

Виды и назначения программ оценки результатов обучения школьников¹

Болотов В.А.,
вице-президент Российской академии образования, д.п.н.

Вальдман И.А.,
зав. лабораторией мониторинга в образовании Института управления образованием РАО,
к.п.н.

Введение

Построение эффективной системы оценки качества образования во многих странах мира является одним из приоритетов национальной образовательной политики. Без использования оценочных процедур невозможно узнать, получают ли необходимые знания и навыки конкретные учащиеся, достигают ли своих целей образовательные реформы, вносит ли система образования вклад в развитие экономики и общества.

Система оценки качества образования призвана обеспечивать обучающихся и их родителей, педагогические коллективы школ и преподавателей учреждений профессионального образования, органы управления образованием всех уровней, институты гражданского общества, работодателей надёжной информацией о состоянии и развитии системы образования на разных уровнях ([4]).

Эта важная задача нашла отражение в государственной программе Российской Федерации "Развитие образования" на 2013-2020 годы. Целью её подпрограммы *«Развитие системы оценки качества образования и информационной прозрачности системы образования»* является «обеспечение надёжной и актуальной информацией процессов принятия решений руководителей и работников системы образования, а также потребителей образовательных услуг для достижения высокого качества образования через формирование общероссийской системы оценки качества образования» ([8]).

Выбор тех или иных оценочных инструментов и процедур для их применения на практике зависит от того, кто и как планирует использовать результаты оценки, для достижения каких целей и для решения каких задач.

Данные о результатах оценки достижений учащихся могут быть использованы для

¹ Данная статья подготовлена на основе учебных материалов Российского тренингового центра ИУО РАО, разработанных авторами при проведении курсов «Проектирование национальных и территориальных систем оценки качества образования: организационные, технологические и содержательные аспекты» (27-29 июня 2011 г.), «Ключевые аспекты построения эффективной системы оценки качества образования и использования результатов оценки учебных достижений школьников» (27-30 марта 2012 г.), «Система оценки качества и образовательная политика: ключевые проблемы и направления развития» (27-30 июня 2012 г.). www.rtc-edu.ru

разработки и реализации самого широкого спектра действий на разных уровнях системы образования – от совершенствования процесса обучения в конкретном образовательном учреждении до проведения масштабных изменений в системе образования страны (введение и совершенствование государственных образовательных стандартов, обновление учебных программ и учебной литературы, изменение программ подготовки и повышения квалификации учителей, разработка мер помощи школам, демонстрирующим низкие результаты, и т.п.). При этом необходимо отметить, что *для достижения разных целей необходимо применять разные оценочные инструменты*. Неправомерно применять один и тот же инструмент для оценки разных результатов, формируемых внутри системы образования.

Последствия игнорирования этого правила хорошо известны многим российским специалистам. Не секрет, что результаты Единого государственного экзамена очень часто используются не только по своему прямому назначению – оценить освоение образовательной программы по предмету выпускником школы и отобрать учащихся для продолжения обучения в вузе, - но и для решения совершенно несвойственных ЕГЭ задач - оценки качества работы педагогов, школ, органов управления образованием. Для преодоления этой проблемы необходимо вводить другие оценочные процедуры (этот вопрос подробно обсуждается в работах [2, 7, 9, 10]). Мы имеем в виду, прежде всего, мониторинги учебных достижений, позволяющие определять образовательные достижения групп детей и результаты работы школ, образовательных систем. ЕГЭ при этом должен быть одним из элементов комплексной общероссийской системы оценки качества образования (ОСОКО).

Задача построения комплексной ОСОКО должна быть решена в ходе реализации государственной программы «Развитие образования». Именно поэтому одним из приоритетов государственной политики в сфере оценки качества образования является «формирование современной и сбалансированной общероссийской системы оценки качества образования, включающей в себя не только национальные экзамены, но, прежде всего, мониторинговые обследования обучения и социализации, процедуры оценки результатов обучения на уровне школы» ([8]);

В рамках данной статьи будет представлена одна из возможных типологий программ оценки учебных достижений школьников и дана их характеристика. Кроме того, с учётом международного и российского опыта в статье будут обсуждаться актуальные научные и практические вопросы, связанные с планированием и проведением программы оценки учебных достижений школьников.

Цели оценивания

Каждая программа оценки учебных достижений школьников, будь то государственный экзамен или международное сравнительное исследование, имеет своё предназначение и зависит от целей, на достижение которых она направлена. Сами цели программы оценки определяются возможными вариантами использования её результатов. Вопрос о том, каким образом могут быть использованы результаты различных программ оценки для повышения качества деятельности образовательной системы и улучшения результатов обучения школьников, является ключевым при формировании набора оценочных процедур в стране.

Способы использования данных программ оценки весьма разнообразны. Они могут быть разделены на две категории: использование для принятий решений о поддержке конкретных учащихся и использование для принятий управленческих решений на разных уровнях системы образования – от конкретной школы до министерства образования страны ([4]).

В первом случае речь идёт о способах использования, которые связаны с корректировкой образовательного процесса (данные оценки используются для диагностики проблем в обучении конкретных школьников и определения мер по их преодолению, а также для планирования индивидуальных образовательных траекторий) и определением уровня освоения образовательной программы учеником.

Во втором случае речь идёт о способах, которые не имеют непосредственного отношения к решениям относительно конкретного учащегося, но имеют влияние на решения относительно образовательных учреждений и их систем. Они могут быть классифицированы в зависимости от объекта оценки следующим образом.

При оценке деятельности образовательного учреждения данные используются по двум основным направлениям: 1) внешняя оценка (например, аккредитация школы и проведение различных контрольных и надзорных мероприятий); 2) проектирование шага развития школы – каковы ближайшие задачи работы образовательного учреждения и каким образом надо их реализовывать.

При оценке образовательных систем данные также могут использоваться в двух аспектах - для осуществления процедур контроля за работой системы образования (муниципальной или региональной) и для разработки образовательной политики и программ действий на разных уровнях образования.

Обобщая выше сказанное, выделим следующие шесть целей, для которых могут использоваться процедуры оценки результатов обучения школьников. Первые три цели ориентированы на ученика, а остальные - на образовательные учреждения и системы.

Цель 1. Определение прогресса в обучении конкретного школьника.

Цель 2. Сертификация - подтверждение завершения ступени обучения.

Цель 3. Селекция – отбор для продолжения обучения или получения рабочего места.

Цель 4. Оценка деятельности образовательных учреждений.

Цель 5. Оценка деятельности образовательных систем.

Цель 6. Оценка деятельности национальной образовательной системы.

Есть и другие, достаточно специфические цели использования результатов оценки учебных достижений (для оценки работы педагога, для оценки образовательных программ, для оценки учебной литературы), но они не являются предметом рассмотрения данной статьи.

Заметим, что от целей оценивания зависит содержание измерительных материалов и процедура проведения оценивания.

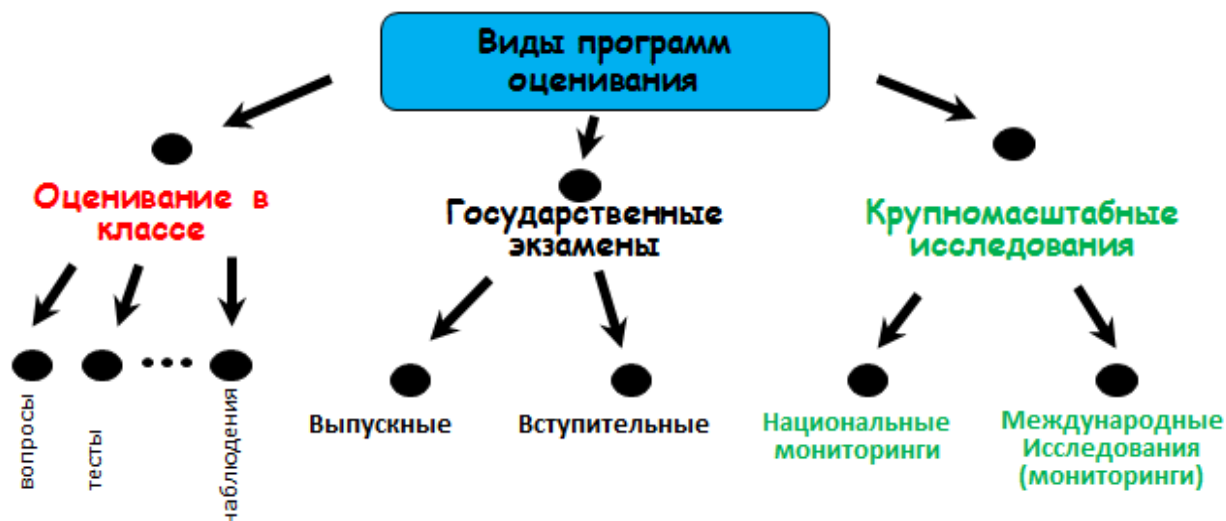
Виды программ оценки

Типология программ оценки учебных достижений может строиться по разным основаниям. Определяющим основанием для классификации могут выступать, например, цели оценки, объекты оценки, возможные последствия по итогам оценки для учащихся и образовательных учреждений (высокие или низкие ставки), периоды времени для проведения оценки (текущее или итоговое оценивание) и т.п.

Мы остановимся на типологии, которая базируется на целях оценки и используется в совместной программе Правительства Российской Федерации и Всемирного банка READ². В большинстве образовательных систем мира системы оценивания преимущественно состоят из трёх основных видов оценочных мероприятий. К ним относятся (см. [10]):

- *оценивание в классе* (или формирующее оценивание), которое встроено в образовательный процесс и позволяет определять достижения и трудности в обучении каждого ученика (Цель 1);
- *государственные экзамены*, которые используются для принятия решений по каждому учащемуся относительно завершения им ступени обучения и возможности продолжения образования (Цели 2 и 3);
- *крупномасштабные исследования* (мониторинги), которые направлены на оценку эффективности работы образовательных систем разного уровня и являются инструментом мониторинга результатов обучения школьников (Цели 4 - 6).

² READ (Russia Education Aid for Development) - Российская программа содействия образованию в целях развития. Реализуется в партнёрстве с Всемирным банком и ориентирована на усиление потенциала развивающихся странах в сфере оценки учебных достижений школьников ([24]).



Рассмотрим эти программы оценки подробнее.

ОЦЕНИВАНИЕ В КЛАССЕ

Наиболее эффективной для организации обучения детей является система оценки, которая приближена к ученику, обеспечивает его постоянное взаимодействие с учителем, служит обратной связью и позволяет модифицировать процесс преподавания и учения. В международной практике такой подход носит название *оценивание в классе* (classroom assessment) или *формирующее оценивание* ([7]). В российской практике этот подход также хорошо известен (как текущее или внутриклассное/внутришкольное оценивание) и используется учителями в своей повседневной работе.



Главная цель оценивания в классе - помочь учителям и учащимся в совершенствовании процесса преподавания и учения, поддержать прогресс ученика.

Базовые вопросы, на которые даёт ответ данная программа оценки, следующие: *Как учится ученик и как лучше его обучать? Каковы сильные стороны конкретного ученика и как их можно развить? В чём ученик испытывает трудности, и как они могут быть преодолены?*

Используются разнообразные методы оценки – учебные задания, тесты, вопросы, наблюдения, собеседования с учителем, самооценка учащегося, оценка учащихся друг другом и т.п.

Внутриклассное оценивание обеспечивает педагогов обратной связью, которая даёт им информацию о том, насколько они эффективны как преподаватели, а детям показывает, насколько они прогрессируют как ученики ([13]).

Пример 1. Оценивание в целях обучения (Великобритания).

Представители системы образования Великобритании определяют оценивание в классе как *оценивание в целях обучения*.

Оценивание в целях обучения (Assessment for Learning) – процесс поиска и интерпретации достоверных данных, которые ученики и их учителя используют для того, чтобы решить, где обучаемые находятся в процессе своего обучения, куда им необходимо двигаться дальше и как сделать это наилучшим образом ([16]).

Оценивание в целях обучения обладает следующими характеристиками ([17]):

- встроено в процесс преподавания и учения и является их существенной частью;
- предполагает обсуждение и согласование учебных целей учителями и учениками;
- нацелено на то, чтобы помочь ученикам осознавать те цели обучения, которых они должны достичь;
- вовлекает учеников в самооценивание или оценивание друг друга;
- обеспечивает обратную связь, которая помогает ученикам осознать, какие следующие шаги в учении им предстоит сделать;
- укрепляет уверенность в том, что каждый ученик может добиться улучшений;
- вовлекает и учителя, и учеников в процесс рассмотрения и рефлексии данных оценивания.

ГОСУДАРСТВЕННЫЕ ЭКЗАМЕНЫ

Государственные экзамены являются оценочной процедурой, которая позволяет определить, освоил ли ученик образовательную программу и/или может ли он продолжать обучение (например, в вузе). В международной практике существуют различные виды экзаменов (на разных ступенях обучения в школе, для поступления в различные учебные заведения, для поступления на работу, получения квалификации и т.п.), но в данной статье мы рассматриваем два вида экзаменов - выпускные (из школы) и вступительные (в вуз).



Целью выпускного экзамена является сертификация учащихся по итогам завершения обучения в школе. Он даёт ответ на вопрос *Как освоил государственный образовательный стандарт /образовательную программу конкретный учащийся?*

Вступительный экзамен обеспечивает отбор претендентов для продолжения обучения в высших учебных заведениях. Это оценочная процедура должна позволить построить рейтинг среди абитуриентов с точки зрения наибольшей готовности продолжить обучение в институтах, университетах, колледжах.

Чаще всего выпускные и вступительные экзамены являются независимыми процедурами, но во многих странах они совмещены.

При проведении экзаменов, особенно вступительных, в большинстве стран мира используются стандартизированные измерительные материалы (тесты). Это делается с целью обеспечения всем учащимся равных возможностей продемонстрировать свои знания и навыки, полученные при обучении в школе.

Пример 2. Экзамены в некоторых странах мира.

ЕГЭ (Россия) – Единый государственный экзамен. Совмещает итоговую аттестацию учащихся и вступительный экзамен в вуз. В ЕГЭ используются стандартизированные тесты, включающие задания с множеством выбором ответа, с кратким свободным ответом и развёрнутым свободным ответом. www.ege.edu.ru

ВАС (Франция) – Baccalauréat. Экзамен на получение квалификации бакалавра после завершения обучения в старшей школе (лицее). Диплом ВАС даёт право для поступления в вузы. Данный экзамен состоит из отдельных экзаменов по нескольким предметам, которые проводятся независимой комиссией в устной или письменной (эссе) формах, стандартизированные тесты не используются. www.education.gouv.fr

GCSE (Великобритания) - General Certificate of Secondary Education, экзамены, которые школьники в возрасте 14-16 лет сдают при окончании первой ступени среднего образования в Англии, Уэльсе и Северной Ирландии. Ученики могут выбирать для сдачи до 10 предметов, из которых часть является обязательными (математика, английский язык, физкультура, естественные науки). Экзамен может сдаваться на двух уровнях сложности (базовый и повышенный), по математике – на трёх (базовый, средний и повышенный). Оценка за экзамен выставляется на основании сдачи стандартизированного теста, результатов промежуточного оценивания по предмету в школе и выполненных в школе курсовых работ и проектов. www.direct.gov.uk/en/EducationAndLearning.

SAT (США) - Scholastic Assessment Test, тест, используемый для поступления в вузы. Существует два варианта этого теста: собственно SAT – тест способностей (оценивает способности абитуриента по чтению, письму и математике к продолжению образования в университетах и колледжах) и SAT Subject Test – тест достижений (оценивает академические знания и навыки по конкретным предметам, требуется для поступления в некоторые учебные заведения). В тесте используются разные тип заданий, но преимущественно это задания с множественным выбором ответа. При приёме вузы кроме результатов SAT часто требуют предъявлять другие данные, характеризующие способности абитуриента (школьные отметки, учительские рекомендации, спортивные достижения и др.) sat.collegeboard.org

КРУПНОМАСШТАБНЫЕ ИССЛЕДОВАНИЯ

Ещё одна важная группа оценочных процедур связана с получением информации о результатах работы образовательных систем. Это так называемые крупномасштабные исследования качества образования или мониторинги учебных достижений школьников. Они могут быть подразделены на международные и национальные мониторинги (этой тематике посвящена работа ([14]).

Определим *мониторинг учебных достижений* школьников как стандартизированную процедуру оценки результатов обучения, которая проводится на регулярной основе и имеет своей целью предоставление актуальной для управленцев и педагогов информации относительно уровня знаний и навыков различных групп учащихся, тенденций их изменений и факторов, оказывающих влияние на результаты обучения. Он может проводиться как на всех учащихся определённого возраста (генеральной совокупности), так и на репрезентативных выборках.

Международные мониторинги.

Международные мониторинги имеют целью сравнение достижений учащихся конкретной страны с достижениями учащихся других стран. Они позволяют дать ответы на следующие вопросы:

- Какой уровень учебных достижений в конкретных областях знаний показывают учащиеся страны в сравнении с учащимися других стран?
- Что школьные системы других стран «делают лучше»?
- Как изменяются результаты с течением времени?

Международные исследования качества образования становятся причиной пересмотра национального учебного плана, введения новых стандартов, изменения программ подготовки учителей. Также они часто стимулируют национальные министерства образования проводить анализ эффективности расходования ресурсов и принимать решения об их увеличении, сокращении или перераспределении для достижения различных целей по развитию образования.

Участие Российской Федерации в международных мониторингах качества образования способствовало формированию культуры педагогических измерений, реформированию содержания образования и созданию федеральных государственных образовательных стандартов нового поколения (ФГОС), разработке новых учебников, а также обновлению программ повышения квалификации учителей ([5, 6]).

Пример 3. Международные мониторинги качества образования.



Исследования Международной Ассоциация по оценке образовательных достижений (IEA – International Association for the Evaluation of Educational Achievement, www.iea.nl/studies).

- *Международное исследование по оценке качества математического и естественнонаучного образования TIMSS (Trends in Mathematics and Science Study).* Целью данного исследования является сравнительная оценка подготовки учащихся 4 и 8 классов школы по математике и естествознанию в странах с различными системами образования. Для выявления факторов, влияющих на качество образования, проводится анкетирование учащихся, их родителей, учителей и администрации школы. В 2011 году в TIMSS приняли участие 63 страны мира. Также в рамках данного исследования периодически оцениваются образовательные достижения выпускников средней школы по углубленному курсу математики и физики (TIMSS Advanced).
- *Международное исследование по изучению качества чтения и понимания текста PIRLS (Progress in International Reading Literacy Study).* Данное исследование сравнивает уровень и качество чтения и понимания текста учащимися начальной школы в различных странах мира. Проводится циклами – один раз в пять лет. В исследовании оцениваются два вида чтения, которые чаще других используются учащимися во время учебных занятий и вне школы: чтение литературных произведений и чтение информационных текстов. В 2011 году в PIRLS приняли участие 49 стран мира.
- *Международное исследование компьютерной и информационной грамотности ICILS (International Computer and Information Literacy Study).* Цель исследования – оценить, насколько хорошо школьники готовы для жизни в информационном обществе, насколько они способны использовать компьютеры для получения знаний и общения в школе, дома и в обществе. Учащиеся 8 классов будут оцениваться на основе компьютерного тестирования. Для сбора контекстной информации применяется анкетирование учеников, учителей и школ. Первое исследование состоится в 2013 году, в нём примут участие 19 стран.




Исследование Организации экономического сотрудничества и развития (OECD - Organisation for Economic Co-operation and Development).


- *Международная программа по оценке образовательных достижений учащихся PISA (Programme for International Student Assessment).* Основной целью PISA является

оценка способности учащихся 15-летнего возраста применять полученные в школе знания и умения в жизненных ситуациях. Областями оценки являются грамотность чтения, естественнонаучная грамотность, математическая грамотность и решение проблем. В 2012 году на добровольной основе проверялись навыки в области финансовой грамотности. В качестве инструментария используются тесты и анкеты для сбора контекстной информации (выявлялись характеристики учащихся и их семей, образовательных учреждений и учебного процесса). Исследование PISA проводится трехлетними циклами. В 2012 году в PISA приняли участие 64 страны мира. www.oecd.org/pisa

Одним из видов международных мониторингов являются межстрановые мониторинги качества образования, которые проводятся для группы стран из одного региона мира.

Пример 4. Межстрановые мониторинги качества образования.

SACMEQ (The Southern and Eastern Africa Consortium for Monitoring Educational Quality) – Южноафриканский консорциум по мониторингу  качества обучения (является неправительственной независимой организацией, включающей 15 африканских стран). Основная цель – всестороннее повышение качества образования, обеспечение достижения планируемых образовательных результатов всеми школьниками, с особым вниманием к грамотности чтения, математике и жизненным навыкам. В мониторинге SACMEQ используются стандартизированные тесты по чтению, математике и естественным наукам с множественным выбором вариантов ответов. www.sacmeq.org

LLECE (The Latin American Laboratory for Assessment of the Quality of Education) –  Латиноамериканская лаборатория по оценке качества обучения. Это сетевая организация, координируемая ЮНЕСКО, которая занимается общей оценкой качества образования в Латинской Америке и странах Карибского бассейна. Цель LLECE состоит в поддержке участия стран в проведении оценки учебных достижений учащихся и использования данных для улучшения качества образования. Участники оценки – учащиеся 3 и 6 классов: 13 стран в 1997 г. и 19 стран в 2007 г. Области оценки – математика, чтение и естественные науки. Инструментарий - тесты достижений с множественным выбором вариантов ответов и анкеты для выявления факторов, влияющих на результаты обучения. www.llece.org

Национальные мониторинги.


Национальный мониторинг качества образования или так называемая национальная оценка³ является стандартизированной оценочной процедурой, предназначенной для определения, насколько эффективно работает система образования страны. Главная задача мониторинга - получение данных, позволяющих судить о том, в какой мере достигаются цели, предусмотренные государственным образовательным стандартом, учебным планом или национальной рамочной структурой образовательных результатов⁴.

Среди базовых вопросов, на которые позволяет получить ответ национальные мониторинги, следует отметить следующие ([11]):

- Насколько эффективно учащиеся овладевают знаниями и навыками в системе образования?
- Какие достижения демонстрируют представители различных групп учащихся?
- Изменяются ли достижения учащихся с течением времени?
- Какие факторы оказывают влияние на результаты обучения?

Данная процедура проводится на периодической основе (1 раз в 2-5 лет), по нескольким предметам (практически всегда по математике и родному языку), на основе выборки учащихся или генеральной совокупности школьников определённого возраста. В качестве инструментария используются стандартизированные тесты и анкеты для сбора контекстной информации. Это даёт возможность изучения факторов, негативно или позитивно влияющих на учебные достижения. Такая информация необходима для разработки политики на различных уровнях образования.

Пример 5. Национальные мониторинги качества образования.

NAEP (The National Assessment of Educational Progress) – национальная оценка прогресса в образовании, США. Проводится не реже 1 раза в два года на  общенациональной выборке учащихся 4, 8, 12-х классов с использованием тестов достижений и анкет для сбора контекстной информации. Оцениваются учебные достижения в ключевых предметных областях (математика, английский язык, социальные науки, экономика, география, история США). Адресатами информации о результатах оценки являются местные сообщества, политики, представители системы образования, широкая общественность. Результаты NAEP публикуются в национальном докладе (The Nation's

³ Национальная оценка (national assessment) - термин, используемый в практике многих зарубежных стран.

⁴ Такая рамочная структура существует в ряде стран (например, США, Австралия, Канада), в которых нет единого государственного стандарта или учебного плана. Национальная рамка является результатом договорённости и консенсуса различных профессиональных групп в системе образования.

Report Card), представляющем обществу информацию об академических достижениях американских школьников. nces.ed.gov/nationsreportcard

NAPLAN (National Assessment Program – Literacy and Numeracy) – национальная программа оценки учебных достижений, Австралия.



Проводится ежегодно для всех учащихся 3, 5, 7 и 9-х классов. Проверяет знание и навыки по четырём областям – чтение, письмо, правила языка (грамматика, пунктуация, орфография) и математические навыки в соответствии с общими требованиями учебных планов всех провинций и территорий страны. Программа направлена на отслеживание индивидуального прогресса каждого учащегося. Результаты NAPLAN общедоступны на национальном сайте «Моя школа», который позволяет сравнить достижения учащихся этой школы со среднестатистическими показателями школ, имеющих статистически сходный контингент учащихся. www.nap.edu.au

SIMCE (Sistema de Medición de Calidad de la Educación) – национальная программа измерения качества образования, Чили. Оцениваются все учащиеся



4, 8 и 10-х классов по математике, испанский язык, социальным и естественным наукам. На основе выборки оцениваются письмо, английский язык, компьютерные навыки и физическое воспитание. Используются анкеты для учащихся, их родителей и учителей с целью сбора дополнительной информации (например, о социальном и экономическом положении, возможности обучаться в школе, квалификации учителей). Программа SIMCE характеризуется комплексным подходом к распространению результатов оценки и преследует следующие цели: информирование лиц, ответственных за политику образования, предоставление педагогической поддержки работникам системы образования, и укрепление подотчётности и ответственности школ. www.simce.cl

Измерения с высокими и низкими ставками

Отдельного внимания и учёта требует вопрос измерений с высокими и низкими ставками. Именно от него зависит, как будут использоваться результаты программ оценки – для жёсткого контроля с соответствующими организационными выводами или для оказания поддержки учащимся и школам.

Хорошо известно, что каждая программа оценки несёт в себе определённые последствия для её участников – учеников, учителей, школ. В одном случае низкие результаты теста не сказываются негативно на ученике, они используются лишь для работы над ошибками и помогают совершенствовать образовательный процесс. Это ситуация измерений с *низкими ставками*. В другом случае неудача по итогам оценки лишает ученика

каких-то важных для него возможностей, приводит к негативным суждениям о качестве работы учителя и школы. Такое положение дел относят к измерениям с *высокими ставками*.

Международные и национальные мониторинги наряду с оцениванием в классе относятся к программам оценки с низкими ставками. В этих программах низкие результаты тестов, полученные школьниками, не могут быть использованы для наложения санкций на учеников, учителей и школы.

Результаты мониторинговых исследований обычно представляются в агрегированном виде, индивидуальные результаты школьников не рассматриваются, а, следовательно, на их основе не могут приниматься решения по конкретному ученику. В случае проведения оценки на выборке учащихся (пример NAEP в США) нет возможностей для оценки качества работы всех школ, так как лишь небольшая доля образовательных учреждений принимает участие в оценочной процедуре.

Экзамены относятся к программам *оценки с высокими ставками*. От их результатов очень сильно зависит будущее учащегося, например, продолжение образования в вузе или получение работы. Кроме того, результаты экзаменов могут использоваться для оценки работы педагогов и школ.

Опыт России и многих стран мира показывает, что ввиду высокой значимости экзаменов возникает серьёзный риск дезориентации образовательного процесса и нацеленности его лишь на те знания и навыки, которые проверяются в рамках экзаменационных тестов (эффект обучения для сдачи теста). При этом многие важные образовательные результаты оказываются вне внимания школ и педагогов, им уделяется гораздо меньше учебного времени.

Боязнь «проиграть», желание избежать негативных последствий, равно как и возможность получить поощрение в виде дополнительных ресурсов (премии и доплаты учителям, гранты школам и т.п.) создают ложные стимулы и становятся причиной нарушений и фальсификаций при проведении измерений с высокими ставками.

При проектировании национальных и региональных систем оценки качества необходимо обеспечивать разумный баланс оценочных процедур с высокими и низкими ставками. Невозможно построить эффективную систему оценки, которая базируется лишь на «жестких» измерениях, порождающих победителей и проигравших.

Основные характеристики программ оценки

В дополнении к сказанному выше, рассмотрим и другие характеристики оценивания в классе, экзаменов, международных и национальных мониторингов. Для удобства представим их в сводной сравнительной таблице⁵.

Характеристики программ оценки	Виды программ оценки			
	Оценивание в классе	Государственные экзамены	Крупномасштабные исследования	
			Международные	Национальные
Цели	Получение обратной связи для выявления успехов и проблем в обучении ученика	Аттестация и отбор учащихся для продолжения образования или трудоустройства	Сравнение результатов работы национальной системы образования с другими странами	Оценка эффективности работы национальной системы образования
Периодичность	постоянно, то часть процесса обучения	ежегодно	регулярно (раз в 3-5 лет)	регулярно (раз в 2-5 лет), иногда ежегодно
Кто оценивается	все учащиеся	все учащиеся, обязанные сдавать экзамен	выборка учащихся определённого возраста или класса	выборка или вся совокупность учащихся определённого возраста или класса
Формат	различные методы оценивания (стандартизир. и не стандартизир.)	Обычно тест с множеств. выбором и кратким ответом, также эссе	обычно тест с множеств. выбором и кратким ответом, также эссе	обычно тест с множеств. выбором и кратким ответом
Охват учебной программы	все учебные темы	охватывает основные учебные темы	обычно ограничена несколькими предметами	обычно ограничена несколькими предметами
Сбор контекстной информации	да, это часть процесса обучения	редко	как правило, да	часто
Возможные последствия	низкие ставки	высокие ставки	низкие ставки	низкие ставки
Влияние на учебный процесс	сильное положительное влияние (даёт обратную связь ученику и учителю)	сильное отрицательное влияние (обучение для сдачи теста)	незначительное	незначительное
Отслеживание тенденций	да	нет	да	да
Методы анализа результатов	как правило, используются неформальные и самые простые методы	от простых до более сложных статистических методов	как правило, используются сложные статистические методы	от простых до более сложных статистических методов

⁵ Таблица составлена с учётом сравнительного анализа экзаменов и мониторингов, проведённого Грини и Келлаган ([11]).

Актуальные задачи в области оценки качества образования

В завершении сформулируем ряд актуальных научно-практических задач, которые, как показывает российский и международный опыт, необходимо решать каждой стране, стремящейся построить сбалансированную и эффективную систему оценки качества образования. Подробное обсуждение таких задач – предмет отдельного рассмотрения и последующих публикаций. Здесь же авторы ограничатся их постановкой.

1. Учёт разнообразных образовательных достижений учащихся.

Это важнейшая задача, она требует использования на практике самых разных «измерительных» процедур. Процедуры оценки должны быть ориентированы на более широкие образовательные результаты, нежели знания в рамках учебных дисциплин (компетентности, здоровье, гражданская позиция, умение позитивно взаимодействовать с другими, самоуважение, забота об окружающей среде и др.) Стандартизированное тестирование – всего лишь один из используемых подходов для определения качества обучения. Также необходимо использовать и другие методы – самооценка, экспертные оценки, социологические опросы и т.п.

2. Минимизация измерений с высокими ставками.

Государственные экзамены должны занимать важное, но не главное место в национальной/региональной системе оценки качества образования. Если мы хотим оценить результаты обучения и при этом понять, чему научила детей школа (увидеть вклад школы), то для такой оценки неправомерно использовать измерения «с высокими ставками». В противном случае начнут работать внешкольные факторы – например, репетиторство, а также могут включаться «криминальные» механизмы повышения результатов теста – нарушения и фальсификации при проведении тестирования. Для получения целостной и неискажённой картины образовательных результатов необходимо вводить инструменты оценки с низкими ставками, прежде всего мониторинги учебных достижений.

3. Оценка достижений в динамике.

Международный опыт показывает, что для справедливой оценки результатов обучения необходимо не только ориентироваться на результаты, показанные учениками в определённый момент времени, но и учитывать динамику их изменений на протяжении нескольких циклов оценки. Здесь главный вопрос – обеспечивает ли учитель и школа прирост в образовательных достижениях ученика, его индивидуальный прогресс? Использование модели «индивидуального прогресса» при оценке результатов обучения школьников и деятельности образовательных учреждений требует введения в практику инструментов, позволяющих проводить

регулярную оценку и фиксировать складывающихся тенденций. Для этих целей этих целей предназначены внутриклассное оценивание и мониторинги учебных достижений.

4. *Корректное сопоставление результатов оценки.*

При анализе деятельности школ по итогам проведения той или иной оценочной процедуры необходимо обеспечивать корректное сопоставление результатов, избегая упрощённых выводов и решений. Очевидно, что школа с большим количеством детей из семей с низким социально-экономическим статусом не сможет демонстрировать результаты, сопоставимые с теми, которые показывают образовательные учреждения, обучающие детей из благополучных семей. Справедливое сравнение должно основываться на кластерном анализе данных, когда школы группируются по ряду схожих характеристик (тип школы, её расположение, социально-экономический статус семьи, доля учеников с неродным языком обучения и др.). Здесь важно изучать опыт зарубежных стран, многие из которых имеют многолетний опыт проведения кластерного анализа и выделения групп статистически подобных образовательных учреждений (например, Англия, Шотландия, Австралия).

5. *Учёт дополнительных данных при принятии управленческих решений по итогам оценки.*

Крайне важно, чтобы при проведении процедур «с высокими ставками», такими как государственные экзамены, оценка эффективности работы школы, оценка качества работы учителя, управленческие выводы о качестве образования делались не только на основе результатов педагогических измерений. Невозможно принимать «сложные» решения на основании данных одного вида. Необходимо учитывать и другие источники данных – статистика, самооценка, портфолио, мнения потребителей образовательных услуг, результаты различных внешних оценок и т.п. Такой подход позволит принимать сбалансированные решения и избегать однобокого взгляда на работу образовательного учреждения. Результаты тестирования являются важной, но далеко не единственной частью информационного портрета школы.

6. *Ориентация на информационные потребности групп пользователей результатов оценки.*

Эффективные системы оценки качества образования предоставляют информацию надлежащего качества и в необходимом количестве для того, чтобы удовлетворить информационные потребности всех заинтересованных групп и тех, кто принимает решения с целью повышения качества обучения учащихся.

Обеспечение максимального влияния результатов оценки на образовательную политику зависит от трёх ключевых условий, которые необходимо учитывать при проведении программы учебных достижений школьников ([3]):

- 1) знание пользователей результатов программ оценки и их информационных потребностей;
- 2) наличие информационных продуктов, способных удовлетворить существующие информационные потребности различных групп пользователей;
- 3) понимание, какие способы распространения информационных продуктов следует использовать.

Передовой международный и российский опыт говорит о том, что умение интерпретировать результаты оценки качества образования и доводить их до представителей заинтересованных сторон является неотъемлемой частью любой программы оценки учебных достижений. От этого зависит её успех или неудача.

Список литературы

1. Болотов В.А. Научно-педагогическое обеспечение оценки качества образования // Педагогика. № 1, 2010.
2. Болотов В.А. Оценка качества образования // Управление школой. № 5, 2012.
3. Болотов В.А., Вальдман И.А. Информирование различных целевых групп как условие эффективного использования результатов оценки учебных достижений школьников. [Электронный ресурс] / В.А. Болотов, И.А. Вальдман // Проблемы современного образования. – 2012. – № 6. – С. 187-202. – Режим доступа: http://www.pmedu.ru/res/2012_6_13.pdf.
4. Болотов В.А., Вальдман И.А. Условия эффективного использования результатов оценки образовательных достижений школьников. // Педагогика. №6, 2012.
5. Болотов В.А., Вальдман И.А., Ковалёва Г.С. Российская система оценки качества образования: чему мы научились за 10 лет? В сб. «Тенденции развития образования: проблемы управления и оценки качества образования». Материалы VIII Международной научно-практической конференции. – М.: Университетская книга, 2012, с. 22-31.
6. Болотов В.А., Вальдман И.А., Ковалёва Г.С., Пинская М.А.. Анализ опыта создания российской системы оценки качества образования. // Управление образованием: теория и практика. Вып. 1-2, 2011. URL: <http://www.iuorao.ru/2010-01-01-14>
7. Вальдман И.А. Ключевые аспекты качества образования: уроки международного опыта. – М.: Московский центр качества образования, 2009.
8. Государственная программа Российской Федерации "Развитие образования" на 2013-2020 годы. Распоряжение Правительства РФ №2148-р от 22 ноября 2012 года.
9. Использование результатов национальной оценки учебных достижений / Грини В., Кэллаган Т., Мюррей С. // Серия «Национальная оценка учебных достижений», Книга 5. – Всемирный банк, 2011.
10. Кларк М. Что является наиболее важным в системах оценки достижений учащихся: основные ориентиры. - Всемирный банк, 2012.
11. Оценка образовательных достижений на национальном уровне / Грини В., Кэллаган Т. // Серия «Национальная оценка учебных достижений», Книга 1. – Всемирный банк, 2011.

12. Оценка качества образования: обзор международных подходов и тенденций. - Всемирный банк, 2005.
13. Пинская М.А. Формирующее оценивание: оценивание в классе: учеб. пособие. – М.: Логос, 2010.
14. Поливанова К.Н. Национальные и международные программы оценки качества образования: учеб. пособие. – М.: Логос, 2010.
15. Хокер Д. Системы оценки в Англии и Уэльсе – четыре вопроса и пять уроков. Журнал руководителя управления образованием. №6, 2011, стр. 47-54.
16. Assessment for learning: 10 principles, Assessment Reform Group, 2002.
17. Assessment for learning: beyond the black box. Assessment Reform Group, 1999.
18. Clarke M. Roadmap for building an effective assessment system. The World Bank, 2010.
19. Erin D. Caffrey. Assessment in Elementary and Secondary Education: A Primer. Congressional Research Service, 2009.
20. Forster, M. (2000) *A Policy Maker's Guide for Managers to International Achievements Studies*, Australian Council for Educational Research, Melburn.
21. Kellaghan, T., Greaney, V. and Murray, T. S. (2009) *Using the results of a national assessment of educational achievement*, The World Bank, Washington, D.C.
22. Murray, T.S. (2004) *Achieving Maximum Impact: A Guide for Managers of National assessment Systems*, Statistics Canada, Ottawa.
23. National Assessment Program - Literacy and Numeracy (NAPLAN), Australia. <http://www.naplan.edu.au/>
24. READ Trust Fund. <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/EXTREAD/0,,menuPK:7526437~pagePK:64168427~piPK:64168435~theSitePK:7526432,00.html>
25. SABER: System Assessment and Benchmarking for Education Results. The World Bank. <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,contentMDK:22710669~menuPK:282391~pagePK:148956~piPK:216618~theSitePK:282386,00.html>