



Структурная модель оценки результатов обучения: основания, принципы, инструментарий

Нежнов П.Г.
Карданова Е.Ю.



Два основания оценки результатов обучения:

- *Измерение*
- *Категоризация*



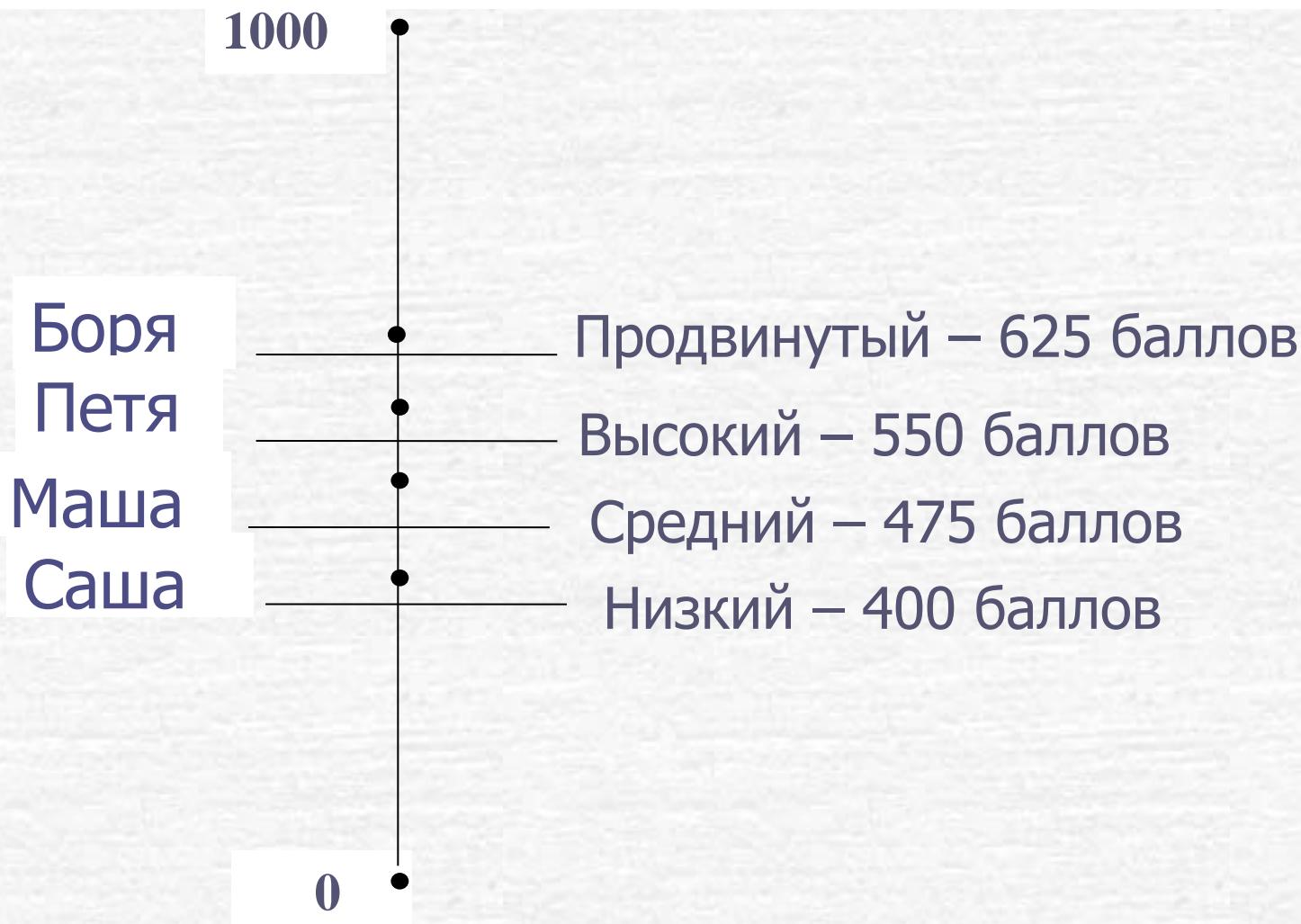
Приоритет измерения TIMSS, PISA ...



Результаты измерения на метрической шкале : сравнительная оценка



Разметка метрической шкалы (TIMSS) : уровневая оценка



Приоритет категоризации

«... Педагогика, занимаясь становлением знаний личности, не может интересоваться только следствиями усвоения, игнорируя процесс усвоения, т.е. самую деятельность усвоения, которая имеет свои этапы, а каждый этап – свой уровень усвоения знаний»

«... При любой организации обучения новому учебному материалу, учащийся проходит три этапа его усвоения. Соответственно можно говорить о трех уровнях усвоения»

Лернер И.Я.

Категоризация результатов обучения в рамках таксономии образовательных целей

	Перенос
Боря	Простейшие умения и навыки
Петя	Понимание
Маша	Запоминание
Саша	Различение

Средства категоризации: таксономии педагогических целей

Б.Блум	В.Симонов	И.Лернер	TIMSS (математика)
	Различение		
Знание	Запоминание	Знание	Знание
Понимание	Понимание		
Применение	Простейшие умения и навыки	Применение	Применение
Анализ	Перенос		Рассуждения
Синтез		Творческое применение	
Оценка			



Культурно-исторически зафиксированы
три целевых ориентира школьного
обучения:

- Повторить культурный образец действия
 - Применить культурный образец с пониманием
 - Функционально владеть образцом
- 

SAM:
баланс измерения и
категоризации

Модель функционального развития (по Л.С.Выготскому)

Функциональный уровень

Свободное действие - ориентация на поле и границы возможностей способа действия

Рефлексивный уровень

Действие с пониманием - ориентация на существенное отношение как основу способа действия

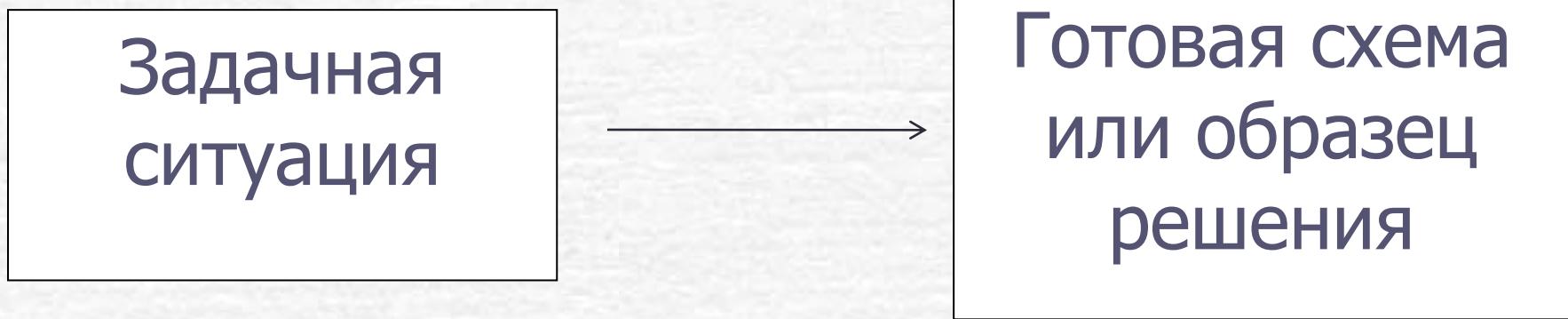
Формальный уровень

Действие по образцу - ориентация на его внешние характеристики

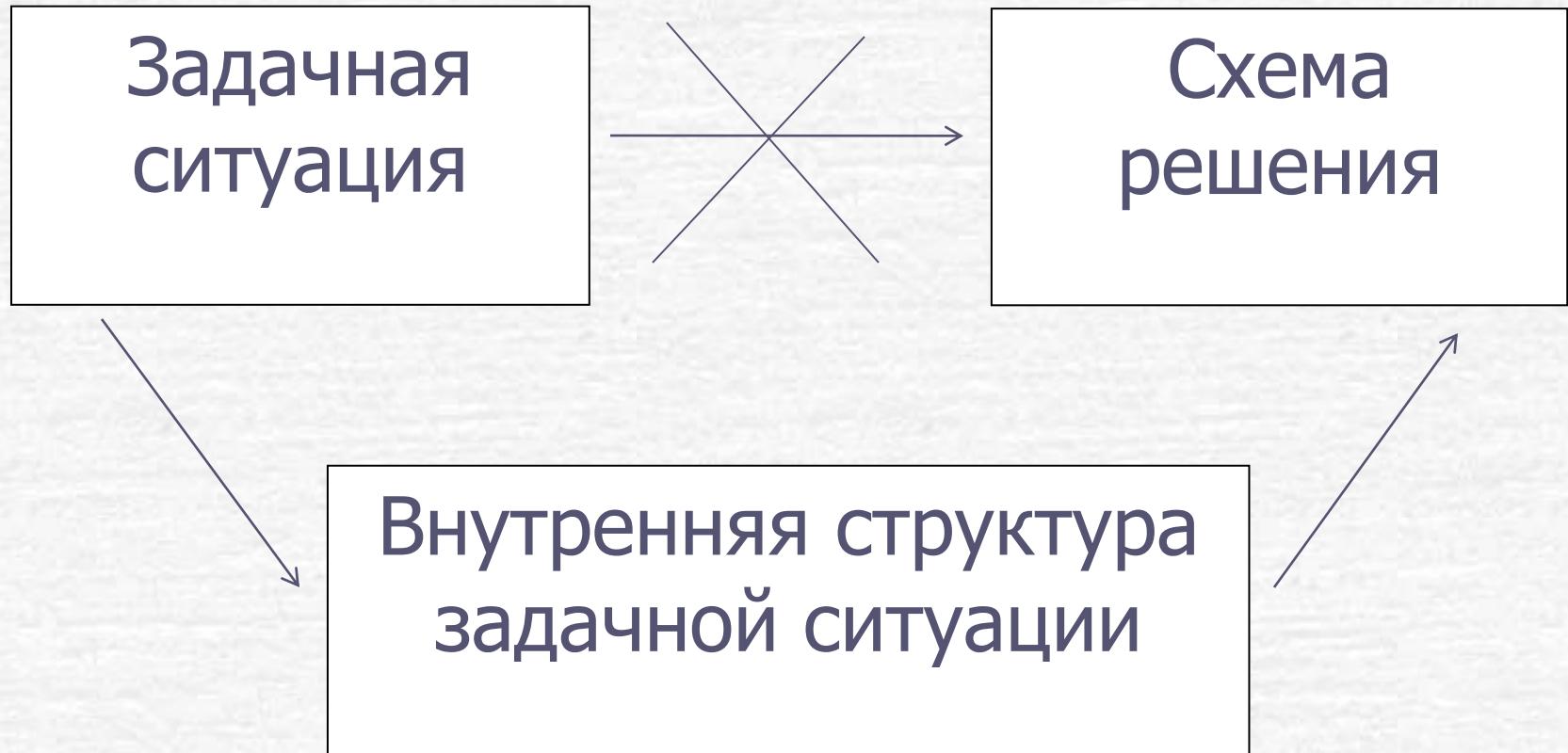
Пройденная учебная программа

Зона
ближайшего
развития

Действие формального уровня



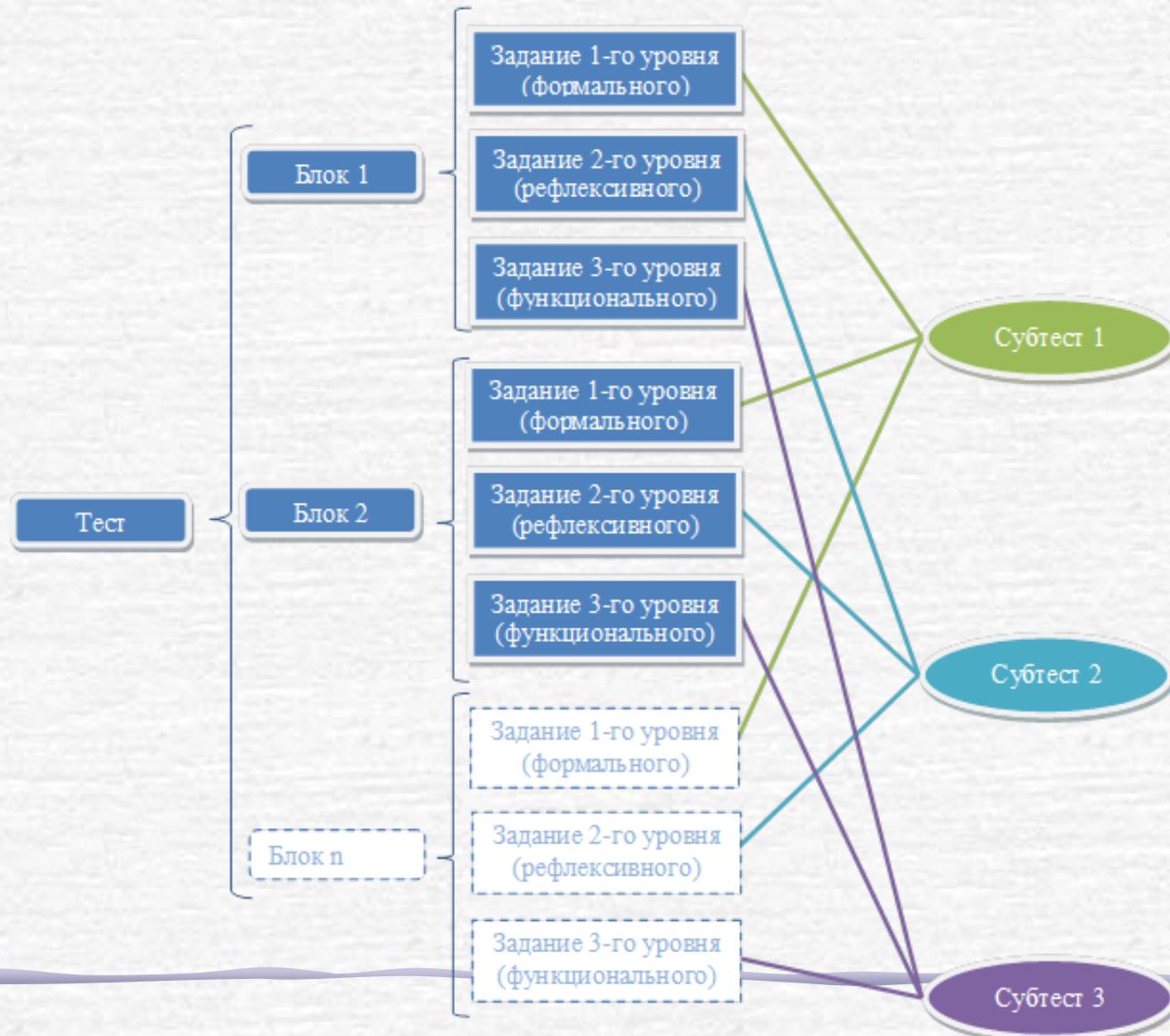
Действие рефлексивного уровня



Действие функционального уровня



Устройство тестов SAM

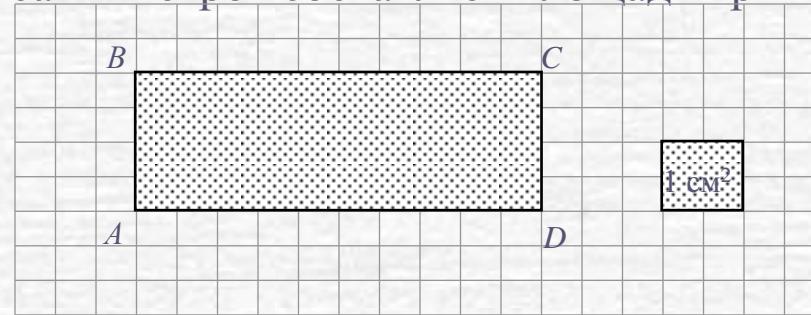


Содержание тестирования (математика/начальная школа)

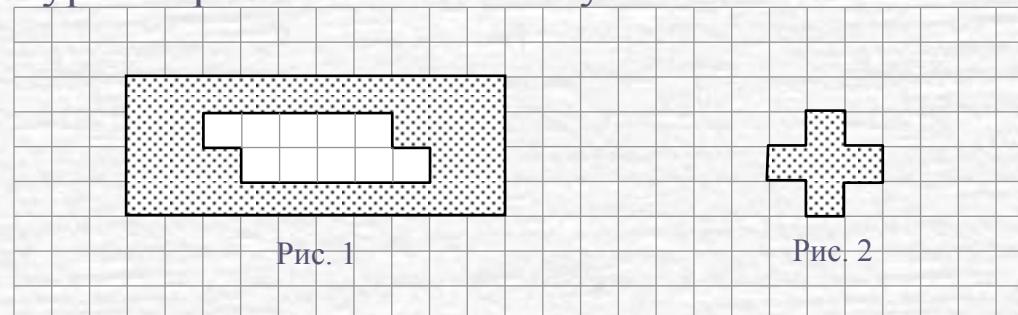
Разделы содержания	Средства ориентировки математических действий
Числа и вычисления	последовательность натуральных чисел; числовая прямая; позиционный принцип; свойства арифметических действий; порядок действий
Измерение величин	отношение между числом, величиной и единицей; отношение «целого и частей»; формула площади прямоугольника
Закономерности	«индукционный шаг»; повторяемость (периодичность)
Зависимости	отношения между однородными величинами (равенство, неравенство, кратности, разностное, «целого и частей»); прямая пропорциональная зависимость между величинами; производные величины: скорость, производительность труда и др.; соотношения между единицами
Элементы геометрии	форма и другие свойства фигур (основные виды геометрических фигур); пространственные отношения между фигурами; симметрия

Блок задач (измерение величин)

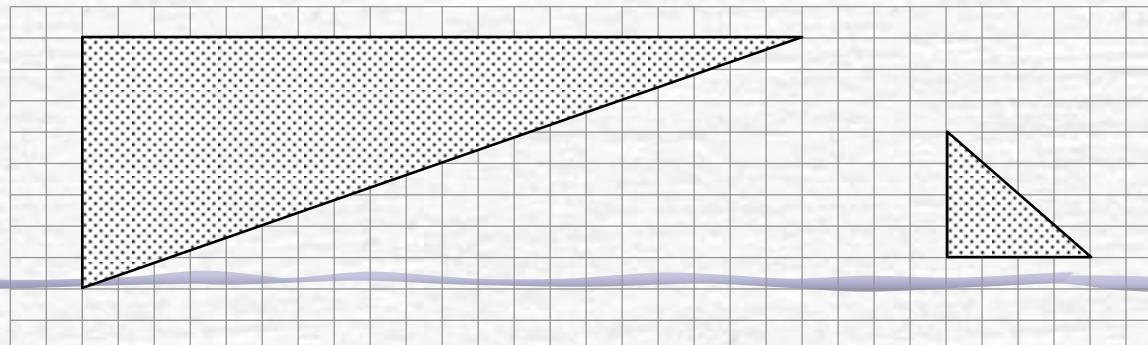
I. Сколько квадратных сантиметров составляет площадь прямоугольника ABCD?



II. Измерь площадь заштрихованной фигуры на рис.1, используя в качестве единицы измерения площадь фигуры на рис. 2. Запиши получившееся число.



III. Измерь площадь большого треугольника, используя в качестве единицы измерения площадь маленького треугольника. Запиши получившееся число.



Блок задач (числа и вычисления)

Какое получится число, если 10472 разделить на 34?

Ответ: _____

Рассеянный Петя, переписал из учебника пример на умножение двух чисел. Первый множитель он записал правильно: 7.

А во втором множителе по ошибке переставил местами цифры. Из-за этого в ответе у него получилось число 147.

Какой ответ должен был получить Петя, если бы он правильно переписал пример?

Ответ: _____

Какой самый большой результат может получиться, если в выражении $AB5 + BC2$ заменить буквы цифрами ?

(разные буквы заменяются разными цифрами)

Ответ: _____

Содержание тестирования (русский язык/начальная школа)

Содержание	Средства действия с языковым материалом (понятия, принципы, приемы, правила, алгоритмы, схемы...)
Слово, его звучание, значение и написание	
Звучание и написание (фонетика, графика и орфография)	правила графики; принципы русского письма; орфографические правила
Состав слова (морфемика и словообразование)	принципы и алгоритмы выделения частей слова; отношение производности; словообразовательные модели
Формы слова (морфология)	принципы выделения частей речи; парадигматические отношения (типы словоизменения: склонение, спряжение); грамматические категории (род, число, падеж, лицо, время)
Значение слова (лексика)	семантические отношения (синонимические, антонимические, родовидовые); структура значения слова (лексическое и грамматическое значение)
Высказывание и его оформление в письменной речи	
Предложение (синтаксис и пунктуация)	синтаксические отношения: сочинение, подчинение принципы и алгоритмы выделения грамматической основы предложения; принципы русской пунктуации

Блок задач (синтаксис и пунктуация)

Отметь предложение, в котором поставлены не все запятые.

1. Осенний листопад. Листья летят, скачут, плывут.
2. Мальчишки просидели в засаде до вечера, но ушли почти ни с чем.
3. Мама сидела за компьютером и писала какой-то доклад.
4. На массивном столе с зелёным сукном лежал ноутбук папки с бумагами, калькулятор.

Поставь знаки препинания в соответствии с описанными ситуациями.

1. Один ученик неверно расставил знаки препинания, и у него получилось, что море выбросило на берег предметы и людей.

Прибоем выбросило на берег корабль_ испанцев_ лодку_ рыбака_ катер.

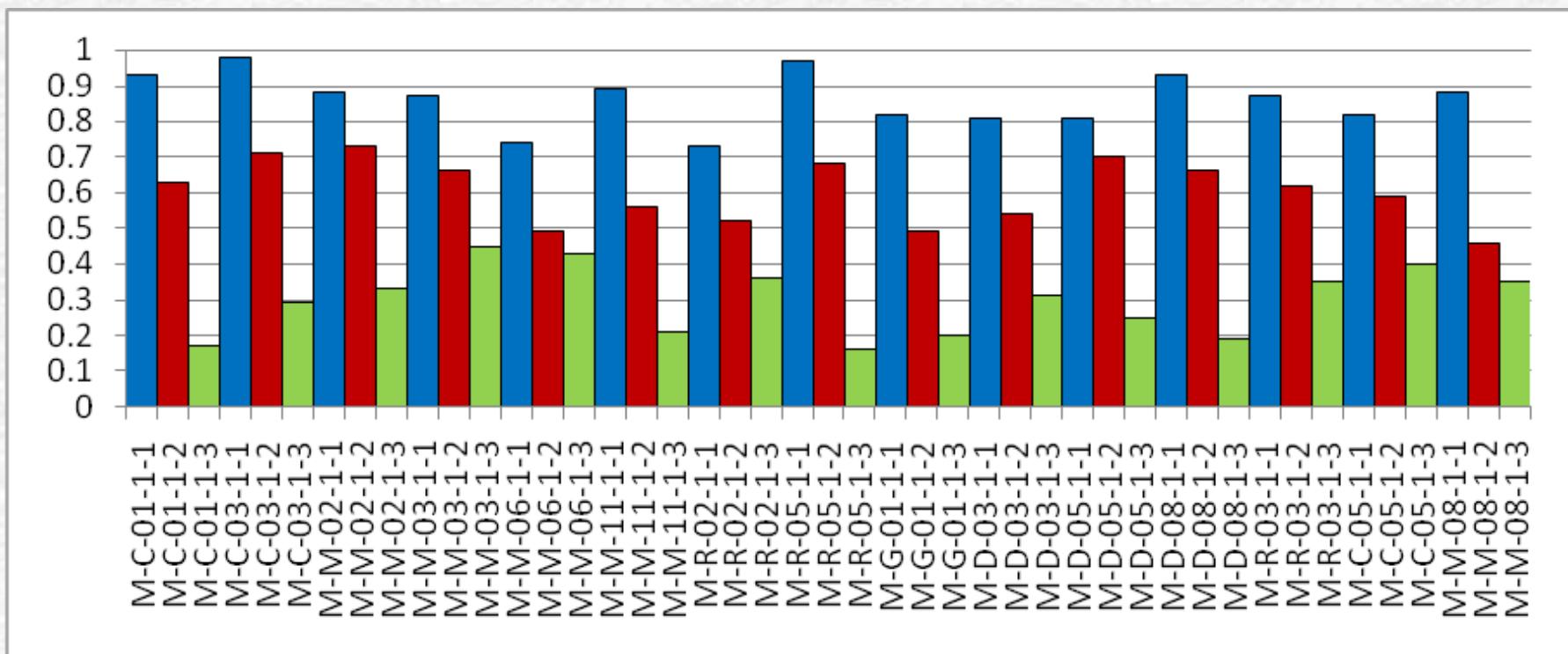
1. Другой ученик верно расставил знаки препинания, и у него получилось, что море выбросило на берег только предметы.

Прибоем выбросило на берег корабль_ испанцев _лодку_ рыбака_ катер.

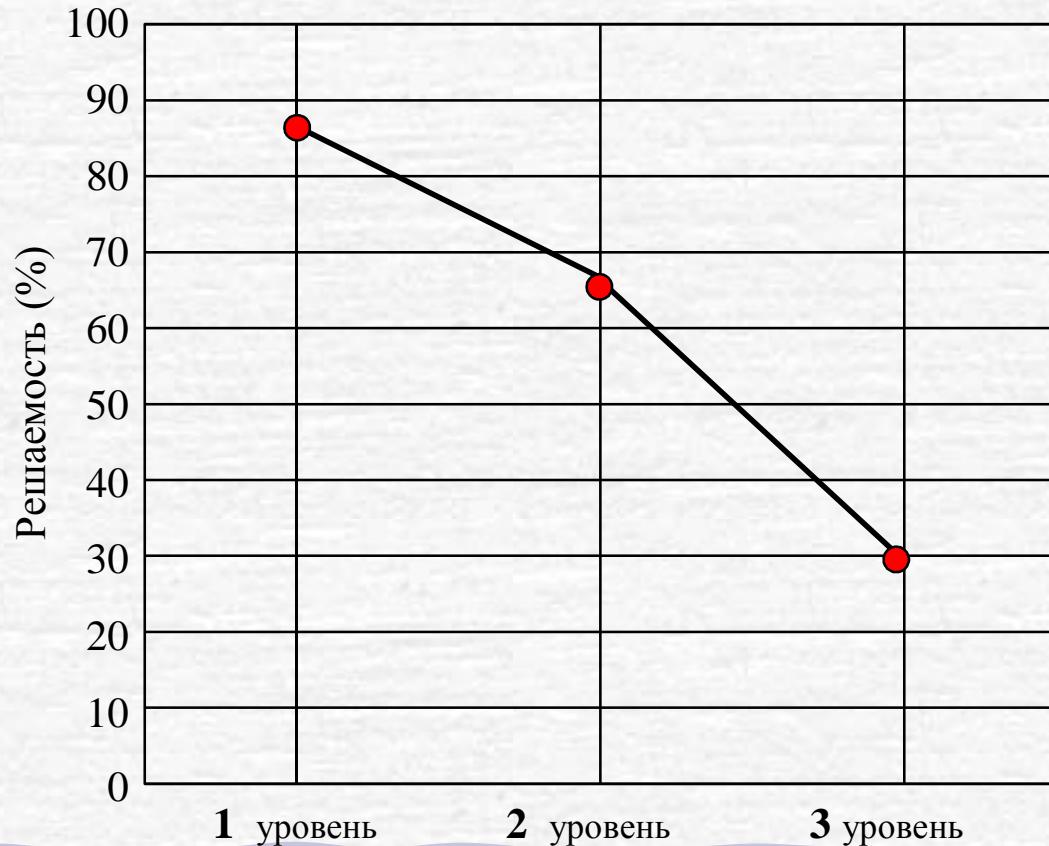
Переформулируй предложение так, чтобы смысл его сохранился, но оно стало простым с однородными членами.

Вода непрестанно трудится, и время от времени края горных уступов обрушиваются.

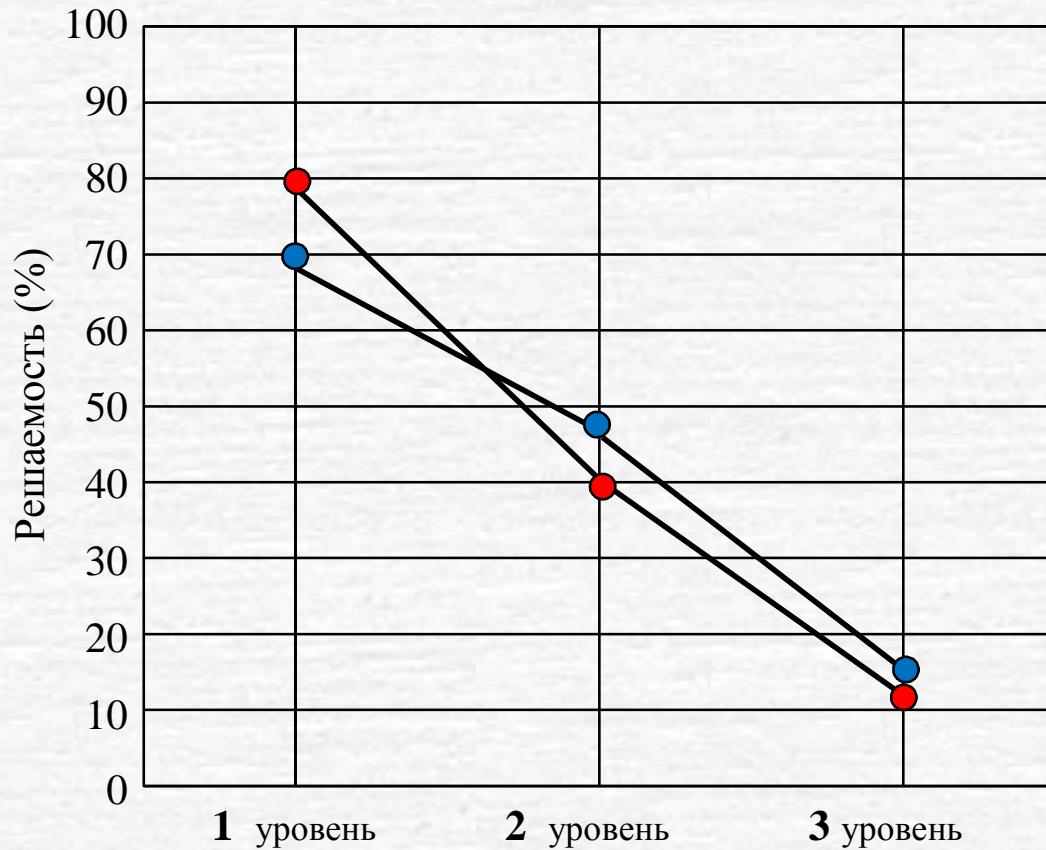
Распределение коэффициентов решаемости (математика)



Индивидуальный профиль достижений учащегося по некоторому предмету



Профили выполнения теста по математике учащимися 4-го класса



Первая ступень: учащийся решает не менее 50% заданий 1-го уровня

Вторая ступень: учащийся решает не менее 50% заданий 2-го уровня

Третья ступень: учащийся решает не менее 50% заданий 3-го уровня

**Пороги ступеней на 1000-балльной шкале
(математика)**

Ступень 1 – 430 баллов

Ступень 2 – 500 баллов

Ступень 3 – 570 баллов

Шкала с разметкой уровней по модели SAM



Реализация модели SAM для начальной школы

- В рамках модели SAM разработаны тесты по двум предметам – математике и русскому языку – имеющие общую структуру
- Тесты рассчитаны на выпускников начальной школы и могут выполняться учащимися 4-го и 5-го классов.
- Структурной единицей теста является блок из трех заданий (1-го, 2-го и 3-го уровня), соответствующий одному разделу предметного содержания. Задачи предъявляются блоками.
- Для оценивания заданий используется дихотомическая оценка: за правильный ответ ученик получает 1 балл, за неправильный ответ (или отсутствие ответа) – 0 баллов.

Валидизация инструмента SAM

- Теоретическая рамка: голландская система СОТАН (Evers, A., 2001; 2010)
- Исследование валидности SAM проводилось в течение 2011-2013 гг. во время пилотных тестирований SAM в различных регионах РФ

Evers, A. (2001). The revised Dutch rating system for test quality. *International Journal of Testing*, 1:2, 155–182.

Evers, A., Sijtsma, K., Lucassen, W. and Meijer, R. R. (2010). The Dutch Review Process for Evaluating the Quality of Psychological Tests: History, Procedure, and Results. *The International Journal of Testing*, 10:4, 295 — 317.

Схема исследования валидности SAM



- Показатель валидности по содержанию характеризует меру адекватности тестовых материалов проверяемой компетенции
- Показатель валидности конструкта устанавливает меру согласованности результатов тестирования с психологической моделью тестируемой способности

Полномасштабная апробация

Май 2012

Объем выборки:
4406 учеников
4-го класса*

* вся генеральная совокупность учащихся 4-ых
классов региона

Новгородская область



Количество учащихся	4406
Количество школ	189
Количество классов	297
Количество населенных пунктов	134

Психометрический анализ (КТТ)

(апробация 2012 г., математика, бланковая форма)

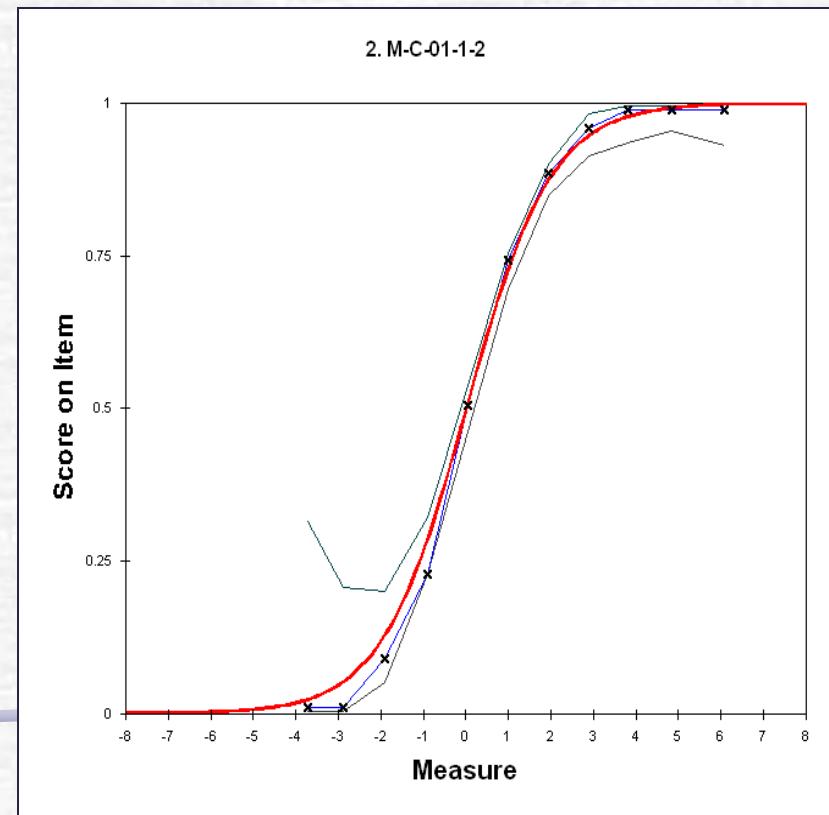
	Вариант 1	Вариант 2
Число испытуемых	3018	2941
Средний первичный балл	26	27
Стандартное отклонение	8.37	8.55
Средний коэффициент решаемости	0.59	0.61
Средний индекс дискриминативности	0.44	0.46
Средний скор. коэф. точ.-бис. корреляции	0.39	0.39
Коэффициент надежности (KR20)	0.90	0.91
Стандартная ошибка измерения	2.61	2.61

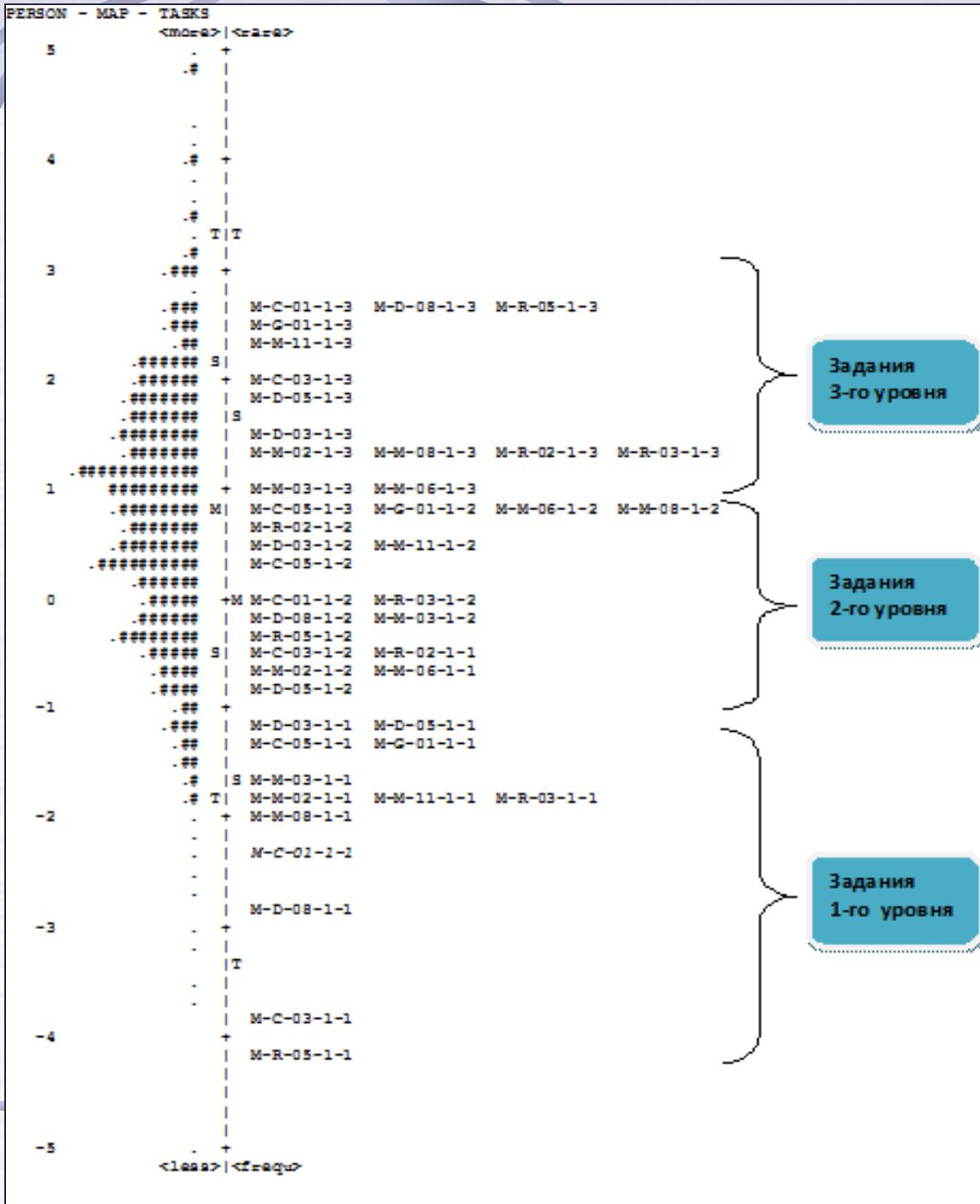
- все задания имеют удовлетворительные показатели
- коэффициенты решаемости заданий находятся в промежутке от 0,16 до 0,98

Психометрический анализ (IRT)

- Тесты - существенно одномерные (измеряют один конструкт)
- Тесты имеют оптимальный уровень трудности для данной совокупности испытуемых

Задания демонстрируют хорошие психометрические показатели и хорошее согласие с используемой моделью измерения (в качестве модели измерения рассматривалась дихотомическая модель Г. Раша)





Анализ в IRT:

Карта переменных (математика, вариант 1)

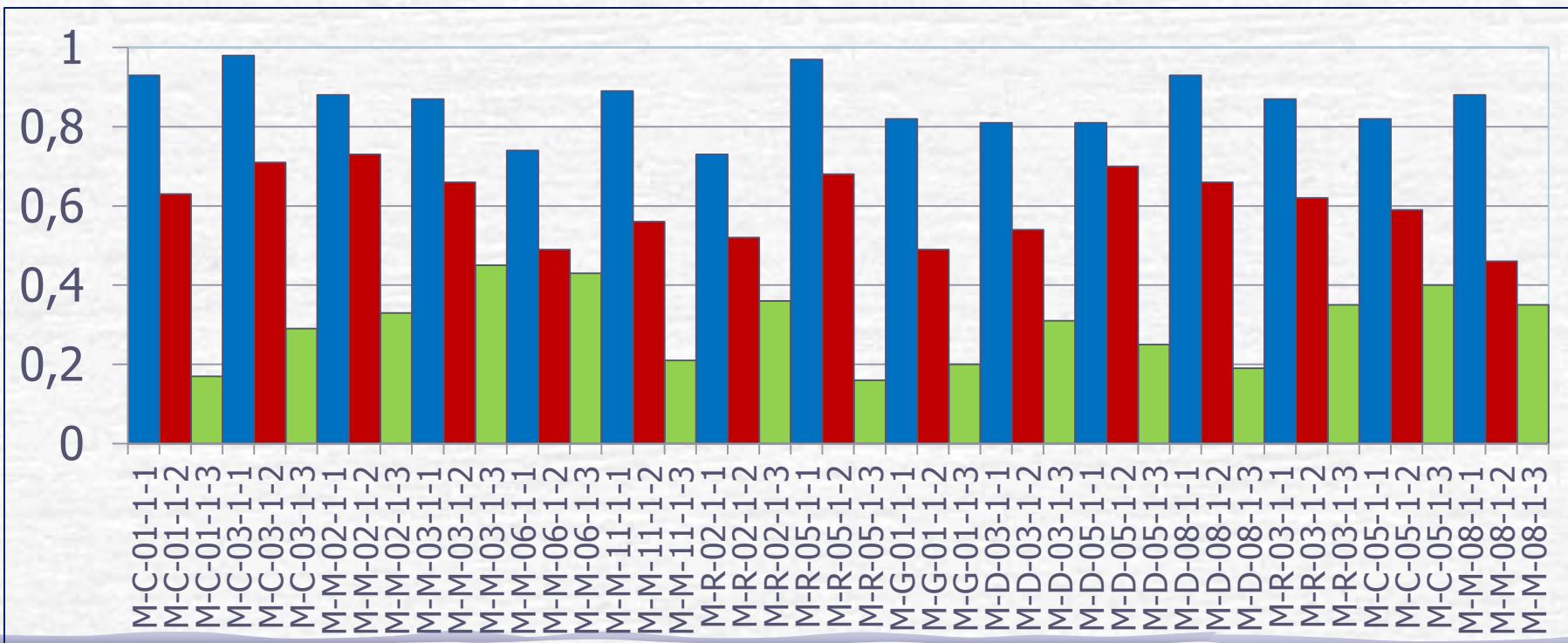
Валидность конструкта

- В модели SAM основным конструктом, подлежащим верификации, является трехуровневая таксономия освоения учебного содержания, вписанная в возрастной контекст.
- Этот конструкт связан как минимум с двумя гипотетическими утверждениями, требующими проверки:
 - 1) Задания трех уровней, относящиеся к одному блоку и отвечающие теоретически заданным критериям трех уровней, должны обнаруживать соответствующую иерархию по трудности.
 - 2) В начальной школе предметный материал в норме может и должен усваиваться на втором (рефлексивном) уровне, т.е. на уровне понимания. Освоение этого материала на третьем (функциональном) уровне может и должно происходить в рамках основной школы.

Гипотеза 1:

Задания трех уровней, относящиеся к одному блоку и отвечающие теоретически заданным критериям трех уровней, должны обнаруживать соответствующую иерархию по трудности.

Распределение коэффициентов решаемости (Математика, вариант1)



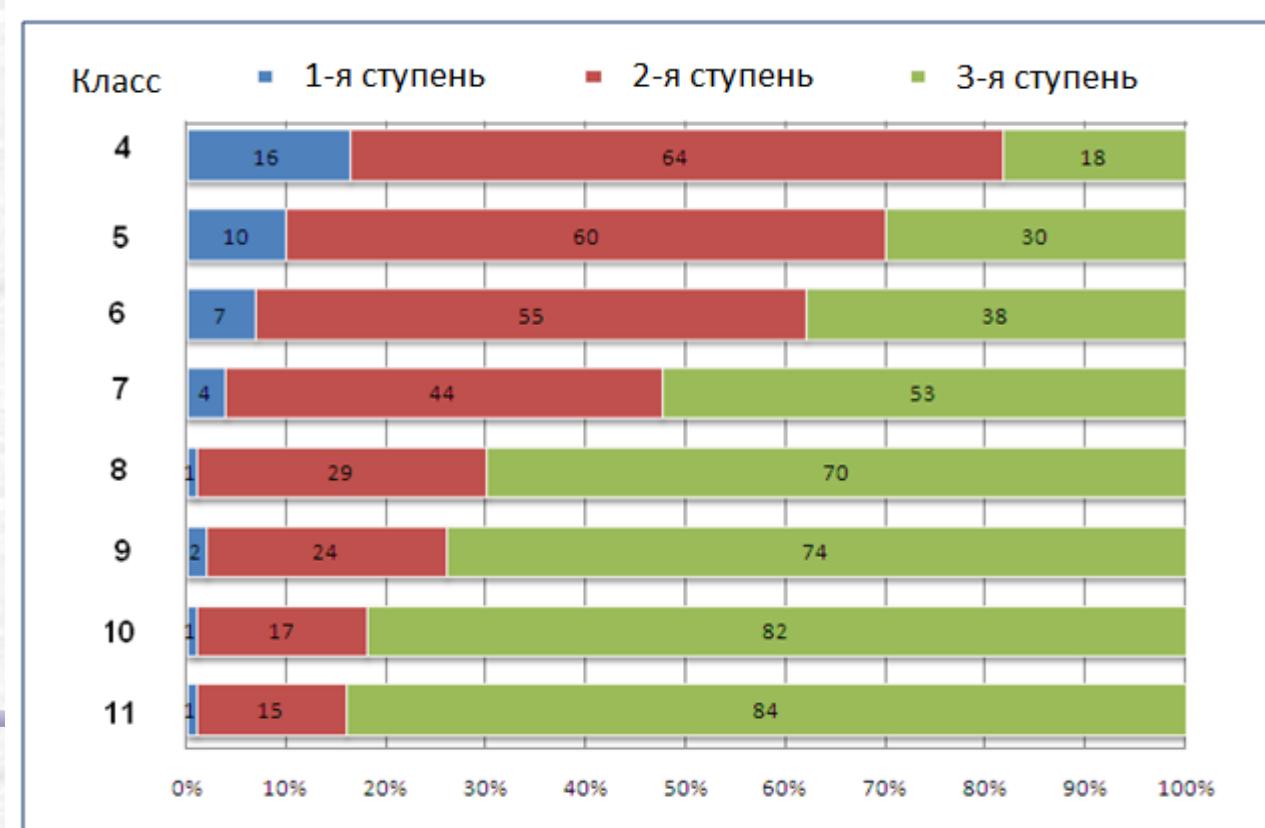
Проверка второй гипотезы

- Специальное исследование в 2011-2012 гг.
- Дизайн исследования: в 2011 г. тесты по математике и русскому языку были предложены учащимся 4-х возрастных групп – 4-х, 6-х, 8-х и 10-х классов. Через год, в 2012 г. те же тесты были предложены тем же учащимся, которые в тот момент были учащимися 5-х, 7-х, 9-х и 11-х классов.
- Тестирование проходило весной, в конце учебного года.
- Выборка составила около 100 человек в каждой параллели.

Гипотеза 2:

В начальной школе при хорошем обучении предметный материал усваивается преимущественно на втором (рефлексивном) уровне, т.е. на уровне понимания. Освоение этого материала на третьем (функциональном) уровне происходит в рамках основной школы.

Распределение учащихся различных классов по ступеням достижений



Критериальная валидность

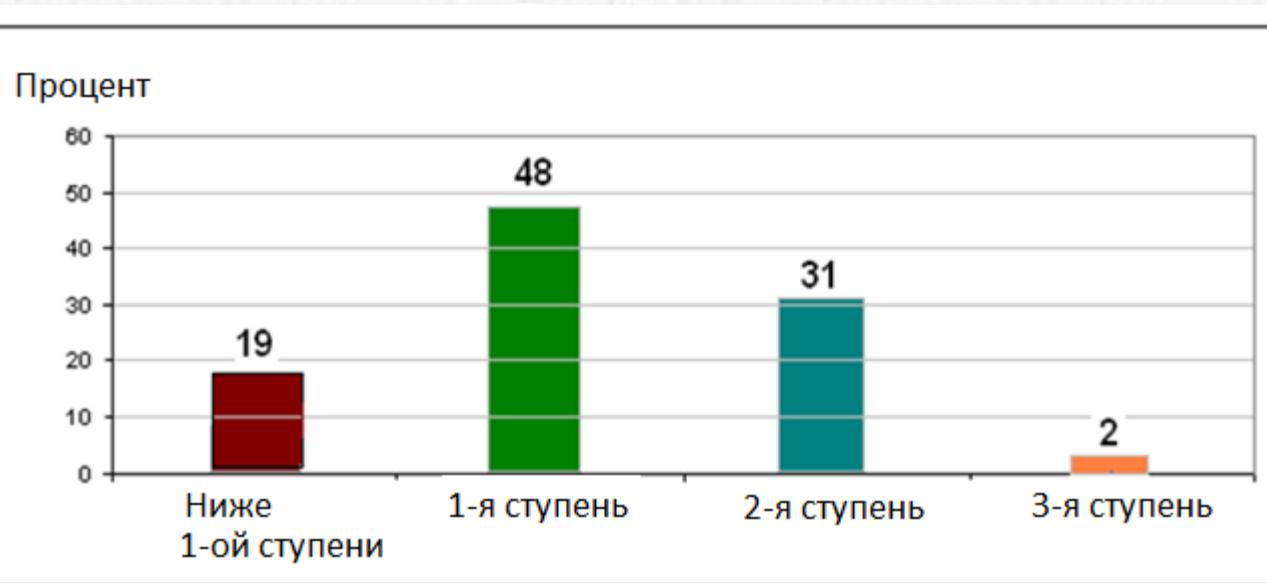
Текущая
валидность

Прогностическая
валидность

Показатель критериальной валидности фиксирует меру согласованности результатов тестирования с известными объективными критериями (в настоящем или в будущем).

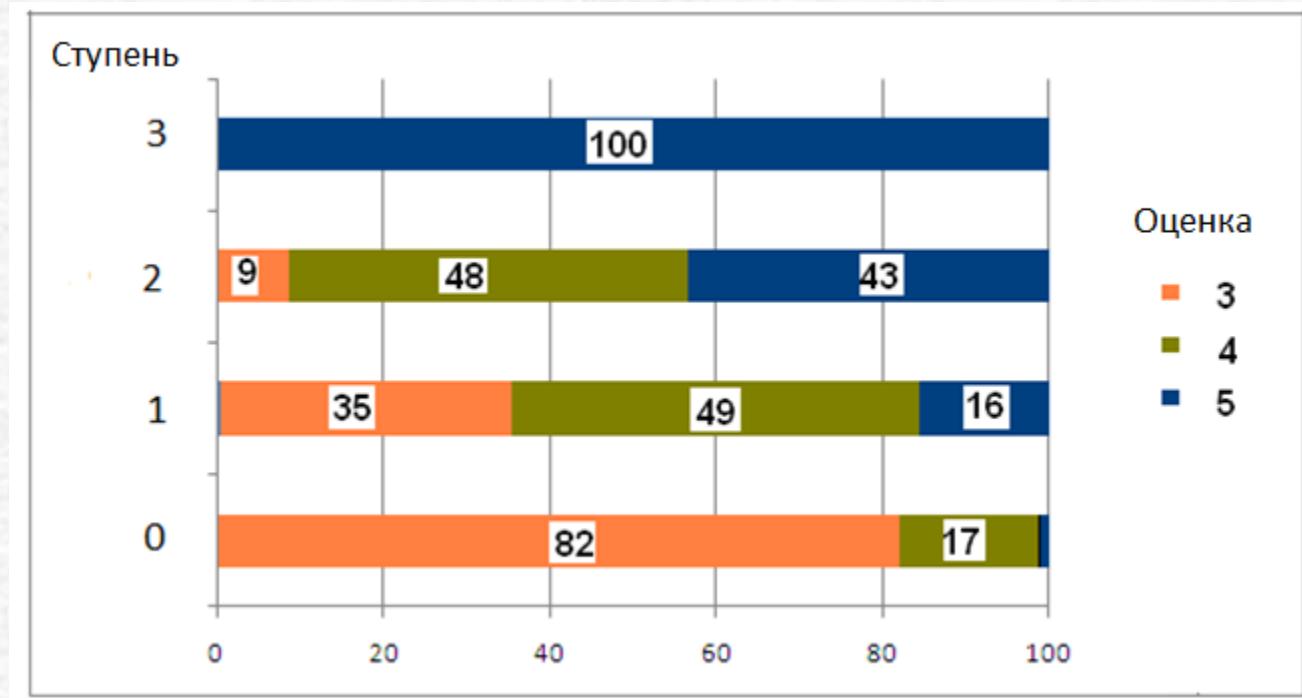
дизайн исследования

- Исследование проводилось на базе данных аprobации тестов SAM в Красноярском крае весной 2011 г.
- Общий объем выборки: 941 учащийся из 12 школ
- Были собраны оценки по математике этих же учащихся в 5-м классе (т.е. через год после проведения тестирования).



Распределение участников тестирования по ступеням достижений (математика)

Распределение оценок учащихся в зависимости от ступени достижений (математика)



Корреляция между тестовым баллом учащихся и их школьными оценками равна 0.6.

Осень 2013 г.

Рецензенты:

1. **Howard T. Everson** (Center for Advanced Study in Education, Graduate School, City University of New York, Professor of Psychology and Senior Research Fellow)
2. **Clancy Blair** (New York University, Steinhardt School of Culture, Education, and Human Development, Professor of Applied Psychology)
3. **Bas Hemker** (Netherlands, Cito National Institute for Test Development, Senior Research Scientist)

Модель оценивания предметных компетенций учащихся на основе SAM

- В качестве модели тестирования используется модель Раша
- Тестовые баллы сообщаются на 1000-балльной шкале со средним около 500 и стандартным отклонением 50
- Тестовые баллы всех участников тестирования находятся на единой метрической шкале, независимо от времени прохождения теста и конкретного набора выполненных заданий.

Подходы к интерпретации результатов



Нормативно-ориентированный

- Результат отдельного участника интерпретируется в зависимости от достижений всей совокупности участников тестирования
- Каждому участнику в результате математической обработки результатов присваивается тестовый балл
- Тестовые баллы всех участников тестирования находятся на единой шкале, независимо от времени прохождения теста и конкретного набора выполненных заданий
- Устанавливаются нормы

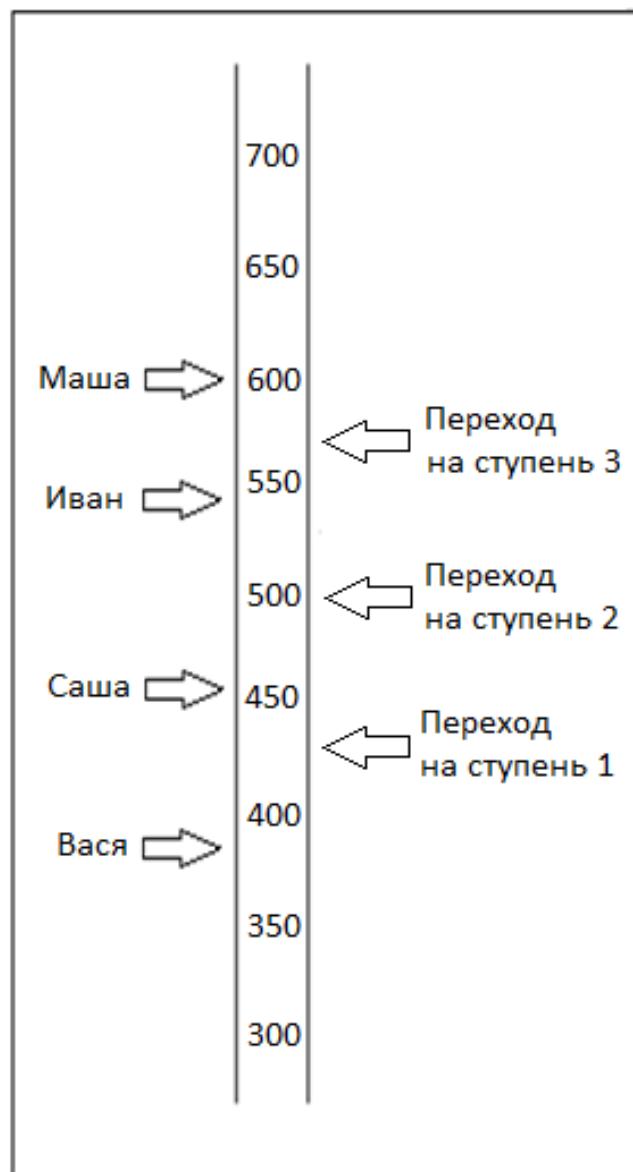


Нормативные ориентиры оценки результатов тестирования

- Среднее значение по генеральной выборке
- Социокультурная норма (среднее значение по выборке школ лидеров)
- Абсолютная культурная норма (стопроцентное выполнение теста)
- Возрастная норма конца начальной школы
(рефлексивный уровень по SAM)



Подходы к интерпретации результатов



Критериально-ориентированный

- ❑ Обеспечивает возможность качественной оценки усвоения предметного содержания
- ❑ Разработан ступенчатый вариант шкалы достижений, основанный на тестовых баллах участников тестирования и пороговых значениях, делящих всех участников на группы, соответствующие различным качественным ступеням достижений

Оценивание испытуемых



Интерпретация пороговых оценок

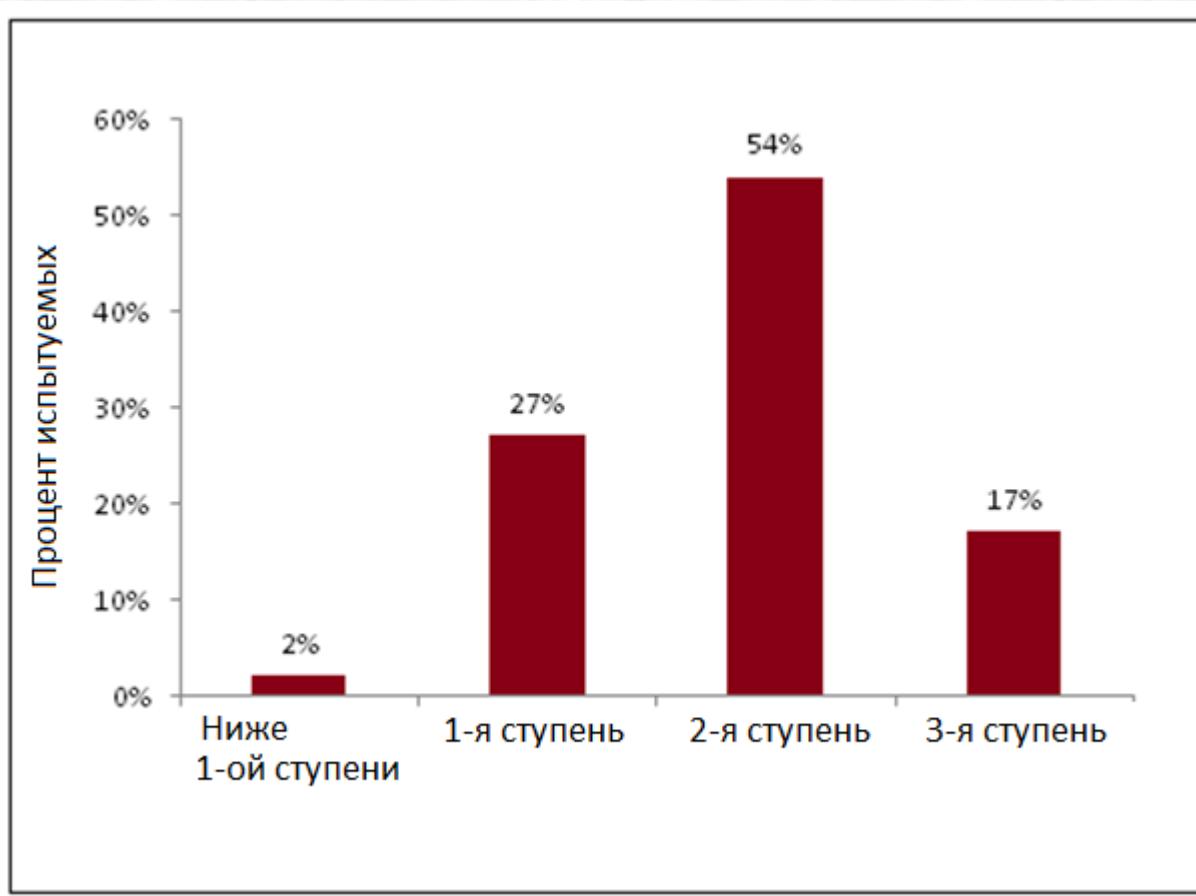
Пороги:

570 (граница между
2 и 3 ступенями)

500 (граница между 1 и 2 ступенями)

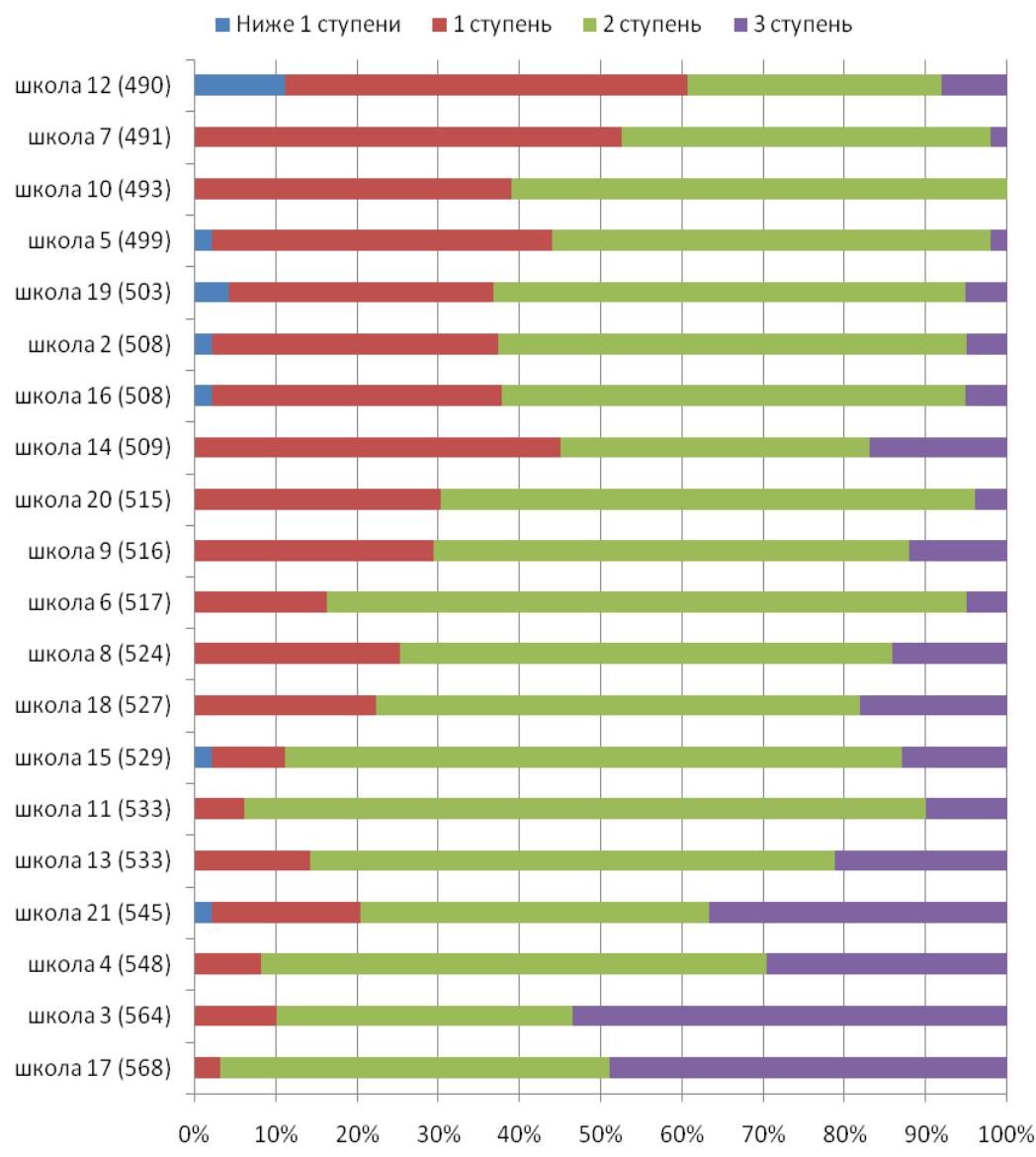
430 (граница между 0 и 1 ступенями)

Распределение участников тестирования по ступеням достижений (математика)



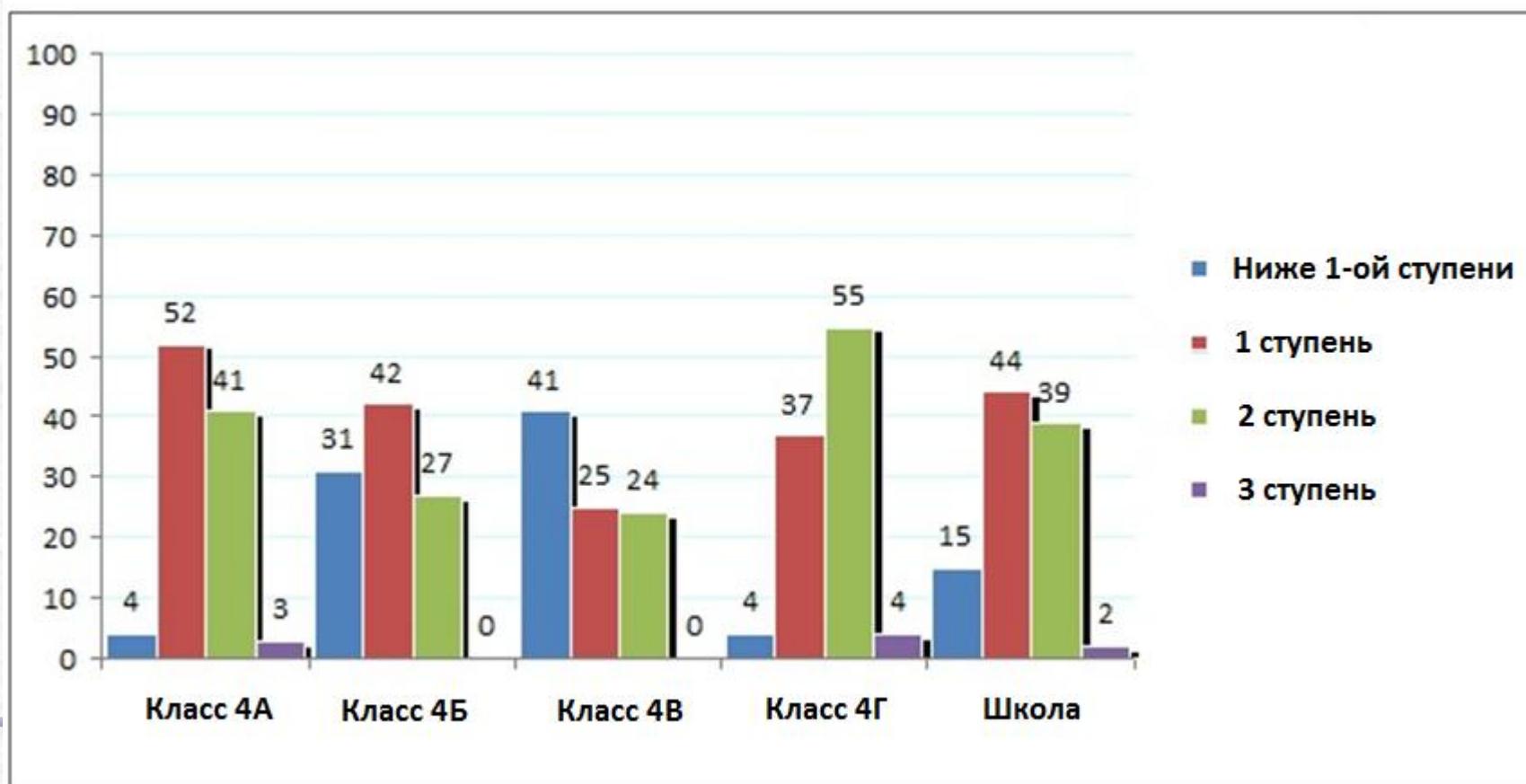
Апробация
2012г.
(выборка около
5000 человек)

Распределение учащихся разных школ региона по ступеням достижений (математика)



Школы упорядочены по возрастанию среднего значения общих тестовых баллов.

Распределение по ступеням достижений учащихся различных классов внутри одной школы (математика)



Проблемы, которые SAM помогает решить:

- Общее понимание образовательных целей учителями, администрацией, учащимися
- Выявление Зоны Ближайшего Развития учащихся для выбора индивидуализированной педагогической стратегии
- Преодоление затруднений в освоении умений высокого порядка (нацеленность на освоение ключевых понятий)
- Мотивация обучения
- Повышение квалификации учителей
- Оценка академической успешности
- Самооценка успешности работы учителями

SAM (Student Achievement Monitoring)

1. Диагностика учебной успешности в начальной школе // Ред. П.Г.Нежнов, И.Д.Фрумин, Б.И.Хасан, Б.Д.Эльконин. М.: ОИРО, 2009.
2. *Нежнов П.Г., Карданова Е.Ю., Эльконин Б.Д.* Оценка результатов школьного образования: структурный подход // «Вопросы образования», 2011, №1. с. 26-43.
3. *Нежнов П.Г., Карданова Е.Ю., Рябинина Л.А.* Исследование процесса присвоения учебного содержания «Вопросы образования», 2013, № 4, с. 168-187.
4. *Nezhnov P.* SAM – Toolkit to assess primary students' academic achievements / CADMO, anno XIX, 1, 2011.
5. *Nezhnov P., Kardanova E., Vasilyeva M., Ludlow L.* Operationalizing Levels of Academic Mastery Based on Vygotsky's Theory: The Study of Mathematical Knowledge / Educational and Psychological Measurement 2015, Vol. 75(2) 235–259



Спасибо за внимание!

