявленными в ходе оценки недостатками: выработка общей политики и помощь в принятии решений в соединении с другой информацией; разработка стандартов; обеспечение школ дополнительными ресурсами (широкомасштабными систематическими или целевыми); пересмотр учебного плана; пересмотр учебников.

ОПИСАНИЕ ДОСТИЖЕНИЙ

Основная цель национальной оценки – предоставить информацию о достижениях учащихся. При этом особенно важно выявить те недостатки, которые служат предпосылками для вмешательства в образование (Aguerrondo, 1992). Эти сведения можно считать уникальными, ведь, как правило, их нельзя получить из других источников. Министерства образования регулярно собирают информацию о входных характеристиках системы образования (число учащихся в школах, материально-техническое обеспечение, учебный план, отношения между учителями и учащимися), национальная оценка способна предоставить данные о результатах образовательных инвестиций. Изучив отчет о результатах национальной оценки (в общих чертах он обрисован в главе 2), разработчики политики получают полную картину результатов учебной деятельности учащихся, которая, даже не будучи явно связанной с ожиданиями, неизбежно вызовет вопросы об адекватности их достижений. Скорее всего, они также узнают о тех областях достижений, в освоении которых учащиеся испытывают определенные трудности, и о достижениях подгрупп совокупности учащихся.

В национальный отчет по оценке обычно помещают информацию относительно достижений в форме средних баллов. Средние баллы сами по себе, однако, обеспечивают только ограниченную информацию для принятия решений. Развер-

нутые описания достижений в форме уровней мастерства (что учащиеся знают и могут сделать на каждом уровне), как описано в главе 2, обеспечивают более информативную основу для решений и действий.

Полезная информация может также быть получена, когда дисперсия оценок достижений разделяется на межшкольный и внутришкольный компоненты (см. главу 2). Большие различия между школами иногда интерпретируются как показатели неравенства в возможностях учебной деятельности учащихся в системе образования. Однако разработчики политики должны относиться к такой интерпретации с осторожностью, поскольку она не учитывает межшкольные различия по факторам, контроль которых со стороны школ, как правило, очень незначителен (характеристики учащихся при поступлении в школу и их влияние на результаты деятельности школ).

Различия между школами в достижениях также привлекают внимание разработчиков политики, поскольку позволяют понять, в какой степени целесообразно предпринять вмешательство. Если различия между школами в регионе относительно низки и финансовые ресурсы ограничены, то попытки изменить ситуацию в отдельных школах не будут оправданы. Напротив, если наблюдаются большие различия, то лучше планировать вмешательство только в тех школах, учащиеся которых продемонстрировали низкие результаты.

ОПИСАНИЕ РЕСУРСОВ

В ходе национальной оценки часто собирается информация относительно имеющихся в школах ресурсов. Например, национальная оценка, проведенная в Индии в 2000 году, помогла установить, что в подавляющем большинстве школ (более 90%) имелись школьный звонок, доска, мел и ластик, в 75% школ – безопасная питьевая вода, однако отдельные туалеты

для девочек были менее чем в 40% средних учебных заведений (Singh and others, n.d.).

Несколько программ национальной оценки в Африке выявили проблему нехватки ресурсов. Например, во многих школах Кении недостает столов и учебников (Nzomo, Kariuki and Guantai, 2001). Как минимум одна треть учащихся не имеет доступа к школьным радиопередачам, так как у них нет возможности слушать школьное радио (табл. 5.1). В Занзибаре оценка выдвинула на первый план проблему серьезной нехватки мебели в классах (например, столов и классных досок) и недостаточных поставок, например, учебников и карандашей (Nassor and Mohammed, 1998).

ТАБЛИЦА 5.1

Школы, обладающие основными материально-техническими средствами: Кения, %

Оборудование	%	Стандартная ошибка
Компьютеры	1,2	0,77
Копировальные аппараты	19,9	3,11
Факсы	0,5	0,35
Проекторы фильмов	0,4	0,27
Диапроектор	0,3	0,30
Фотокопировальные устройства	1,1	0,70
Радио	66,4	4,31
Магнитофоны	10,9	2,38
Телевидение	3,2	1,74
Пишущие машинки	27,5	3,70
Кассетные видеоплееры	1,3	0,77

Источник: Nzomo, Kariuki and Guantai, 2001, табл. 3.1.4.

ТАБЛИЦА 5.2

**Школы, обладающие различным оборудованием, 1990–2002 годы:
Малави, %**

Оборудование	SACMEQ I	SACMEQ II
Мел	95,2	96,4
Классные библиотеки	13,3	20,4
Буфет	17,8	51,2
Одна или несколько книжных полок	14,7	17,6
Стул для учителя	42,3	50,5
Стол для учителя	40,7	47,7
Письменные принадлежности, пригодные к употреблению	84,8	94,5
Настенные диаграммы	56,6	58,2

Примечание: Южно- и Восточноафриканский консорциум по мониторингу качества образования (SACMEQ) – международная некоммерческая развивающаяся организация. В нее входят 15 министерств образования стран Южной и Восточной Африки. Они обмениваются опытом и проводят совместную экспертизу деятельности специалистов по планированию образования для выявления уровня применения научных методов в мониторинге и оценки условий обучения и качества образования. Консорциум SACMEQ имеет два завершённых основных исследовательских проекта в сфере образовательной политики (SACMEQ I и SACMEQ II).

Источник: Postlethwaite, 2004b. Воспроизводится с разрешения правообладателя. EFA. Программа глобального мониторинга, ЮНЕСКО.

В немногих школах Нигерии имелись карты (13%), плакаты (15%) и спортивное оборудование (5%). Ответы учителей на задания анкет свидетельствуют о высоком уровне озабоченности по поводу нехватки обучающих материалов и недооценки роли учителей, плохих условий обслуживания и нерегулярной выплаты зарплат (Федеральное министерство образования Нигерии, 2000).

В Малави мониторинг изменений в ресурсном обеспечении классов между 1998 и 2002 годами выявил прогресс по всем

видам оборудования (табл. 5.2). В Зимбабве обзор, проведенный в 1990 году, показал, что в школах провинции Южный Мателеленд имеется меньше ресурсов, чем в других областях страны. В ходе исследования 1995 года позитивных изменений обнаружено не было (Postlethwaite, 2004b).

МОНИТОРИНГ ДОСТИЖЕНИЙ

Данные о результатах оценочных процедур, проводившихся в разное время, показывают определенные тенденции – улучшаются ли достижения, остаются на том же уровне или ухудшаются (см. главу 2). Эта информация иногда использовалась для проведения мониторинга эффектов изменений в достижениях учащихся в системе образования (например, изменения в языке обучения или увеличение размеров класса). В США данные NAEP использовались для мониторинга основных положений реформы «Ни один ребенок не останется без внимания». В дополнение к мониторингу данных программы NAEP каждый штат обязан проводить мониторинг прогресса всех учащихся с 3-го по 8-й класс с помощью собственных тестов по чтению, математике и естественным наукам. За прошедшие годы многие штаты сообщали о значительных позитивных изменениях в достижениях учащихся, что, однако, не подтверждалось результатами программы NAEP. Если судить по тестам штатов, рост достижений с 2003 по 2005 год, особенно у учащихся 8-го класса, был намного значительнее, по сравнению с тестом NAEP. Например, в штате Мэриленд результаты учащихся 8-х классов по математике, в соответствии с отчетом по итогам тестирования, проведенного в штате, выросли на 12%, а согласно тестам NAEP, никакого позитивного роста не наблюдается (de Vise, 2005). Такая разница, вероятнее всего, объясняется высокой мотивацией учащихся в силу той важности, которая отводится тестированию в отдельных штатах, результаты которого влекли за собой определенные санкции. Именно поэтому

учителя сосредоточили внимание в обучении на содержании тестов штатов, что привело к увеличению тестовых баллов без сопутствующего совершенствования умений, для измерения которых первоначально предназначался тест (см. Madaus and Kellaghan, 1992).

В программах национальной оценки Уругвая, осуществлявшихся между 1996 и 2002 годами, сообщалось о значительном росте результатов выполнения тестов учащимися 6-х классов. Рост результатов был особенно заметен в школах, находящихся в «очень неблагоприятном положении», где число учащихся, достигших приемлемого уровня по тесту, выросло с 37,1 до 54,8%. Это больше, чем в школах, относящихся в плане социального контекста к категории «очень благоприятных» – там прирост составил от 57,1% до 66,5% (Ravela, 2006).

В некоторых странах национальная оценка проводится каждый год по одной и той же области учебного плана и на одной и той же категории учащихся. Однако если цель состоит в простом мониторинге стандартов, то в такой частоте нет необходимости, ведь исследование является довольно дорогим. В большинстве индустриально развитых стран, которые проводят мониторинг уровней достижений, оценка осуществляется менее часто. Скажем, в США, где в отличие от развивающихся стран вряд ли можно столкнуться с недостатком финансовых и технических ресурсов, программа NAEP по математике и чтению выполняется каждые два года.

Результаты программ национальной оценки, выполнявшихся в течение нескольких десятилетий, показывают, что значительные изменения в достижениях за короткий промежуток времени невозможны, даже если решению выявленных проблем уделяется большое внимание. Учитывая данный факт, вполне разумно проводить оценку раз в четыре года или пять лет. Если бы оценку осуществлялась чаще и результаты менялись бы значительно, то неизбежно бы встал вопрос об эквивалентности используемых тестов и процедур. Во многих

странах основной источник различий в результатах, проявляющихся со временем, связывают с изменением состава учащихся, посещающих школу. Из-за увеличивающегося уровня участия в образовании сравнение результатов разных школ в области учебных достижений требует более тщательного анализа. Эта и другие проблемы при измерении прироста результатов учебной деятельности могут возникать из-за изменений в учебных планах, языке и планируемых результатах, а также из-за трудностей методического характера, например, если не выполняются требования для использования моделей измерения (особенно при использовании современной теории тестов) или если баллы учащихся имеют тенденцию регрессировать к среднему (Goldstein, 1983).

АНАЛИЗ СИСТЕМЫ ОБРАЗОВАНИЯ

Во многих странах результаты национальной или региональной оценки были упомянуты в обзорах образовательной политики или использовались для поддержки важных инициатив по реформированию системы образования (см. табл. 5.3).

Хорошим примером использования результатов национальной оценки в основном обзоре образовательной системы, который сопровождался рядом амбициозных стратегий повышения качества образования, является Доминиканская Республика (дополнительный материал 5.1). Обзор послужил основой для обсуждений, проводимых в рамках встреч команд чиновников и членов местных сообществ по связям со школами, которые изучали причины низких результатов деятельности школ и предлагали стратегии, направленные на их улучшение.

ВЫРАБОТКА ОБЩЕЙ ПОЛИТИКИ И ПОМОЩЬ В ПРИНЯТИИ РЕШЕНИЙ

Валидные объективные данные об условиях образования, которые обеспечивает грамотно разработанная и воплощенная

ТАБЛИЦА 5.3**Примеры использования результатов национальной оценки в анализе систем образования различных стран**

Страна	Примеры использования
Аргентина	Установление программ инспекции школ
Боливия	Установление связи данных оценки с программами льготного питания школьников
Буркина Фасо	Обеспечение входных данных для анализа страны
Куба	Разработка программ улучшения заботы о детях в период дошкольного и раннего развития
Кения	Выбор точек отсчета для решения об обеспечении дополнительного школьного оборудования
Кувейт	Обеспечение поддержки для политики организации библиотек в классах
Малави	Обеспечение входных данных для программы реформ
Маврикий	Использование данных для поддержки сектора национальных исследований
Намибия	Использование данных национальной комиссией
Непал	Поддержка основной программы правительственных реформ
Нигер	Обеспечение входных данных для анализа страны
Шри-Ланка	Обеспечение входных данных для национального сектора стратегий в образовании
Уганда	Использование данных для подготовки программы реформ в образовании
Уругвай	Использование данных для поддержки политики расширения программ работы круглогодичных школ
Вьетнам	Использование данных для установления точек отсчета при обеспечении школьного оборудования (столов на каждого учащегося, книг на каждого учащегося)
Занзибар (Танзания)	Использование данных в обзоре образовательной политики для установления стандартов и точек отсчета
Зимбабве	Использование данных в обзоре комиссии

Источник: разработки авторов, основанные на: Arregui and McLauchlan, 2005; Bernard and Michaelowa, 2006; Ferrer, 2006; Министерство образования Кувейта, 2008; Murimba, 2005; Nzomo and Makuwa, 2006; Ravela, 2002.

в жизнь национальная оценка, должны способствовать повышению объективности управленческих решений и обеспечению гарантий того, что представления о системе образования будут базироваться на достоверных данных, равно как и личностные смещенные оценки, неподтвержденная информация, мифы и другие формы «накопленной мудрости» не будут следовать из эмпирических доказательств (см. дополнительный материал 5.2). Кроме того, конкретные и убедительные свидетельства, которые обеспечивает оценку, выдвигают на первый план проблемы, связанные с доступностью, качеством, эффективностью или равенством возможностей, которые остались бы незамеченными. Однако сведения, полученные благодаря национальной оценке, не приведут к политическим изменениям, конкретным действиям со стороны разработчика политики или другого потенциального пользователя. С другой стороны, как уже было отмечено, национальная оценка обеспечивает только часть информации, которую необходимо рассматривать в контексте множества других факторов, таких как наличие подходящих ресурсов (кадровых и материальных) и законные права заинтересованных лиц.

В отчетах по итогам проведения национальной оценки обнаруживаются некоторые различия, касающиеся того, насколько широко описываются последствия, вытекающие из данных оценки, и рекомендации для дальнейших действий или же предлагается выполнить эту работу пользователям результатами оценки. Некоторые отчеты являются фактически подходящими независимо от того, что кажется оправданным на основе результатов оценки (например, в определении деталей вмешательства для учащихся в сельских школах или в предложении подходов к обучению чтению или математике). Желательно, чтобы такие рекомендации подкреплялись и другими свидетельствами.

Разработчики политики и руководители отделов министерств образования нуждаются в помощи при интерпретации

Дополнительный материал 5.1**Использование результатов оценки для продвижения реформ: Доминиканская Республика**

В марте 2005 года чиновники Доминиканской Республики – представители каждого из 17 национальных регионов страны – встретились для обсуждения результатов учащихся, полученных по самым новым национальным тестам для завершения основного уровня обучения (8-й класс) и уровня средней школы, или бакалавриата (12-й класс), а также результатов, полученных учащимися в конце 4-го класса. Они рассмотрели средние национальные результаты по разным предметам учебного плана и темам, или содержательным разделам, и уделили основное внимание различиям между результатами их собственного региона и других регионов или тех и других в сравнении с национальными средними результатами. Были проанализированы различия между частным, государственным и общественным финансированием школ. Хотя чиновникам не были доступны стандарты, по отношению к которым оценивались достижения учащихся, однако региональным руководителям и директорам школ предлагались некоторые определения компетенций, умений или проверяемого содержания.

Результаты оценки стали одним из главных источников информации о входных данных для недавно инициированного процесса стратегического планирования системы образования, в котором 17 региональных команд, 101 команда районных офисов и члены различных сообществ совместно рассмотрели основные достоинства и недостатки обучения для управления изменениями, направленными на повышение качества образования. Теперь они разрабатывают новые планы на районном и региональном уровнях, нацеленные на совершенствование учебной деятельности учащихся. Некоторые районы за очень короткий промежуток времени значительно продвинулись в установлении связей между средними учебными заведениями. Так, сотрудники школ периодически собираются вместе для анализа региональных результатов в национальном тестировании и для обсуждения возможных причин неудовлетворительной работы школ. Хотя не существует никаких убедительных данных о воздействии различных определяющих факторов, некоторые решения на региональном уровне были приняты довольно быстро на основе наиболее достоверной информации, полученной опытным путем, в частности решение о распределении небольшого числа доступных компетентных учителей по естественным наукам среди нескольких связанных

Дополнительный материал 5.1

школ. Еще один пример – основание сообществ учителей для обсуждения претензий, встречающихся в ходе реализации усилий по улучшению способностей к пониманию в ходе учебной деятельности учащихся.

Слишком рано говорить о том, будет ли влияние региональных и районных планов, основанных на оценочной информации, долгосрочным. Однако нынешнему руководству уже сейчас ясно, что такие встречи оказались более полезными и способствовали большей мобилизации средних учебных заведений, чем прежняя рассылка отчетов отдельным школам и учащимся по всей стране.

Источник: Arregui and McLauchlan, 2005: 32–33. Воспроизводится с разрешения правообладателя.

результатов национальной оценки. Главная цель Южно- и Восточноафриканского консорциума по мониторингу качества образования состояла в том, чтобы укрепить потенциал системы образования путем предоставления разработчикам образовательных планов методических материалов для повышения их способности осуществлять мониторинг и оценку обучения и качества образования в государствах, входящих в консорциум. Специфическая особенность этого подхода – «обучение на практике» специалистов по планированию, чтобы привлечь их к участию в исследовании (SACMEQ, 2007). Всемирный банк осуществил подобную программу в Африке и Южной Азии.

Хотя программы национальной оценки вызвали дискуссии по вопросам образовательной политики и реформ, независимых объективных данных, указывающих на то, что результаты оценки имели реальное влияние на политику, практически нет. Исключением является Чили. В 1997 году проведенное там исследование показало, что 40% учащихся не поняли смысла прочитанного. Это убедило Национальный Конгресс в том, что система образования нуждается в существенных изменениях (Schiefelbein and Schiefelbein, 2000). Впоследствии результаты

Дополнительный материал 5.2**Мифы об образовании в США**

Национальная оценка помогает развенчать следующие мифы.

1. За прошедшие 25 лет достижения по чтению в США снизились.
2. 40% американских детей не умеют читать на базовом уровне.
3. 20% американских детей не способны научиться читать.
4. Дети, родившиеся во время демографического взрыва, читают лучше, чем нынешнее поколение учащихся.
5. Учащиеся в США читают хуже всех в мире.
6. Число хорошо читающих учащихся постоянно уменьшалось.
7. Тестовые оценки в Калифорнии резко снизились из-за использования в обучении метода «целостного языка» .

Источник: McQuillan, 1998.

национальной оценки играли важную роль в принятии нескольких политических решений (Meckes and Carrasco, 2006). Среди них отметим следующие.

- Предоставление методической и финансовой поддержек национальным правительствам для тех категорий учащихся, которые находятся в наименее выгодном положении (что определяется по результатам учебной деятельности), для обеспечения школ ресурсами и оказания другой помощи учащимся из бедных семей (программа, на которую выделяется 5% от полного бюджета государственного сектора образования).
- Определение критериев для целевого вмешательства со стороны министерства образования.
- Разработка программы повышения качества образования и обеспечения равенства прав на получение образования.
- Определение стимулов и целей для проведения улучшений образования.
- Оценка определенных политических действий и программ.
- Предоставление данных для целей исследований в образовании.

ОПРЕДЕЛЕНИЕ СТАНДАРТОВ

В последние годы разработчики образовательной политики и ответственные за учебные планы во многих странах делают все больший акцент на установлении стандартов и определении требований к результатам обучения, которых учащиеся должны достигать в конце этапов обучения в школе. Результаты национальной оценки могут помочь в решении этой задачи путем операционализации стандартов или целей достижений в ключевых предметных областях и с помощью обеспечения основных показателей для мониторинга прогресса достижений с течением времени. Например, правительство Румынии, частично из-за плохих результатов учащихся 4-х классов, в национальной оценке использовало полученные данные с целью обеспечения информации для мониторинга будущих уровней достижений (Bethell and Mihail, 2005). Также неважные результаты учащихся сельских школ послужили обоснованием для развития Проекта сельского образования, поддержанного Всемирным банком.

Необходимо, чтобы разработчики образовательной политики, спонсорские агентства и другие заинтересованные лица принимали участие в установлении стандартов, выражая свою точку зрения с позиции некоторого предостережения. В частности, они должны ставить реалистичные цели для желаемых норм совершенствования. В Перу некоторые разработчики политики предложили стопроцентный прирост средних баллов в девятимесячный период (Arregui and McLauchlan, 2005). Эта цель недостижима, особенно если она ставится для школ оценок, у которых нет нулевых точек. В столице Колумбии Боготе был установлен пороговый балл для «приемлемого» уровня, которого в то время достигли меньше чем 2% школ (Arregui and McLauchlan, 2005). В программе США «Ни один ребенок не останется

без внимания» заявленная цель «100% учащихся, достигших “высокого” уровня по тестам штатов к 2014 году» является нереалистичной. В 2003 году ни в одном штате или большом районе 100% учащихся 4-х и 8-х классов не достигли стандартов по чтению или математике даже на базовом уровне, который намного ниже высокого (Linn, 2005a).

Из-за нового понимания терминов «чтение на высоком уровне» и «математика на высоком уровне» делать выводы о том, что программы оценки предоставляют доказательства достижений учащихся, стало намного сложнее. Например, данные в США указывают, что термин «высокий уровень» в программе NAEP и программах оценки в штатах трактуется по-разному. Например, в соседних Мэриленде и Вирджинии доли учащихся, оцененных на высоком уровне по чтению и математике по тестам NAEP и в государственных тестах не совпадали (de Vise, 2005). Как видно из табл. 5.4, согласно тесту штата, высокого уровня достигли целых 85% учащихся Делавэра, но только 34% показали этот результат в тесте NAEP.

ОБЕСПЕЧЕНИЕ РЕСУРСОВ ДЛЯ ШКОЛ

Одно из возможных следствий национальной оценки состоит в том, что оно может способствовать увеличению ресурсов школ. Иногда этот процесс охватывает всю систему образования, а иногда затрагивает только отдельные школы или категории учащихся.

Обеспечение ресурсами на уровне всей системы образования по результатам национальной оценки было характерной особенностью действий, предпринятых в целом ряде стран. Так, в Кении, где, как показали исследования, во многих школах наблюдался недостаток столов и учебников, правительство решило ввести специальные нормы для обо-

ТАБЛИЦА 5.4

Ранги учащихся 4-го класса, оцененных на высоком уровне или выше по тестам штата и в национальной оценке, 2005: США

Штат	Штат (% на высоком уровне) (1)	NAEP (% на высоком уровне) (2)	Расхождение (1) – (2)
Делавэр	85	34 (1,2)	51
Айдахо	87	33 (1,4)	54
Северная Каролина	83	29 (1,4)	54
Орегон	81	29 (1,5)	52
Южная Дакота	87	33 (1,3)	54

Примечание: стандартные ошибки для данных программы NAEP заключены в круглые скобки.
Источник: Stoneberg, 2007. Воспроизводится с разрешения правообладателя.

рудования классных комнат, гарантирующие, что все школы будут обеспечены оборудованием (Nzomo and Makuwa, 2006). В Зимбабве были созданы специальные фонды для управления поставками в классы и школьные библиотеки и начаты программы обучения (Madamombe, 1995). В Доминиканской Республике (см. дополнительный материал 5.1) учителя по естественным наукам были перераспределены среди школ (Arregui and McLaughlan, 2005).

Целевое вмешательство осуществляется в тех случаях, когда результаты оценки показывают связи между ресурсами школы и достижениями учащихся. Таким образом, если установлено, что достижения учащихся в маленьких сельских школах ниже, чем национальный средний балл, то принимается решение предоставить этим школам дополнительные

ресурсы. Например, в австралийском Квинсленде школам, учащиеся которых получили плохие оценки (нижняя граница – 15%) в тестах по грамотности и действиям с числами, были выделены ассигнования за счет дополнительных фондов (Forster, 2001). В Кувейте результаты, согласно которым испытуемые, чье обучение проводилось в кабинетах, оборудованных библиотекой или специальными «уголками для чтения», получили более высокие оценки в тесте по грамотности (тест для программы PIRLS), использовались для поддержки политики Министерства образования по созданию классных библиотек (Министерство образования Кувейта, 2008).

Хотя оценка, основанная на выборке, и не обеспечивает данные по всем школам, она может предоставить информацию по отдельным их категориям (например, по школам с различным местоположением, различных видов или обслуживающих учащихся из семей с различными социально-экономическими уровнями). Эта информация служит важнейшим основанием для целенаправленного вмешательства. Такое вмешательство кажется вполне обоснованным, хотя очевидно требуются дополнительные меры для идентификации школ, которые получают выгоду от поставок оборудования. В большинстве образовательных систем основное внимание уделяется школам, обслуживающим учащихся из не самых благополучных семей (например, школы с низким уровнем образования родителей учащихся). В некоторых странах ключевыми факторами для выбора средних учебных заведений, которым необходимо оказать поддержку, являются их виды или местоположение. В этом отношении также могут быть полезны данные всеобщей переписи населения, данные информационной системы управления образованием и информация, полученная от школьных инспекторов, наблюдателей и советников.

Хотя национальная система оценки в Чили основана на генеральной совокупности учащихся, обеспечение ресурсами школ с низкими достижениями может быть релевантным обеспечению, которое осуществляется в соответствии с результатами оценки, основанной на выборке, если при ее формировании имеет место адекватная система идентификации школ. Чилийским школам, находящимся в областях с высокой долей социально неблагополучных семей, учащиеся которых показали неудовлетворительные результаты, были выделены дополнительные ресурсы, включая учебники, библиотеки для классов и педагогические материалы (González, Mizala and Romaguera, 2002; Wolff, 1998). Также была обеспечена внешкольная деятельность, укреплены связи внутри школьного сообщества, а провинциальные и центральные службы от 8 до 16 раз в год проводили контроль школ (Arregui and McLauchlan, 2005).

Когда разработчики политики рассматривают проблему выделения дополнительных ресурсов школам по результатам национальной оценки, они должны решить, будет ли распределение ресурсов вестись на постоянной основе или носить временный характер. Ограничение по времени можно считать предпочтительным по ряду причин. Во-первых, учитывая, что все ресурсы ограничены, временные ограничения на условия предоставления ресурсов определенным школам в какой-то мере учитывают интересы других школ, которые не получили дополнительных ресурсов и поэтому чувствуют себя ущемленными в правах. Во-вторых, вполне вероятно, что поставки ресурсов на временной основе приведут к более интенсивному вмешательству, которое могло бы включать, например, помощь в использовании ресурсов. В-третьих, международная поддержка спонсоров тяготеет к традиционным и уже одобренным инициативам, которые являются ясно определенными и ограниченными по времени (Chapman and Snyder, 2000).

ПОДДЕРЖКА ПЕРЕСМОТРА УЧЕБНОГО ПЛАНА

Учебный план чаще всего пересматривается по итогам международной, а не национальной оценки. Ведь именно международная оценка позволяет увидеть, как учащиеся данной страны выполняют тесты по сравнению с учащимися из других стран (принятый учебный план). Кроме того, благодаря ему можно сопоставить намеченный учебный план, который работает на уровне системы образования (зафиксированный в документах), и реализованный учебный план, который работает на уровне класса (предоставляет информацию о времени и основных особенностях по областям достижений).

Вполне очевидно, что национальная оценка тоже может предоставить информацию о выполнении учебного плана и проведении реформ. Она позволяет сравнить фактические достижения учащихся с планируемыми, которые определяются или подразумеваются в документах учебного плана, и выявить несоответствия. Такое сравнение показывает, как учебный план воплощается в жизнь, от каких факторов зависит его успешное выполнение, ставит ли учебный план адекватные цели относительно запланированных достижений учащихся.

Результаты национальной оценки и оценки в штатах способствовали изменению учебных планов в ряде стран. Так, в Бразилии в штате Парана по итогам оценки и анализа выявленных проблем было составлено руководство для учителей по учебным стратегиям (Ferrer, 2006). В Таиланде из-за неудовлетворительных результатов учащихся в тестах по математике и естественным наукам составители учебных планов стали уделять больше внимания умениям по применению знаний (Pravalpruk, 1996). В Гвинее благодаря оценке началось развитие национальных инициатив, касающихся обучения чтению и переноса начала обучения на более ранний возраст (R. Prouty, информация получена в ходе личной беседы, которая происходила в Вашингтоне, округ Колумбия, 15 мая 2005 года). Пример достаточно

многообразного использования результатов оценки в контексте реформы учебного плана можно найти в Панаме. В этой стране одной из важнейших задач национальной оценки является проверка влияния нового учебного плана на результаты учащихся (Ferrer, 2006).

Проблемы, характерные для большинства систем национальной оценки, вызваны тем, что адекватность намеченного учебного плана изначально не подвергается сомнению, а это в корне неверно. Задачи оценки основываются именно на учебном плане, поэтому предусматривает анализ результатов, направленных на выявление тех знаний и умений, которые в нем определены, и, возможно, того, в какой степени учителя реализуют учебный план. Однако во многих развивающихся странах образование ориентируется на учащихся – представителей элиты, следовательно, для обычных учащихся учебный план оказывается неадекватным (Glewwe and Kremer, 2006).

В ряде стран по итогам анализа результатов национальной оценки положения учебного плана подверглись сомнению и пересмотру. Например, в Бутане в отчете о национальной оценке учащихся 6-го класса в 2003 году была выявлена предусмотренная учебным планом неоправданная загруженность обучаемых, которая была обозначена как проблема. Рекомендовалось осуществить стратегический пересмотр программы по математике так, чтобы «делать на алгебре и геометрии меньший акцент (или не делать вообще) или по крайней мере не преподавать их на уровне формальных знаний, ожидаемых от учащихся в настоящее время» (Комиссия Бутана по экзаменам 2004 года: viii; см. также Powdye1, 2005). Интерпретация результатов оценки заинтересованными лицами также может подчеркнуть потребность реформирования учебного плана. Так, в Эфиопии наряду с анализом выполнения тестов учащимися в национальной оценке для принятия соответствующих решений об

изменении учебного плана использовалась обратная связь между учителями, учащимися и их родителями (Z. Gebrekidan, информация получена в ходе личной беседы, которая состоялась в Аддис-Абебе 30 июня 2008 года).

В результате консультаций с учителями команда по национальной оценке в Уругвае при составлении тестов вышла за пределы, официально установленные учебными планами. Это спровоцировало дебаты относительно валидности и релевантности содержания преподаваемых в школах предметов. В течение долгого времени концептуальная структура теста (вместе с учебниками для учащихся и руководствами по преподаванию) стала альтернативой национальному учебному плану для учителей (Fetter, 2006).

Значимость результатов национальной оценки для анализа учебных планов и их пересмотра, вероятно, возрастет, если специалисты по учебному плану будут участвовать в установлении оцениваемых умений. Например, они могут оказать большую помощь при разъяснении сложных терминов, таких как, например, «чтение». Это понятие имеет множество значений – от умения произносить простые слова до способности читать и постигать смысл простых предложений и параграфов, извлекать смысл из различных текстов, включая расширенные литературные пассажи и информационные сообщения (такие как меню и расписания). Кроме того, более продуктивному использованию результатов национальной оценки для пересмотра учебного плана будет способствовать привлечение вышеупомянутых специалистов к определению уровней мастерства и видов анализа данных. Так, в случае оценки умений работы с числами очень полезным был бы анализ того, насколько хорошо учащиеся выполняют задания, которые оценивают простые вычислительные навыки по работе с целыми числами или концептуальное понимание простых и десятичных дробей. Наконец, члены команды по национальной оценке должны

спроектировать краткий отчет по применению результатов оценки для пересмотра учебного плана и обсудить его с руководителями.

ПЕРЕСМОТР УЧЕБНИКОВ

В идеале учебники должны точно отражать содержание учебного плана. Однако исследование, проведенное Американской ассоциацией по продвижению в естественных науках, показало, что содержание самых популярных учебников по математике и естественным наукам для средней и старшей школы не было в достаточной мере приведено в соответствие индикаторам, содержащимся в образовательных стандартах большинства штатов (Kulm, Roseman and Treistman, 1999), которые также используются для установления стандартов оценки. Результаты программ национальной (и международной) оценки выявляют эту проблему, идентифицируют аспекты учебного плана, которые заслуживают большего освещения или внимания в новых или переработанных учебниках и руководствах для учителя.

Пересмотр учебников в результате национальной оценки, как правило, начинается уже после того, как в официальный учебный план внесены существенные изменения. Однако в Латинской Америке известны случаи, когда изменения в учебниках и деятельность по оценке осуществляются параллельно. Например, в Уругвае методические команды, отвечающие за национальную оценку и пересмотр учебников, работали на основе похожих, хотя и не идентичных, концепций и педагогических перспектив (P. Agregui, информация получена в ходе личной беседы, которая состоялась в Лиме 22 сентября 2008 года). В области языка новые теории обучения, которые требуют обращения к более широкому диапазону литературных жанров, были отражены в содержании тестов для национальной оценки и способствовали выпуску широкого спектра материалов для чтения.

Интересное предложение, касающееся учебников, последовало за национальной оценкой во Вьетнаме (Всемирный банк, 2004). В свете результатов, согласно которым менее половины учащихся владели всеми предписанными и рекомендованными учебниками и учебными материалами, члены команды по национальной оценке предложили утвердить в учебном плане более короткий и реальный список и провозгласить гарантии, что каждый учащийся будет обеспечен необходимой литературой.

Результаты национальной оценки чаще используется для пересмотра учебников в тех случаях, когда:

- в разработке теста принимают активное участие специалисты по составлению учебных планов;
- результаты национального тестирования представлены в подходящей форме, позволяющей рассмотреть качество охвата учебного плана в учебниках;
- среди разработчиков стандартов оценки или уровней выполнения тестов есть люди с опытом проведения экспертизы при разработке учебного плана, педагогических измерений или оценке персонала.

ВЫВОДЫ

Использование результатов национальной оценки зависит прежде всего от характера информации, которую она обеспечивает. Лица, ответственные за политику и управление, должны определить свои информационные потребности в перспективном планировании оценки.

Все системы национальной оценки описывают достижения учащихся. Однако выбор пути использования полученных сведений во многом зависит от деталей, в которых представлены достижения, например, для отдельных разделов или содержательных областей математики, и от той информации,

которую обеспечивает оценка при выявлении причин низких достижений. Будут ли данные, полученные в итоге оценки, идентифицировать проблемы, связанные с пересмотром учебных планов или учебников, зависит от уровня детализации содержания тестов учебных достижений.

Не все системы национальной оценки помогают собирать информацию относительно ресурсов по обеспечению учебного процесса. Это происходит только в тех случаях, когда есть подозрения на существование глобальных проблем во всей образовательной системе.

Результаты оценки играют важную роль в установлении стандартов для оценки достижений учащихся в системе образования. Конечно, мониторинг тенденций будет возможен только тогда, когда сопоставимые данные доступны в различные моменты времени.

Теоретически результаты оценки могут использоваться для разработки новой политики, анализа проблем существующей системы образования и принятия решений о предоставлении школам дополнительных ресурсов. Однако степень их использования будет зависеть от ряда факторов. В число этих факторов входят: наличие отчетов необходимого объема на соответствующем языке; связь с интересами и потребностями разработчиков политики в сфере образования и лиц, принимающих решения; доступность адекватных каналов коммуникации; наличие бюджетных ресурсов; интерес, способность и готовность лиц, принимающих решения, понимать, интерпретировать и применять результаты оценки на практике.

РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ОБУЧЕНИЕ

Первичная цель национальной оценки состоит в том, чтобы предоставить информацию разработчикам политики и управленцам в образовании. Однако вряд ли следует ожидать какого-либо прогресса в достижениях учащихся, если результаты национальной оценки не будут использоваться для развития политики и стратегий, направленных на повышение эффективности деятельности школы и практики обучения. Один из возможных вариантов – вмешательство в деятельность школ, хотя бы тех из них, которые особенно нуждаются в поддержке (обеспечение дополнительными материально-техническими ресурсами или обучение учителей для повышения их квалификации, приобретения новых знаний или умений). В этой ситуации руководство должно ясно определить сроки вмешательства. Но вмешательство в деятельность школ обычно не следует за национальной оценкой. Чаще всего бывает так, что сначала публикуются результаты, а решения, которые должны повлиять на достижения учащихся, принимаются позже. Как правило, их оставляют за ключевыми заинтересованными лицами – учителями, инспекторами, советниками и наблюдателями. В этой главе освещается как информация, полученная в национальной оценке,

основанной на выборке, может способствовать повышению качества обучения в отдельных школах и классах.

Рассматриваются два подхода к решению этой задачи. Первый основан на профессиональном развитии учителей путем организации для них курсов повышения квалификации до начала или в течение работы. Согласно второму подходу, школы принимают решение относительно релевантности результатов национальной оценки исходя из их специфики, а также той или иной ситуации. Если значимость результатов не подвергается сомнению, разрабатываются стратегии для решения выявленных проблем. Эти два подхода не всегда могут быть четко разделены. Так, курсы для преподавателей могут быть организованы в школе, а некоторые формы обучения в процессе работы или на семинарах должны полностью основываться на результатах национальной оценки и демонстрировать реакцию школ на них.

ПРОФЕССИОНАЛЬНОЕ РАЗВИТИЕ УЧИТЕЛЕЙ

Результаты национальной оценки обеспечивают обширные данные для министерств образования, региональных структур и тех учебных заведений, которые готовят учителей школ в штатном режиме и обеспечивают их повышение квалификации. Вначале будут описаны четыре источника таких данных, которые могут использоваться для руководства деятельностью по профессиональному развитию учителей в сфере обучения, а затем приводится ряд примеров из опыта многих стран.

Источники данных национальной оценки для руководства профессиональным развитием учителей

Для профессионального развития учителей доступны четыре источника данных: структура оценки и типовые вопросы; результаты выполнения учащимися тестов при оценке;

ответы на анкеты; выполнение учителями тестов учебных достижений (принятое в некоторых системах национальной оценки).

Структура национальной оценки и типовые вопросы. Анализ области учебного плана или конструкта, содержащегося в структуре национальной оценки, может помочь учителям лучше понять результаты оценки и выявить новые перспективы совершенствования своей работы. Так, учителя, посещающие курсы повышения квалификации, в свете анализа достижений по математике, содержащихся в структуре теста, могут по-иному взглянуть на свои методы обучения. Они должны задаться следующими вопросами: «Не слишком ли много внимания я уделяю репродуктивным знаниям (таким как, например, умножение чисел) в ущерб задачам на более высоких познавательных уровнях» или «Уделяю ли я достаточно внимания некоторым аспектам учебного плана, в частности разделам “Форма” и “Пространство”?». Курсы повышения квалификации помогут ответить на типовые вопросы, которые возникают после национальной оценки и, как правило, служат моделью для оценки учащихся в классе (Департамент образования штата Джорджия).

Выполнение учащимися тестов при проведении оценки. На курсах по профессиональному развитию учителей основное внимание следует уделять тем аспектам учебного плана, которые, судя по результатам национальной оценки, являются наиболее проблемными. Развернутая экспертиза результатов оценки, дает информацию о конкретных областях чтения (например, обнаружение неустановленного значения в тексте) или математики (например, решение задач), где учащимся необходимо оказать поддержку (см. главу 2).

Ответы на анкеты. Важными источниками информации, которую предоставляет национальная оценка для проектирования курса повышения квалификации учителей, являются не только тесты. Весьма заметную роль играет также анкетирование. Оно может пролить свет на совокупность факторов, относящихся к обучению: социально-экономические условия, в которых работают школы; проблемы, с которыми сталкиваются учителя; оценка профессиональной пригодности учителей; качество и полноценность имеющихся материалов – технических условий (включая учебники и руководства для учителей); отношение родителей и учащихся к образованию; мотивация учащихся (Griffith and Medrich, 1992).

В американских программах национальной оценки последовательный анализ информации, полученной в ходе анкетирования, проводился на основе рассмотрения ее корреляции с достижениями учащихся. Полученные результаты вызвали бурную реакцию со стороны учителей. Анкеты для учащихся содержали сведения об уровне образования родителей и условиях для дополнительных занятий. В анкетах для учителей запрашивалась подробная информация относительно особенностей обучения математике, например, соответствие методу, включая обращение к нестандартным и типовым задачам, использование учебников и рабочих тетрадей, наличие заданий на решение реальных жизненных проблем и характеристики склонности к работе в коллективе (Wenglinsky, 2002).

Пример из программы TIMSS 2003 года показывает, как вопрос о готовности к преподаванию различных разделов учебного плана по математике помог выявить недостатки в подготовке учителей, влияющие на результаты обучения. В отчете приводятся проценты учащихся, учителя которых готовы подтвердить свою компетентность по определенным темам преподаваемого предмета (Mullis and others, 2004). В табл. 6.1 представлена выдержка из отчета для двух из пяти содержательных областей математики, включая арифметику

и алгебру. Результаты показывают, что систему повышения квалификации (также как и обучение учителей перед началом работы) можно сделать более эффективной, если уделить больше внимания обучению определенным умениям, например умению анализировать свойства графиков. Подобные вопросы можно включить в национальную оценку в ряде стран, чтобы установить те области учебного плана, которые следует осветить особенно ярко в курсах подготовки учителей.

Анкеты для учителей также обеспечивают информацию о том, имеется ли у них опыт обучения на курсах повышения квалификации и о тех потребностях, которые определили необходимость их пройти. Так, в Индии и Бутане, согласно данным анкет, многие учителя не посещали курсы повышения квалификации либо обучались при них (Бутанская комиссия по экзаменам, 2004; Powdyel, 2005; Singh and others). Необходимость пройти курсы повышения квалификации по вопросам преподавания по проблемам, возникающим при обучении детей письму, была отмечена в Лесото (член Конгресса М.С. Ramoкоена, информация получена в ходе личной беседы, которая состоялась в Масеру 6 апреля 2009 года).

Эти сведения являются очень ценными, так как определяют направления работы курсов повышения квалификации. Тем не менее получить их отнюдь не просто. Например, могут возникнуть проблемы, аналогичные с теми, с которыми столкнулись при проведении национальной оценки в Шотландии, где анкетирование не выявило данные, устанавливающие связь между низкими достижениями учащихся и методами преподавания, которые следует исключить из практики (Robertson, 2005).

Выполнение тестов учебных достижений учителями. Среди курсов повышения квалификации наиболее распространены те, где обучение направлено на развитие педагогических умений, а не на совершенствование знаний учителей по предметам. Между тем потребность в последнем направлении обучения

ТАБЛИЦА 6.1

Число учащихся, учителя которых сообщили о готовности преподавать математику, данные TIMSS, 8-й класс, %

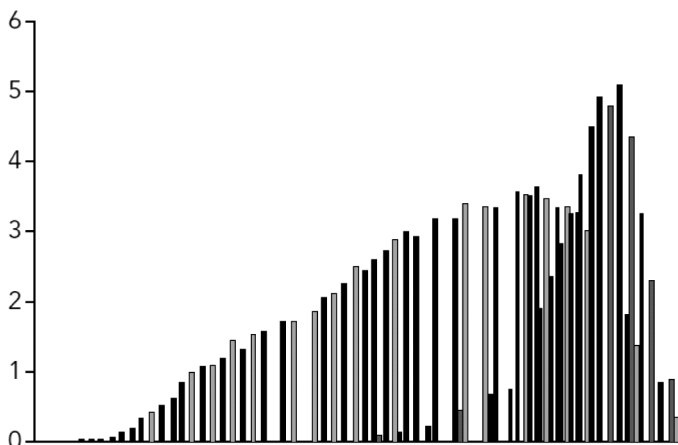
Выбранные страны	Содержательная область					
	Арифметика			Алгебра		
	Представление десятичных и простых дробей, используя словесную форму, цифры и числовые оси	Целые числа, включая словесную форму, числа, числовые оси, упорядочивание целых чисел и операции (+, -, × и ÷) с целыми числами	Числовые алгебраические и геометрические примеры последовательностей	Простые линейные уравнения и неравенства, уравнения с двумя переменными	Эквивалентные представления функций в виде упорядоченных пар значений, таблиц, графиков, словесных описаний или уравнений	Свойства графиков, такие как точки пересечения с осями и интервалы
Болгария	100	100	99	100	100	100
Исламская Республика Иран	98	98	90	98	94	87
Ливан	98	100	93	96	95	95
Саудовская Аравия	96	100	86	95	94	80
Сербия	91	90	93	90	90	90
Тунис	99	98	87	71	74	71

Источник: Mullis and others, 2004: 256.

намного выше, хотя и не всегда признается в педагогическом сообществе. Некоторые программы национальной оценки, которые обеспечивают информацию о знании учителями своего предмета, выявили серьезные пробелы в их профессиональной подготовке. На рис. 6.1 показано распределение оценок по тесту на понимание прочитанного для учащихся (более плоское распределение) и учителями (более крутое распределение в правой части графика) по результатам программы национальной оценки учащихся после окончания 5-го класса во Вьетнаме. В ходе этой программы тестовые задания, выполнявшиеся учащимся, предъявлялись и учителям. Полученные данные свидетельствуют о том, что результаты 12% лучших учащихся, превысили наиболее низкие результаты 30% учителей.

Рис. 6.1

Распределение оценок по чтению учащихся и учителей: Вьетнам



Примечание: оценки по чтению для учащихся представлены более плоской кривой, а преподавателей – более «крутой» кривой. (Имеется в виду ломанная кривая, которую можно провести, соединив верхние точки столбиков гистограмм. – *Примеч. ред.*)

Источник: Всемирный банк, 2004, том 1: рис. 2.

В системе национальной оценки образования Пакистана одни и те же задания предъявлялись учащимся 4-го класса и учителям. Как и во Вьетнаме, распределения оценок перекрываются. Приблизительно 3% учителей получили баллы ниже среднего балла учащихся. Очень незначительное число заданий было выполнено верно 90% учителей (А. Таууаб, информация получена в ходе личной беседы, которая состоялась в Оксфорде 24 июля 2008 года). Углубленный анализ показывает, что учителя плохо выполнили те задания, которые предназначались для оценки по разделам: геометрия, поместное размещение (цифры) и измерения. Например, только 42% учителей правильно ответили на задание, аналог которого представлен в дополнительном материале 6.1. В нем требуется установить соответствие между четырехзначным числом и его словесным эквивалентом.

Похожие результаты продемонстрировали программы национальной оценки в Африке (Duthilleul and Allen, 2005). В табл. 6.2 приведены данные о процентах учителей в программе национальной оценки в Мозамбике, которые были размеще-

Дополнительный материал 6.1

Задание на установление соответствия между числом и его словесным эквивалентом: Пакистан

Какое из этих словосочетаний эквивалентно числу 4256?

- A) Четыре тысячи двести пятьдесят шесть
- B) Четыре тысячи пятьсот двадцать шесть
- C) Четыре тысячи двести пятьдесят
- D) Четыреста двадцать пять тысяч шесть

Источник: А. Таууаб, информация получена в ходе личной беседы, которая состоялась в Оксфорде 24 июля 2008 года.

ны на шкале с восьмью уровнями мастерства в соответствии с результатами выполнения теста по математике для учащихся 6-х классов. Содержательная трактовка уровня мастерства определяет те знания и умения учителей, которые они приобрели на каждом уровне, тем самым позволяя косвенным образом определять то, чего им недостает в профессиональной подготовке. Все учителя достигли уровней мастерства 1 и 2, но примерно каждый четвертый не сумел правильно выполнить проблемные задания, соответствующие уровням 7 и 8.

ТАБЛИЦА 6.2

Учителя, оцененные на каждом уровне профессионализма по математике: Мозамбик, %

Уровень	Описание	%
Уровень 1: Предварительная способность к действиям с числами	<ul style="list-style-type: none"> • Умения применять одношаговые операции сложения или вычитания. Распознавание простых фигур. • Установление соответствий между словесной формой чисел и их цифровым отображением. • Умение совершать действия с целыми числами 	0,0
Уровень 2: Способность к действиям с числами на стадии становления	<ul style="list-style-type: none"> • Умение применять двухшаговые операции сложения и вычитания, приблизительно проверить полученный результат с помощью прикидки, преобразовать цифровую в словесную форму числа. • Умение приблизительно оценить длину знаковых объектов. • Умение распознавать простейшие двумерные фигуры 	0,0
Уровень 3: Базовая способность к действиям с числами	<ul style="list-style-type: none"> • Умение перевести вербальную информацию, представленную в виде предложения, в простой график или таблицу, используя одну арифметическую операцию при нескольких повторяющихся шагах. • Умение перевести графическую информацию в дробные числа 	0,3

	<ul style="list-style-type: none"> • Умение интерпретировать значение цифр при их поместном размещении в целых числах до тысячи. • Умение интерпретировать простые, общеприняты повседневные единицы измерения 	
Уровень 4: Высокая способность к действиям с числами	<ul style="list-style-type: none"> • Умение перевести словесную или графическую информацию в цифровую форму. • Многовариантное использование различных арифметических операций (в правильном порядке) с целыми числами, простыми и десятичными дробями 	2,9
Уровень 5: Компетентность в действиях с числами	<ul style="list-style-type: none"> • Умение перевести словесную, графическую или табличную информацию в цифровую форму, чтобы решить поставленную задачу. • Решение задач с многошаговыми операциями (используя правильный порядок арифметических действий) с привлечением основных единиц измерения, а также целых и смешанных чисел. • Умение перевести основные единицы измерения из одной размерности в другую (например, метры в сантиметры) 	
Уровень 6: Математически квалифицирован	<ul style="list-style-type: none"> • Решение задач с многошаговыми операциями (используя правильный порядок арифметических действий) с дробями, пропорциями и десятичными числами. • Умение переводить вербальную и графическую информацию в форму символов, алгебраических выражений и уравнений для решения данной математической задачи. • Умение проверить и приблизительно оценить ответы, используя внешние знания (не предоставленные в пределах постановки задачи) 	16,3
Уровень 7: Способен к конкретному решению задач	<ul style="list-style-type: none"> • Умение извлечь и конвертировать (например, относительно единиц измерения) информацию из таблиц, диаграмм и визуальных и символических представлений для идентификации и последующего решения многошаговых проблем 	44,3
Уровень 8: Способен к абстрактному решению задач	<ul style="list-style-type: none"> • Умение идентифицировать природу неустановленной математической задачи, включенной в вербальную или графическую информацию, а затем перевести ее в форму символов, алгебраических выражений и уравнений для последующего решения 	31,7

Источник: Bonnet, 2007. Воспроизводится с разрешения правообладателя.

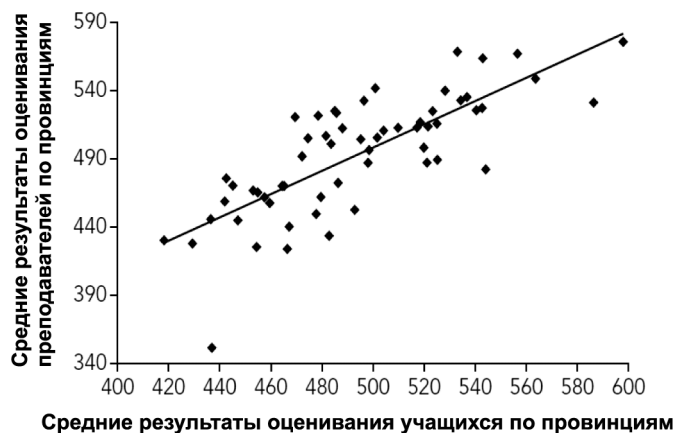
Исследование, проведенное во Вьетнаме по математике, показывает степень, в которой знание учителями своего предмета влияет на результаты выполнения тестов учащимися (рис. 6.2). Если средний балл по результатам тестирования учителей (основанный на среднем результате для провинции) увеличивается, средний балл учащихся по тесту также повышается (Всемирный банк, 2004). Эти результаты привели к разработке рекомендаций о том, что учителей с плохим уровнем владения предметом нужно поощрять к прохождению курсов повышения квалификации для улучшения своих знаний. Это должно привести к сокращению числа учащихся с низкими достижениями.

Примеры деятельности, направленной на повышение профессионализма учителей по результатам национальной оценки

Разработчики политики и специалисты по планированию образования далеко не всегда используют результаты программ

Рис. 6.2

Программа национальной оценки по математике во Вьетнаме: корреляция между средними баллами учителей и учащихся в провинциях



национальной оценки для определения необходимых акцентов в программах подготовки учителей. Исключение составляют некоторые латиноамериканские страны – Аргентина, Бразилия, Чили, Куба, Доминиканская Республика, Эквадор, Гватемала, Гондурас, Уругвай и Боливарианская Республика Венесуэла (Fetter, 2006). Некоторые страны использовали результаты как основу для акций массового обучения учителей с помощью национального телевидения (Куба) или с помощью обзоров по национальной оценке, подготовленных опытными педагогами, включая представителей педагогических вузов (Доминиканская Республика). Ярким примером использования данных оценки для улучшения подготовки преподавателей служит бразильский штат Минас-Жерайс (дополнительный материал 6.2). Однако имеется очень мало сведений о том, насколько эффективными оказались эти усилия по повышению квалификации учителей (Arregui and McLaughlan, 2005).

В Уругвае, где основной аудиторией для обсуждения результатов тестирования являются школы, основное внимание уделяется знаниям и способностям учителей. Особый акцент делается на обучении чтению и математике в школах, обслуживающих учащихся из семей с неблагоприятным социальным положением. Этот подход оказался весьма эффективным, о чем свидетельствует заметный рост результатов в программах оценивания на протяжении длительного времени (см. главу 5).

К существенным особенностям подхода можно отнести следующие (Ravela, 2006):

- анализ заданий тестов и типичных ошибок учащихся;
- отсутствие ответственности учителей за результаты оценки учащихся при принятии административно-управленческих решений;
- обратная связь со школами, основанная на результатах тестирования;
- встречи учителей для обмена идеями по вопросам профессионального развития;
- целевая поддержка школ, исходя из социального положения учащихся, которых они обслуживают, а не их достижений.

Дополнительный материал 6.2**Использование результатов национальной оценки для повышения уровня подготовки учителей: Минас-Жерайс, Бразилия**

В бразильском штате Минас-Жерайс в рамках проекта по использованию результатов национальной оценки для повышения качества подготовки учителей были осуществлены следующие шаги.

- Оценка предназначалась для предоставления педагогическим образовательным учреждениям полезной информации.
- В системе педагогического высшего образования для проектирования и предъявления тестов и обработки результатов и предоставления их школам было организовано официальное сотрудничество между 29 институтами.
- Каждый университет прилагал основные усилия, пользуясь поддержкой других.
- Были проанализированы данные для установления проблем учащихся по чтению, выявленные по результатам оценки.
- От учащихся, посещавших подготовительные курсы, ожидалось, что они изучат результаты оценки для ознакомления с образовательными проблемами.
- Результаты использовались для разработки программ обучения учителей.

Источник: Ferrer, 2006.

Программа крупномасштабной оценки ежегодно привлекала приблизительно 3000 работающих учителей из 300–500 бедных городских школ (дополнительный материал 6.3). Она выявила те характеристики программ профессионального развития, которые привели к совершенствованию знаний и умений учителей. К ним относятся: акцент на знании содержания; широкие возможности для активной учебной деятельности учащихся; установление правильной последовательности действий, коллективное участие учителей одной и той же школы; выделение достаточного времени на реализацию программы

для оказания влияния на достижения учащихся (Garet and others, 2001).

Коммуникационные и политические проблемы могут свести на нет все усилия по улучшению качества педагогического образования. Так произошло в Аргентине, где стало известно о недостаточно хорошо налаженных связях между агентствами, ответственными за педагогическое образование и оценку

Дополнительный материал 6.3

Основные особенности программы для работающих учителей, основанной на результатах национальной оценки: Уругвай

В Уругвае крупномасштабная программа для работающих учителей имела следующие особенности.

- По субботам проводились платные курсы (стоимость – примерно 25% начальной зарплаты учителей).
- Посещение было добровольным.
- Учителя допускались к обучению на курсах только в тех случаях, если в школе, где они работают, не менее половины сотрудников (включая руководство) согласились принимать участие в программе.
- Сессии (курсы) включали приблизительно 60 учителей из 5–10 школ, что обеспечивало возможность обмена опытом при повышении квалификации как внутри одной школы, так и между школами.
- Основное внимание уделялось школам, находящимся в невыгодном положении по контекстным факторам. Однако достижения учащихся не учитывались, поэтому в программе могли принять участие учителя школ, учащиеся которых показали как высокие результаты выполнения тестов, так и низкие.
- Группа наблюдателей и профессоров из педагогических образовательных учреждений, отобранных и обученных для осуществления программы, проводили учебные сессии.
- Программа была сосредоточена на обучении подходам к преподаванию по математике, языку, естественным наукам и общественным наукам.

Источник: Ravela, 2005.

(Ferrer, 2006). Проблемы в Чили объясняются тем, что педагогические учреждения часто настроены идеологически против национальной оценки (Sistema de Medición de la Calidad de la Educación, или SIMCE) (Meckes and Carrasco, 2006).

Курсы повышения квалификации и программы для работающих учителей, проведенные по результатам национальной оценки, часто способствуют признанию той роли, которую играет в жизни школы ее руководящий состав. В Уругвае, например, в дополнение к семинарам для учителей и инспекторов, после национальной оценки были организованы семинары специально для руководящих лиц, ориентированные на их проблемы (Benveniste, 2002). Считается, что обучение руководства в процессе работы позволит охватить широкий диапазон тем, разработанных с целью обеспечения знаний и развития умений, способствующих созданию окружающей среды, благоприятствующей учебной деятельности учащихся. Анкеты, предъявляемые в ходе национальной оценки, помогают сформировать содержание этого вида обучения в процессе работы, побуждая руководство в образовании организовать курсы, охватывающие одно и более направлений развития навыков и способностей (Murphy, Yff and Shipman, 2000):

- способность руководящих лиц осуществлять управление школьными ресурсами, чтобы обеспечить сбережение, эффективность (сводится к обеспечению эффективности в условиях оптимизации результатов и затрат на их получение. – *Примеч. ред.*) и создание эффективной окружающей среды для учебной деятельности учащихся;
- навыки лидерства, способствующие поддержке учебной деятельности учащихся;
- способность помогать преподавателям развивать их знания и умения;
- способность анализировать проблемы учебных предметов, школьного обслуживания и моральных качеств сотрудников школ;
- способность к развитию сотрудничества с семьями и сообществами в работе школы (особенно, если сообщество

- поддерживает сложившееся в обществе отрицательное восприятие школы);
- способность отвечать на вопросы и влиять на широкий политический, социальный, экономический, юридический и культурный контексты.

ОСНОВНОЕ ВНИМАНИЕ ШКОЛАМ И ОБУЧЕНИЮ В КЛАССЕ

Гарантировать, что результаты национальной оценки влияют на практику деятельности учителей в классе и способствуют улучшению учебной деятельности учащихся, довольно сложно. Об эффективных стратегиях использования результатов, которые, скорее всего, будут меняться от одной системы образования к другой и от школы к школе, известно очень немного.

В этом разделе рассматривается ситуация, когда отдельные школы отказываются от интерпретации результатов национальной оценки при их распространении с целью анализа релевантности результатов и развития стратегий, направленных на установление проблем в обучении. Для эффективного использования результатов важны три специфических процедурных шага, которые описаны ниже: предоставление результатов учителям; интерпретация результатов учителями и оценка их релевантности по отношению к своей школе; применение результатов в стратегиях, разработанных для улучшения учебной деятельности учащихся. После рассмотрения этих шагов и принимаются решения о вмешательстве в деятельность школ.

Сообщение результатов

Существует множество вариантов для сообщения результатов национальной оценки учителям по всей системе образования: печатные материалы и отчеты; семинары и платные школы-семинары; посещения инспекторов, наблюдателей и совет-

ников. В некоторых странах у учителей есть доступ к отчетам в сети Интернет.

Иногда отчеты содержат комментарии общего характера и дают весьма обширные рекомендации. Например, результаты национальной оценки в Уганде были представлены общими сообщениями, разосланными в форме постеров учителям школ для демонстрации в классе (дополнительный материал 6.4). В других странах готовят более детальные отчеты в первую очередь для учителей, которые описывают образовательные достижения по тем областям учебного плана, где учащиеся продемонстрировали наиболее высокие и низкие результаты. Включение в отчеты практических предложений, которые шко-

Дополнительный материал 6.4

Отрывок постера, опубликованного после национальной оценки: Уганда

В Уганде среди школьных учителей был распространен постер со следующим текстом.

Мы умеем выполнять следующие действия:

- Воспринимать обучение на английском языке и правильно отвечать на вопросы.
- Читать короткие отрывки текстов на английском языке.
- Складывать, вычитать и умножать.
- Писать некоторые английские слова и предложения.

В ходе дальнейшего обучения помогите нам:

- Расширить словарный запас.
- Научиться читать быстро и плавно.
- Писать ясно и аккуратно.
- Правильно выполнять деление чисел.
- Использовать ранее изученный материал в новых ситуациях в повседневной жизни.

Источник: Национальная экзаменационная комиссия Уганды.

лы и учителя могут использовать для улучшения достижений учащихся, способствует повышению их ценности.

Отчет для учителей в Британской Колумбии (провинция на западе Канады), служит примером детального описания выполнения тестов учащимися. Он содержит ряд таблиц, в которых указаны доли учащихся на районном уровне, которые выбрали правильные варианты ответов на каждое задание, и тех, кто ответил неправильно (в %). Также в отчете приводится интерпретация типичных действий учащихся при выборе неправильных ответов. В табл. 6.3 представлен пример для теста по математике (описание примеров заданий, на которые чаще всего давались неправильные ответы).

Выполняя математические вычисления, включающие более одного шага и возведение чисел в квадрат, учащиеся зачастую ошибаются. Пример задания, которое можно использовать в групповом обсуждении для помощи учителям в ходе установления наиболее распространенных ошибок, представлен в дополнительном материале 6.5.

Числа в круглых скобках после каждого ответа обозначают долю учащихся, которые выбрали данный вариант. В ходе обсуждения результатов учителями были сделаны следующие выводы.

Дополнительный материал 6.5

Задание по математике

$$5^2 - 3(4 - 1) =$$

A) -4 (13%)

B) 10 (22%)

C) 14 (27%)

D) 16 (38%)

Источник: разработки авторов.

ТАБЛИЦА 6.3

Доля учащихся, ответивших на задания теста по математике правильно: Британская Колумбия, Канада, %

Число заданий	Правильные ответы, %	Описание задания	Комментарии к неправильным ответам
2	53	Текстовая формулировка проблемы, включающая деление с остатком (деление числа объектов на число объектов другого множества той же размерности и определение того, насколько левая часть превышает делитель)	Более четверти (28%) лучше справились с вычитанием, а не с делением; 13% правильно выполнили деление, но не определили остаток
21	55	Текстовая формулировка проблемы, включающая умножение, вычитание и деление (выявлено наибольшее число заданий, которые учащиеся выполняют правильно по сравнению с более ранними результатами)	Наиболее распространенные ошибки: неправильные вычисления (19%), пропуск части информации (10%), и использование только части информации в вычислениях (10%)

Источник: Министерство образования Британской Колумбии, 1999.

- Большинство учащихся выполнили задание неправильно, поскольку не выбрали вариант D.
- Учащиеся, выбравшие вариант A, возможно, не знают смысла символа 2 (возведения в степень).
- Учащиеся, выбравшие вариант B, возможно, изменили знак внутри скобки на противоположный из-за минуса, который стоит перед скобкой.

- Учащиеся, выбравшие вариант С, возможно, умножили 3 на 4 и затем вычли 1, вместо того, чтобы вначале вычесть 1 из 4.

Хотя отчеты предоставляют потенциально ценную информацию для учителей, простое распределение материалов по школам вряд ли будет эффективно по следующим причинам.

- Учителя, заваленные разными документами по учебному плану, вряд ли увидят в отчете по национальной оценке ценный материал, заслуживающий их первоочередного внимания. Так, исследование, проведенное в одном из городов Новой Зеландии, установило, что большинство учителей практически не имели понятия о том, каковы потенциальные возможности отчетов по национальной оценке для повышения качества их работы. Копии отчетов распространялись по школам, однако очень немногие учителя знали об их существовании или прочитали их (Lovett, 1999).
- Отчеты, даже специально предназначенные для школ и учителей, не всегда оказываются достаточно легкими в использовании (Arregui and McLauchlan, 2005; ОЭСР, 2004).
- Учителям нужно помочь понять смысл предлагаемых реформ и сделать выводы относительно их уместности в конкретных образовательных условиях.
- Учителям нужно оказать поддержку и в том случае, если для достижения целей, предложенных в нововведениях, они должны изменить сам процесс обучения, например, сместить акцент с механического запоминания и группового воспроизведения материала на процедуры, развивающие ответственность учащихся и способствующие их более активному участию в учебной деятельности (Chapman and Snyder, 2000; Grant, Peterson and Shojgreen-Downer, 1996; Ravela, 2005; Snyder and others, 1997).

Интерпретация результатов и оценка их релевантности

Даже без национальной оценки руководители и учителя школ, особенно те, которые обслуживают учащихся из семей с неблагополучными фоновыми социальными характеристиками, должны четко представлять, будет ли испытывать трудности в обучении значительное число учащихся. В этой ситуации результаты национальной оценки могут послужить не более чем стимулом к конкретным действиям при реагировании на проблемы. Однако в то же время они, наряду с коррекционными мерами (например, идентификацией отдельных категорий школ как проблемных), могут объяснить происхождение трудностей и выявить факторы, которые препятствуют или, наоборот, способствуют учебной деятельности учащихся. Тем не менее школы должны четко представлять, что фактические условия, в которых они работают, зачастую значительно меняются в зависимости от местоположения школы.

В этом отношении внимания заслуживают множество факторов. Школы отличаются по уровням достижений их учащихся. Причины появления низких результатов тоже разные. Учителя не одинаково знают предмет, который они будут преподавать, обладают разными педагогическими умениями и отношением к работе. Школы отличаются по доступным ресурсам и способам их использования, по ограничениям, которые существуют в их работе (например, очень большие классы) и по своей способности к изменениям. Они также отличаются по социально-экономическим условиям, в которых работают, по готовности учащихся к учебной деятельности при поступлении в школу и по той поддержке, которую они получают от семьи и сообществ. Учителям следует учитывать эти данные при анализе релевантности результатов оценки применительно к своей ситуации. В частности, они должны определить знания

и умения, которые нуждаются в наиболее интенсивном преподавании, и необходимые изменения, которые, вероятно, будут самыми эффективными.

Оценка, основанная на выборке, не предоставляет информации по стандартам достижений в отдельных школах. Поэтому необходимо определить, каким способом сравнивать достижения учащихся в одних школах с достижениями учащихся в других, если результаты национальной оценки являются релевантными в отношении существующих проблем в обучении. Уругвай обратился к решению этой задачи путем разработки инструментария оценки, доступного школам, которые не участвовали в национальной оценке. Этот подход позволяет всем школам сравнивать достижения своих учащихся с национальными нормами. Однако у него есть недостаток, связанный с тем, что возможности осуществления мониторинга достижений в системе образования в течение долгого времени были поставлены под угрозу, так как широко доступные задания не могут использоваться несколько раз. Альтернативный подход должен основываться на разработке и обеспечении доступности для всех школ ряда стандартизированных тестов достижений, которые учителя могли бы предъявлять своим собственным учащимся, чтобы определить их положение относительно национальных норм, которые обеспечивают стандартизированные тесты. (Стандартизация теста заключается в определении норм его выполнения для выборочных совокупностей или генеральной совокупности учащихся в стране. – *Примеч. ред.*)

Ценность стандартизированных тестов резко возрастет, если они будут предоставлять диагностическую информацию о достижениях учащихся, обеспечивающую основания для корректирующих воздействий.

Применение результатов

Когда установлена релевантность результатов национальной оценки для отдельной школы, ей необходимо подробно рассмотреть достижения своих учащихся (идентифицируя пробелы

и слабые стороны в учебной деятельности) и проанализировать обстоятельства, которые могут сыграть существенную роль в данной ситуации. Это должно способствовать разработке стратегий для решения выявленных проблем. Многие учителя сочтут данную задачу очень трудной и потребуют значительной поддержки. Например, министерство образования может дать руководящие принципы и примеры эффективной практики обучения. Она может быть представлена также семинарами для руководителей школ и учителей или вылиться в форму локальной помощи со стороны инспекторов или наблюдателей. Все они могут внести свой вклад в обсуждения, организуемые на встречах штатного состава школ.

Учителям необходима специальная помощь для выявления тех аспектов своей практической деятельности, которые нуждаются в изменении; описания предпочтительных методов обучения; описания различий между существующими и желательными методами обучения; определения тех стадий, которые покажут продвижение к рекомендуемым методам обучения. Это продвижение будет расти, и можно надеяться, что по ходу этого роста учителя адаптируют существующие методы обучения путем перехода к тем, которые, вероятно, окажут положительное влияние на учебную деятельность учащихся.

Некоторые отчеты по национальной оценке включают специальные рекомендации к действию (см. дополнительный материал 6.6). Эти рекомендации могут носить специфический характер применительно к определенной стране и учебному плану и, возможно, не будут релевантными для других целевых ориентиров в образовании. Однако существует точка зрения, согласно которой если национальная оценка включает рекомендации, то она может служить оперативным руководством или контрольным списком проблем для проведения ряда обсуждений с сотрудниками школ относительно того, как директор и учителя анализируют практику обучения и в какой мере она соответствует практике, рекомендуемой в отчете по национальной оценке.

Дополнительный материал 6.6**Рекомендации по результатам национальной оценки по математике, 5-й класс: Ирландия**

Следующие рекомендации взяты из руководства, которое было издано по результатам национальной оценки Министерством образования и науки Ирландии.

- Учителя должны поддерживать учащихся при изучении раздела «Форма и пространство», привлекая их к решению задач, требующих рассуждений о форме и пространстве.
- Учителя должны активизировать работу, касающуюся сбора и анализа данных, построения и интерпретации графической информации в области географии и естественнонаучных дисциплин. Необходимо изыскать возможности для применения знаний по разделам «Числа», «Меры» и «Формы и пространство» в других предметных областях и в реальных контекстах.
- Учителя должны сделать значительный акцент на разделе «Меры», обеспечивая возможности учащихся переносить приобретенные знания и умения в практические действия на новые проблемы.
- При обучении математике учителя должны усилить акцент на формировании у учащихся умений более продуктивного уровня, включая умения по применению знаний и решению проблем.
- Руководители и учителя школ должны поддерживать участие родителей в развитии подготовки своих детей по математике с помощью: предоставления информации относительно изменений в учебном плане и обучающих методах; советов привлечении детей к занятиям математикой дома; руководств по использованию домашней работы (включая количество времени, которое необходимо уделять домашней работе), чтобы активизировать учебную деятельность учащихся.
- Руководители и учителя школы должны расширить использование калькуляторов, а также информационных и коммуникационных технологий в преподавании математики. Калькуляторы, а также информационные и коммуникационные технологии должны использоваться не только для развития таких умений, как основные вычисления, но и для развития умений рассуждать и решать проблемы.

Источник: Surgenor and others, 2006; информация также доступна на сайте <<http://www.erc.ie>>.

Очень часто отчеты по национальной оценке не содержат специальных рекомендаций. В этом случае, обращаясь к проблемам в достижениях учащихся, учителя должны полагаться на свои собственные ресурсы и изобретать свои стратегии улучшения обучения. В Уганде рассмотрение результатов оценки (например, низких результатов при чтении большого текста, письме и геометрии) привели к тому, что руководящий состав ряда школ определил способы улучшения методов и условий для учебной деятельности в своих школах. Эти методы включали: совершенствование пошаговых инструкций к ответам для выявления продвижения учащихся (или его отсутствия); проверку планов работы учителей и планов урока; вовлечение учащихся в создание учебных материалов; проведение соревнований среди учащихся по чтению и решению задач на действия с числами (Acana, 2008).

Рассмотрение результатов исследований в образовании может предоставить информацию для обсуждения на встречах штатных учителей школ и других форумах, а также обратиться к проблемам, установленным в национальной оценке. Поскольку учителя вряд ли знакомы с методами исследования результатов, ниже мы приводим основные данные, которые можно представить на обсуждение. Как правило, в ходе дискуссии анализируются результаты исследований: эффективности деятельности школ (дополнительный материал 6.7); эффективности деятельности учителей (дополнительный материал 6.8); связей между результатами учебной деятельности учащихся; связей результатов каждого учащегося с семейными факторами (дополнительный материал 6.9).

Рассматривая результаты исследований, учителя должны учитывать целый ряд факторов. Во-первых, результаты в основном базируются на данных, полученных в индустриально развитых странах, поэтому они редко совпадают с аналогичными данными в развивающихся странах. Исключение составляет тесная связь между достижениями учащихся и качеством рабо-

Дополнительный материал 6.7**Темы для обсуждения: результаты национальной оценки и школьные переменные для анализа эффективности деятельности школ**

Необходимо рассмотреть следующие факторы, которые, как показало исследование, влияют на эффективность деятельности школ.

- Наличие в школах структурированной среды, содействующей учебному процессу.
- Наличие отличительных качеств лидера в деятельности директора школы.
- Создание позитивной школьной культурной среды в школе и хорошего, микроклимата (единство целей с акцентом на учебной деятельности учащихся).
- Высокие ожидания по достижениям и поведению учащихся.
- Строгая дисциплина в школе.
- Высокие моральные качества учителей.
- Наличие возможности повышения квалификации сотрудников.
- Активное участие родителей в жизни школы.
- Координация последовательных этапов обучения во всех классах.

Источники: Lockheed and Verspoor, 1991; Reynolds and Teddlie, 2000.

Дополнительный материал 6.8**Темы для обсуждения: результаты национальной оценки и переменные для анализа эффективности деятельности учителей**

Необходимо рассмотреть следующие факторы, которые, как показало исследование, влияют на эффективность деятельности учителей.

- Знание предмета.
- Умение управлять деятельностью учащихся и организовать работу в классе (например, для обеспечения гарантий минимальных нарушений дисциплины).
- Знание методов обучения (наличие четких ориентиров в преподавании; продвижение применительно к уровню учащихся; предоставление учащимся возможности практиковаться и применять на практике то, что они изучили; адекватное освещение учебного плана и предоставление учащимся возможности осуществления учебной деятельности).

Дополнительный материал 6.8

- Наличие способности обучать умениям продуктивного уровня.
- Стремление к индивидуализации обучения.
- Способность к осуществлению учебной деятельности в сотрудничестве с учащимися, которое облегчает их групповую работу.
- Высокие ожидания относительно достижений и поведения учащихся при общении.
- Пригодность и адекватность учебных ресурсов (учебники и другие учебные материалы).
- Проведение мониторинга и оценка прогресса учащихся (использование и частота повторения слабоизученного материала; умение сформулировать соответствующие вопросы, которые задаются на подходящем уровне трудности и требуют, чтобы учащиеся отображали, структурировали и применяли знания, а не просто воспроизводили информацию).
- Частота и контроль домашней работы.
- Отслеживание посещаемости.
- Время, затраченное на подготовку к обучению в классе.
- Уверенность, восприятие и предположения (включая ожидания относительно достижений учащихся).

Источники: Brophy and Good, 1986; Fuller, 1987; Teddlie and Reynolds, 2000; Wang, Haertel and Walberg, 1993; Wenglinsky, 2002.

ты учителей, а также с качеством учебного оборудования, в частности, учебников (Reynolds, 2000). Во-вторых, по одним областям учебного плана деятельность учителя может быть более эффективной, чем по другим. Например, для обучения в области искусства или проведения лабораторной работы по естественным наукам требуются совсем другие умения, чем для преподавания языка или математики (области учебного плана, проанализированные в большинстве исследований школьной эффективности). В-третьих, учителя могут более эффективно работать с одними категориями учащихся, чем с другими (Campbell and others, 2004).

Дополнительный материал 6.9**Темы для обсуждения: результаты национальной оценки, индивидуальные факторы учащегося и факторы семьи, влияющие на учебную деятельность учащегося**

Исследование показывает, что существенное влияние на достижения учащихся оказывают следующие факторы.

- Характеристики учащегося (мотивация, ответственность, качество питания и здоровье).
- Посещаемость.
- Образование родителей и уровень дохода семьи.
- Семейная поддержка (ожидание высоких учебных достижений, помощь при подготовке домашнего задания, чтение детям, общение и взаимодействие ребенка с учителем).
- Требования к учащимся вне школы (работа или обязанности присматривать за детьми).

Источник: компиляция авторов.

Эффективность деятельности школ. Общие факторы, связанные с эффективностью на уровне школы (см. дополнительный материал 6.7), подчеркивают важность роли директора и развития школьной культуры, которая способствует росту учебной деятельности учащихся. Такие условия, как соблюдение школьной дисциплины и норм морали, также играют немаловажную роль. Обсуждения штатного преподавательского состава школ должны сосредоточиться на связи между результатами национальной оценки и школьными факторами эффективности, которые являются наиболее релевантными ситуации в школе.

Эффективность деятельности учителя. На протяжении ста лет эмпирические доказательства той ключевой роли, которую играют фоновые факторы семьи учащихся в развитии их способности учиться, время от времени вынуждали исследователей

недооценивать эффекты воздействия характеристик школы, в частности методов обучения в классе. Недавнее исследование, основанное на данных, собранных для программы NAEP и рассмотренных с помощью сложных методов статистического анализа, опровергает эти сложившиеся представления. Результаты исследования указывают, что методы обучения в классе (например, ориентация на интеллектуальные умения продуктивного уровня, применение методик решения одних проблем к новым проблемам и индивидуализация обучения) повлияли на достижения учащихся по математике, полученные в 8-м классе, по крайней мере так же сильно, как и фоновые факторы (Wenglinsky, 2002).

Факторы, связанные с эффективностью деятельности учителей (см. дополнительный материал 6.8), идентифицируют их знание предмета и педагогические умения. Наиболее важные условия обучения включают: возможности для осуществления учебной деятельности; время, отведенное для выполнения заданий; структурирование учебного материала; оценка и осуществление обратной связи с учащимися; уровень адекватности ресурсов и использование множественных путей решения проблем для приспособления обучения к тем вариациям в знаниях и опыте учебной деятельности учащихся, которые можно обнаружить в любом классе (Scheerens, 1998; Wenglinsky, 2002). Ориентация в контроле на высокие достижения учащихся в классе укрепляет потребность в высоких достижениях на школьном уровне (см. дополнительный материал 6.7). Связи между эффективностью деятельности школы и эффективностью деятельности учителей также находят свое отражение в потребности скоординировать обучение ряда классов. Поэтому сотрудники школ в ходе обсуждения результатов национальной оценки должны рассматривать их роль для преподавания во всех классах, в частности в младших классах по сравнению с тем уровнем, на котором осуществлялась оценка.

В идеале механизм для улучшения методов преподавания должен включать обзоры, подготовленные сотрудниками школы, с целью исследования и интерпретации данных по выполнению тестов учащимися и для развития обучающихся стратегий роста достижений учащихся. В этом процессе важна поддержка со стороны инспекторов, наблюдателей или советников, поскольку учителя обычно стремятся применять в своих классах только то, что они уже умеют. Наиболее вероятно, что изменения произойдут и будут поддержаны в тех случаях, когда будут соблюдаться следующие специальные условия (Elmore and Burney, 1999; Fullan, 2001; Reezigt and Creemers, 2005).

- Ясно определена природа той или иной проблемы (например, низкие результаты выполнения заданий, относящихся к разделу «Десятичные дроби») и ее возможные причины.
- Выявлены потребности усовершенствования и установлены четкие цели (например, установление связи достижений в определенных областях, повышение позитивного отношения учащихся к своей роли в учебном процессе, рост посещаемости).
- Разработан пошаговый план, в котором выделены приоритетные области для обучения, что позволяет обратиться к установленным недостаткам.
- Установлены ожидаемые результаты для достижений учащихся.
- Разработаны стратегии на основе анализа расхождений между существующей и желательной практикой обучения.
- Мониторинг прогресса учащихся осуществляется на регулярной основе.
- Центр внимания в обучении сосредоточен на конкретном применении общих представлений.
- Учителя заняты реальным учебным процессом, а не описанием практики обучения.
- Для учителей обеспечены возможности взаимодействия друг с другом и с инспекторами или наблюдателями в совместной работе по совершенствованию процесса обучения.

- Ведется и рефлексруется оценка прогресса осуществления плана (например, при определении того, были ли достигнуты заявленные в нем цели).

Должны быть также отмечены другие, менее явные, аспекты некоторых заданий, предназначенных для эффективной деятельности учителей, которые приводятся в дополнительном материале 6.8. Если система взглядов, лежащая в основе ценностей обучения, целей учебного плана, ожиданий учителей и их стремления помочь учащимся более важны, чем изменение деталей учебного плана или обучающих стратегий (Anderson, 2002; Robertson, 2005), тогда стратегии для улучшения учебной деятельности учащихся должны будут охватывать более широкий диапазон обстоятельств, чем это обычно принято на практике.

Факторы, касающиеся учащегося, семьи и сообщества. В рамках дискуссий о результатах национальной оценки руководители и учителя школ должны проанализировать факторы, влияющие на результаты учебной деятельности, касающиеся каждого отдельного учащегося, его семьи и сообщества (см. дополнительный материал 6.9). Если в ходе анализа выявляются проблемы, то действия, которые помогут их решить и косвенно повлиять на познавательное развитие учащихся, должны осуществляться не на уровне класса или школы, на более высоком уровне. Например, данные на уровне учащегося могут включать информацию об обеспечении питания школьников или о необходимости дополнительного питания в случае пищевого дефицита. Однако последняя ситуация обязательно потребует помощи извне, например привлечения родителей, потому что вряд ли будет возможно покрыть необходимые расходы средствами школьного бюджета. Школы не в состоянии получить точной информации об образовании родителей учащихся, но они могли бы консультировать родителей по вопросам создания тех условий в семье, которые бы благоприятствовали учебной

деятельности их детей. Так, в результате национальной оценки во Вьетнаме (см. главу 2) родителям посоветовали обеспечить полноценное личное пространство для учащихся дома и приложить все усилия, чтобы их дети как можно реже пропускали занятия (Всемирный банк, 2004).

В обзоре результатов национальной оценки учителями школ могут привести свои соображения о том, как усилить взаимодействие между школой и домом. Во многих странах школы поддерживают связь с семьями учащихся, хотя в некоторых случаях она существует чисто номинально. Чтобы родители вновь начали играть активную роль в учебной деятельности своих детей, эту связь необходимо восстанавливать и развивать. Диапазон стратегий в этом отношении весьма широк: привлечение родителей к управлению и защите школы (например, участие в ассоциациях, членами которых являются и родители, и учителя, или в школьном комитете по управлению); помощь родителей во внеучебной деятельности школы; проведение лекций и семинаров для обсуждения школьной политики или домашних условий, которые способствуют учебной деятельности; развитие программы привлечения родителей (или других взрослых) к учебной деятельности учащихся дома, включающей направленность на развитие критического мышления, умений решать проблемы и вести беседу, наблюдения за выполнением домашней работы (Kellaghan and others, 1993).

На учебную деятельность учащегося в школе могут также повлиять более широкие контекстные факторы, такие как социально-экономические характеристики сообщества, которое обслуживает школа. Однако эти факторы в значительной степени находятся вне ее контроля. Тем не менее поддерживать сотрудничество школы с местными представителями административного и политического сообществ необходимо. Сообщества могут взять на себя заботу о школьных помещениях или об обеспечении учебными ресурсами. Особенно важно

предоставить сообществу сведения о том, что школы делают все возможное для развития детей. В своих сообщениях руководители и учителя школы должны четко отразить, что они верят в результаты осуществляемых ими мероприятий.

Планирование вмешательства

Договариваясь о конкретных действиях после обсуждения результатов национальной оценки, директор школы и учителя должны прежде всего составить план по реализации ряда стратегий, направленный на улучшение учебной деятельности учащихся. В нем необходимо определить роль каждого исполнителя. Кроме того, время от времени нужно его проверять, тем самым осуществляя мониторинг его выполнения. Например, в плане может быть определено, что учителю, который обладает большим опытом в обучении письму, следует поделиться своими идеями с коллегами на собрании сотрудников школы. План должен быть реалистичным, учитывать размер школы, пригодность учебных ресурсов, способности отдельных учителей и вероятность получения поддержки от министерств образования и членов местного сообщества. В идеале в плане должны быть определены следующие позиции (Reezigt and Creemers, 2005):

- время, необходимое для действий;
- приоритеты и порядок действий;
- стратегии (например, семинары, тренинги);
- вовлеченность внешних агентов, их роли и задачи, если это необходимо;
- сотрудники, которые будут привлечены, и их задачи;
- роль и сфера полномочий сотрудников, которые будут активно привлекаться (чтобы предотвратить появление более поздних проблем);
- роль учащихся, родителей и сообщества.

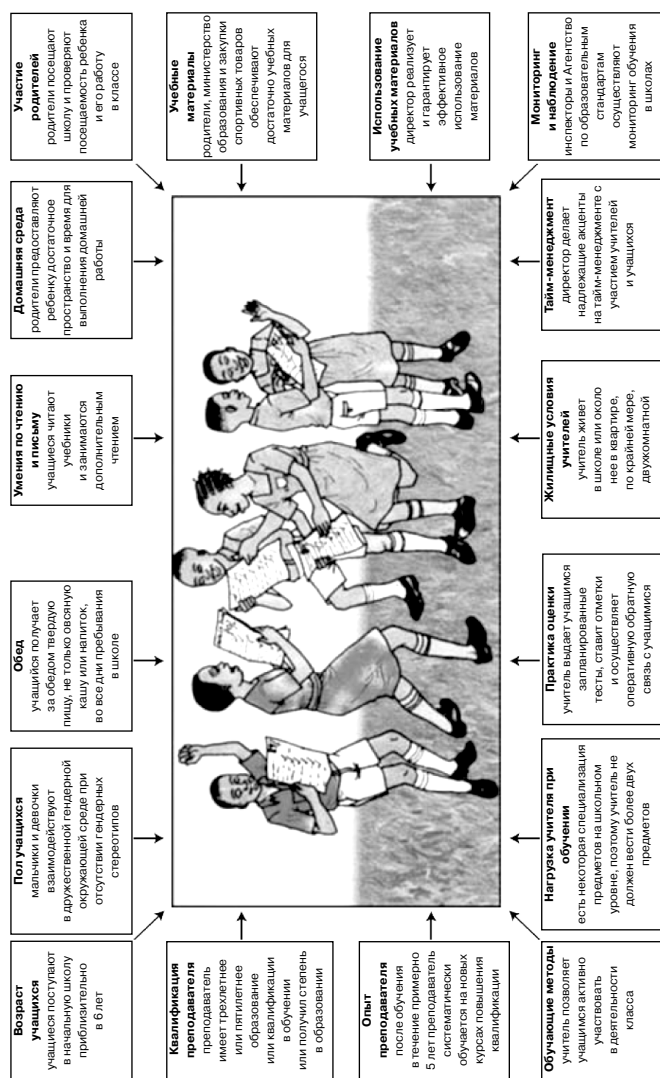
ВЫВОДЫ

Многочисленные и весьма сложные факторы, влияющие на достижения учащихся и представленные на схеме, подготовленной Национальной экзаменационной комиссией Уганды по итогам ряда программ национальной оценки (рис. 6.3), выдвигают на первый план проблемы, с которыми регулярно сталкиваются школы и учителя. Свою роль в устранении недостатков, выявленных в знаниях учителей своего предмета или в педагогических навыках, очевидно должны сыграть курсы повышения квалификации. Однако, вероятно, потребуются и другие инициативы в школах, которые должны учитывать результаты национальной оценки и реагировать на них. В случае оценки, основанной на выборке учащихся, первоначальная задача состоит в том, чтобы выявить релевантность результатов оценки проблемам, существующим в отдельных школах. После этого, если результаты отнесены к числу релевантных или если школа была включена в генеральную совокупность для оценки, она сталкивается с проблемой выбора стратегий, которые могут повысить достижения учащихся зачастую в условиях ограниченных ресурсов. В данной главе основной акцент сделан на значении результатов национальной оценки для преподавания. Если национальная оценка не приводит к непосредственным действиям, то именно школа должна руководить разработкой стратегий таких действий с помощью исследования эффективности своей деятельности и деятельности учителя, а также роли дома и окружения в поддержке учебной деятельности учащихся.

Иногда результаты национальной оценки могут указать на необходимость принятия определенных мер, со стороны и учителей (например, предоставление большего времени на изучение отдельных тем, более заботливое отношение к проверке домашней работы и усиление обратной связи с учащимися). Многое зависит и от родителей. Например, они могут поощрять детей за регулярное посещение школы и отсутствие

Рис. 6.3

Факторы, которые влияют на достижения: Уганда



Источник: Асапа, 2006. Воспроизводится с разрешения правообладателя.

пропусков без уважительной причины или следить за тем, чтобы дети выполняли домашнее задание полностью. Отделы министерства образования иногда вынуждены предоставить школам персональные и финансовые ресурсы, чтобы определить значение результатов оценки для подготовки учителей и совершенствования практики обучения в классе; сообщить результаты оценки школам; помочь школам в интерпретации результатов и понимании их релевантности существующим проблемам; помочь школам в реализации стратегий, направленных на повышение низких достижений учащихся; гарантировать, что при реализации стратегий школам будет оказана непрерывная поддержка.

Чтобы увеличить влияние национальной оценки на практику обучения в классе и подготовку учителей, рекомендуется выполнить следующие действия.

- Привлечь к национальной оценке педагогов, осуществляющих подготовку учителей.
- Собрать релевантную информацию о методах работы в классе с помощью анкет, предъявленных в национальной оценке.
- Установить значение результатов национальной оценки для обучения учителей и их информирования до начала и во время работы.
- Поощрять использование результатов оценки для подготовки учителей и учащихся.
- Подготовить легкие для чтения отчеты о результатах оценки для учителей.
- Поощрять школы в подготовке обзора национальной оценки и анализе его значения.
- Рекомендовать учителям рассмотреть школьные и внешкольные факторы, которые влияют на результаты деятельности учащихся.
- Поощрять школы к увеличению участия родителей в образовании своих детей.

- Поощрять школы и учителей к проведению необходимого числа изменений в школьной организации и в подходах к обучению.
- Проводить мониторинг и оценку эффектов тех изменений, которые были предприняты.

РЕЗУЛЬТАТЫ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ОСВЕДОМЛЕННОСТЬ ОБЩЕСТВЕННОСТИ

Описание деятельности системы образования может сконцентрировать общественное внимание на образовательных проблемах и стимулировать национальные общественные дебаты. Например, на общественность, несомненно, произвело бы шокирующее впечатление сообщение о том, что достижения по чтению большей доли учащихся 5-го класса оцениваются на уровне ниже чем «высокий». Хотя политические деятели иногда могут обнаружить, что управление потоками этого вида информации можно использовать в своих интересах, долгосрочные преимущества открытой информационной системы, вероятно, перевесят краткосрочные недостатки. Информация о достижениях учащихся может быть полезной для дебатов об образовании, которые в свою очередь будут способствовать росту общественной поддержки усилий по улучшению системы (Greaney and Kellaghan, 1996; Vegas and Petrow, 2008). Фактически публикация результатов национальной оценки даже в тех случаях, когда она основана на выборочных данных, придает операции оценки статус «высокой ставки» в том смысле, что она представляет вниманию общественности информацию о качестве деятельности школ,

таким образом привлекая их внимание к проблемам ответственности за результаты образования (Charman и Snyder, 2000).

Как бы там ни было, результаты исследований (включая программы национальной оценки) будут определяться политическими соображениями, установленными методами и культурой политической элиты (см. главу 1). В некоторых ситуациях власти предоставляют агрегированную информацию по национальному и региональным уровням достижений, но не по школам или достижениям на уровне отдельных учащихся. В других случаях они обеспечивают доступ к обоим видам информации. Однако иногда власти сопротивляются обеспечению доступа к любому виду информации. В некоторых странах доступ может быть получен в условиях свободы информационных процедур (см. дополнительный материал 7.1). В отсутствие свободы в соответствии с информационным законодательством члены правительств, возможно, не считают, что общественность имеет право знать результаты национальной или международной оценки. В Перу правительство Фухимори отказалось предоставить общественности результаты своей первой программы национальной оценки по политическим причинам (Arreguı and McLauchlan, 2005). В число других стран, которые отказались или задержали публикацию результатов национального уровня по программам международной оценки, входят Мексика (Fegger and Arreguı, 2003) и Румыния (Gilmore, 2005). Для отказа в публикации информации было множество причин, включая опасения по поводу того, что результаты могут быть использованы для межшкольных сравнений в случае оценки генеральной совокупности учащихся, что родители не будут в состоянии понять информацию об оценке и что результаты оценки не были методически адекватными (Fegger, 2006).

Традиции и предшествующие соглашения с группами, такими как союзы учителей, могут ограничить степень обнародования результатов. В Уругвае, например, союзы учителей, которые были первоначально настроены решительно против самой

концепции национальной оценки, но в результате обсуждений с Министерством образования согласились на ряд шагов после оценки, основанной на генеральной совокупности учащихся (дополнительный материал 7.2). Эти шаги не способствовали широкому осведомлению общественности, но действительно помогли в том, чтобы обеспечить полезную информацию для штатного состава системы педагогического образования.

В случае проведения выборочной национальной оценки могут потребоваться специальные усилия для привлечения общественного внимания. Конечно, этот тип оценки не привлечет такое же внимание средств массовой информации, как международное исследование, которое включает комментарии результатов выполнения тестов в стране относительно других стран, сопровождаемые заголовками типа «Швеция на уровне развивающихся стран» (Husén, 1987) и «SA [житель Южной Африки] Нижние ранги учащихся по математике и естественным наукам» (Monage, 2006). Оценка, основанная на генеральной совокупности учащихся, как правило, имеет тенденцию привлекать больше внимания средств массовой информации, чем выборочная оценка в тех случаях, когда информация опубликована (как это часто бывает) по результатам выполнения тестов отдельными школами.

Дополнительный материал 7.1

Законы о свободе информации

Не во всех странах законодательство дает общественности право на доступ к информации, находящейся в распоряжении правительства. Только в 68 из 195 стран в 2006 году существовали законы о свободе информации.

Источник: Banisar, 2006.

Дополнительный материал 7.2**Соглашения по релизу результатов национальной оценки:
Уругвай**

В Уругвае союзы учителей согласились предпринять следующие шаги по разрешению релиза результатов национальной оценки.

- Результаты национальной оценки обнародованы на семинаре для инспекторов начальной школы.
- СМИ могут посещать семинар и получать результаты национального уровня.
- Результаты по отдельным школам не публикуются и доступны только для инспекторов и директоров школ.
- Некоторые родители могут попросить результаты на школьном уровне у директора школы.
- Публикация результатов ранжирования школ не разрешена.
- Результаты быстро возвращают школьным властям (обычно спустя 40 дней после предъявления).
- Учителя в классе не должны нести ответственность за достижения своих учащихся.
- Оценки должны интерпретироваться, принимая во внимание подходящие социокультурные факторы.

Источники: Benveniste, 2000; Ravela, 2005.

**ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ РЕЗУЛЬТАТОВ ОЦЕНКИ ДЛЯ
СООБЩЕНИЯ ОБЩЕСТВЕННОСТИ**

Министерства образования с нарастающей интенсивностью признают важность распространения результатов программ национальной оценки. Они полагают, что результаты оценки могут служить функции «просвещения» не только политиков и лиц, принимающих решения (см. главу 5), но также и всей общественности при выявлении проблем в образовании и размышлениях о них. В Латинской Америке информирование общественности об уровнях достижений учащихся рассматривается как важный результат национальной оценки (Arregui

and McLauchlan, 2005). В Аргентине, используя результаты достижений учащихся, национальное правительство намеревалось предоставить общественности доказательство того, что система образования была в кризисном состоянии, которое требовало широкого проведения масштабных реформ (Benveniste, 2002).

В Чили, где результаты национальной оценки используются для создания сети поддержки в гражданском обществе, ежегодная их публикация актуализировала проблему перестройки системы образования (Benveniste, 2002). Опять же в Чили придание гласности связи между социально-экономическим статусом семей учащихся и их достижениями повысило социальные требования к равенству. В частности, оно сформировало базис для политики, сконцентрированной на том, чтобы выявлять несправедливость в системе образования (Meckes and Carrasco, 2006).

В других частях мира команды по национальной оценке тратят значительные силы, чтобы сообщить общественности результаты исследований во всех деталях. Например, и в Англии, и в США, где стратегии оценки отличаются радикально, в Интернете имеются подробные сайты, посвященные программам оценки. Правительства в обеих странах используют сайты для поддержки своей политики оценки и сообщения общественности данных о качестве образования. Сайт программы NAEP в США (<http://nces.ed.gov/nationsreportcard>), основанный на данных выборочной оценки, включает: раздел с информацией о текущих и прошлых программах национальной оценки; раздел, предназначенный родителям; детализацию результатов оценки по отдельным предметам, включая структуру оценки; порядок оценки и типовые задания (см. дополнительный материал 3.6). Руководство для родителей предоставляет информацию о том, как отбираются учащиеся для оценки и сколько времени требует тест. В нем также заявляется, что информация об отдельных учащихся и школах не будет размещена в открытом доступе. Сайт для оценки, основанной на генеральной со-

вокупности учащихся в Англии (<http://www.direct.gov.uk/en/Parents/Schoolslearninganddevelopment/ExamsTestsAndTheCurriculum/>), описывает тесты и оценку учителей.

Последовательное выполнение национальной оценки вместе с регулярной публикацией ее результатов, вероятно, будет способствовать повышению осведомленности общественности (Meckes and Carrasco, 2006). Комментарии по поводу значимых цифр в результатах оценки также, вероятно, повысят осведомленность общественности. В США после появления релиза по результатам повторной программы TIMSS для учащихся 8-х классов госсекретарь США по образованию сказал следующее: «Мы не можем ожидать значительного продвижения в математике и естественным наукам по сравнению с результатами других стран, если обучать геометрии будут учителя истории, а химии – учителя по физкультуре» (Riley, 2000). Хотя эти комментарии последовали после публикации результатов международной оценки, они несомненно отражают наблюдения существующей национальной практики в сфере образования и, вполне возможно, были вызваны программой национальной оценки.

РОЛЬ СРЕДСТВ МАССОВОЙ ИНФОРМАЦИИ

Важным источником информации о национальной оценке являются СМИ. Например, в Германии большую роль для роста интереса в политических кругах к проблемам знаний учащихся и деятельности системы образования в целом в программе PISA сыграли газеты. В течение одного месяца после релиза результатов PISA-2000 немецкие газеты примерно 687 страниц посвятили вопросу существования более высоких результатов, полученных в других странах (OECD and INES Project, Network A, 2004). Хотя исследование было международным, результаты использовались

для идентификации основных различий в достижениях между 16 федеральными землями (*Länder*) внутри страны, что вызвало дебаты по поводу роли немецкого среднего образования. Результаты оценки также интерпретировались как доказательства социальной несправедливости. Общественная и политическая реакция привела к реформам, которые включали введение образовательных стандартов и норм по всей стране, подготовку общего годового отчета по образованию и увеличению количества часов обучения в школе (Rubner, 2006).

СМИ играют важную роль в распространении результатов исследований, однако иногда возникают проблемы. Некоторые репортеры, сообщая о результатах национальной (и международной) оценки, уделяют больше внимания отрицательным данным о системе образования, совершенно забывая о положительных. Другие стремятся вызвать сенсацию или исказить некоторые аспекты. Потенциальная проблема состоит в том, что исследования низкого качества могут привлечь такое же большое внимание журналистов, как и высококачественные исследования. Таким образом, результатам оценки, проведенной в нескольких школах (или даже в одной), СМИ могут придать такую же важность, как и результатам оценки, основанной на репрезентативных выборках учащихся.

Освещение программы национальной оценки в СМИ с течением времени усложняется по терминологии и интерпретируемым данным. Ранее данные приводились в процентах, а теперь, как правило, используются термины уровней мастерства учащихся (например, «адекватный», «высокий», «превосходный»). Соответствующим образом сообщают о тенденциях в достижениях, а результаты детализируются по отдельным категориям, по регионам, гендерным признакам и социально-экономическому статусу.

Дополнительный материал 7.3**Меры для придания результатам национальной оценки максимального эффекта воздействия на образование**

Следуя приведенным ниже рекомендациям, можно придать результатам национальной оценки максимальную эффективность.

- Выработайте соглашения по ограниченному числу ключевых сообщений, которые будут последовательно переданы официальным докладчикам.
- Предугадайте вопросы и запланируйте стандартные ответы.
- Помните, что журналисты и заинтересованные группы пользователей нередко склонны искажать ключевые сообщения, поэтому заранее подумайте о том, как решить эту проблему.
- Разработайте механизмы для мониторинга общественного и частного освещения результатов исследования, чтобы в намеченном сообщении было дано объяснение неправильному использованию результатов.
- Разработайте стратегии, помогающие противостоять искажениям интерпретации результатов.

Источник: разработки авторов.

РАЗВИТИЕ СТРАТЕГИИ КОММУНИКАЦИИ

Для повышения гарантий того, что результаты национальной оценки принесут максимальную пользу, необходимо использовать специальную стратегию коммуникации, включающую несколько шагов (см. дополнительный материал 7.3).

Поскольку системы национальной оценки являются источниками для множества отдельных результатов, первый шаг в развитии стратегии коммуникации может показаться несколько аномальным. Прежде всего, сверните все результаты в несколько ключевых сообщений, которые могут быть последовательно переданы всем докладчикам и которые будут включены во все итоговые документы, такие как пресс-релизы. Тщательно обдуманное и реализованное распространение результатов повысит готовность потенциальных клиентов

к усвоению соответствующей информации о национальной оценке и поощрит их к большему пониманию. Целью стратегии является тщательная подготовка к формированию первых впечатлений от результатов национальной оценки.

Предполагая, что у общественности возник устойчивый интерес к результатам оценки, второй шаг должен гарантировать, что целевым аудиториям дадут грамотные ответы на любые имеющие смысл непредусмотренные вопросы по поводу ключевых результатов оценки и стандартные журналистские вопросы: кто, что, когда, где, почему, как и сколько? Для обеспечения методической корректности ответы должны быть подготовлены заранее и предложены для апробации людям, не подготовленным в методических аспектах национальной оценки. Это будет способствовать лучшему пониманию планируемых ответов и трактовке всеми официальных докладов до выхода релиза.

Члены команд по национальной оценке также должны быть готовы к тому, что и журналисты, и заинтересованные группы лиц попытаются исказить ключевые сообщения в своих собственных интересах. Журналисты могут уговорить докладчиков сделать необоснованные выводы о результатах оценки. Чтобы пресечь такую возможность, официальных докладчиков следует специально обучить методу общения со СМИ. Желательно даже устраивать инсценировки интервью, где в роли «враждебных» журналистов выступят коллеги докладчиков.

Ключевые сообщения должны быть представлены в виде, который будет препятствовать необоснованным вопросам, утверждениям и искажениям со стороны групп, заинтересованных лиц и журналистов. Если членов команды по национальной оценке просят прокомментировать утверждения, имеющие отношение к третьей стороне, с намерением возбудить конфликт, то самая безопасная реакция – воздержаться от комментария этих результатов и использовать представившийся случай для повторения ключевых сообщений. В этой ситуации есть одно

исключение, когда представители третьей стороны, журналисты, или заинтересованные группы намеренно искажают факты. В этих случаях члены команды по национальной оценке должны отреагировать настолько оперативно, насколько это возможно.

Реакция на неправильную интерпретацию предполагает, что механизм ее корректировки устроен таким образом, чтобы отследить ответы, исключающие искажения, мог каждый. Фактически члены команды по национальной оценке должны разместить в целях информирования две параллельных системы: одну – для мониторинга общественного освещения исследования и вторую – для мониторинга освещения результатов для частных лиц.

Мониторинг общественного освещения результатов можно вести несколькими способами. Во многих странах предусмотрены услуги пресс-клиппинга, который сканирует ключевые СМИ для упоминания и обеспечивает детали в сообщении. Альтернативный вариант предполагает, что есть члены команды, которые отслеживают все сообщения в СМИ.

Проведение мониторинга освещения результатов оценки для частных лиц сопровождается большим количеством трудностей, но зачастую является более важным. Его лучше всего реализовывать посредством личных коммуникаций с несколькими ключевыми пользователями, по телефону или лично, так, чтобы получить достоверные ответы. Полезным будет создание списков возникающих проблем. Эти списки должны сопоставляться и обсуждаться на регулярной основе, желательно сразу после подготовки релиза ключевых продуктов. Цель обзора состоит в том, чтобы решить, каким образом приспособить ключевые сообщения для обеспечения сверхактивной и быстрой реакции на искажение фактов. Менеджеры по национальному оцениванию должны назначить специалиста для управления этой деятельностью.

ВЫВОДЫ

В наши дни не вызывает сомнений тот факт, что результаты национальной оценки имеют огромный потенциал для формирования общественного мнения, осознания общественностью образовательных проблем и увеличения поддержки мер по улучшению системы образования. Однако по ряду причин страны все еще значительно отличаются по степени своих усилий для передачи результатов оценки общественности. Во-первых, традиции в стране могут поддерживать закрытый, а не открытый подход к обнародованию информации, полученной правительственным учреждением. Во-вторых, оценка может быть непосредственно предназначено для предоставления информации управленцам в образовании и учителям, а не общественности. И в-третьих, недостаточное распространение результатов может просто свидетельствовать не более чем об отсутствии интереса со стороны общественности, быть может, вследствие отсутствия плана выполнения таких работ, либо о недостатке средств на распространение, не заложенных при формировании бюджета.

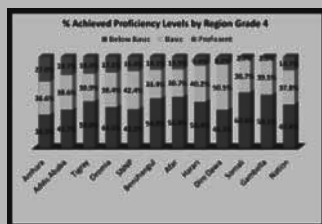
Даже когда результаты оценки распространены, цель информирования общественности может быть не достигнута, потому что отчеты по оценке будут перенасыщены методическими материалами, статистическими данными, вообще непонятными для многих читателей. Когда результаты национальной оценки распространяются СМИ, отдельные авторы и редакторы будут играть ведущую роль в определении проблем и содержания и в конечном счете того, как общественность и разработчики политики интерпретируют результаты оценки (Stack, 2006). Обзор освещения американской программы NAEP за период с 1990 по 1998 год в газетах, выявил, что авторы публикаций склонны: полагаться в значительной степени на информацию в пресс-релизах; отдавать предпочтение детальному обсуждению ряда позиций, а не общему освещению проблем; ранжи-

Дополнительный материал 7.4

Обложка брошюры для национальной оценки: Эфиопия

**ОСНОВНЫЕ МОМЕНТЫ
НАЦИОНАЛЬНОЙ ОЦЕНКИ
УЧЕБНОЙ ДЕЯТЕЛЬНОСТИ
УЧАЩИХСЯ (ETNLA)**

Третья национальная оценка результатов учебной деятельности учащихся в Эфиопии было выполнено в 1999 Е.С. (2006/2007) учебном году для учащихся 4-х и 8-х классов на национальном уровне во всех областях. Более чем 500 школ и приблизительно 23 000 учащихся и их учителей участвовали в исследовании. Здесь представлен общий взгляд, образно говоря, «с высоты птичьего полета» на проект, связанный с этой целью, и его основные результаты.

**КРАТКИЙ ОБЗОР**

Основные цели двух исследований состояли в том, чтобы предоставить информацию об общих достижениях в учебной деятельности учащихся 4-го и 8-го классов и установить главные факторы, которые потенциально влияют на учебные достижения. Чтобы достичь этих целей, использовались и количественные, и качественные подходы. Текущее исследование использует инструментарий и процедуры, более или

менее схожие с теми, которые были использованы во второй национальной оценке результатов учебной деятельности учащихся в Эфиопии. В 4-м классе в исследовании приняли участие в общей сложности 11 373 учащихся, включенных в выборку, из 305 школ во всех областях страны. Из 8-х классов были включены в выборку в общей сложности 10 806 учащихся из 280 школ, которые тестировались по всем предметам. С целью получения данных по факторам, которые определяют учебные достижения, 832 учителя и 305 директоров школ были привлечены для получения данных по 4-му классу. Для сбора информации по 8-му классу были привлечены 1 242 учителя и 280 директоров школ. В качественном исследовании участвовали в обсуждениях фокус-группы из 312 учащихся, 311 учителей и 286 родителей.

ЦЕЛИ ИССЛЕДОВАНИЯ

1. Определение учебных достижений учащихся 4-го класса по английскому языку, математике, естественным наукам об окружающей среде и чтению на родном языке.
2. Определение учебных достижений учащихся 8-го класса по английскому языку, математике, биологии, химии и физике.
3. Оценка отношения учащихся к некоторым из аффективных факторов системы образования.
4. Предоставление высшим чиновникам и заинтересованным лицам результатов и рекомендаций по учебным достижениям учащихся 4-х и 8-х классов по исследованным предметам.

Источник: Агентство по проверке качества образования и проведению экзаменов Эфиопии, 2007. Воспроизводится с разрешения правообладателя.

ровать штаты по результатам; использовать многочисленные диаграммы и графики; испытывать незначительный интерес к проблемам валидности тестов; испытывать трудности в понимании и трактовке понятий, касающиеся уровней освоения чтения, а также некоторых статистических терминов таких как статистическая значимость, процентиля (Hambleton and Meara, 2000).

Поскольку при назначении ответственных за распространение результатов национальной оценки необходимо понимать, что проблемы освещения оценки в СМИ крайне важны, следует подготовить пресс-релизы, четко сфокусированные на главных проблемах, и принять меры, чтобы персонал, выполнявший оценку, был доступен для интервьюирования. Чтобы предотвратить упрощенные интерпретации, как в случаях, когда достижения приписывают влиянию единственной переменной (например, размеру класса, посещению частной школы или количеству времени, потраченного на просмотр телевизора), команда по национальной оценке должна обстоятельно объяснить ограничения на данные.

В будущем доступ к результатам национальной оценки станет более открытым благодаря проведению брифингов СМИ, использованию других форм общественного представления результатов, разработке видеофильмов и размещению результатов в Интернете. С этой целью можно использовать разнообразные СМИ, включая центральные и местные газеты, парламентские дебаты, статьи журналов, брошюры, радиопередачи и телевизионные программы (см., например, дополнительный материал 7.4). Публикации учителей, семинары и встречи, а также информационные бюллетени, брошюры, журналы и конференции могут удовлетворить потребности наиболее заинтересованных групп пользователей, таких как учителя.

ГЛАВА

8

ОПТИМИЗАЦИЯ ИСПОЛЬЗОВАНИЯ ПРОГРАММ НАЦИОНАЛЬНОЙ ОЦЕНКИ И ИХ ЗНАЧЕНИЕ

Краткий обзор использования результатов национальной оценки, представленный на страницах этой книги, показывает, что программы оценки во многих странах обеспечили эмпирические доказательства успешности систем образования при выполнении своих функций: способствовать развитию всех учащихся и обеспечивать фундаментальные знания и умения, которые потребуются им в будущем для карьеры или продолжения образования. Программы национальной оценки дают общее представление о характере распределения достижений в системе образования, а также поднимают вопросы о степени, в которой основные принципы, связанные с доступностью, качеством, эффективностью и равенством прав, регламентируют деятельность системы. Тем самым они показывают, в какой степени уровень достижений в целом оказывается ниже ожидаемого; объясняют неудачи многих учащихся в освоении базовых компетенций после нескольких лет обучения и различия в их достижениях, связанные с полом учащихся, местом жительства и семейным фоном. Кроме того, сравнение входных данных и результатов обучения при анализе качества образования способствует сосредоточению внимания не только образовательного сообщества, но также

и политических деятелей и общественности на важности того, что получили учащиеся в результате своего образовательного опыта.

В результате изучения отчетов по национальной оценке во всех странах напрашивается вывод о том, что во многих странах уровень достижений учащихся не является удовлетворительным. Однако некоторые страны не предпринимают после оценки никаких корректирующих действий. В других странах предпринимаются некоторые действия, например обеспечиваются дополнительные ресурсы и возможности для профессионального развития учителей. В нескольких странах были применены санкции по результатам выполнения тестов учащимися, несмотря на вероятные нежелательные, если непреднамеренные, последствия, связанные с такой практикой. Выбор подходящей реакции на результаты национальной оценки остается основной проблемой для разработчиков политики и управленцев в сфере образования во всем мире.

Учитывая сложность политических и методических процедур, используемых в ходе оценки, совсем не удивительны те трудности, которые возникают при выборе адекватной реакции на результаты национальной оценки. Действия могут быть обозначены на нескольких уровнях системы образования: политические деятели в правительстве, центральная (региональная и местная) администрации, школы и учителя. Могут быть также установлены более специфические и сложные задачи. Проблемы, связанные с качеством данных национальной оценки и доверием к ее результатам, остаются нерешенными во многих странах. Способностей и политического желания или и того и другого для использования результатов оценки может быть недостаточно. Результаты оценки не всегда сообщают заинтересованным лицам, и даже если сообщают, они могут не повлиять на формирование политики и принятие решений на макроуровне или в повседневной практике в классе на микроуровне. Некоторые из этих проблем, если не все, могут

быть вполне разумно отнесены к равнодушию ключевых разработчиков образовательной политики в проектах по оценке, новизне национальной оценки, нехватке методических навыков, которые требуются при проведении оценки (эти навыки находятся на недостаточном уровне развития в очень многих странах), и к нехватке способности ответить на вызовы, заложенные результатами национальной оценки.

В данной главе приводится перечень задач (его можно использовать как контрольный список), которые ставятся в национальной оценке, чтобы оптимизировать использование ее результатов. Далее предлагаются девять способов, с помощью которых можно изменить или расширить будущую деятельность по национальной оценке, чтобы лучше описать образовательные достижения, которые, в свою очередь, должны обеспечить более адекватный базис для формирования политики и принятия решений.

ОПТИМИЗАЦИЯ ИСПОЛЬЗОВАНИЯ РЕЗУЛЬТАТОВ НАЦИОНАЛЬНОЙ ОЦЕНКИ

Представленный в книге анализ программ национальной оценки предполагает, что для оптимизации использования его результатов разработчики политики в министерствах образования и управленцы должны обратить пристальное внимание на следующие задачи.

Миссия

- Обеспечение с самого начала ясности целей национальной оценки, которая заключается в том, чтобы получить информацию о результатах функционирования системы образования в виде оценок достижений учащихся. Для этого:
 - информируйте политиков и практиков;
 - повышайте качество учебной деятельности учащихся с акцентом на развитие познавательных умений продук-

тивного уровня (включающих рассуждения, способность к идентификации и решению проблемы и способность к решению нетрадиционных задач), принимая во внимание потребности экономики, основанной на информации и глобальной конкуренции;

- идентифицируйте проблемы в существующей системе обеспечения образовательных условий и способы обращения к ним.

Контекст использования

- Развивайте способность организаций усваивать и использовать информацию, предоставленную национальной оценкой.
- Интегрируйте программы национальной оценки в существующие структуры, политику и процессы принятия решений и каналы распределения ресурсов для системы образования.
- Сформируйте у общественности и заинтересованных лиц понимание того, что программы национальной оценки предоставляют информацию, которая может быть использована для повышения качества учебной деятельности учащихся.

Проведение национальной оценки

- Привлеките политиков и лиц, принимающих решения, а также других заинтересованных лиц к участию в национальной оценке, чтобы гарантировать, что она ориентирована на решение их проблем.
- Постарайтесь гарантировать, что ответственные за методические аспекты национальной оценки обладают соответствующими компетенциями, связанными с разработкой инструментария, формированием выборки, анализом данных и подготовкой письменных отчетов.

- Постарайтесь гарантировать, что ответственные за проведение национальной оценки в школах предпринимают верные меры по коррекции учебного процесса.

Описание результатов

- Опишите достижения учащихся в деталях, чтобы удовлетворить потребности потенциальных пользователей, сделав акцент на диагностировании проблем в системе образования.
- Выявите факторы, которые связаны с высокими или низкими достижениями.

Сообщение результатов

- Предоставьте информацию о достижениях учащихся политическим деятелям, разработчикам политики и управленцам в сфере образования своевременно – тем способом и в той форме, которая им понятна и которая мотивирует их включать данные, полученные в результате оценки, в формирование политики и принятие решений.
- Подготовьте отдельные отчеты по национальной оценке, которые адаптированы к потребностям потенциальных пользователей: инспекторов и советников; преподавателей педагогических вузов, авторов учебников; администрации школ; учителей.
- Предоставьте информацию о результатах национальной оценки СМИ и общественности.

Формулирование политики и программ вмешательств

- Привлеките заинтересованных лиц к исследованию результатов оценки, установлению приоритетов для политики и формулированию программы, сосредоточенной на учебной деятельности учащихся, развитию стратегий для отражения этой политики, принимая во внимание другие соответствующие свидетельства (например, результаты исследований в об-

разовании и суждения школьных инспекторов, советников и наблюдателей).

При формулировании политики примите во внимание приоритеты, давление и ограничения, накладываемые имущественными правами, и используйте существующие структуры в системе образования (например, механизм для управления при проведении переговоров с профсоюзами), чтобы достигнуть согласия или компромисса при установлении курсов действий.

Осуществление политики и программ или вмешательств

- Разработайте критерии для отбора школ или учащихся, нуждающихся в целенаправленной поддержке, основанной на результатах национальной оценки. Проведите работу со школами и сообществами, привлеченными для развития программ и стратегий по совершенствованию учебной деятельности учащихся.
- При выборе программ вмешательств примите во внимание «передовой опыт» (например, обеспечение повышения квалификации без отрыва от работы для учителей).
- Выберите агентов для осуществления программ или вмешательств, разработанных после национальной оценки, и опишите их задачи (инспекторы, штатные советники, директора школ).
- Обеспечьте непрерывную поддержку персоналу, привлеченному к развитию и выполнению стратегий, разработанных для обращения к проблемам, установленным при оценке.

Мониторинг эффектов

- Проведите мониторинг изменений в политике и результатов вмешательств, чтобы определить, повлияли ли они на совершенствование учебной деятельности учащихся.

РАЗВИТИЕ НАЦИОНАЛЬНОЙ ОЦЕНКИ ДЛЯ ПОВЫШЕНИЯ ЕЕ ЗНАЧИМОСТИ

В этом разделе описаны девять способов, которыми деятельность по национальной оценке могла бы быть изменена или расширена для усиления ее значения.

Более подробная информация о достижениях

Как отмечалось в главе 1, одна из целей национальной оценки состоит в том, чтобы обеспечить свидетельства относительно сильных и слабых сторон в знаниях и умениях учащихся. Однако число заданий в тестах, используемых во многих программах национальной оценки, является слишком маленьким, чтобы получить приближенные данные об освоении определенных областей достижений. Чтобы увеличить число заданий, с помощью которых в оценке можно получить также данные без неоправданного увеличения нагрузки на отдельных учащихся, многие программы национальной (и международной) оценки используют буклет тестов с ротируемым дизайном (Beaton, 1994). (Для дизайна с ротацией в смежные варианты тестов включают блоки общих заданий – якорей, чередование которых по вариантам идет согласно особой схеме, предусмотренной в общем дизайне теста. – *Примеч. ред.*) Возможны несколько вариантов такого дизайна, но существенные особенности каждого состоят в том, что: тест, который выполняет каждый учащийся, включает только часть всех заданий, используемых при оценке; включение общих заданий во все буклеты позволяет связать результаты выполнения учащимися каждого теста и привести их к единой шкале. (Эффект перевода баллов в общую шкалу метрического характера достигается только при условии использования современной теории тестов IRT – *Примеч. ред.*)

Администраторы, особенно в первые годы проведения оценки, могут предпочесть, чтобы все учащиеся отвечали на один и тот же тестовый буклет и отказаться от использования буклетов тестов с ротируемым дизайном из-за сложностей, неизбежных при подготовке и предъявлении инструментария для оценки и при выполнении исследований. Однако по мере того, как они приобретают опыт, следует стремиться к дизайну тестов с ротацией, который обеспечит более подробную информацию о достижениях.

Информация о характеристиках учащихся, связанных с их достижениями

Некоторые программы национальной (и международной) оценки собирали информацию относительно характеристик учащихся, связанных с достижениями в различных областях учебного плана. Например, в национальной оценке в Чили были получены данные по Я-концепции учащихся, их отношению к школе и учебной деятельности и социальных связях. Однако в дальнейшем эти данные не были признаны достаточно полезными, поэтому практика их сбора была прекращена (Himmel, 1996).

В главе 1 упоминалось, как трудно получить надежные оценки таких переменных, а также говорилось о проблемах при установлении направлений причинной связи между ними и мерами учебных достижений. Однако недавние исследования при установлении переменных, характеризующих особенности стратегии учебной деятельности учащихся, включая, например, саморегуляцию, уверенность в себе, обязательность и мотивацию (см. Artelt and others, 2003; Postlethwaite, 2004a), установили, что они значимо связаны с достижениями учащихся в учебных областях. Эти переменные заслуживают большего внимания в программах национальной оценки, которые стремятся установить важные корреляты учебной деятельности.

Информация относительно условий учебной деятельности в классах

Большинство программ национальной оценки собирает контекстные данные по школьному и внешкольному опыту учащихся. В некоторых случаях были получены данные, весьма сомнительные по своей уместности; в других случаях собранные данные были использованы в анализе не полностью. В большинстве программ национальной оценки различные аспекты опыта учащихся, наиболее значимые для их учебной деятельности в школе, установлены не были.

Эффект воздействия программ национальной оценки на образование можно повысить, если уделить внимание проектированию анкет по контекстным факторам. Например, желательно установить потребности потенциальных пользователей в информации, полученной по результатам оценки. Необходимо также больше внимания уделить получению информации относительно повседневной работы учащихся в классе. До настоящего времени эта практика должным образом не распространена. Например, нормы охвата содержания обучения при тестировании, которые должны были внести свой вклад в трактовку достижений учащихся в международных исследованиях, не получили надлежащего распространения в программах национальной оценки.

Выявление школ, нуждающихся во вмешательстве по результатам национальной оценки

Проблема выявления школ с низкими результатами, полученными по результатам национальной оценки на основе данных по выборке учащихся, заставила разработчиков политики в некоторых странах рассмотреть возможность обращения к оценке, основанной на генеральной совокупности. Эта проблема особенно характерна для стран, где национальная оценка была введена, чтобы обеспечить информацию об учебной деятельности учащихся после отмены государственных экзаменов в конце начальной школы. Такое увеличение числа участников национальной оценки путем включения всех (или большинства) школ

и учащихся могло бы переложить на нее некоторые функции государственных экзаменов. Однако перед тем, как пойти по этому пути, отделы министерства должны серьезно рассмотреть возможные трудности и дополнительные затраты, неизбежные при оценке генеральной совокупности учащихся, и эффекты, которые можно ожидать, если должны быть применены санкции к школе, учителям или учащимся. Отделы должны также рассмотреть альтернативные варианты, включающие развитие аналитических способностей персонала в системе образования (особенно инспекторов, советников или наблюдателей) для выявления школ, нуждающихся в помощи. Более формальные процедуры, чем те, что в настоящее время используются во многих странах, вероятно, потребовали бы привлечения к тестированию учащихся (на основе стандартизированных тестов, если это возможно) учителей и использования стандартизированных рейтинговых шкал контролирующим персоналом.

Вторичные исследования

Первичные отчеты по национальной оценке, как правило, готовятся в условиях нехватки времени. Кроме того, для лучшего понимания результатов оценки пользователями могут быть выбраны менее сложные модели и методы анализа, чем используемые в более фундаментальных исследованиях. По любой из этих причин при первичном анализе вряд ли будут полностью задействованы все возможности данных, которые может обеспечить национальная оценка. Ясно, что если обеспечение доступности данных для дальнейших исследований предоставляет много потенциальных преимуществ (см. главу 3). Поэтому идеально, когда в проекте по оценке предусматриваются вторичные исследования.

Использование дизайна сбора лонгитюдных данных в национальной оценке

Во многих программах национальной оценки контекстные данные (по учащимся, школам, семьям и сообществам) собирают в то же самое время, что и данные по достижениям учащихся.

У этого подхода есть два недостатка. Во-первых, трудно сделать выводы о том, как взаимосвязаны данные о достижениях и контекстные данные. Во-вторых, «прямое» воздействие на учащихся опыта образовательной деятельности, которое представляют результаты оценки и которое непосредственно относится к этому опыту, нельзя отличить от «опосредованного» воздействия, которое отражает эффект влияния других факторов на достижения учащихся (например, генетическое наследие учащихся, достижения при поступлении в школу, а также поддержка и помощь, получаемые дома и в ближайшем окружении). Эти проблемы могут иметь адресное решение (по крайней мере до некоторой степени), если данные собираются по учащимся чаще, чем один раз (лонгитюдные данные). Тогда при оценивании достижений в последующие моменты времени принимаются во внимание характеристики отдельных учащихся, которые были оценены до этого, включая более ранние достижения и контекстные данные, чтобы определить «величину» опыта учебной деятельности каждого учащегося в школе, «добавленную» в процессе их продвижения (Linn, 2005b).

Во франкоязычной Африке (Программа анализа систем образования, или PASEC), учащиеся оценивались в начале и в конце учебного года, что позволило получить приближенные данные о приросте их достижений. Затем этот прирост связывался с внутришкольными (образование учителей, размеры классов, пригодность учебников) и внешкольными (образование родителей, расстояние, которое должен преодолеть учащийся при поездке в школу, язык, используемый в семье) факторами (Kulpo and Coustère, 1999; Michaelowa, 2001; см. также Greaney and Kellaghan, 2008, приложение С.2). Дальнейшей особенностью исследований по программе PASEC было установление ряда проблем, представляющих интерес в национальном масштабе (например, существующая политика занятости учителей), и попытка оценить эффект альтернативных мер, также с помощью лонгитюдных данных. При исполь-

зовании исследований с «приростом величины» достижений был выявлен ряд проблем, в число которых вошли неполные данные об учащихя, являющиеся результатом прекращения ими обучения, перехода в другие школы, пропуска занятий, регрессии к среднему в статистическом анализе и ненадежности мер, когда число учащихя в школе является маленьким.

Последующие социологические исследования

С целью усиления базиса для получения причинных выводов из данных национальной оценки в некоторых последующих исследованиях анализировались факторы, связанные с высокими (или низкими) достижениями учащихя. Например, в США Национальная экспертная группа по постановке целей образования провела анализ данных одного из штатов с наиболее высокими оценками (Коннектикут), чтобы идентифицировать образовательную политику и методы, которые, возможно, внесли свой вклад в получение высоких оценок по чтению. Результаты выполнения тестов в штате были связаны с тремя политическими линиями в образовании: обеспечение подробной информации относительно выполнения тестов учащимися для широкого диапазона заинтересованных лиц (районы, школы, учителя, родители); проведение мониторинга достижений учащихя в течение длительного времени; обеспечение дополнительных ресурсов для наиболее нуждающихся школьных районов, например, обеспечение профессионального развития для учителей; поддержка и оценка всех, начиная с учителей (Baron, 1999).

Дополнительная информация была получена после национальной оценки по математике (которое включало задания из Второго международного исследования по математике Международной ассоциации оценки образовательных достижений) для учащихя 13 и 17 лет в Доминиканской Республике, проведенного вследствие того, что низкие результаты выполнения

тестов учащимися вызвали значительную обеспокоенность. В этом случае исследование, включающее наблюдения в классе и интервью с учителями, было выполнено на уроках математики на выборке 8-х классов с целью выявления возможностей и учебных процедур. Проблемы идентификации были адресованы: центру по настройке и развитию учебного плана для подготовки материалов, чтобы способствовать улучшению обучения по математике; программам повышения квалификации для учителей (Luna, 1992).

Особую осторожность следует соблюдать при интерпретации результатов выполнения тестов группами учащихся с высокими уровнями достижений, особенно когда такая группа представляет школу с небольшим количеством учащихся. Поскольку вполне вероятно, что любая небольшая группа учащихся покажет значительный разброс в тестовых баллах из года в год относительно распределений межшкольных различий, то любые «причины», предложенные школами для объяснения своих превосходных результатов выполнения тестов, могут быть поддельными (Kane and Staiger, 2002).

Последующие экспериментальные, или квазиэкспериментальные, исследования

Эксперименты, которые включают взаимное рассмотрение и контроль измерения переменных, обеспечивают наиболее информативный базис для получения выводов относительно причинных связей. Их использование потребовало бы дополнительных исследований для проверки гипотезы о возможности совершенствования результатов учащихся на основе применения стратегий, выбранных благодаря результатам национальной оценки, в частности предоставления дополнительных ресурсов отдельным школам, ряду школ или отдельным предметным областям (*экспериментальная группа*). После завершения выбранного периода исследования эффект

осуществления стратегии должен быть оценен путем сравнения достижений экспериментальной группы с достижениями контрольной группы (т.е. группы школ или предметных областей, которые не получали дополнительных ресурсов, но были по другим признакам подобны школам или предметным областям в экспериментальной группе). Лучший способ гарантировать сопоставимость экспериментальной и контрольной групп состоит в том, чтобы отобрать их участников в случайном порядке. Однако из-за практических трудностей часто используется квазиэкспериментальный дизайн, который не предполагает случайного отбора. В этом случае группа сравнения выбирается таким образом, чтобы быть, по возможности, подобной экспериментальной группе (в частности, должны быть предприняты усилия для сокращения возможности появления систематических ошибок измерения). Информация по переменным, представляющим интерес в исследовании, собирается до и после эксперимента. Для установления постэкспериментальных норм может использоваться ковариационный или регрессионный анализ в тех случаях, когда начальные различия между характеристиками исследуемых объектов не слишком велика.

Несколько рандомизированных исследований были проведены в развивающихся странах (Индия, Кения, Никарагуа), хотя и не в контексте национальной оценки. В этих исследованиях эксперимент (например, выделение дополнительных учебных пособий и учебников или введение корректирующих программ) был осуществлен в случайным образом отобранных школах, и его эффект был оценен на основе сравнения достижений учащихся в экспериментальных школах с достижениями учащихся в подобных школах, которые не участвовали в эксперименте (Glewwe and Kremer, 2006).

По информации, которой обладает автор этой книги, в контексте национальной оценки эксперименты или квазиэксперименты не использовались. По крайней мере отчасти их отсутствие может быть объяснено трудностями, возникающими при проведении экспериментов и мониторинга их хода. Более серьезная проблема с точки зрения политиков или лиц,

принимающих решения в образовании, может быть связана со стоимостью и временем, которые необходимы для выполнения такого исследования. Однако можно смело утверждать, что затраты на эксперименты будут значительно меньше по сравнению с затратами на реализацию политики, которая плохо работает или, возможно, могла бы стать значительно более эффективной, если будет доступна информация, которая получена из эксперимента или квазиэксперимента (Postlethwaite, 1975). Следует учитывать, что планирование, выполнение и анализ, необходимые в экспериментальных или квазиэкспериментальных исследованиях, могут лежать за пределами способностей многих агентств, которые осуществляют национальную оценку. В этом случае такие исследования лучше всего предоставлять более специализированным исследовательским агентствам.

Оценка детей, не охваченных системой образования

Поскольку программы национальной оценки ориентированы на детей, посещающих учебные заведения, они не проливают свет на неохваченные количественные характеристики, необходимые для целостного анализа доступности системы образования в своих странах, за исключением очень ограниченных данных (например, предоставление информации о расстоянии от школы до дома учащегося). Однако во многих развивающихся странах большое количество детей школьного возраста не посещают школу или присутствуют в ней только в течение короткого промежутка времени. Учет обязательств этих стран по отношению к всеобщему начальному образованию, полученные данные о достижениях этих детей, не посещающих школы, сбор информации о препятствиях при получении доступа к образованию, об использовании альтернативных ресурсов (например, неофициального образования) и о нереализованных потребностях в получении образования должны представлять значимый интерес в исследованиях. Получение такой информации, конечно, потребовало бы применения специальных методов формирования

выборки и других процедур оценки, кроме тех, которые в настоящее время используются в программах национальной оценки, основанных на выборках школ. Например, такое исследование, в котором оцениваются дети от 6 до 16 лет (по чтению, математике и английскому языку), обучающиеся дома, проводится в Индии (<http://www.asercentre.org>). Другие модели, которые обеспечивают данные по детям не посещающим школы, существуют в специальных исследованиях (см., например, Greaney, Khandker and Alam, 1999) и в контексте домашних обзоров.

ВЫВОДЫ

Поскольку практика национальной оценки устойчиво внедряется в системы образования, разумно ожидать, что сфера использования ее результатов также будет расширяться. Есть несколько условий, которые будут способствовать нарастанию этих тенденций. С одной стороны, это развитие методических навыков и способности к осуществлению коммуникационных стратегий у лиц, ответственных за национальную оценку; повышение способности политиков и лиц, принимающих решения, к интерпретации результатов оценки и разработке стратегий для коррекции выявленных недостатков в системе образования, а также привлечение большого числа заинтересованных лиц, которые должны способствовать установлению более восприимчивой окружающей среды для использования результатов. С другой стороны, необходимо рассмотреть, удовлетворяет ли информация, которую в настоящее время предоставляет национальная оценка, потребности пользователей, должны ли быть изменены процедуры, нужны ли дополнительные исследования, чтобы обеспечить информационные основы для политиков и лиц, принимающих решения. Выполнение этих условий может потребовать дальнейшего финансирования и новых обязательств. Дополнительные расходы не будут большими, если рассмотреть их в контексте стоимости выполнения программ национальной оценки, однако они значительно повысят ценность ее результатов.



СПИСОК ЛИТЕРАТУРЫ

Acana, S. 2006. "Reporting Results of National Assessment: Uganda Experience." Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore, May 22–26.

———. 2008. "Using Assessment Results to Improve Learning: A Message to Headteachers." Paper presented at the 34th Annual Conference of the International Association for Educational Assessment, Cambridge, U.K., September 7–12.

Aguerrondo, I. 1992. "Educational Reform in Latin America: A Survey of our Decades." *Prospects* 22 (3): 353–65.

Ainley, J., J. Fraillon and C. Freeman. 2007. *National Assessment Program: ICT Literacy Years 6 and 10 Report, 2005*. Carlton South, Victoria, Australia: Ministerial Council on Education, Employment, Training and Youth Affairs. http://www.mceetya.edu.au/verve/_resources/NAP_ICTL_2005_Years_6_and_10_Report.pdf.

Altinok, N. 2008. "An International Perspective on Trends in the Quality of Learning Achievement (1965–2007)." Background paper prepared for *Overcoming Inequality: Why Governance Matters—EFA Global Monitoring Report 2009*, United Nations Educational, Scientific, and Cultural Organization, Paris. <http://unesdoc.unesco.org/images/0017/001780/178009e.pdf>.

Amrein, A.L., and D.C. Berliner. 2002. "High-Stakes Testing, Uncertainty, and Student Learning." *Education Policy Analysis Archives* 10 (18). <http://epaa.asu.edu/epaa/v10n18/>.

Anderson, P., and G. Morgan. 2008. *Developing Tests and Questionnaires for a National Assessment of Achievement*. Washington, DC: World Bank.

Anderson, S.E. 2002. "The Double Mirrors of School Improvement: The Aga Khan Foundation in East Africa." In *Improving Schools through Teacher Development: Case Studies of the Aga Khan Foundation Projects in East Africa*, ed. S.E. Anderson, 1–20. Lisse, Netherlands: Swets & Zetlinger.

Arregui, P., and C. McLauchlan. 2005. "Utilization of Large-Scale Assessment Results in Latin America." Report prepared for the Partnership for Educational Revitalization in the Americas and the World Bank Institute, Washington, DC.

Artelt, C., J. Baumert, N. Julius-McElvany, and J. Peschar. 2003. *Learners for Life: Student Approaches to Learning – Results from PISA 2000*. Paris: Organisation for Economic Co-operation and Development.

Banisar, D. 2006. *Freedom of Information around the World 2006: A Global Survey of Access to Government Records Laws*. London: Privacy International. http://www.freedominfo.org/documents/global_survey2006.pdf.

Baron, J.B. 1999. "Exploring High and Improving Achievement in Connecticut." Document 1233, National Educational Goals Panel, Washington, DC. <http://govinfo.library.unt.edu/negp/issues/publication/othpress/body.pdf>.

Báthory, Z. 1992. "Hungarian Experiences in International Student Achievement Surveys." *Prospects* 22 (4): 434–40.

Beaton, A.E. 1994. "Item Sampling in Testing." In *The International Encyclopedia of Education*, 2nd ed., ed. T. Husén and T. N. Postlethwaite, 3055–61. Oxford, U.K.: Pergamon.

Beaton, A.E., and N.L. Allen. 1992. "Interpreting Scales through Scale Anchoring." *Journal of Educational Statistics* 17 (2): 191–204.

Beaton, A. E., and E. G. Johnson. 1992. "Overview of the Scaling Methodology Used in the National Assessment." *Journal of Educational Measurement* 29 (2): 163–75.

Benveniste, L. 2000. "Student Assessment as a Political Construction: The Case of Uruguay." *Education Policy Analysis Archives* 8 (32). <http://epaa.asu.edu/epaa/v8n32.html>.

- . 2002. “The Political Structuration of Assessment: Negotiating State Power and Legitimacy.” *Comparative Education Review* 46 (1): 89–118.
- Bernard, J.M., and K. Michaelowa. 2006. “How Can Countries Use Cross-National Research Results to Address ‘the Big Policy Issues?’” In *Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact*, ed. K.N. Ross and I.J. Genevois, 229–40. Paris: International Institute for Educational Planning.
- Bethell, G., and R. Mihail. 2005. “Assessment and Examinations in Romania.” *Assessment in Education* 12 (1): 77–96.
- Bhutan Board of Examinations. 2004. *National Educational Assessment in Bhutan: A Benchmark of Student Achievement in Literacy at Class 5, 2003*. Thimphu: Ministry of Education.
- Binkley, M., and K. Rust, eds. 1994. *Reading Literacy in the United States: Technical Report of the U.S. Component of the IEA Reading Literacy Study*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- Blalock, A.B. 1999. “Evaluation Research and the Performance Management Movement.” *Evaluation* 5 (2): 117–49.
- Bonnet, G. 2007. “What Do Recent Evaluations Tell Us about the State of Teachers in Sub-Saharan Africa.” Background paper for *Education for All: Will We Make It – EFA Global Monitoring Report 2008*, United Nations Educational, Scientific, and Cultural Organization, Paris. <http://unesdoc.unesco.org/images/0015/001555/155511e.pdf>.
- Braun, H. 2004. “Reconsidering the Impact of High-Stakes Testing.” *Education Policy Analysis Archives* 12 (1). <http://epaa.asu.edu/epaa/v2n1/>.
- Braun, H., A. Kanjee, E. Bettinger, and M. Kremer. 2006. *Improving Education through Assessment, Innovation, and Evaluation*. Cambridge, MA: American Academy of Arts and Sciences.
- British Columbia Ministry of Education. 1999. *Interpreting Your District’s Assessment Results, 1999*. Victoria: Province of British Columbia.
- Brophy, J.E., and T.L. Good. 1986. “Teacher Behavior and Student Achievement.” In *Third Handbook of Research on Teaching*, ed. M. Wittrock, 328–75. New York: Macmillan.

- Campbell, J., D.L. Kelly, I.V.S. Mullis, M.O. Martin, and M. Sainsbury. 2001. *Framework and Specifications for PIRLS Assessment 2001*. 2nd ed. Chestnut Hill, MA: Boston College.
- Campbell, J., L. Kyriakides, D. Muijs, and W. Robinson. 2004. *Assessing Teacher Effectiveness: Developing a Differentiated Model*. London: Routledge Falmer.
- Carroll, D. 1996. "The Grade 3 and 5 Assessment in Egypt." In *National Assessments: Testing the System*, ed. P. Murphy, V. Greaney, M.E. Lockheed, and C. Rojas, 157–65. Washington, DC: World Bank.
- Chabbott, C., and E. J. Elliott, eds. 2003. *Understanding Others, Educating Ourselves: Getting More from International Comparative Studies in Education*. Washington, DC: National Academies Press.
- Chapman, D.W., and C.W. Snyder. 2000. "Can High Stakes National Testing Improve Instruction? Reexamining Conventional Wisdom." *International Journal of Educational Development* 20 (6): 457–74.
- Cheng, K., and H. Yip. 2006. *Facing the Knowledge Society: Reforming Secondary Education in Hong Kong and Shanghai*. Washington, DC: World Bank.
- Clegg, S.R., and T.F. Clarke. 2001. "Intelligence: Organizational." In *International Encyclopedia of the Social and Behavioral Sciences*, vol. 11, ed. N.J. Smelser and P.B. Baltes, 7665–70. Amsterdam: Elsevier.
- Clotfelter, C.T., and H.L. Ladd. 1996. "Recognizing and Rewarding Success in Public Schools." In *Holding Schools Accountable: Performance-Based Reform in Education*, ed. H.F. Ladd, 23–63. Washington, DC: Brookings Institution.
- DataAngel Policy Research. 2007. *A Tool for Understanding Performance in Science Instruction in Qatar*. Doha: Supreme Education Council.
- Davies, I.C. 1999. "Evaluation and Performance Management in Government." *Evaluation* 5 (2): 150–59.
- de Vise, D. 2005. "State Gains Not Echoed in Federal Testing: Results Fuel Criticism of Md., Va. Education." *Washington Post*, October 24, B01.
- Duthilleul, Y., and R. Allen. 2005. "Which Teachers Make a Difference? Implications for Policy Makers in SACMEQ Countries." Paper presented at the Educational Policy Research Conference, International Institute for Educational Planning, Paris, September 28–30.

Eivers, E., G. Shiel, R. Perkins, and J. Cosgrove. 2005a. *Succeeding in Reading? Reading Standards in Irish Primary Schools*. Dublin: Department of Education and Science.

———. 2005b. *The 2004 National Assessment of English Reading*. Dublin: Educational Research Centre.

Elley, W.B. 1992. *How in the World Do Students Read? IEA Study of Reading Literacy*. The Hague, Netherlands: International Association for the Evaluation of Educational Achievement.

Elmore, R., and D. Burney. 1999. "Investing in Teacher Training." In *Teaching as the Learning Profession*, ed. L. Darling-Hammond and G. Sykes, 236–91. San Francisco, CA: Jossey-Bass.

Ethiopia Education Quality Assurance and Examinations Agency. 2007. *Highlights of the Findings: Ethiopian Third National Learning Assessment*. Addis Ababa: Ethiopia Education Quality Assurance and Examinations Agency.

Ferrer, G. 2006. *Educational Assessment Systems in Latin America: Current Practice and Future Challenges*. Washington, DC: Partnership for Educational Revitalization in the Americas.

Ferrer, G., and P. Arregui. 2003. "Las Pruebas Internacionales de Aprendizaje en America Latina y Su Impacto en la Calidad de la Educación: Criterio para Guiar Futuras Aplicaciones." Documento de Trabajo 26, Partnership for Educational Revitalization in the Americas, Santiago.

Finn, J.D., and C.M. Achilles. 1990. "Tennessee's Class Size Study: Findings, Implications, Misconceptions." *Educational Evaluation and Policy Analysis* 21 (2): 97–110.

Forster, M. 2001. *A Policy Maker's Guide to Systemwide Assessment Programs*. Camberwell, Victoria: ACER Press.

Frederiksen, J., and A. Collins. 1989. "A Systems Approach to Educational Testing." *Educational Researcher* 18 (9): 27–32.

Fullan, M. 2001. *The New Meaning of Educational Change*. New York: Teachers College Press.

Fuller, B. 1987. "What School Factors Raise Achievement in the Third World?" *Review of Educational Research* 57 (3): 255–92.

Garet, M. S., A. C. Porter, L. Desimone, B. F. Birman, and K. S. Yoon. 2001. "What Makes Professional Development Effective? Results from a

National Sample of Teachers.” *American Educational Research Journal* 38 (4): 915–45.

Gebrekidan, Z. 2006. “Ethiopian Second National Learning Assessment.” Paper presented at the National Assessment Capacity Building Workshop, Using National Assessment Results, Kampala, January 30 – February 2.

Georgia Department of Education. n.d. “Standards, Instruction, and Assessment.” Georgia Department of Education, Atlanta. http://www.doe.k12.ga.us/ci_testing.aspx?PageReq=C1_TESTING_NAEP.

Gilmore, A. 2005. “The Impact of PIRLS (2001) and TIMSS (2003) in Low and Middle-Income Countries: An Evaluation of the Value of World Bank Support for International Surveys of Reading Literacy (PIRLS) and Mathematics and Science (TIMSS).” International Association for the Evaluation of Educational Achievement, Amsterdam. http://www.iea.nl/fi_leadadmin/user_upload/docs/WB_report.pdf.

Glewwe, P., and M. Kremer. 2006. “Schools, Teachers, and Education Outcomes in Developing Countries.” In *Handbook of the Economics of Education*, vol. 2, ed. E.A. Hanushek and F. Welch, 945–1017. Amsterdam: Elsevier.

Goldstein, H. 1983. “Measuring Changes in Educational Attainment over Time: Problems and Possibilities.” *Journal of Educational Measurement* 20 (4): 369–77.

González, E.J. 2002. *Evaluation Systems in Latin America*. Brasilia: National Institute for Educational Studies and Research, Ministry of Education.

González, P., A. Mizala, and P. Romaguera. 2002. “Recursos Diferenciados a la Educación Subvencionada en Chile.” Serie Economía 150, Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago.

Grant, S., P. Peterson, and A. Shojgreen-Downer. 1996. “Learning to Teach Mathematics in the Context of System Reform.” *American Educational Research Journal* 33 (2): 500–43.

Greaney, V., and T. Kellaghan. 1996. *Monitoring the Learning Outcomes of Education Systems*. Washington, DC: World Bank.

———. 2008. *Assessing National Achievement Levels in Education*. Washington, DC: World Bank.

- Greaney, V., S. R. Khandker, and M. Alam. 1999. *Bangladesh: Assessing Basic Learning Skills*. Dhaka: University Press. Griffith, J.E., and E.A. Medrich. 1992. "What Does the United States Want to Learn from International Comparative Studies in Education?" *Prospects* 22 (4): 476–84.
- Gucwa, B., and M. Mastie. 1989. "Pencils Down!" Michigan State Board of Education, Lansing. <http://www.ncrel.org/sdrs/areas/issues/methods/assment/as6penc2.htm>.
- Guilfoyle, C. 2006. "NCLB: Is There Life Beyond Testing?" *Educational Leadership* 64 (3): 8–13.
- Gvirtz, S., and S. Larripa. 2004. "National Evaluation System in Argentina: Problematic Present and Uncertain Future." *Assessment in Education* 11 (3): 349–64.
- Haertel, E.H., and J.L. Herman. 2005. "A Historical Perspective on Validity Arguments for Accountability Testing." In *Uses and Misuses of Data for Educational Accountability and Improvement: The 104th Yearbook of the National Society for the Study of Education, Part 2*, ed. J. L. Herman and E.H. Haertel, 1–34. Malden, MA: Blackwell.
- Hambleton, R.K., and K. Meara. 2000. "Newspaper Coverage of NAEP Results, 1990 to 1998." In *Student Performance Standards on the National Assessment of Educational Progress: Affirmations and Improvements*, ed. M.L. Bourque and S. Byrd, 131–55. Washington, DC: National Assessment Governing Board.
- Hambleton, R.K., and S.C. Slater. 1997. "Are NAEP Executive Summary Reports Understandable to Policy Makers and Stakeholders?" CSE Technical Report 430, Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
- Himmel, E. 1996. "National Assessment in Chile." In *National Assessments: Testing the System*, ed. P. Murphy, V. Greaney, M. E. Lockheed, and C. Rojas, 111–28. Washington, DC: World Bank.
- Hopmann, S.T., and G. Brinek. 2007. "Introduction: PISA According to PISA – Does PISA Keep What It Promises?" In *PISA Zufolge PISA: PISA According to PISA*, ed. S.T. Hopmann, G. Brinek, and M. Retzl, 9–19. Vienna: LIT.

Howie, S. 2002. “English Proficiencies and Contextual Factors Influencing Mathematics Achievements of Secondary School Pupils in South Africa.” PhD thesis, University of Twente, Enschede, Netherlands.

Howie, S., and T. Plomp. 2005. “International Comparative Studies of Education and Large-Scale Change.” In *International Handbook of Education Policy*, ed. N. Bascia, A. Cumming, A. Datnow, K. Leithword, and D. Livingstone, 75–99. Dordrecht, Netherlands: Springer.

Husén, T. 1984. “Issues and Their Background.” In *Educational Research and Policy: How Do They Relate?* ed. T. Husén and M. Kogan, 1–36. Oxford, U.K.: Pergamon.

———. 1987. “Policy Impact of IEA Research.” *Comparative Education Review* 31 (1): 29–46.

Kane, T.J., and D.O. Staiger. 2002. “Volatility in School Test Scores: Implications for Test-Based Accountability Systems.” In *Brookings Papers on Education Policy*, 2002, ed. D. Ravitch, 235–83. Washington, DC: Brookings Institution Press.

Kellaghan, T., and V. Greaney. 1992. *Using Examinations to Improve Education: A Study in Fourteen African Countries*. Washington, DC: World Bank.

———. 2001. *Using Assessment to Improve the Quality of Education*. Paris: International Institute for Educational Planning.

———. 2004. *Assessing Student Learning in Africa*. Washington, DC: World Bank.

Kellaghan, T., and G.F. Madaus. 2000. “Outcome Evaluation.” In *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, 2nd ed., ed. D.L. Stufflebeam, G.F. Madaus, and T. Kellaghan, 97–112. Boston: Kluwer Academic. Kellaghan, T., K. Sloane, B. Alvarez, and B. S. Bloom. 1993. *The Home Environment and School Learning*. San Francisco, CA: Jossey-Bass.

Kirsch, I., J. De Jong, D. Lafontaine, J. McQueen, J. Mendelovits, and C. Monseur. 2002. *Reading for Change: Performance and Engagement across Countries – Results from PISA 2000*. Paris: Organisation for Economic Co-operation and Development.

Kulm, G., J. Roseman, and M. Treistman. 1999. “A Bench-Based Approach to Textbook Evaluation.” *Science Books and Films* 35 (4): 147–53.

Kulpoo, D., and P. Coustère. 1999. "Developing National Capacities for Assessment and Monitoring through Effective Partnerships." In *Partnerships for Capacity Building and Quality Improvements in Education: Papers from the ADEA 1997 Biennial Meeting, Dakar*. Paris: Association for the Development of Education in Africa.

Kupermintz, H., M.M. Ennis, L.S. Hamilton, J.E. Talbert, and R.E. Snow. 1995. "Enhancing the Validity and Usefulness of Large-Scale Educational Assessments: I. NELS:88 Mathematics Achievement." *American Educational Research Journal* 32 (3): 525–54.

Kuwait Ministry of Education. 2008. *PIRLS 2006: Kuwait Technical Report*. Kuwait City: Kuwait Ministry of Education.

Leimu, K. 1992. "Interests and Modes in Research Utilisation: The Finnish IEA Experience." *Prospects* 22 (4): 425–33.

Linn, R.L. 2000. "Assessments and Accountability." *Educational Researcher* 29 (2): 4–16.

———. 2005a. "Conflicting Demands of No Child Left Behind and State Systems: Mixed Messages about School Performance." *Education Policy Analysis Archives* 13 (33). <http://epaa.asu.edu/epaa/v13n33/>.

———. 2005b. "Issues in the Design of Accountability Systems." In *Uses and Misuses of Data for Educational Accountability and Improvement: The 104th Yearbook of the National Society for the Study of Education, Part 2*, ed. J. L. Herman and E. H. Haertel, 78–98. Malden, MA: Blackwell.

Linn, R.L., and E. Baker. 1996. "Can Performance-Based Student Assessments Be Psychometrically Sound?" In *Performance-Based Student Assessment: Challenges and Possibilities: 95th Yearbook of the National Society for the Study of Education, Part 1*, ed. J.N. Baron and D.P. Wolf, 84–103. Chicago: National Society for the Study of Education.

Lockheed, M.E., and A.M. Verspoor. 1991. *Improving Primary Education in Developing Countries*. Oxford, U.K.: Oxford University Press.

Lovett, S. 1999. "National Education Monitoring Project: Teachers Involvement and Development—Professional Development from NEMP." Paper presented at New Zealand Association for Research in Education and Australian Association for Research in Education Conference, Melbourne, Australia, November 29 – December 2.

Luna, E. 1992. "Dominican Republic: The Study on Teaching and Learning in Mathematics." *Prospects* 22 (4): 448–54.

Madamombe, R.T. 1995. "A Comment on the Analysis of Educational Research Data for Policy Development: An Example from Zimbabwe." *International Journal of Educational Research* 23 (4): 397–402.

Madaus, G.F., and T. Kellaghan. 1992. "Curriculum Evaluation and Assessment." In *Handbook of Research on Curriculum*, ed. P.W. Jackson, 119–54. New York: Macmillan.

Madaus, G.F., M. Russell, and J. Higgins. 2009. *The Paradoxes of High Stakes Testing: How They Affect Students, Their Parents, Teachers, Principals, Schools, and Society*. Charlotte, NC: Information Age Publishing.

Martin, M.O., I. V.S. Mullis, and S.J. Chrostowski. 2004. *The Trends in International Mathematics and Science Study 2003: Technical Report*. Chestnut Hill, MA: International Study Center, Boston College.

Mauritius Examinations Syndicate. 2003. *Monitoring Learning Achievement: Joint UNESCO/UNICEF Project – A Survey of 9-Year-Old Children in the Republic of Mauritius*. Reduit, Mauritius: Mauritius Examinations Syndicate.

McDonnell, L. M. 2005. "Assessment and Accountability from the Policy Maker's Perspective." In *Uses and Misuses of Data for Educational Accountability and Improvement: 104th Yearbook of the National Society for the Study of Education, Part 2*, ed. J.L. Herman and E.H. Haertel, 35–54. McQuillan, J. 1998. "Seven Myths about Literacy in the U.S." *Practical Assessment, Research, and Evaluation* 6 (1). <http://pareonline.net/getvn.asp?v=6&n=1>.

Meckes, L., and R. Carrasco. 2006. "SIMCE: Lessons from the Chilean Experience in National Assessment Systems of Learning Outcomes." Paper presented at the Conference on Lessons from Best Practices in Promoting Education for All: Latin America and the Caribbean, hosted by the World Bank and the Inter-American Development Bank, Cartagena de Indias, Colombia, October 9–11.

Messick, S. 1989. "Validity." In *Educational Measurement*, 3rd ed., ed. R. Linn, 13–103. New York: American Council on Education and Macmillan.

Michaelowa, K. 2001. "Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations." *World Development* 29 (10): 1699–716.

- Michaelowa, K., and A. Wechtler. 2006. *The Cost-Effectiveness of Inputs in Primary Education: Insights from the Literature and Recent Student Surveys for Sub-Saharan Africa*. Hamburg: Institute of International Economics.
- Monare, M. 2006. "SA Pupils Rank Last in Maths, Science Study." *Star* (Johannesburg), December 6.
- Moncada, G., R. Hernández Rodríguez, M. C. Aguilar, D. Orellana, M. Alas Solís, and B. Hernández. 2003. *Uso e Impacto de la Información Empírica en la Formulación y Ejecución de Política de Educación Básica en Honduras en el Período 1990–2002*. Tegucigalpa: Dirección de Investigación, Universidad Pedagógica Nacional Francisco Morazán.
- Mullis, I. V. S., A. M. Kennedy, M. O. Martin, and M. Sainsbury. 2006. *PIRLS 2006: Assessment Framework and Specifications*. Chestnut Hill, MA: International Study Center, Boston College.
- Mullis, I.V.S., M.O. Martin, E.J. González, and S.J. Chrostowski. 2004. *Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: International Study Center, Boston College.
- Murimba, M. 2005. "The Impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)." *Prospects* 35 (1): 91–108.
- Murphy, J., J. Yff, and N. Shipman. 2000. "Implementation of the Interstate School Leaders Licensure Consortium Standards." *International Journal of Leadership in Education* 3 (1): 17–39.
- Nassor, S., and K. A. Mohammed. 1998. *The Quality of Education: Some Policy Suggestions Based on a Survey of Schools – Zanzibar*. Paris: International Institute for Educational Planning.
- Nepal Educational and Developmental Service Centre. 1999. *National Assessment of Grade 5 Students*. Kathmandu: Nepal Educational and Developmental Service Centre.
- Nigeria Federal Ministry of Education. 2000. *Education for All: The Year 2000 Assessment*. Abuja: Nigeria Federal Ministry of Education.
- Nzomo, J., M. Kariuki, and L. Guantai. 2001. "The Quality of Education: Some Policy Suggestions Based on a Survey of Schools." Southern and Eastern Africa Consortium for Monitoring Educational Quality Policy Research Paper 6, International Institute for Educational Planning, Paris.

Nzomo, J., and D. Makuwa. 2006. "How Can Countries Move from Cross-National Research Results to Dissemination and Then to Policy Reform? *Case Studies from Kenya and Namibia.*" In *Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact*, ed. K. N. Ross and I. J. Genevois, 213–28. Paris: International Institute for Educational Planning.

OECD (Organisation for Economic Co-operation and Development). 2004. Chile: *Reviews of National Policies for Education*. Paris: OECD.

OECD (Organisation for Economic Co-operation and Development) and INES (International Indicators of Education Systems) Project, Network A. 2004. "Attention Getting Results." *Review of Assessment Activities* 16 (February–March): 2.

Ogle, L.T., A. Sen, E. Pahlke, D. Jocelyn, D. Kastberg, S. Roey, and T. Williams. 2003. "International Comparisons in Fourth Grade Reading Literacy: Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001." U.S. Department of Education, National Center for Education Statistics, Washington, DC. <http://nces.ed.gov/pubs2003/2003073.pdf>.

Olivares, J. 1996. "Sistema de Medición de la Calidad de la Educación de Chile: SIMCE – Algunos Problemas de la Medicion." *Revista Iberoamericana de Educación* 10: 117–96. <http://www.rieoei.org/oeivirt/rie10a07.htm>.

Pérez, B.A. 2006. "Success in Implementing Education Policy Dialogue in Peru." United States Agency for International Development, Washington, DC.

Perie, M., W. Grigg, and G. Dion. 2005. "The Nation's Report Card: Mathematics 2005." National Center for Education Statistics, U.S. Department of Education, Washington, DC. <http://nces.ed.gov/nationsreportcard/pdf/main2005/2006453.pdf>.

Porter, A., and A. Gamoran. 2002. "Progress and Challenges for Large-Scale Studies." In *Methodological Advances in Cross-National Surveys of Educational Achievement*, ed. A.C. Porter and A. Gamoran, 3–23. Washington, DC: National Academies Press.

Postlethwaite, T. N. 1975. "The Surveys of the International Association for the Evaluation of Educational Achievement (IEA)." In *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement*, ed. A.C. Purves and D.U. Levine, 1–32. Berkeley, CA: McCutchan.

- . 1987. “Comparative Educational Achievement Research: Can It Be Improved?” *Comparative Education Review* 31 (1): 150–58.
- . 1995. “Calculation and Interpretation of Between-School and Within-School Variation in Achievement (ρ).” In *Measuring What Students Learn*, 83–94. Paris: Organisation for Economic Co-operation and Development.
- . 2004a. *Monitoring Educational Achievements*. Paris: International Institute for Educational Planning.
- . 2004b. “What Do International Assessment Studies Tell Us about the Quality of School Systems?” Background paper for *The Quality Imperative: EFA Global Monitoring Report 2005*. United Nations Educational, Scientific, and Cultural Organization, Paris. <http://unesdoc.unesco.org/images/0014/001466/146692e.pdf>.
- Postlethwaite, T. N., and T. Kellaghan. 2008. *National Assessments of Educational Achievement*. Paris: International Institute for Educational Planning; Brussels: International Academy of Education.
- Powdyel, T.S. 2005. “The Bhutanese Education Assessment Experience: Some Reflections.” *Prospects* 35 (1): 45–57.
- Pravalpruk, K. 1996. “National Assessment in Thailand.” In *National Assessments: Testing the System*, ed. P. Murphy, V. Greaney, M. E. Lockheed, and C. Rojas, 137–45. Washington, DC: World Bank.
- Ravela, P. 2002. “Cómo Presentan Sus Resultados los Sistemas Nacionales de Evaluación Educativa en América Latina?” *Partnership for Educational Revitalization in the Americas*, Washington, DC.
- . 2005. “A Formative Approach to National Assessments: The Case of Uruguay.” *Prospects* 35 (1): 21–43.
- . 2006. “Using National Assessments to Improve Teaching ... and Learning: The Experience of UMRE in Uruguay.” Paper presented at the Conference on Lessons from Best Practices in Promoting Education for All: Latin America and the Caribbean, hosted by the World Bank and the Inter-American Development Bank, Cartagena de Indias, Colombia, October 9–11.
- Reezigt, G.J., and B.P.M. Creemers. 2005. “A Comprehensive Framework for Effective School Improvement.” *School Effectiveness and School Improvement* 16 (4): 407–24.

- Reimers, F. 2003. "The Social Context of Educational Evaluation in Latin America." In *International Handbook of Educational Evaluation*, ed. T. Kellaghan and D.L. Stufflebeam, 441–63. Boston: Kluwer Academic.
- Reynolds, D. 2000. "School Effectiveness: The International Dimension." In *The International Handbook of School Effectiveness Research*, ed. C. Teddlie and D. Reynolds, 232–56. London: Falmer.
- Reynolds, D., and C. Teddlie. 2000. "The Processes of School Effectiveness." In *The International Handbook of School Effectiveness Research*, ed. C. Teddlie and D. Reynolds, 134–59. London: Falmer.
- Riley, R.W. 2000. Remarks prepared for a TIMSS-R press conference, Washington, DC, December 5. <http://www.ed.gov/Speeches/12-2000/120500.html>.
- Robertson, I. 2005. "Issues Relating to Curriculum, Policy, and Gender Raised by National and International Surveys of Achievement in Mathematics." *Assessment in Education* 12 (3): 217–36.
- Robitaille, D.F., A.E. Beaton, and T. Plomp, eds. 2000. *The Impact of TIMSS on the Teaching and Learning of Mathematics and Science*. Vancouver, BC: Pacific Educational Press.
- Rojas, C.C., and J.M. Esquivel. 1998. "Los Sistemas de Medición del Logro Académico en Latinoamérica." Education Paper 25, World Bank, Washington, DC.
- Rubner, J. 2006. "How Can a Country Manage the Impact of 'Poor' Cross-National Research Results? A Case Study from Germany." In *Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact*, ed. K.N. Ross and I.J. Genevois, 255–64. Paris: International Institute for Educational Planning.
- Rust, V.L. 1999. "Education Policy Studies and Comparative Education." In *Learning from Comparing: New Directions in Comparative Education Research – Policy, Professionals, and Development*, vol. 2, ed. R. Alexander, P. Broadfoot, and D. Phillips, 13–39. Oxford, U.K.: Symposium Books.
- SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality). 2007. SACMEQ Web site. <http://www.sacmeq.org/>.
- Scheerens, J. 1998. "The School Effectiveness Knowledge Base as a Guide to School Improvement." In *International Handbook of Educational Change*,

- ed. A. Hargreaves, A. Lieberman, M. Fullan, and D. Hopkins, 1096–115. Boston: Kluwer Academic.
- Schiefelbein, E., and P. Schiefelbein. 2000. “Three Decentralization Strategies in Two Decades: Chile 1982–2000.” *Journal of Educational Administration* 38 (5): 412–25.
- Schubert, J. 2005. “The Reality of Quality Improvement: Moving toward Clarity.” In *The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa*, ed. A.M. Verspoor, 53–68. Paris: Association for the Development of Education in Africa.
- Shiel, G., R. Perkins, S. Close, and E. Oldham. 2007. *PISA Mathematics: A Teacher’s Guide*. Dublin: Department of Education and Science.
- Singh, A., V.K. Jain, S.K.S. Guatam, and S. Jumjar. n.d. *Learning Achievement of Students at the End of Class V*. New Delhi: National Council for Educational Research and Training.
- Snyder, C.W., B. Prince, G. Lohanson, C. Odaet, L. Jaji, and M. Beatty. 1997. *Exam Fervor and Fever: Case Studies of the Influence of Primary Leaving Examinations on Uganda Classrooms, Teachers, and Pupils*. Washington, DC: Academy for Educational Development.
- Sri Lanka National Education Research and Evaluation Centre. 2004. *Achievement after Four Years of Schooling*. Colombo: National Education Research and Evaluation Centre.
- Stack, M. 2006. “Testing, Testing, Read All about It: Canadian Press Coverage of the PISA Results.” *Canadian Journal of Education* 29 (1): 49–69.
- Stoneberg, B. 2007. “Using NAEP to Confirm State Test Results in the No Child Left Behind Act.” *Practical Assessment Research and Evaluation* 12 (5): 1–10.
- Surgenor, P., G. Shiel, S. Close, and D. Millar. 2006. *Counting on Success: Mathematics Achievement in Irish Primary Schools*. Dublin: Department of Education and Science.
- Teddlie, C., and D. Reynolds. 2000. “School Effectiveness Research and the Social and Behavioural Sciences.” In *The International Handbook of School Effectiveness Research*, ed. C. Teddlie and D. Reynolds, 301–21. London: Falmer.

Uganda National Examinations Board. n.d. National Assessment of Progress in Education. Poster, Uganda National Examinations Board, Kampala.

UNESCO (United Nations Educational, Scientific, and Cultural Organization). 1990. *World Declaration on Education for All: Meeting Basic Learning Needs*. Adopted by the World Congress on Education for All. New York: UNESCO.

———. 2000. *The Dakar Framework for Action: Education for All – Meeting Our Collective Commitments*. Paris: UNESCO.

UNICEF (United Nations Children’s Fund). 2000. “Defining Quality of Education.” Paper presented at a meeting of the International Working Group on Education, Florence, Italy, June.

U.S. National Center for Education Statistics. 2005. *National Assessment of Educational Progress (NAEP), Selected Years, 1971–2004: Long-Term Trend Reading Assessments*. Washington, DC: U.S. National Center for Education Statistics.

———. 2006a. “The NAEP Mathematics Achievement Levels by Grade.” U.S. National Center for Education Statistics, Washington, DC. <http://nces.ed.gov/nationsreportcard/mathematics/achieveall.asp>.

———. 2006b. *National Indian Education Study, Part 1: The Performance of American Indian and Alaska Native Fourth- and Eighth-Grade Students on NAEP 2005 Reading and Mathematics Assessment – Statistical Analysis Report*. Washington, DC: U.S. National Center for Education Statistics. <http://www.xmission.com/~amauta/pdf/ienationarep.pdf>.

———. 2007. *The Nation’s Report Card: Reading–2007 State Snapshot Report*. Washington, DC: U.S. National Center for Education Statistics. <http://nces.ed.gov/nationsreportcard/pubs/stt2007/200774978.asp>.

———. 2008. “The Nation’s Report Card.” U.S. National Center for Education Statistics, Washington, DC. <http://nces.ed.gov/nationsreportcard/sitemap.asp>.

Vanneman, A. 1996. “Geography: What Do Students Know and What Can They Do?” NAEPfacts 2 (2). <http://nces.ed.gov/pubs97/web/97579.asp>.

Vegas, E., and J. Petrow. 2008. *Raising Student Learning in Latin America: The Challenge for the 21st Century*. Washington, DC: World Bank.

- Wang, M.C., G.D. Haertel, and H.J. Walberg. 1993. "Toward a Knowledge Base for School Learning." *Review of Educational Research* 63 (3): 249–94.
- Watson, K. 1999. "Comparative Educational Research: The Need for Reconceptualisation and Fresh Insights." *Compare* 29 (3): 233–48.
- Weiss, C. H. 1979. "The Many Meanings of Research Utilization." *Public Administration Review* 39 (5): 426–31.
- Wenglinsky, H. 2002. "How Schools Matter: The Link between Teacher Classroom Practices and Student Academic Performance." *Education Policy Analysis Archives* 10 (12). <http://epaa.asu.edu/epaa/v10n12/>.
- Willms, D. 2006. "Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems." UIS Working Paper 5, UNESCO Institute for Statistics, Montreal.
- Wolff, L. 1998. *Educational Assessments in Latin America: Current Progress and Future Challenges*. Washington, DC: Partnership for Educational Revitalization in the Americas.
- World Bank. 2004. *Vietnam Reading and Mathematics Assessment Study*. 3 vols. Washington, DC: World Bank.
- Zhang, Y. 2006. "Urban-Rural Literacy Gaps in Sub-Saharan Africa: The Roles of Socioeconomic Status and School Quality." *Comparative Education Review* 50 (4): 581–602.

Научное издание

Национальная оценка учебных достижений

Келлаган Томас
Грини Винсент
Мюррей Т. Скотт

**ИСПОЛЬЗОВАНИЕ РЕЗУЛЬТАТОВ
НАЦИОНАЛЬНОЙ ОЦЕНКИ
УЧЕБНЫХ ДОСТИЖЕНИЙ**

Ответственный редактор *Ю.П. Голобокова*
Редактор *Е.В. Комарова*
Корректор *Е.Б. Гракина*
Компьютерная верстка *А.Б. Дунаевой*
Оформление *Н.А. Андреева*

Подписано в печать 21.11.2013. Формат 70х100/16,
Печать офсетная. Бумага офсетная. 15,5 печ. л.
Тираж 3000 экз. Заказ

Издательская группа «Логос»
111024, г. Москва, ул. Авиамоторная, д. 55, корп. 31, офис 305
Тел. (495) 981-51-12; 955-78-30
Электронная почта: universitas@mail.ru
Дополнительная информация на сайте:
<http://www.logosbook.ru>