



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Современные подходы к оцениванию в образовании



**Карданова Елена Юрьевна**

Директор Центра психометрики и измерений в образовании  
Институт образования  
Национальный исследовательский университет  
«Высшая школа экономики»

**25 апреля 2021 г.**

# Центр психометрики и измерений в образовании НИУ ВШЭ

- Центр компетенций в области психометрики и разработки инструментов нового поколения, нацеленных на измерение сложных конструктов с использованием комплексных поведенческих индикаторов
- Основная функция - разработка инновационных инструментов оценивания в образовании: сложные конструкты / новые формы заданий / новые технологии тестирования / сложное психометрическое моделирование

Подготовка специалистов по измерениям и психометрике: магистерская программа «Измерения в психологии и образовании», PhD программа «Educational Measurement and Evaluation», ежегодная международная летняя школа «Applied Psychometrics in Psychology and Education»





# Центр психометрики и измерений в образовании НИУ ВШЭ



## Штат 27 сотрудников, из них:

- ✓ 50% - выпускники магистерской программы «Измерения в психологии и образовании»
- ✓ 7 аспирантов
- ✓ 2 постдока
- ✓ 13 – преподаватели МП и аспирантуры

## + 14 стажеров:

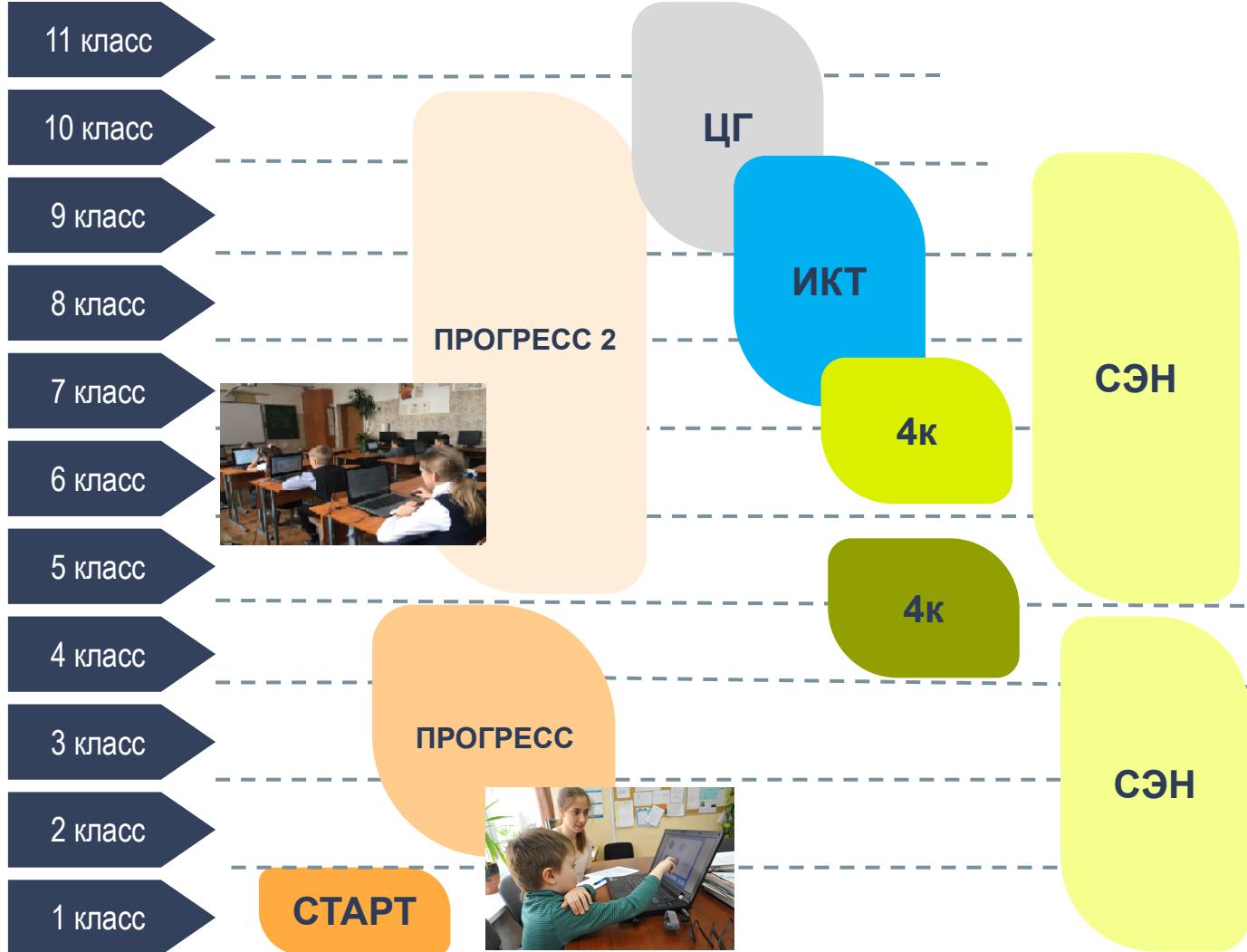
- ✓ студенты магистерских программ института
- ✓ студенты с ярмарки проектов вышки

# Современные тенденции в оценке качества образования

- Переход от разовых срезов к измерению прогресса → лонгитюдные исследования – отслеживание динамики, изменений, тенденций
- Изменение предмета оценки: метапредметные и личностные результаты обучения; компетенции → усложнение конструктов и психометрических моделей тестирования (многомерность)
- Изменение технологий оценки → компьютерное и компьютерное адаптивное тестирование
- Использование технологий машинного обучения → автоматический скоринг заданий в цифровых инструментах
- Проведение вторичных исследований, изучение факторов → оценка не ради оценки
- Повышение требований к инструментарию оценивания

# Исследовательская повестка центра

- ✓ Какими инструментами можно измерять новые грамотности и компетенции? Как можно оценить метапредметные и личностные результаты обучения учащихся?
- ✓ Какие психометрические модели и методы необходимо использовать для современных инструментов оценивания и исследований в области качества образования? В частности, для измерения индивидуального прогресса, роста.
- ✓ Как использовать коллатеральную информацию, получаемую из цифровой среды оценивания и других источников, для справедливого, точного и несмещенного оценивания прогресса учащихся?
- ✓ Какие психометрические модели позволяют улучшить надежность, точность, справедливость оценивания, а также помогают в интерпретации результатов?
- ✓ Каков вклад социально-эмоциональных навыков учащихся в образовательные и другие жизненные достижения, а также в их благополучие?
- ✓ Какова динамика развития ребенка в начальной школе? Как связаны когнитивные предикторы с достижениями в математике и чтении в начальной школе?
- ✓ Каков уровень сформированности цифровой грамотности у граждан страны? Готовы ли они к жизни в условиях цифровой экономики?



Центр  
психометрики и  
измерений в  
образовании  
НИУ ВШЭ:  
инструменты  
независимой  
оценки в  
образовании

# Инструмент СТАРТ

- Индивидуальное оценивание**
  - И наблюдение, и работа с ребенком
  - Комфортный для ребенка режим процедуры оценивания
- Оцениваются базовые когнитивные и социально-эмоциональные навыки**
  - Словарный запас
  - Фонематическая грамотность
  - Базовые навыки в математике и чтении
  - Адаптация к школе, личностное развитие, социальное развитие



- Компьютерное адаптивное тестирование**
  - Подстраивание алгоритма под ребенка
  - Посторонние факторы меньше сказываются на результатах тестирования (утомление, беспокойство, неаккуратность)
  - Эффективное время тестирования

# СТАРТ: Когнитивная батарея

Комплекс тестов для оценки когнитивных навыков, которые являются важными предикторами успешности в изучении математики

1. **Тест на несимволическое чувство числа** - способность оценить количество объектов, не прибегая к их точному подсчету
2. **Цифровой тест Струпа** - оценка когнитивного контроля.
3. **Тест “Цветок и Сердце”** - оценка произвольного контроля и особенностей переключения внимания.
4. **Тест на кратковременную и рабочую память.**
5. **Тест на навыки транскодинга** - оценка способности переводить числовую информацию между разными символическими форматами (цифры и числовые слова).

# ПРОГРЕСС

2,3,4 классы

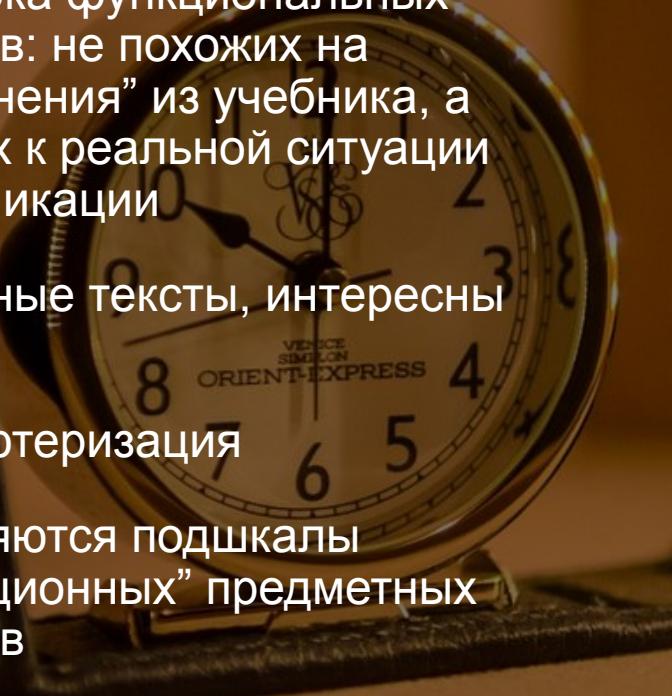


- Инструментальное, психометрическое продолжение СТАРТ на середину и конец начальной школы
- Измерение базовых грамотностей
- Компьютерное адаптивное тестирование
- Единая платформа на всю начальную школу



# Языковая грамотность

- Проверка функциональных навыков: не похожих на “упражнения” из учебника, а близких к реальной ситуации коммуникации
- Сюжетные тексты, интересны детям
- Компьютеризация
- Выделяются подшкалы “традиционных” предметных навыков



Анна сомневается в выборе некоторых слов. Помоги ей определить, какие слова лучше употребить.

День 3

Сегодня мы с Артёмом ходили **в** мастер-класс по стрелянию на стрельбищу из лука. **на** стрельбе **про** Стрельбище

Это **олимпиадный** инструкции мы узнали, что в Средние олимпийский инструктажа века лучники стреляли на 90-100 метров. Мы бегаем на физкультуре олимпийский инструктора **вид** спорта. От **до скольки** стреляли лучники когда-то.

стометровку, легко представить, **как** далеко стреляли лучники когда-то. **до каких пор** **лучше** Конечно, лучники могли стрелять и вдвоем **далее** на 200 метров, но **дальше**

# Оценивание «сложных конструктов» в образовании

- Универсальные компетентности, мягкие навыки, навыки 21 века - непредметные компетентности, необходимые человеку для успешной адаптации в современном мире
- Включены в ФГОС всех уровней образования
- Это сложные (комплексные) латентные конструкты
  - ✓ Стандартные инструменты оценивания не подходят для оценивания таких сложных конструктов.
  - ✓ Чтобы их оценить, необходимо провести наблюдение за тем, как учащиеся принимают решения и действуют в сложных ситуациях в реальной жизни.
- Нужны инструменты на основе заданий сценарного типа, игр, симуляций.

Примеры инструментов измерения сложных конструктов (Центр психометрики и измерений в образовании НИУ ВШЭ):

- Для уровня школьного образования: информационно-коммуникационная компетентность, “4К” (креативность, критическое мышление, коммуникация и коопeração), цифровая грамотность.
- Для уровня высшего образования: инструмент оценивания критического мышления студентов вузов.

# Пример: измерение метапредметных результатов

Освоить «способы решения проблем творческого и поискового характера» и овладеть «логическими действиями сравнения, анализа, синтеза, обобщения, классификации»

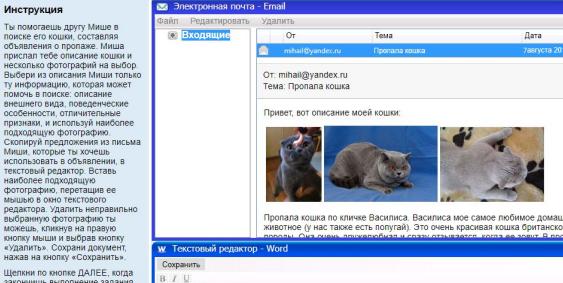
ФГОС: <https://fgos.ru/>



- Это сложные комплексные конструкты
- Традиционные инструменты: использование экспертов, наблюдение
- Инструмент измерения сложных конструктов должен:
  - позволять ученику **поведенчески** проявлять сложную структуру навыка
  - приближать ситуацию тестирования к реальной жизни (DiCerbo, 2014)
  - мотивировать на честное и увлеченное прохождение тестирования (Sundre & Wise, 2003)
  - учитывать процесс решения задания (Davier et al., 2019)
- Новый формат: сценарные задания с фокусом на процесс и продукт работы в компьютерной форме (computerized performance-based) (Bennett, 1999).

# Как измерить сложные конструкты?

- Инструмент измерения — компьютерные задания сценарного типа: тестируемый погружается в подготовленный сценарий и решает ряд задач.
- Методология — метод доказательной аргументации (ECD, Evidence-Centered Design; Mislevy, Steinberg, & Almond, 2003): выстраивание системы аргументации от тех действий, которые совершают тестируемый в процессе решения заданий, к выводу о конструкте.
- Чтобы тестируемый мог **проявить** навыки, задания делаются “**интерактивными**”.
- Каждый сценарий обеспечивает сбор наблюдаемых действий
- **Действия ложны: а) быть проявлением навыка; б) встраиваться в логику сценария**



# Мониторинг компетенций 4К

- Измерение 4 компетенций в конце начальной школы и середине основной школы:
  - 1) Критическое мышление
  - 2) Креативность
  - 3) Коммуникация
  - 4) Кооперация
- Мониторинг разработан на основе методологии Evidence-Centered design
- Компьютерная форма, реализация через интернет**
- Задания-сценарии игрового типа**
- Тестирование длится 30-40 минут
- Моментальная обработка результатов измерения и обратная связь с помощью Байесовских сетей**
- 3 уровня сформированности компетенций:  
продвинутый, базовый, развивающийся



## Отличительные особенности инструмента:

- Современная форма и интересный контент
- Автоматическая обработка результатов
- Отсутствие нагрузки на сотрудников школы

**Цифровая грамотность** - способность использовать цифровые технологии, инструменты коммуникации и сети для поиска, анализа, создания и управления информацией с целью удовлетворения личных, образовательных и профессиональных потребностей, сотрудничества и коллективной работы в цифровой среде, учитывая основы информационной безопасности, а также этические и правовые нормы работы с информацией.



# ТЕСТОВЫЕ ЗАДАНИЯ ИНСТРУМЕНТА ОЦЕНКИ ЦИФРОВОЙ ГРАМОТНОСТИ

**Тестовые задания различной сложности  
сценарного типа в электронном виде**

**Контекст заданий:** учебный (академический), личный (для учащихся)

Каждое задание представляет собой ситуацию, максимально приближенную к повседневной жизни тестируемого, содержащую в себе задачу, которую надо решить.

**Широкий спектр цифровых технологий,**  
в том числе электронная почта, текстовые редакторы, программы создания презентаций, электронные таблицы, графические редакторы, базы данных, средства мультимедиа, Интернет-сообщества

The screenshot displays a digital assessment interface. On the left, a browser window shows a travel website with images of the Antarctic and text about it being the coldest, driest, and windiest continent. On the right, a messaging sidebar shows several messages from users:

- Иван Дымков**: Помогите Степану, что куда мы сажаем мороженое. Вот сюда пару ссылок, которые могут быть полезными, посмотрите их: [www.coolantarctica.org](http://www.coolantarctica.org) и [www.icebergsguide.com](http://www.icebergsguide.com). Так же отыщите картинки, которых можем использовать в презентации. А здесь ты можешь посмотреть советы по оформлению презентации: [www.yourproject.ru](http://www.yourproject.ru)
- Вера Иванова**: Привет! Помоги сделать часть нашего с ребятами обсуждения в классиков формата. Давно. Беря в презентацию все, что посчитаешь нужным
- Денис Малышев**: Ребята, я собираю все материалы
- Иван Дымков**: Несмотря на то, что она - самый южный континент, согласно (иногда и не согласно) Гринвичу. Антарктида является самым холодным местом на Земле. Вот одно из свидетельств: самой низкой температурой, когда либо зарегистрированной на Земле, была

Below the messaging sidebar, there are two more browser windows. The top one shows a search for 'universalbank' on Google, displaying the official website of Universal Bank. The bottom window shows a detailed view of a credit card offer from the bank, including terms like 'Кредит до 5%' and 'Кредит до 10%'. The right side of this window also contains a sidebar with additional card options and terms.

Спасибо за внимание!  
Приходите к нам учиться

---

Карданова Е.Ю.  
[ekardanova@hse.ru](mailto:ekardanova@hse.ru)



Центр психометрики и измерений в образовании  
Институт образования  
НИУ Высшая школа экономики  
<http://ioe.hse.ru/monitoring/>

