

# Применение моделей когнитивной диагностики в мониторинговых исследованиях школьного образования

Саляхутдинова Д.Р., Федерякин Д.А.

Центр психометрики и измерений в образовании

Института образования НИУ ВШЭ

# Критика традиционных мониторингов

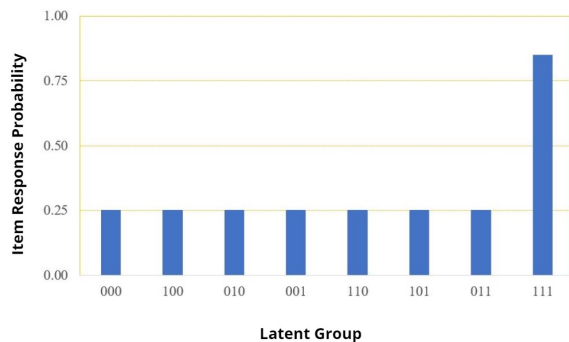
- Ранжирует респондентов – сравнивает их друг с другом
  - То же самое для групп респондентов – не дает прямой информации о гетерогенности групп
- Требуем выстраивания одномерной\* шкалы, что чрезвычайно трудно достичь при больших банках заданий
- Не дает респондентам информации о том, как именно им развиваться дальше
  - То же самое для групп респондентов – администраторы и практики слабо представляют себе, какие шаги нужно принимать непосредственно по результатам мониторинга для улучшения ситуации

# Модели когнитивной диагностики

- Вместо *ранжирования* респондентов направлены на их классификацию (*ранкирование*)
- Декомпозируют единый тестовый балл на отдельные суб-компетенции\* (von Davier, Lee, 2019)
- Облегчают интерпретацию результатов
- Облегчают агрегирование индивидуальных результатов на уровень групп респондентов

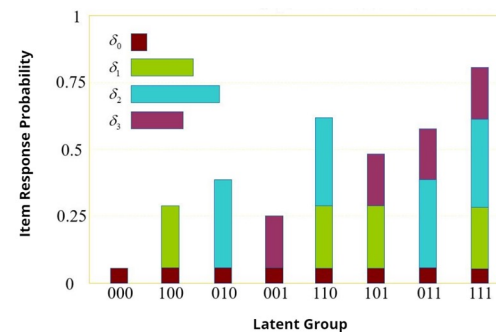
Задание	Сложение	Умножение	Порядок действий
3+5*4	1	1	1
4+7	1	0	0
3*6	0	1	0
4*8+3	1	1	0

Результаты	Сложение	Умножение	Порядок действий
Освоение компетенции	1	0	1
Вероятность освоения	.83	.34	.56



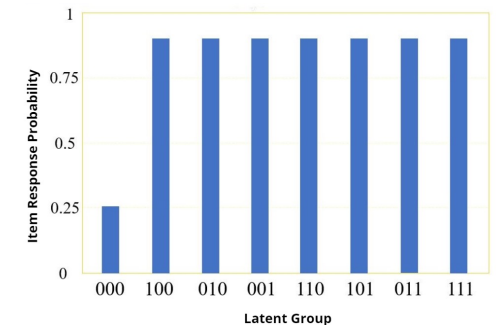
DINA model

(Macready Dayton, 1977;  
Junker, Sijtsma, 2001)



ACDM/RRUM

(de la Torre, 2011)



DINO model

(Macready Dayton, 1977;  
Junker, Sijtsma, 2001)

# Почему нельзя использовать наблюдаемый тестовый балл?

- Каждый балл за задание учитывается столько раз, сколько суб-компетенций оно нагружает, поэтому происходит переоценка способностей
- Суб-компетенции становятся более скоррелированными, чем они есть на самом деле (Wilson, de Boeck, 2004)
- Надежность, не позволяющая делать хоть сколько-нибудь достоверные выводы (He, 2009; Wilson, Gochyyev, 2020)

# Требования для CDM моделей

- Суб-компетенций должно быть меньше, чем заданий (или порогов категорий в политомическом случае)
- Четкая однозначная таксономия заданий, выраженная в спецификации теста, и транслированная в дихотомическую Q-матрицу
- Каждое задание должно проверять несколько суб-компетенций
- Каждая суб-компетенция должна проверяться несколькими заданиями
- Профили нагрузок суб-компетенций на задания не должны повторяться

# Тест и выборка

- Тест:

- 25 заданий по математике для учащихся 5-х классов: 15 заданий базового уровня, 10 повышенного
- Все задания теста разработаны в соответствии с ФГОС
- Форматы заданий: с выбором ответа и кратким ответом
- Использовался в начале учебного года в рамках исследования «Исследование готовности общеобразовательных организаций субъектов Российской Федерации к реализации технологий смешанного обучения»

- Выборка:

- $N = 7439$
- Школьники 4 регионов России
- Нерепрезентативная для регионов выборка

# Используемые Q-матрицы: области содержания

Области Содержания	Количество заданий
Числа и вычисления	6
Текстовые задачи	7
Элементы статистики и теории вероятностей	6
Наглядная геометрия	6
Знание/понимание	2
Алгоритмы	6
Решение задач	10
Практическое применение	7
Базовый уровень	15
Повышенный уровень	10

# Используемые Q-матрицы: когнитивные уровни

Уровни освоения	Количество заданий
Базовый	15
Повышенный	10

Уровень задания	Базовый	Повышенный
1	1	0
2	<b>0</b>	<b>1</b>

Уровень задания	Базовый	Повышенный
1	1	0
2	<b>1</b>	<b>1</b>

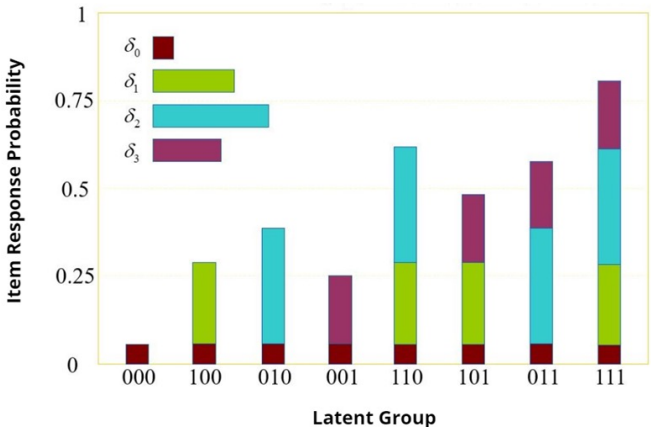


# Сравнение моделей

	Не-вложенные когнитивные уровни			Вложенные когнитивные уровни		
	DINO	DINA	ACDM	DINO	DINA	ACDM
AIC	192471.0	192155.6	188750.5	192774.4	192171.2	188545.2
BIC	193203.0	192887.6	189827.7	193506.3	192903.1	189691.5
MADcor	0.038	0.037	0.012	0.039	0.037	0.010
SRMSR	0.046	0.045	0.016	0.049	0.045	0.013
MADQ3	0.025	0.026	0.032	0.029	0.026	0.031
MADaQ3	0.024	0.026	0.023	0.028	0.026	0.020

Все задания хорошо согласуются с моделью (у всех заданий  $RMSD < 0.03$ , средний  $RMSD = 0.017$ )

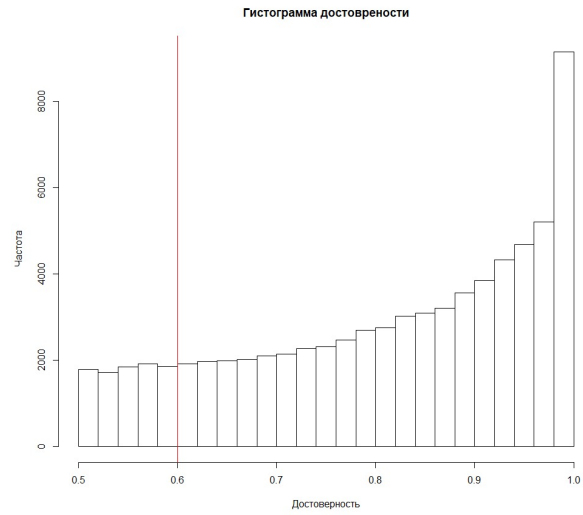
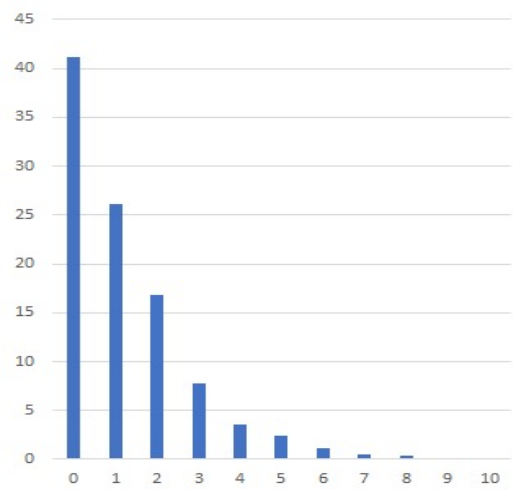
ACDM с вложенными  
когнитивными  
уровнями



Уровень задания	Базовый	Повышенный
1	1	0
2	1	1

# Надежность классификации

- 12.35% всех выводов сделаны неоднозначно (вероятность освоения от 0.4 до 0.6)
- Средняя надежность классификации 0.811

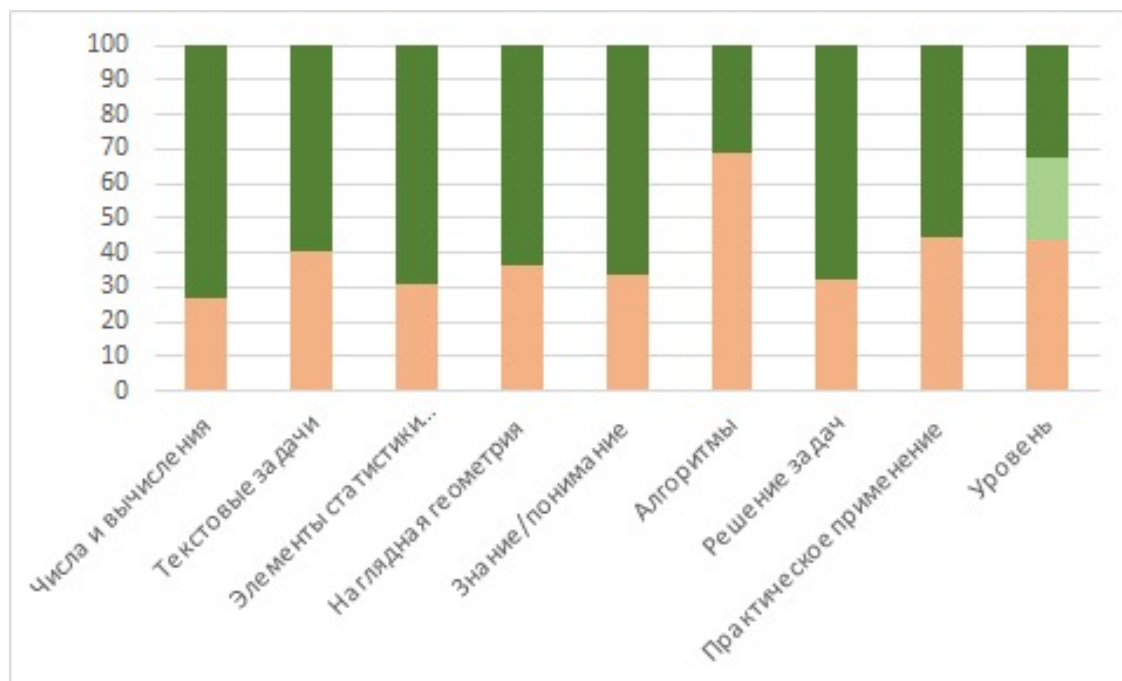


Компетенция	Достоверно классифицированные респонденты (%)	Точность (%)	Состоятельность (%)
Числа и вычисления	88.89	78.95	73.86
Текстовые задачи	86.62	78.01	68.62
Элементы статистики и теории вероятностей	89.02	80.53	73.53
Наглядная геометрия	86.85	78.31	71.66
Знание/понимание	81.01	74.94	68.09
Алгоритмы	82.27	74.43	67.01
Решение задач	91.21	82.91	75.21
Практическое применение	84.72	76.26	66.32
Базовый уровень	94.60	90.32	84.59
Повышенный уровень	91.26	84.91	76.90

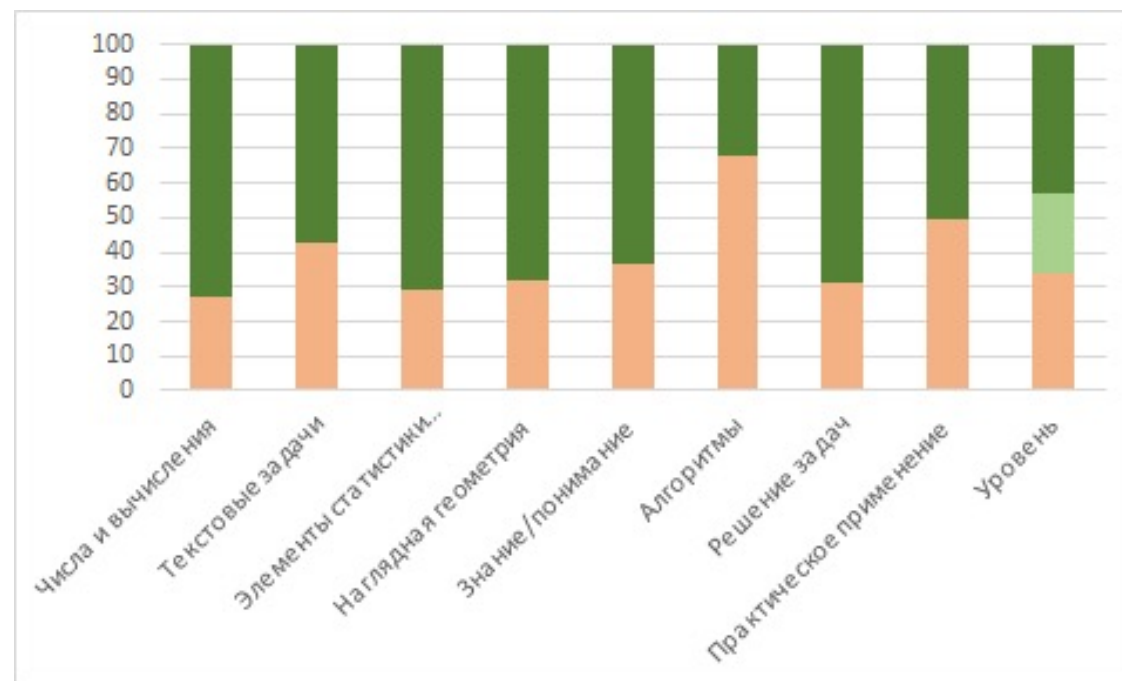
# Индивидуальная обратная связь

		Числа и вычисления	Текстовые задачи	Элементы стат истики и теории вероят ностей	Наглядная гео метрия	Знание/поним ание	Алгоритмы	Решение задач	Практическое применение	Базовый уровень	Повышенный у ровень
Ученик 1	Освоенность	1	1	1	1	1	1	1	1	1	1
	Вероятность освоения(%)	72.48	77.25	71.63	85.79	77.24	98.41	83.36	91.69	100	74.7
Ученик 2	Освоенность	1	1	0	0	1	1	0	1	0	0
	Вероятность освоения(%)	76.93	89.6	14.65	11.4	86.37	73.31	12.64	95.74	3.95	1.73
Ученик 3	Освоенность	0	1	0	0	0	1	0	1	0	0
	Вероятность освоения (%)	17.83	55.84	22.27	17.55	36.16	78.14	5.54	56.13	0.09	0.05
Сколько знают (%)		31.75	40.15	30.74	35.91	41.83	64.36	32.85	51.22	59.76	36.63

# Групповая обратная связь



Регион 1



Регион 2

# Заключение

- Модели когнитивной диагностики
  - Позволяют декомпозировать общий тестовый балл на суб-компетенции, что является основой для доказательного формирующего оценивания
  - Позволяют проводить ранжирование групп респондентов
  - Дают респондентам информацию о том, какие компетенции им необходимо развивать дальше
  - Повышают надежность и достоверность выводов, что важно для дальнейшего формирования образовательной политики
  - Легки в использовании результатов, диагностике качества и интерпретации параметров
  - Напрямую связывают теоретические представления о структуре измеряемого конструкта с данными

# Спасибо за внимание!

Диана Расимовна Саляхутдинова: [dsalyahutdinova@hse.ru](mailto:dsalyahutdinova@hse.ru)

Денис Александрович Федерякин: [dafederiakin@hse.ru](mailto:dafederiakin@hse.ru)

Центр психометрики и измерений в образовании  
Института образования НИУ ВШЭ