

МУЛЬТИПЛЕКСНОЕ ИЗМЕРЕНИЕ КАЧЕСТВА ОБРАЗОВАНИЯ

Н.Н. Найденова (Москва)

Рассматривая проблему измерения качества образования в школе с психологических, социальных, культурных, педагогических и др. позиций, можно попытаться объективно, надежно и валидно измерить уровень качества образования учащегося, класса внутри одной параллели, между классами, на разных ступенях обучения, в целом по школе. Такое измерение должно быть мультипредметным, многовариативным, полиинструментальным, многофакторным, мультишкальным и др. (поли- и много-) – т.е. оно является мультиплексным. В результате мониторинга на основе мультиплекса формируются индивидуальная оценка качества образования учащихся, учителей, классов, школы, что позволяет оценить свои индивидуальные оценки в школе и в мониторинге и идентифицировать себя на шкалах международного, наднационального и национального инструментария с учетом социально-педагогических и психологических факторов. Кроме того, согласно современной теории эвалюации измерение качества образования проводится в рамках реалистической стадии эвалюации, включая международный инструментарий, оценку ключевых компетенций, оценивание уровня информационной грамотности и др.

Поэтому мониторинг качества образования в школе мультиплексным методом Найденовой в условиях современного глобализирующегося и изменяющегося образования, когда модернизация общероссийской, региональной, муниципальной и школьной систем оценки качества образования и текущая информационная поддержка позитивного образа педагогических измерений (ЕГЭ, ГИА, стандарты и т.п.) просто необходимы, представляется особенно актуальным, так как позволяет получить надежную и валидную информацию о реальном уровне качества образования в школе и выявить факторы, способствующие повышению качества образования.

Для оценивания качества образования в школе были использованы:

- *Мультиплексный метод измерения качества образования в школе*, относящийся к классу проективных методов, который дает возможность школе и самому обучающемуся проецировать свою успешность и по-своему ее интерпретировать,
- *Анкеты учащихся, их родителей и учителей, администрации школы, интервью, анкеты экспертов по кодировке открытых заданий и др.*, позволяющие учитывать социально-педагогические и психологические факторы, которые влияют на уровень качества образования в школе,
- *Протоколы наблюдения, руководства, инструкции, шаблоны* и др. инструктивные материалы, по результатам которых можно оценить процесс измерения,
- *Тетради с тестовыми заданиями* в различных формах предъявления, по разным предметам, видам деятельности и компетенций, разным инструментальным средствам, на основании которых проводится

латентная оценка и заданий, и результатов учащихся на единой шкале Раша****.

В ходе мониторингового исследования качества образования в школе были использованы различные подходы к мониторингу: тотальный (ежегодный), срезовый (одноразовый), периодичный (интервальный). Использовались международный, наднациональный и национальный инструментарий во всех параллелях: II-VIII, X классы.

В результате исследования сформировался подход к определению профильного состава классов. На основании мультиплекса появилась возможность оценивания учителей по результатам их учащихся. Каждый ученик имеет свою индивидуальную линию обучения с прогностической функцией. Средние результаты оценивались градиентом.

Большой уровень качества образования демонстрировали те учащиеся и классы, которые имели более высокий уровень социально-педагогических и психологических факторов, но имелись и исключения. На рис.1 приведена градиентная оценка средних результатов классов одной параллели. Рис.2 показывает средние результаты тех же классов, измеренные традиционным методом, но разными инструментариями. Полный профиль мультиплекса по предметности и инструментарию приведен на рис. 3 для одного из учеников выпуска 2010 г. Ее же сертификат, врученный ей вместе с аттестатом приведен на рис. 4. Кросс-оценивание двух учащихся, включая результаты предыдущей девочки, приведены на рис.5 (текущие оценки в X кл., сертификат мультиплекса, аттестат и сертификат ЕГЭ по обязательным предметам).

Учащиеся с высоким и средним уровнем качества образования имеют результат выше 450 баллов, а также проявили большую активность при ответе на задания в открытой форме.

Метод прошел апробацию в нескольких регионах РФ, в том числе и на репрезентативной выборке школ региона. Результаты по мультиплексу близки, как не странно, и с ПИЗА, ПИРЛС и ТИМС, а также и с ЕГЭ и ГИА.

Таким образом, за 2008-10 гг. на экспериментальной площадке были проведены исследования в разных режимах по оценке качества образования методом мультиплексного измерения. Результаты такого оценивания использовались в школе с разными целями для улучшения учебно-воспитательного процесса и повышения эффективности деятельности учителей.

Использование мультиплекса позволило провести сравнительное исследование оценок результатов учащихся на единой шкале, включая мультиплекс, текущие оценки за X класс, средний балл по аттестату и тестовые баллы ЕГЭ по обязательным предметам, что доказало объективность мультиплекса, который обладает достаточно хорошей прогностической функцией для результатов ЕГЭ. Такая информация полезна не только в целях грамотного принятия управленческих решений, но и для родителей, так как начиная с X класса, когда проводится мультиплекс, можно

получить надежные, объективные и достоверные рекомендации по совершенствованию качества образования как учащегося, так и школы.

Введение нового метода оценки качества образования, в конечном счете, повышает в целом качество образования в школе, так как администрация школы получает полную интегрированную картину качества образования в школе, классе, параллели. Валидность измерения подтверждается и кросс-оценками на разных шкалах, и заданиями разного уровня и вида, чаще всего юнитного характера мультипредметного и межпредметного содержания в рамках авторской концепции измерения «Читаю. Пишу. Размышляю». Эта концепция измерения относится ко всем предметным областям. Ниже приводятся показатели мультиплекса (А-Ж):

А. Инструментарий измерения:

1. Международный (ПИЗА, ПИРЛС, ТИМСС и др.).
2. Наднациональный (адаптированный инструментарий стран с высоким качеством образования в мире или переработанные отечественные задания из ЕГЭ и ГИА по международной технологии).
3. Национальный (отечественные задания из ЕГЭ, ГИА, итоговых работ за разные годы обучения, вступительные для приема в специализированные школы и т.д.)

Б. Предметность:

- Чтение как межпредметная деятельность (адаптированные тексты содержат информацию всех предметных областей измерения, указанных ниже).
- Математика.
- Естествознание (физика, биология, химия, экология).
- Русский язык.
- Иностранный язык (в основном английский).
- Социальный цикл (история, право, обществознание, экономика, география).

В. Концепция измерения:

1. Читательские компетенции на всех предметах, включая физкультуру.
2. Письменные компетенции на всех предметах.
3. Креативные компетенции в разных предметах (умение составить дизайн эссе и написать по нему эссе).

Г. Виды деятельности:

- Репродукция.
- Интерпретация.
- Рефлексия.

Д. Грамотность:

1. Информационная.
2. Естественно-математическая.
3. Коммуникативная.

Е. Эссе

- Дизайн.
- Создание.

Ж. Профильность (измерение компетенций):

1. Точные науки.
2. Гуманитарная.
3. Социальная.
4. Информационная.

По всем показателям качества образования (А-Ж) каждый учащийся получает рейтинговую и балльную оценки на единой шкале Раша. Учащиеся школ, участвующих в мониторинге по мультиплексу, имеют портфолио за все годы участия с такими внешними по отношению к школе оценками.

Нельзя забывать и о том, что нынешняя школа функционирует в эпоху глобализации, поэтому в инструментарии оценки качества образования должны быть задания не только национального, но и международного уровней.

Следовательно, чтобы оценить качество образования в школе необходимо:

1. Иметь валидный и надежный КИМ, желательно мультиплекс.
2. Иметь штат сотрудников с соответствующей квалификацией.
3. Сочетать в одной процедуре оценивание международным, наднациональным и национальным инструментариями.
4. Оценивать не только учебные достижения, но и личные возможности как школьника, так и учителя, класса и школы в целом.
5. Обязательно включать в оценку качества образования социально-психологические измерения (тип мышления, скорость и вид восприятия, усидчивость, уровень социализации обучающегося и его семьи, экономический статус и др.).
6. В режиме мониторинга менять фокус измерения в волновом режиме, сочетая с изменениями в сложности инструментария измерения и оценивая качество образования по кривой роста.
7. Сравнивать по массовому градиенту, а не по средним баллам.
8. Сделать эвалюацию (оценивание оценок) в школе доступной родителям, ученикам, учителям и работникам административного звена.
9. Регулярно публиковать и вывешивать итоги эвалюации на своем сайте.
10. Шире участвовать в мультиплексных измерениях на международном, наднациональном и национальном уровнях.
11. Использовать кросс-оценки из разных исследований.

Метод мультиплексного измерения качества образования можно использовать в образовательных учреждениях разного вида и различных ступеней обучения. Так, пилотная апробация метода была проведена в учреждениях дошкольного образования и среди студентов образовательных учреждений высшего образования.

Как любой метод оценивания качества образования мультиплекс является измеряемым методом, а не экспертным, хотя при измерении

качества образования мультиплексом используются и экспертные заключения по некоторым пунктам. В качестве основных целей при внедрении мультиплекса в школьную систему оценки качества образования были определены следующие:

1. Способствовать внедрению современных методик и технологий измерения для улучшения текущей деятельности школы.
2. Организовывать проведение внешнего и внутреннего оценивания в школе полным по мощности измерения инструментарием.
3. Оценка качества образования должна быть систематической, чтобы повышать измерительную квалификацию школы.
4. Главный принцип измерения – не доказательство измеряемых характеристик, а рекомендации по улучшению этих характеристик.
5. Измерение является целостным и включает как формирующую, так и итоговую оценки.
6. При оценивании применяются как количественные, так и качественные методы анализа собранных измерительных данных.
7. Дизайн и методология педагогического измерения должны быть полезными для управления качеством в школе, легкими по выполнению, не нарушающими радикально учебный процесс, соответствовать принятым стандартам как обучения, так и точности измерения.

Для выполнения этих целей необходимо гарантировать объективность измерения, что достигается ежегодным оцениванием качества образования квалифицированными в области педагогических измерений специалистами, являющимися для школы третьими лицами.

Кроме перечисленных целей педагогическое измерение качества образования содержит и дополнительные цели:

1. Обеспечение валидной информацией об эффективности деятельности школы.
2. Предоставление другим школам анонимных результатов школ, прошедших эвалюацию качества образования.
3. Предоставление учащимся и родителям эвалюационной информации, включая рейтинговые и ранговые оценки по школе, среди школ, по региону и стране.

Опытная апробация мультиплексного измерения качества образования проводилась в течение пяти лет среди учащихся начальной, основной и средней ступеней обучения в различных школах как по профилю обучения, так и в рамках альтернативной педагогики.

Условия участия школы в мультиплексе были едиными:

1. Школы участвовали добровольно.
2. Школы участвовали в разные отрезки времени и не пересекались друг с другом.
3. Все школы работали в рамках пошаговой модели.
4. Первые восемнадцать шагов являются базовыми и выполняются в каждой школе.

Пошаговая модель мультиплекса состоит из двух моделей: модели разработки инструментария из десяти шагов и модели проведения мультиплексного измерения из сорока шагов.

Модель проведения мультиплекса включает:

1. Подготовку оценочного портфолио всех учащихся.
2. Разработку репрезентативной выборки учащихся для распределения вариантов теста.
3. Анкетирование учащихся и их родителей; учителей, преподающих в классах тестирования; завуча и директора школы.
4. Проведение трехминутного теста проверки информационной грамотности.
5. Верификация и составление схемы оценивания заданий теста персоналом школы.
6. Верификация тем же персоналом экспертизы заданий и теста по разработанной форме. Этот шаг выполняется в школах, недостаточно владеющих тестологическими умениями.
7. Подбор персонала тестирования и организация самого тестирования, включая план рассадки учащихся.
8. Размножение материалов педагогического измерения.
9. Формирование минимальной выборки для верификации данных.
10. Формирование группы интервью по итогам апробации.
11. Разработку протоколов интервью.
12. Проведение тестирования по руководству и заполнение протокола.
13. Проведение интервью.
14. Разработку шаблонов ввода.
15. Ввод информации по шаблону.
16. Первичный анализ данных и верификация.
17. Тестологический анализ.
18. Разработку презентации по мультиплексу.
19. Вторичный анализ, включая IRT-анализ.
20. Анализ по разным параллелям.
21. Разработку презентации по школе.
22. Анализ по годам.
23. Разработку презентации по годам.
24. Формирование мониторинговой выборки.
25. Мониторинговый анализ.
26. Разработку презентации по мониторингу.
27. Регрессионный, дисперсионный, кластерный, факторный и т.п. анализ.
28. Построение прогностической модели.
29. Определение профиля класса.
30. Построение индивидуальной траектории обучения.
31. Разработку интерпретации результатов.
32. Сравнение школьного портфолио и мультиплекса.

33. Разработку рекомендаций учащимся, родителям, учителям, администрации.

34. Кросс-оценки, рейтинги, ранги, прогнозы, градиенты, оценки ключевых компетенций, оценивание по деятельности и др.

35. Разработку прогностической презентации.

36. Оценивание образовательной среды, квалификации учителя, управления качеством, применения современных методик обучения, материально-технического состояния, социального статуса, экономического статуса, культурного статуса, уровня здоровья и др.

37. Расчет надежных и валидных характеристик, объясняющих результаты по качеству образования с учетом взвешенной факторной нагрузки.

38. Разработку интерпретационной презентации.

39. Анализ по ступеням обучения от дошкольной до высшей по итогам мониторинга.

40. Разработку непрерывной презентации по ступеням обучения.

Так как инициатива мультиплексного измерения качества образования в школе является новой не только для нашей страны, но и для всего мирового образовательного пространства и школы находятся на различных стадиях развития в сфере педагогических измерений, план оценки включает элементы и формирующей, и итоговой оценок.

Цель формативной оценки состоит в том, чтобы предоставить школе информацию для улучшения качества обучения и воспитания в школе. Такое измерение проводится непрерывно в течение учебного года (2-4 раза) и в течение всей ступени обучения. Типичные вопросы, которые возникают в пределах контекста формирующих оценок, представлены ниже:

1. Управление в школах выстроено в рамках модели?

2. Есть ли в школе специалисты по педагогическим измерениям, которые могут выступать как методисты по формирующей оценке?

3. Текущие планы управления школ сопровождаются или приспособлены под цели формирующей оценки? □

4. Наблюдается ли перемещение учащихся между классами одной параллели в соответствии с результатами мультиплекса?

5. Какие элементы/факторы школ помогают учащимся двигаться к индивидуальным целям?

6. Есть ли соответствующие ресурсы для выполнения целей мультиплекса?

В контексте суммативной (итоговой) оценки появляются следующие вопросы:

1. Является ли школа открытой для инициатив по мультиплексу?

2. Хотят ли участвовать в мультиплексе сами учащиеся и все образовательное сообщество?

3. Может ли школа продолжать участвовать в мультиплексе при существующих условиях функционирования?

4. Были ли в результате мультиплекса итоговые оценки непредсказуемыми и непреднамеренными? Что же является ценным и заслуживающим внимание при формировании итоговой оценки по мультиплексу?

Перед школой, участвующей в мультиплексе, ставится много вопросов:

1. Насколько эффективно участие у школы с недостаточным финансированием по сравнению со школой, имеющей больше средств для запуска мультиплекса?

2. Может ли сказаться нехватка помещений?

3. Является ли участие в мультиплексе полезным и эффективным для учащихся? Полезно ли оно для улучшения эмоционального и культурного фона конкретного учащегося при оценивающих процедурах мультиплекса?

4. Появились ли новые профессиональные возможности для учителей как в обучении и воспитании, так и в тестологии? Отразились ли результаты участия в мультиплексе на образовательные программы учителя?

5. Подсчитывается ли уровень образовательной деятельности учащегося на школьном уровне?

6. В какой степени осуществляется мультиплекс в школе?

7. Как воспринимают учащиеся и родители результаты мультиплекса?

8. Способствует ли климат школы, участвующей в мультиплексе, лучшему обучению и индивидуальному продвижению по мнению учащегося и родителей?

9. Может быть, такой благоприятствующий климат формируется совсем без участия в мультиплексе, а определяется только местоположением школы и селективным приемом в первый и другие классы?

10. Почему локальные органы управления не используют мультиплекс в других школах?

11. Какие административные препятствия со стороны органов управления были у школ, участвующих в мультиплексе?

12. Какие инновационные методы оценивания качества образования применялись в школе, которая участвовала в мультиплексе и могут ли они транслироваться в других школах?

13. До какой степени родители вовлечены в процесс оценивания результатов своих детей, и какие пути такого вовлечения предложены мультиплексом?

В дальнейшем школа может получить статус в мультиплексе в виде: 1) образцовая школа; 2) успешная школа с указанием показателей успешности, которые могут использовать другие школы для повышения качества образования; 3) школа вне мультиплекса с указанием того, что требует обязательного вмешательства с целью улучшения качества обучения.

Среди общих фаз эвалюации в мультиплексе были использованы три фазы: планирование, сбор данных и написание отчета.

Планирование и подготовка к участию осуществляется за 1-2 года до основного тестирования и состоит из развития инструментария, подготовки контрольно-измерительных материалов, апробации и предварительного верификационного тестирования.

Сбор данных осуществляется в соответствии с пошаговой моделью, включая модель разработки и модель проведения. Также включаются различные интервью и итоги встреч с участниками мультиплекса. Анализ данных для каждой школы ведется в рамках базовой части модели проведения, для некоторых школ – в рамках продвинутой модели проведения, для школ, систематически участвующих несколько лет в мультиплексе – в рамках мониторинговой модели, для научной общественности – в рамках полной модели.

Написание отчета состоит также из нескольких видов: первичного, вторичного, мониторингового, интерпретационного, прогностического, сравнительного и др. Первичный отчет предоставляется каждой школе после завершения сбора информации через месяц. Если завершение сбора информации задерживается более полугода от даты тестирования, остальные отчеты школа не получает. При написании отчета участвует координатор школы по мультиплексу и координатор самого мультиплекса, а также привлекаются наблюдатели, учителя и менеджеры мультиплекса.

Таким образом, мультиплекс является дружелюбным инструментарием измерения общего качества образования для обучающегося и является повышающим квалификацию в сфере оценочно-измерительной деятельности для обучающего.

Кроме того, каждый испытуемый доволен тестированием, так как он самостоятельно выстраивает стратегию выполнения теста. Стратегия выполнения оценивается как определенной этап зрелости в принятии решений. Он не обязан в отведенное время выполнить все задания, что невозможно в принципе в мультиплексе. А так как обучающиеся резко ограничены во времени, то практически пропадает проблема списывания. При этом успешные учащиеся стараются набрать больше баллов и выбрать выгодную для себя стратегию, а каждая минута отвлечения приводит к снижению результатов. Слабые учащиеся при очень низком уровне читательской компетентности в средней школе и вузе в состоянии за это время лишь прочитать и выбрать для себя по силам соответствующие задания. Такой подход в целом не годится для текущего оценивания, но для итогового очень полезен, так сокращает время тестирования: за одно время, не более трех часов, сразу формируются оценки по нескольким предметам. Выставление оценочных характеристик ведет не преподаватель оцениваемого предмета, что приводит к снижению субъективности оценок, нет любимчиков и заведомых нелюбимчиков. Поэтому расхождения в один балл по школьной шкале оценок между итоговым результатом по традиционному измерению и по мультиплексу бывает в верхней части шкалы, то есть любимчик может получить хорошую оценку вместо отличной (чаще у девочек), а нелюбимчик – наоборот из хорошей получает отличную

(чаще у мальчиков). Кроме того, результат мультиплекса близок к объективному результату и в ЕГЭ, и в международных исследованиях, то есть обладает прогностической функцией.

Результатом мультиплекса для обучающегося является индивидуальная траектория образования (от А до Ж), для образовательного учреждения – индивидуальные траектории групп или классов, параллелей или курса, факультетов или ступеней обучения, самого учреждения. А также приводятся градиенты, а не средние оценки для различных подгрупп мультиплекса.

Ниже представлены образцы индивидуальной траектории учащегося (рис.1), сертификат конкретного выпускника (рис. 2) и градиентные оценки по классам конкретной школы (рис. 3).

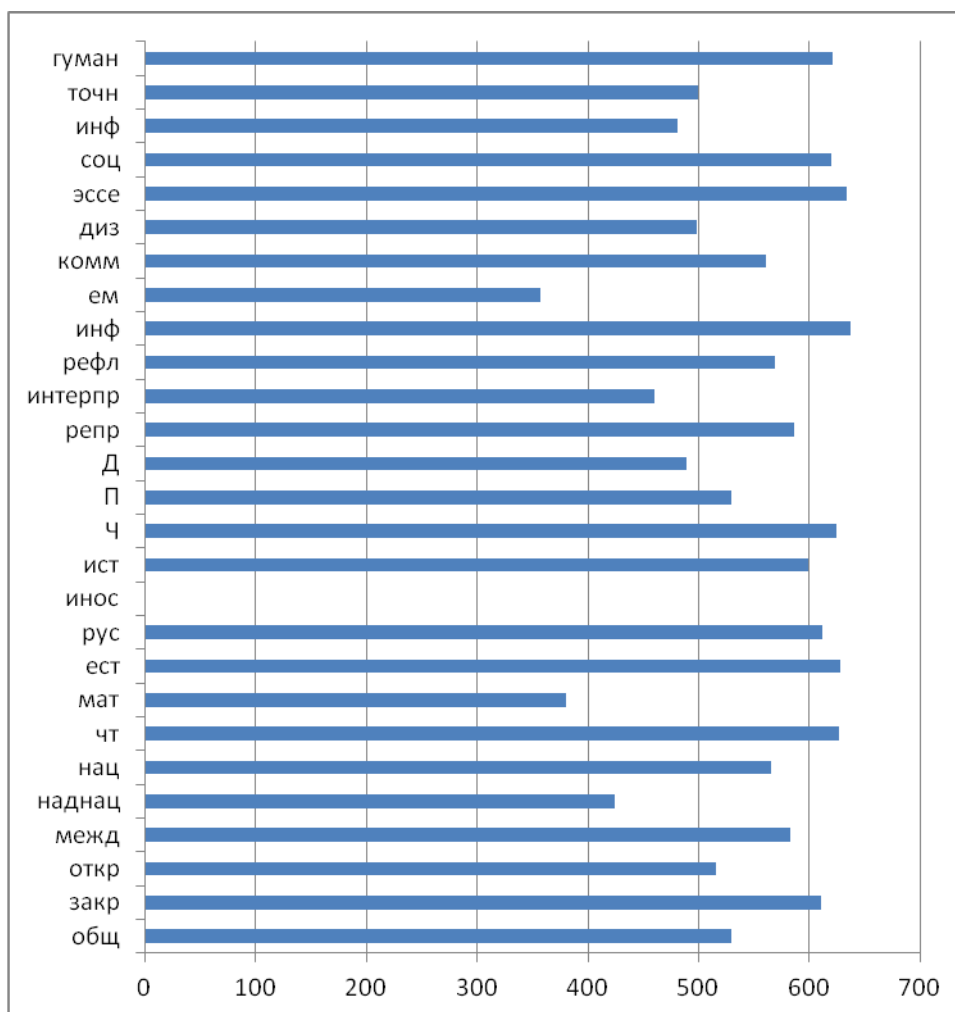


Рис. 1. Индивидуальная траектория образования

На рис. 1 представлены все критерии оценки качества образования мультиплексом, указанные выше от А до Ж, а также приводится шкала выполнения заданий в целом и по их формам предъявления: открытым и закрытым. Данный ученик построил стратегию выполнения теста таким образом, что исключил все задания из предметной области по иностранным языкам.

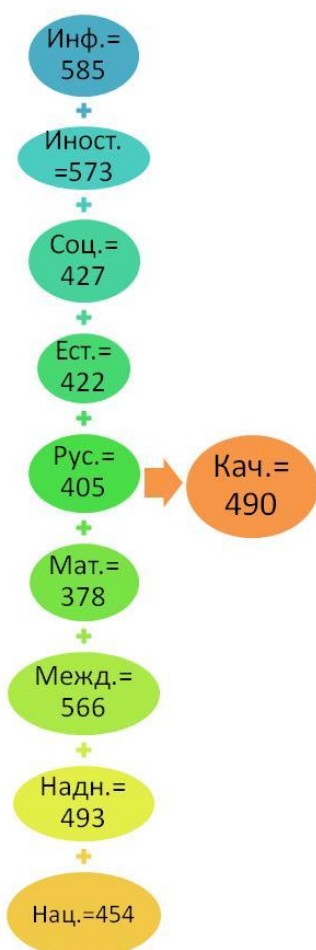


Рис. 2. Оценки мультиплекса в сертификате для выпускника школы.

На рисунке 2 приводится общий результат по качеству образования с детализацией по некоторым позициям мультиплекса: инструментальные оценки для дальнейшего прогноза, основные компетенции: математика, естествознание, русский и иностранный языки, информационная грамотность, предметы социального цикла. Общий результат – средний (500 плюс / минус 10 баллов). Очень высокий результат на международной шкале – 566 баллов. Крайне низкий результат по математике – 378 баллов. В целом информационная грамотность – высокая (585 баллов).

Из рис. 3 следует, что результаты шестых классов в конкретной школе среди 50% учащихся со средними результатами сильно различаются. В массовом градиенте не учитываются по 10% очень слабых и очень сильных учащихся, чьи результаты чаще обусловлены генетически. А также не учитываются по 15% слабых и сильных учащихся, характеризующих избирательность обучения, факторы социализации и др. Таким образом, массовый градиент показывает успешность обучения типичного ученика, что характеризует в наибольшей степени именно эффективность обучающей деятельности школьных педагогов. Улучшение прежде всего массового градиента в первую очередь приводит к повышению качества образования в таких странах, как Финляндия и Южная Корея.

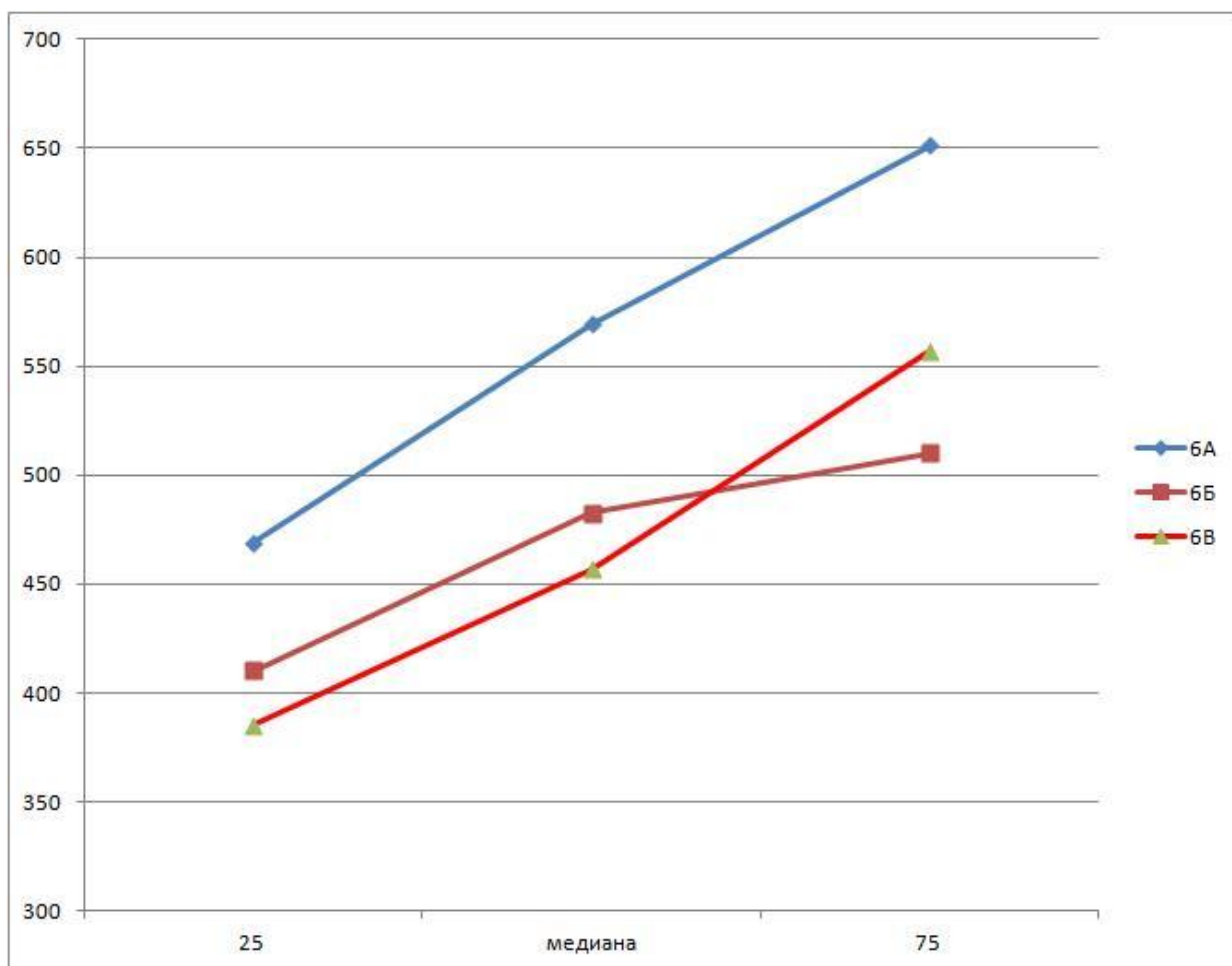


Рис. 3. Массовый градиент среди шестых классов

Данный метод измерения качества образования является авторским. По поводу реализации мультиплекса в образовательном учреждении следует обращаться в ИТИП РАО в лабораторию педагогических измерений. Электронный адрес лаборатории: edumeasure@rambler.ru. Электронный адрес ИТИП РАО: itiprao@mail.ru.