

# Разработка системы оценивания

Ричард П. Фелпс

Международная научно-практическая  
конференция  
АОО «Назарбаев Интеллектуальные школы»

Астана, Казахстан

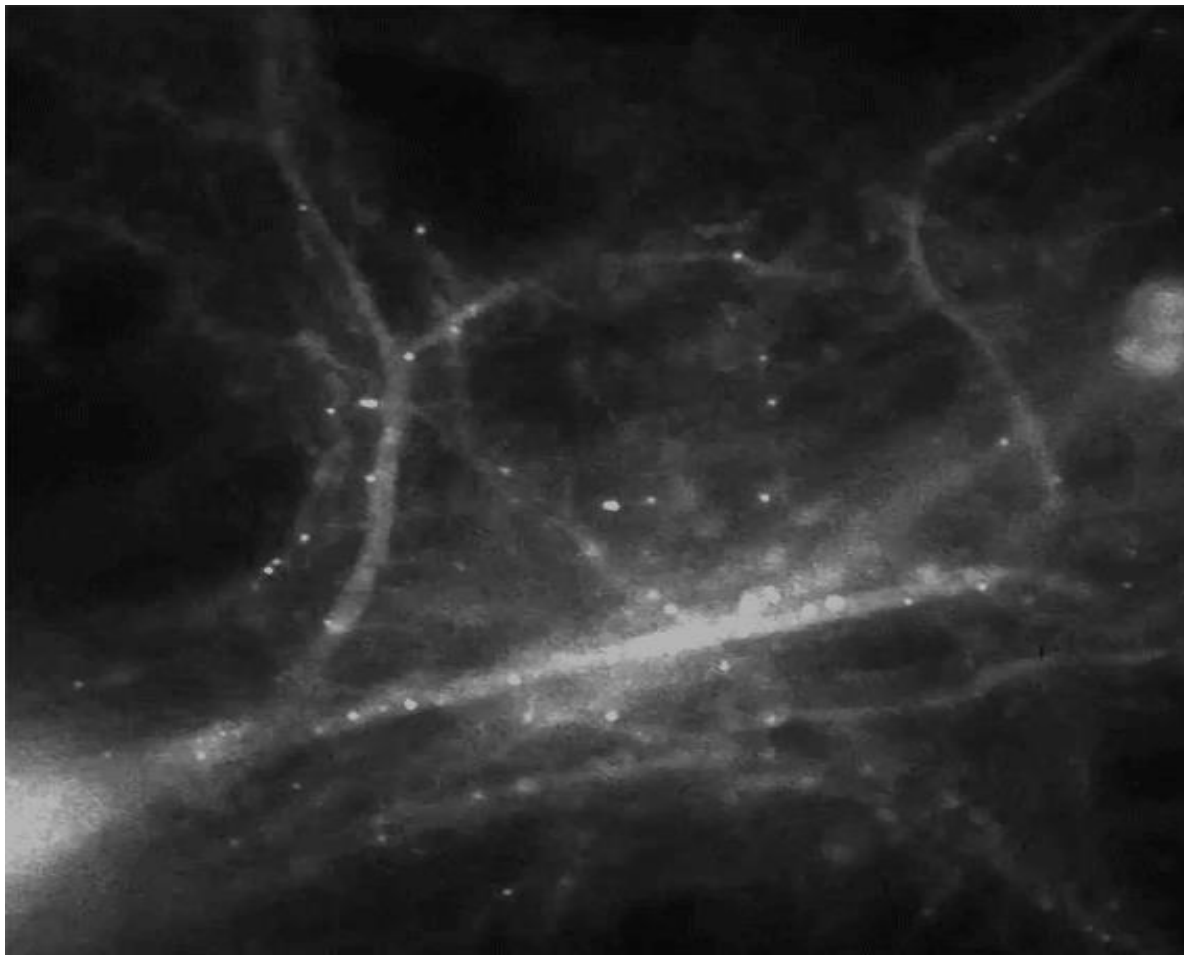
Октябрь, 2016



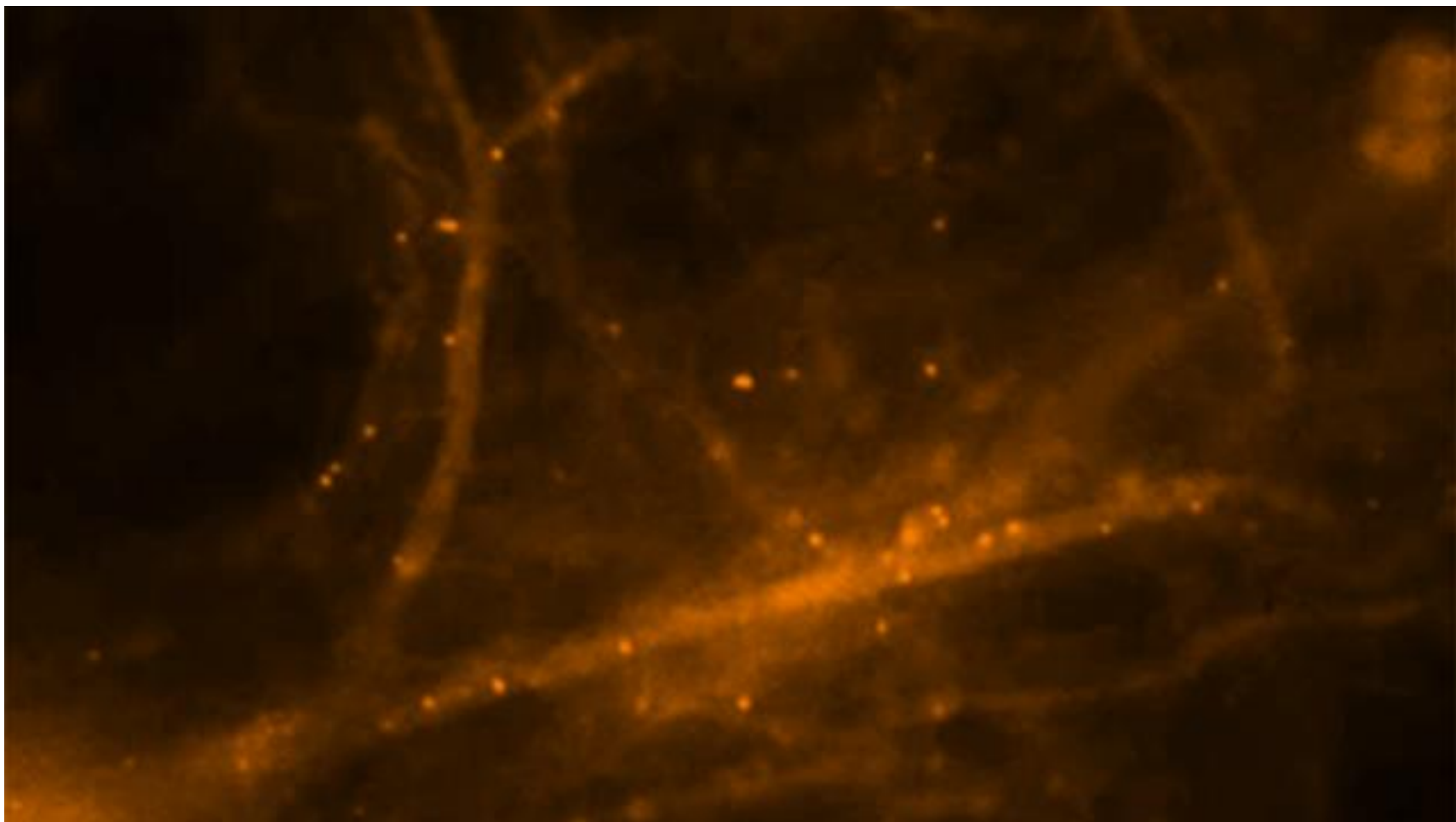
“Если что-то существует, то оно существует в определенном количестве. Если оно существует в определенном количестве, то это можно измерить.”

--Рене Декарт,  
*Первоначала философии*, 1664

Изображение белковых молекул, формирующих память  
Медицинский колледж им. Альберта Эйнштейна, Нью-Йорк, январь  
2014



Изображение белковых молекул, формирующих память  
Медицинский колледж им. Альберта Эйнштейна, Нью-Йорк, январь 2014



# Кривая обучения



## Кривая забывания (1870-е годы)





Эббингауз:

«Обучение, как  
правило, требует  
повторения»



Герман Эббингауз  
(1850-1909)





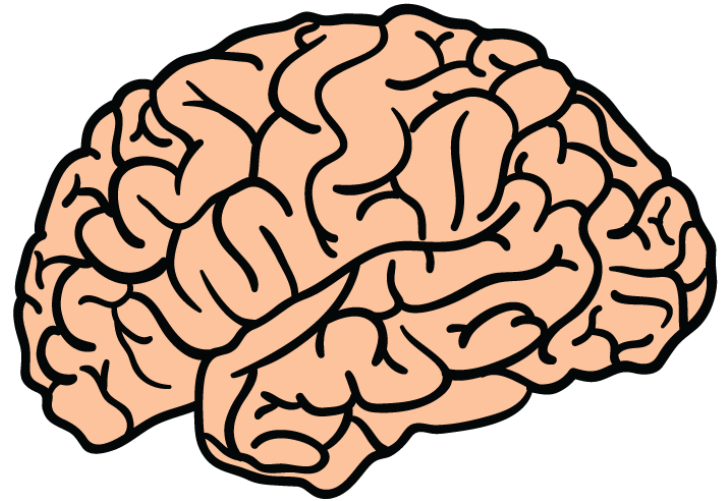
Объем рабочей памяти  
Джордж Миллер, 1950-е годы



Теория когнитивной нагрузки  
Джон Свеллер, 1980-е годы

## Рабочая память:

Способность временного хранения и управления информацией для когнитивных задач



Работа рабочей памяти стимулируется:

новой, незнакомой информацией и

количеством дискретной информации

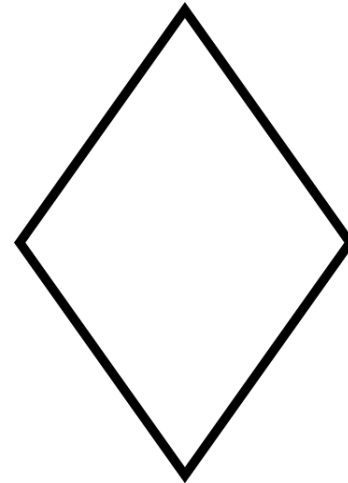
# О каком объекте я думаю?

Описание 1:

Это фигуры, геометрические плоские  
фигуры, многоугольники, четырехугольники,  
и параллелограммы с противоположными  
равными острыми углами, с  
противоположными равными тупыми  
углами, и равными сторонами

**О каком объекте я думаю?**

Описание 2:



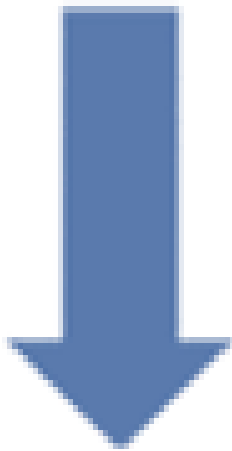
**Естественная  
нагрузка**

+

**Посторонняя  
нагрузка**

+

**Относительная  
нагрузка**



**Управлять**



**Уменьшить**



**Увеличить**

# Два столетия исследования по обучению приводит к выводу о том, что...



“...повторный поиск и извлечение информации из памяти во время обучения является ключом к долговременному запоминанию.”

— Генри Родигер

# 6 стратегии ученых-когнитивистов об эффективном обучении



Content by Yana Weinstein [University of Massachusetts Lowell] & Megan Smith [Rhode Island College] | Illustrations by Oliver Caviglioli [teachinghow2s.com/cogsci]  
Funding provided by the APS Fund for Teaching and Public Understanding of Psychological Science

Практика поиска и  
извлечения информации  
из памяти

Интервальное  
повторение

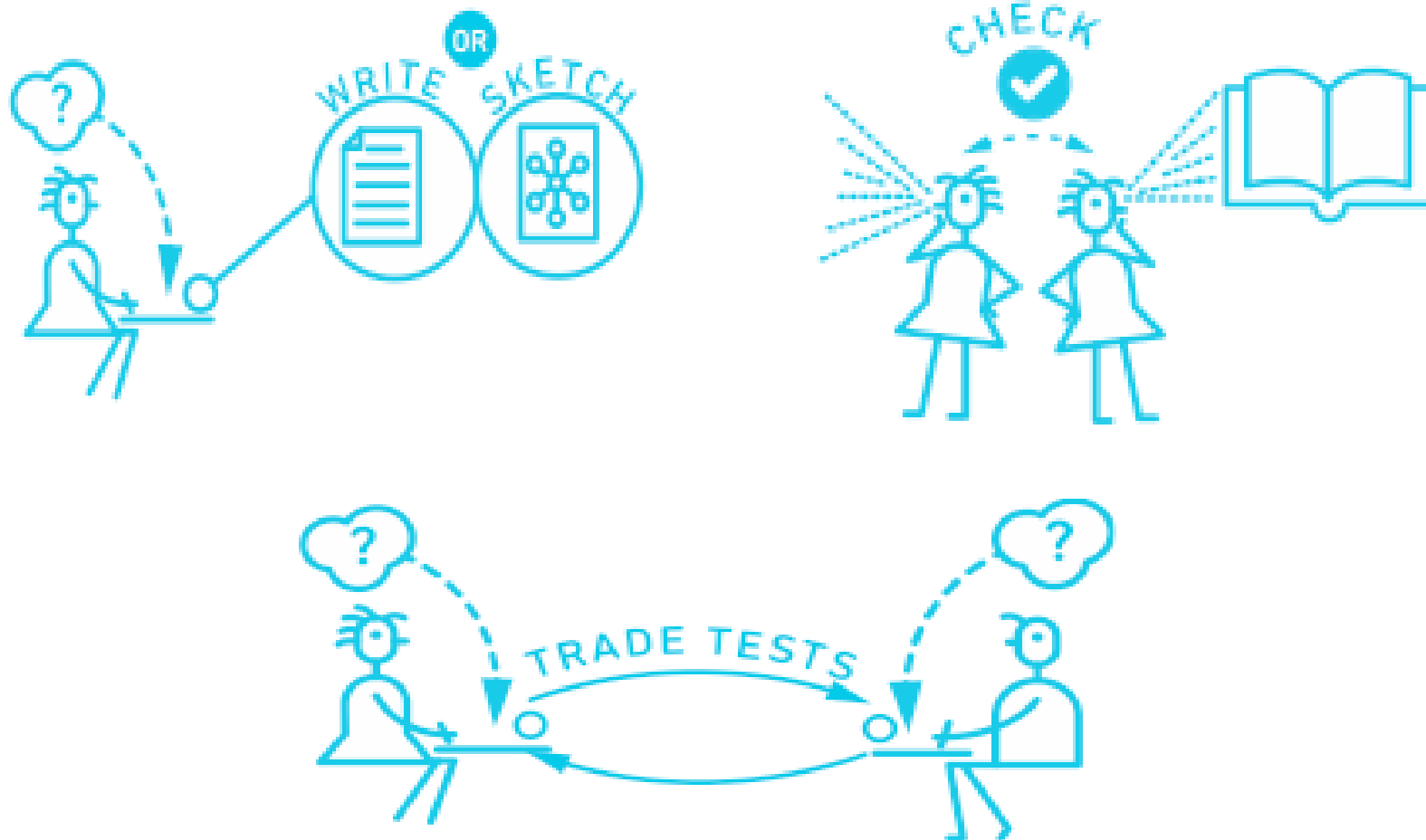
Двойное кодирование

Чередование

Конкретные примеры

Переработка

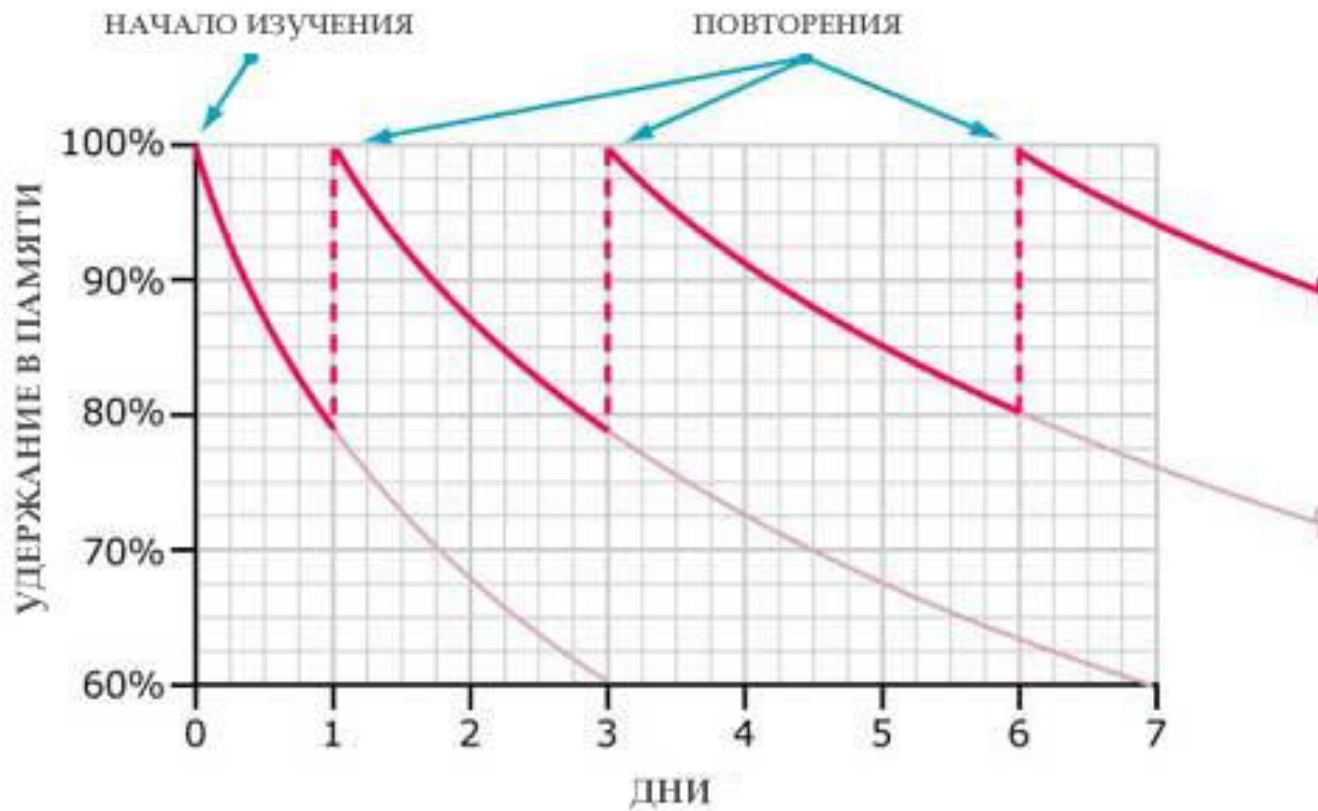
# Практика поиска и извлечения информации из памяти

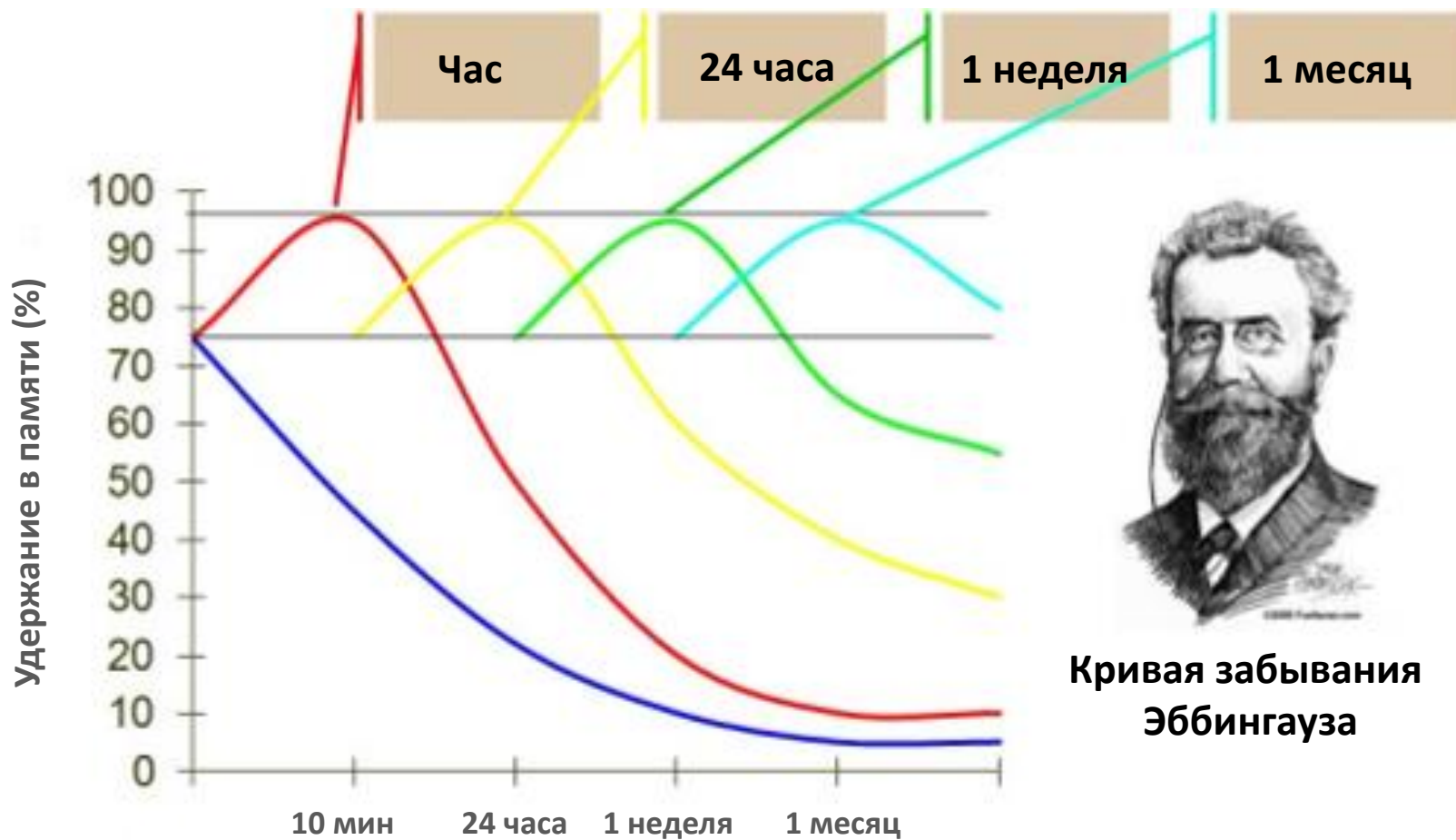




# Кривая забывания с интервальным повторением

## КРИВАЯ ЗАБЫВАНИЯ





**Кривая забывания  
Эббингауза**

# Выводы для учителей 1

Большинство учителей должны тестировать чаще, ... используя небольшие, короткие и простые тесты



Понимать, что эффективное оценивание может быть коротким и простым.

## Выводы для учителей 2



Имеет ли значение формат теста?

- многовариантный тест?
- эссе?
- короткий ответ?
- устный?
- иллюстрация?
- ...и т.д.?

Not so much.

## Выводы для учителей 3

Тесты предоставляют учителям обратную связь о том, что работает и не работает



Так же, как учащиеся могут учиться друг у друга путем тестирования, учителя могут помогать друг другу проверять свои тесты

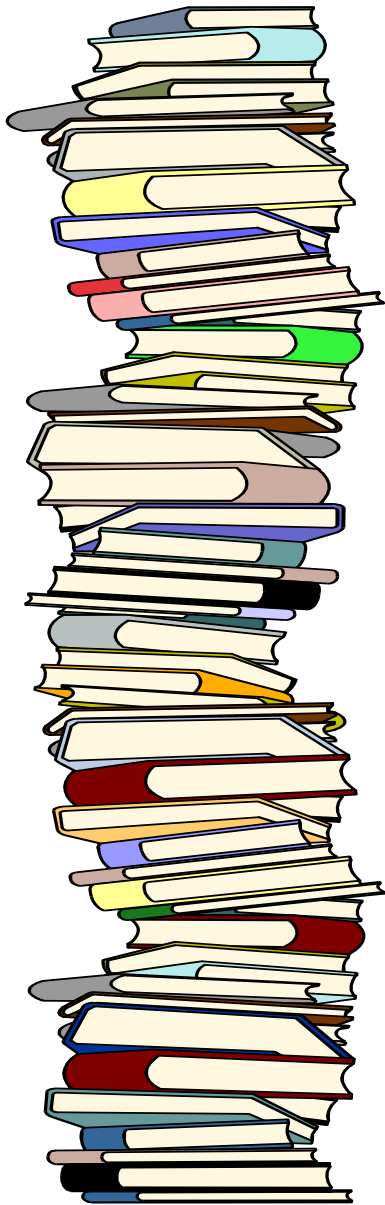
Эксперименты  
когнитивной психологии  
были проведены путем  
«формативного»  
тестирования в школах и  
классах



# Как насчет общесистемных, масштабных тестов?

Высокий приоритет:  
не навредите системе  
формативного оценивания  
в школах и классах





## Влияние тестирования на обучение учащихся

- 12-летнее исследование, прочитано >3,000 документов
- проанализировано около 700 отдельных исследований, а также более 1600 отдельных влияний
- 2000 других исследований были проверены и признаны неполными или несоответствующими
- Сотни исследований должны быть проанализированы



# Влияние тестирования на обучение учащихся

245 Качественных исследований

813 Опросов или голосований

640 Количественных исследований:

    Эксперименты:

        На уровне школы и класса

    Многомерные исследования:

        Программы масштабных  
        тестирований



# Мета-анализ

Систематизированный обзор научной литературы, объединяющий данные нескольких похожих исследований в единый статистический результат



( 0.5 величина эффекта  $\approx$  1 grade level of learning )

## Результаты исследования Фелпса (2012):

- Средняя величина эффекта исследования  $>1.0$
- Более 90% положительных качественных исследований
- В количественных исследованиях, положительное влияние на величины эффекта при:
  - Более частом тестировании
  - Тестировании с обратной связью
  - Тестирование по степени важности



## Результаты исследования Фелпса & Сильва (2015)

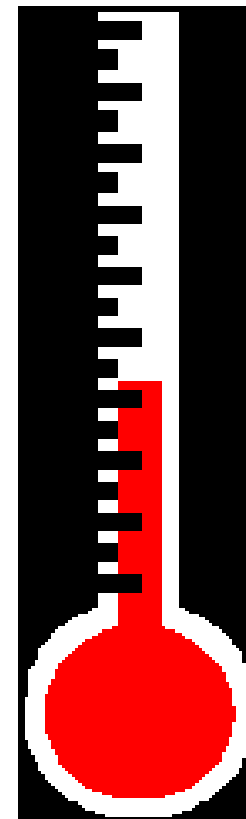
Для количественных исследований, величина эффекта колеблется между 0,55 и 0,88:



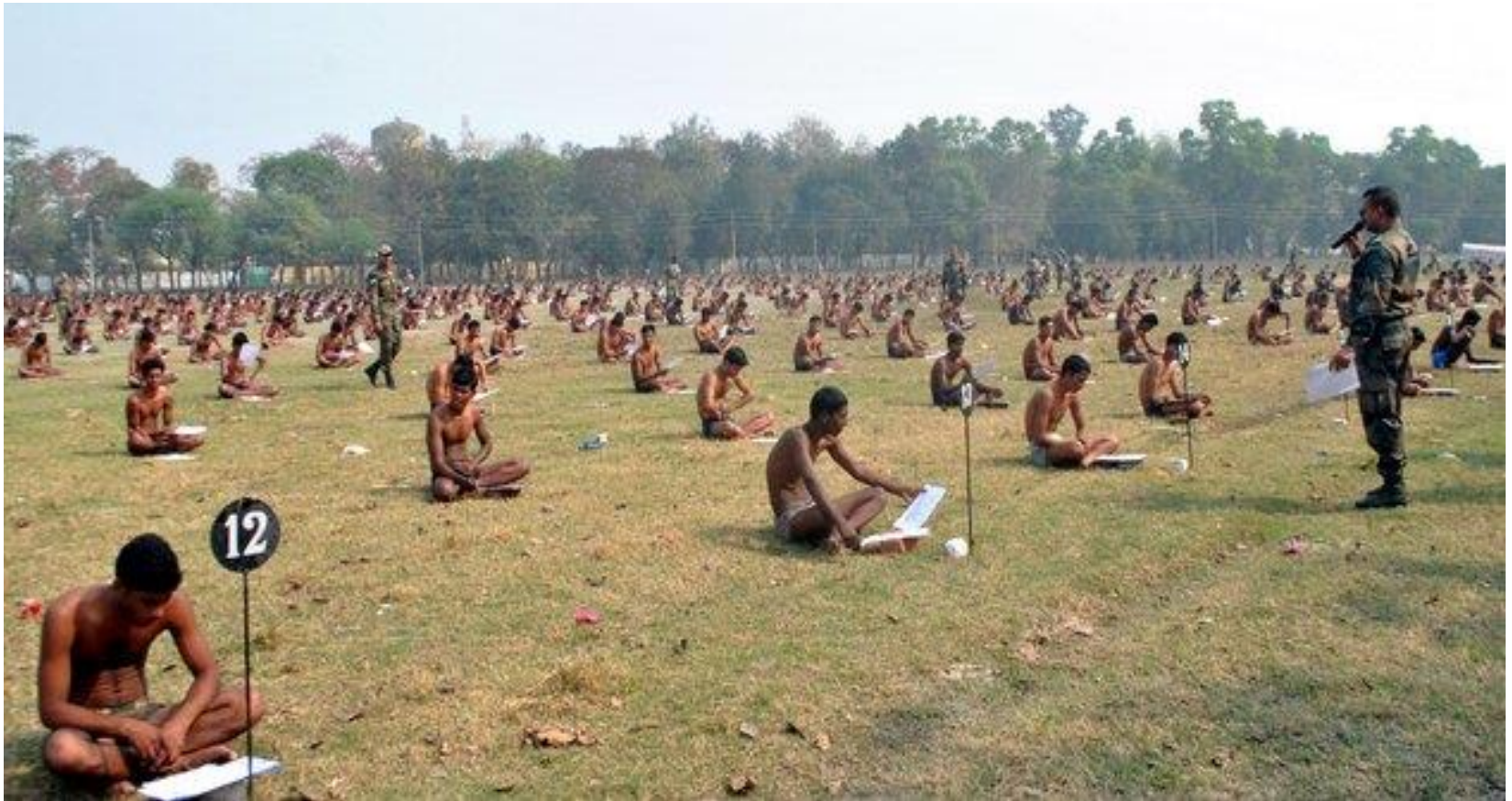
- +++ более частое тестирование
- ++ тестирование по степени важности
- + тестирование с обратной связью

# Влияние масштаба на преимущества тестирования

- величина исследуемой группы
  - небольшое +0.34 над большим
- масштаб проведения тестирования
  - небольшое +0.14 над масштабным
- ответственный уровень правления
  - местные тесты +0.29 над государственными тестами



# Масштабное тестирование, меры безопасности



# Масштабное тестирование, недостаточная безопасность



## Кроме того, общесистемные тесты необходимы для других целей, таких как...

...отбор для обучения в программы с ограниченным  
количеством мест

...мониторинг и система диагностики

...кадровое планирование

...подотчетность

...аттестация

Этого достаточно!





# Некоторые преимущества масштабного тестирования

На индивидуальной основе, недорого

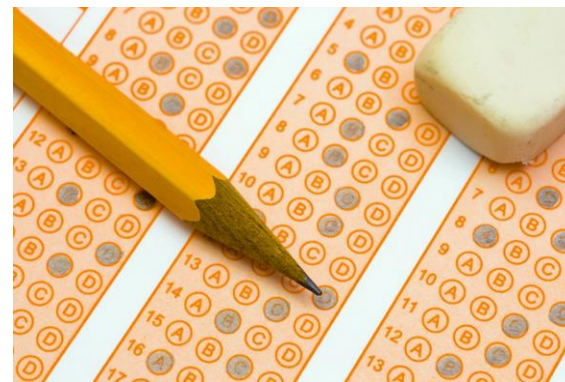
Возможность когнитивного тестирования в когнитивной лаборатории

Стандартизация предлагает сравнение между школами и регионами.

Возможность получения продуктов высокого качества для дальнейшего пользования школами и учителями.

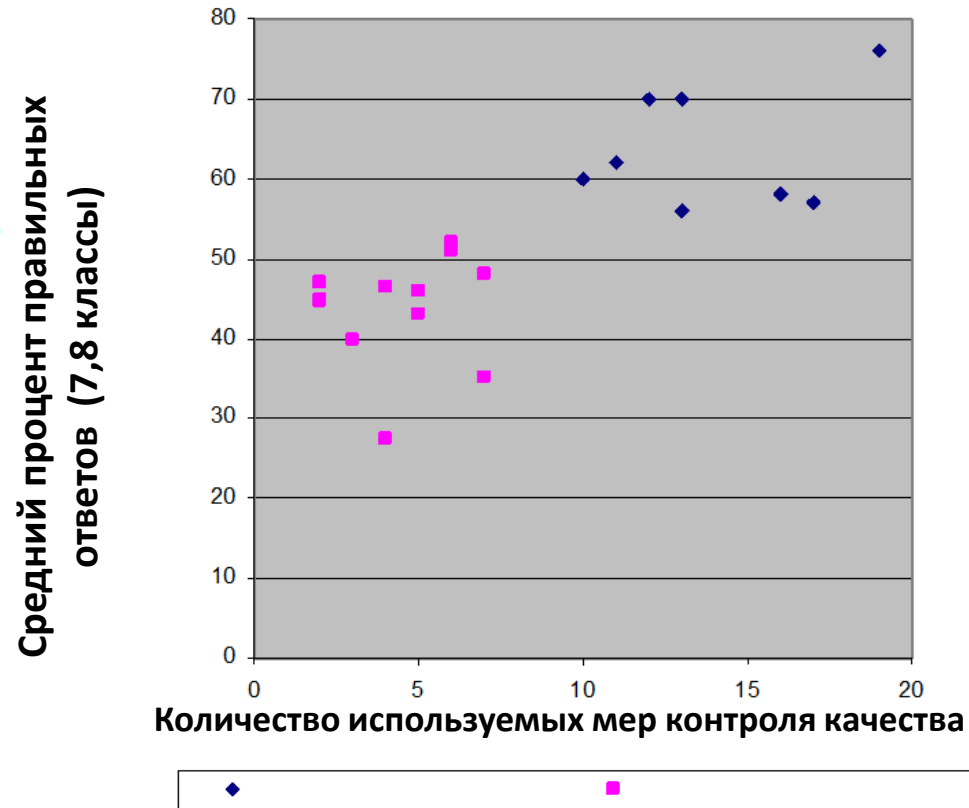
## **САМОЕ ВАЖНОЕ:**

обеспечивает надежную, сравнительную информацию для всех тех, кто не участвует в определенной школе



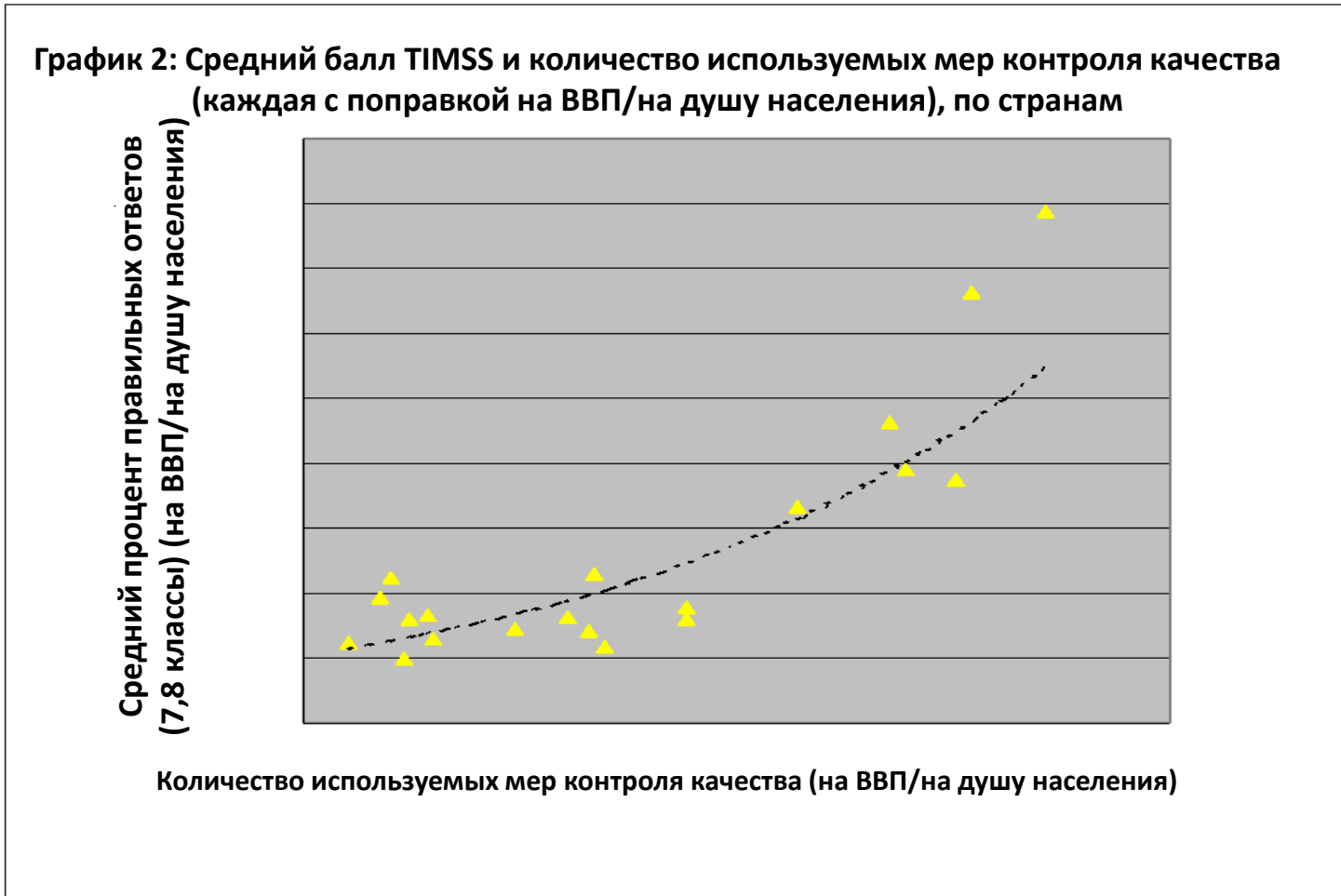
# Чем больше общесистемных этапов принятия решений, тем лучше?

График 1: Средний балл TIMSS и количество используемых мер контроля качества, по странам



SOURCE: Phelps, Benchmarking to the best in mathematics, *Evaluation Review*, 2001

# Контроль качества имеет пропорционально больший эффект в более бедных странах



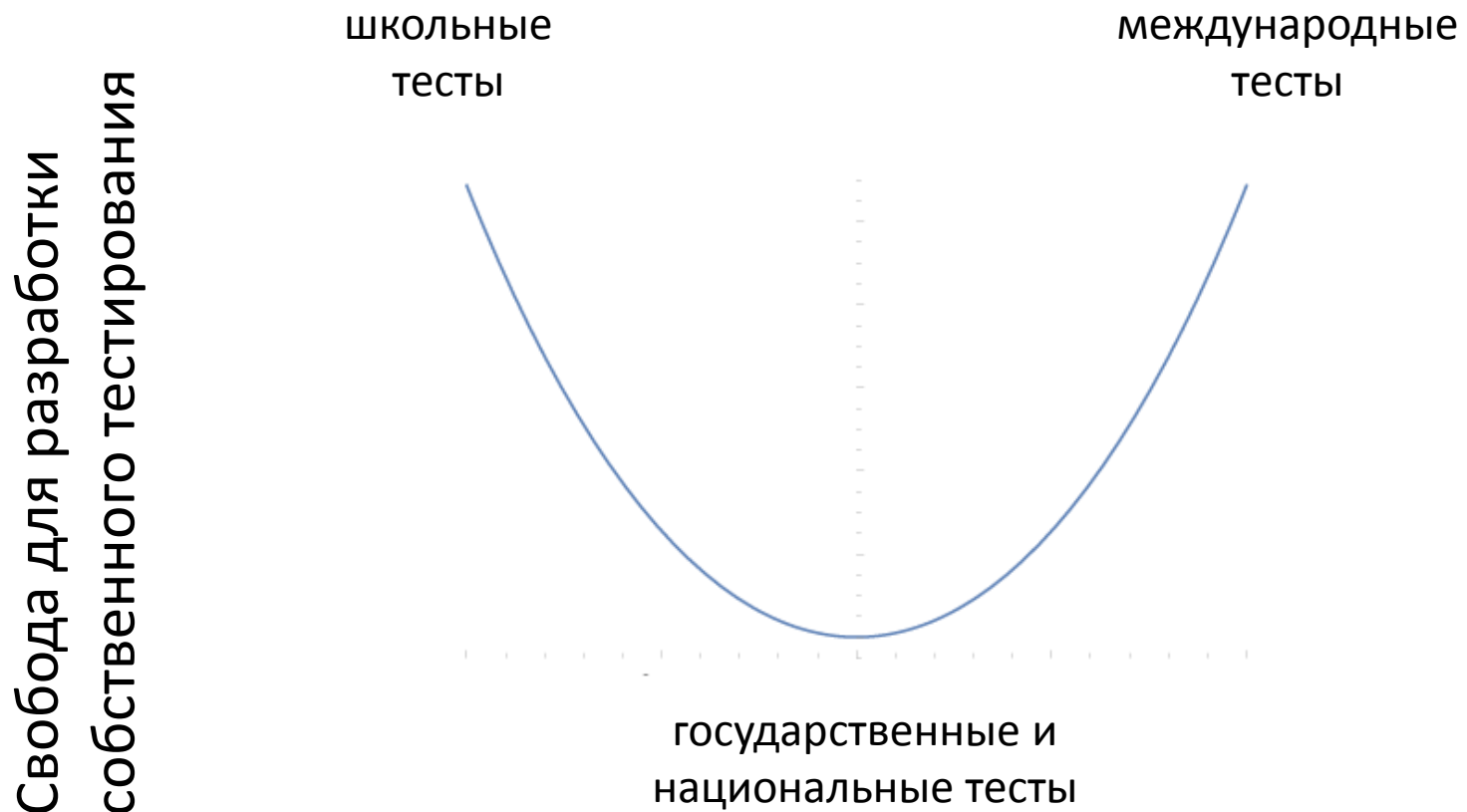
SOURCE: Phelps, Benchmarking to the best in mathematics, *Evaluation Review*, 2001

**IEA:** TIMSS, PIRLS, CIVED, SITES, ICILS,  
PPP, ECES, TEDS

**ОЭСР PISA:** PISA, PISA для школ  
PISA для развития

**Всемирный  
банк:** READ, SABER  
...предусматривает  
финансирование для PISA

# Влияние международных программ тестирования



# Всемирным банком и ОЭСР руководят экономисты

Насколько хорошо экономисты разбираются в психометрике?

Интересные примеры:

Национальная программа тестирования Чили, финансируемая Всемирным банком

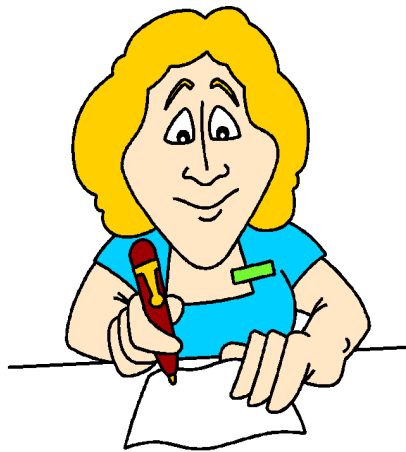
Проект ОЭСР «Взаимодействие для более эффективного обучения»



# Интересные факты:



# Разработка системы оценивания



richard {at} nonpartisaneducation {dot} org