



Болотов Виктор Александрович,
доктор педагогических наук, академик Российской академии образования



Вальдман Игорь Александрович,
кандидат педагогических наук, заведующий лабораторией мониторинга
в образовании Института управления образованием РАО

КАК ОЦЕНИВАТЬ

ОБРАЗОВАТЕЛЬНЫЕ ДОСТИЖЕНИЯ ШКОЛЬНИКОВ

ВВЕДЕНИЕ

Данная статья подготовлена на основе учебных материалов Российского тренингового центра ИУО РАО, разработанных авторами в 2011–2012 годах. В ней будет представлена одна из возможных типологий программ оценки учебных достижений школьников и дана их характеристика.

Ее актуальность обусловлена тем, что построение эффективной системы оценки качества образования во многих странах является приоритетом 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.

Выбор оценочных инструментов и процедур зависит от того, кто и как планирует использовать результаты оценки.

Данные о результатах оценки достижений учащихся могут быть использованы для 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.

Это может быть:

- 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...

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. 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».



Государственная программа Российской Федерации «Развитие образования» на 2013–2020 годы. Распоряжение Правительства РФ от 22 ноября 2012 года № 2148-р.

ЦЕЛИ ОЦЕНИВАНИЯ

Цели оценивания (исходя из того, как его применять) условно могут быть разделены на две категории:

- для принятий решений о поддержке конкретных учащихся;
- для принятий управленческих решений на разных уровнях системы образования — от конкретной школы до министерства образования страны.

В первом случае речь идет о способах использования, которые связаны с корректировкой образовательного процесса и (или) определением уровня освоения образовательной программы учеником.

Во втором случае речь идет 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 и определение ближайших задач образовательного учреждения и того, каким образом надо их решать.

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.

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 пред-

методом 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 READ.

В большинстве образовательных систем мира системы оценивания преимущественно состоят из трех основных видов оценочных мероприятий.

READ (Russia Education Aid for Development) — Российская программа содействия образованию в целях развития. Реализуется в партнерстве с Всемирным банком и ориентирована на усиление потенциала развивающихся стран в сфере оценки учебных достижений школьников.

К ним относятся:
— оценивание в классе (или формирующее оценивание), которое встроено в образовательный процесс и позволяет определять

достижения и трудности в обучении каждого ученика;

— государственные экзамены, которые используются для принятия решений по каждому учащемуся относительно завершения им ступени обучения и возможности продолжения образования;

— крупномасштабные исследования, которые направлены на оценку эффективности работы образовательных систем разного уровня и являются инструментом мониторинга результатов обучения школьников.

1. 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.

Базовые вопросы, на которые дает ответ данная программа оценки, следующие:

— как учится ученик и как лучше его обучать?

— каковы сильные стороны конкретного

ученика и как их можно развить?

— в чем ученик испытывает трудности и как они могут быть преодолены?

Вальдман И. А. Ключевые аспекты качества образования: уроки международного опыта. М.: Московский центр качества образования, 2009.

Внутриклассное оценивание обеспечивает педагогов обратной связью, которая дает им информацию о том, насколько они эффективны как преподаватели, а детям показывает, насколько они прогрессируют как ученики.

2. 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).

См., например: Пинская М. А. Формирующее оценивание: оценивание в классе: учеб. пособие. М.: Логос, 2010.

Целью выпускного экзамена является сертификация учащихся по итогам завершения обучения в школе. Вступительный экзамен обеспечивает отбор претендентов для продолжения обучения в высших учебных заведениях.


Чаще всего выпускные и вступительные экзамены являются независимыми процедурами, но в ряде стран они совмещены.

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

3. Крупномасштабные исследования

Еще 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 проводиться как на всех учащихся определенного возраста (генеральной совокупности), так и на репрезентативных выборках.



Под мониторингом здесь мы понимаем 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.

3.1. 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.

3.2. Национальные мониторинги

Национальный мониторинг качества образования, или так называемая национальная оценка, 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 результатов.




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

Среди базовых вопросов, на которые позволяет получить ответ национальный мониторинг, следует отметить следующие:

- 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),



См., например: Оценка образовательных достижений на национальном уровне / Грини В., Кэллаган Т. // Серия «Национальная оценка учебных достижений». Кн. 1. Всемирный банк, 2011.

по нескольким предметам (практически всегда по математике и родному языку), на основе выборки учащихся или генеральной совокупности школьников определенного возраста.

В качестве инструментария используются стандартизированные тесты и анкеты для сбора контекстной информации. Это дает возможность изучения факторов, негативно или позитивно влияющих на учебные достижения. Такая информация необходима для разработки политики на различных уровнях образования.

ИЗМЕРЕНИЯ С ВЫСОКИМИ И НИЗКИМИ СТАВКАМИ

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

Хорошо известно, что каждая программа оценки несет в себе определенные последствия для ее 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 показывает, что ввиду высокой значимости экзаменов возникает серьезный 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 дополнение к сказанному выше рассмотрим и другие характеристики оценивания в классе, экзаменов, международных и национальных мониторингов. Для удобства представим их в сводной сравнительной таблице.

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, это часть процесса обучения	Ежегодно	Регулярно (раз в 3–5 лет)	Регулярно (раз в 2–5 лет), иногда ежегодно
Кто оценивается	Все учащиеся	Все учащиеся, обязательные сдавать экзамен	Выборка учащихся определенного возраста или класса	Выборка или вся совокупность учащихся определенного 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 (дает обратную связь ученику и учителю)	Сильное отрицательное влияние (обучение для сдачи теста)	Незначительное	Незначительное
Отслеживание тенденций в обучении	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

АКТУАЛЬНЫЕ ЗАДАЧИ В ОБЛАСТИ ОЦЕНКИ КАЧЕСТВА ОБРАЗОВАНИЯ

В завершение 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. Учет разнообразных образовательных достижений учащихся.

Это важнейшая задача, она требует использования на практике самых разных «измерительных» процедур, ориентированных на более широкие образовательные результаты, нежели знания в рамках учебных дисциплин. 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 неискаженной 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 определенный 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 складывающихся тенденций. Для 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 упрощенных 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.**

Таковыми являются тип школы, ее 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. Учет дополнительных данных при принятии управленческих решений по итогам оценки.

Крайне важно, чтобы при проведении процедур «с высокими ставками», таких 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.

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 ее успех или неудача.