



В поисках "идеальной" системы онлайн-тестирования.

Ким И.А. Доцент департамента теоретической
экономики НИУ ВШЭ, к.э.н.

Moscow, 2017

Мотивация к исследованию



+



- Хорошо ли работает такая комбинация?
- Насколько можно «доверять» задачи промежуточного и итогового контроля онлайн-компоненте смешанного курса?

Примеры моделей смешанного обучения в НИУ ВШЭ

- Факультет социальных наук, программа «Политология», курс «Макроэкономика»: семинары (32 акад. часа) и экзамен – оффлайн, лекции + тесты – онлайн, на платформе Курсера (120 акад. часов самостоятельной работы)
- Факультета бизнеса и менеджмента, программа «Менеджмент», курс «Макроэкономика»: лекции + семинары (100 акад. часов), контрольные, экзамен - оффлайн, подготовка эссе, доп. материалы и тесты - онлайн (100 акад. часов самостоятельной работы)

Макроэкономика на программе «Политология»

Деятельность под
контролем преподавателя



Семинарские
занятия

Письменный
экзамен

Самостоятельная работа онлайн



Оценка за тесты на
платформе Курсера

Корреляция оценок с кумулятивным
рейтингом студентов:

0,55

0,33

0,25

Макроэкономика на программе «Менеджмент»

Деятельность под
контролем преподавателя



Семинарские занятия
Контрольные 1ая 2ая
Письменный экзамен

Самостоятельная работа,
в том числе онлайн



Эссе
Тесты на платформе LMS

Корреляция оценок с кумулятивным
рейтингом студентов:

0,72	0,59	0,60	0,70	0,28	0,48
------	------	------	------	------	------

Результаты использования курса «Экономика для неэкономистов» для ряда других программ

Образовательная программа	Кампус НИУ ВШЭ	Корреляция оценки за курс с рейтингом
Реклама и связи с общественностью	Москва	0,17
Юриспруденция	Нижний Новгород	0,35
Востоковедение	Санкт-Петербург	0,23

Что такое «Кумулятивный рейтинг»?

- Это взвешенная сумма оценок за курсы, ранее пройденные студентом (6-10 шт. для первокурсников).
- Чем больше образовательных кредитов начисляется за курс, тем выше вес его оценки в рейтинге.
- Во внимание принимаются все пройденные курсы (в т.ч. иностранный язык, безопасность жизнедеятельности и т.д.).

Макроэкономика на программе «Политология» (2)

Деятельность под контролем преподавателя



Семинарские занятия

Письменный экзамен

Самостоятельная работа онлайн



Оценка за тесты на платформе Курсера

Корреляция оценок по макроэкономике с оценкой по микроэкономике:

0,59 (0,55)

0,65 (0,33)

0,42 (0,25)

Макроэкономика на программе «Менеджмент» (2)

Деятельность под контролем преподавателя



Самостоятельная работа, в том числе онлайн



Семинарские занятия

Контрольные
1ая 2ая

Письменный экзамен

Эссе

Тесты на платформе LMS

Корреляция оценок по макроэкономике с оценкой по микроэкономике:

0,70 (0,72)	0,47 (0,59)	0,52 (0,60)	0,69 (0,70)	0,31 (0,28)	0,46 (0,48)
----------------	----------------	----------------	----------------	----------------	----------------

В любом случае для видов работ, выполняемых вне контроля преподавателя, корреляция хуже. Почему?

- Некорректное поведение студентов?
- Плохая различающая способность тестовых заданий?

Некорректное поведение?

Деятельность под контролем преподавателя



Самостоятельная работа, в том числе онлайн



Семинарские занятия

Контрольные
1ая 2ая

Письменный экзамен

Эссе

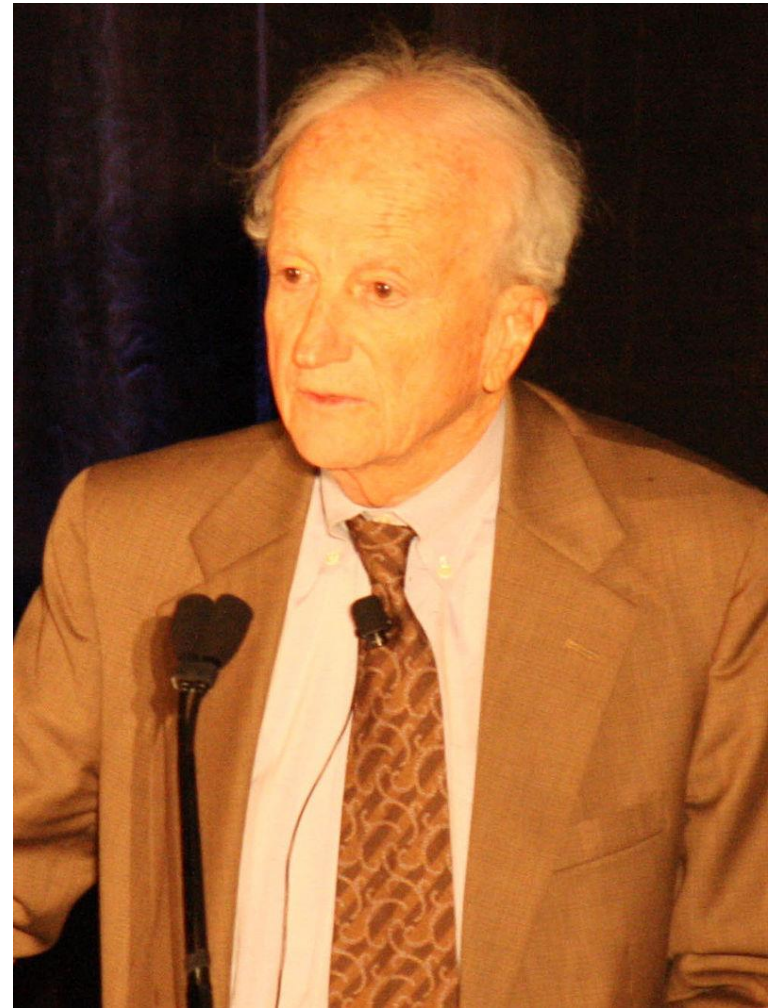
Тесты на платформе LMS

Корреляция оценок по макроэкономике с оценкой по микроэкономике:

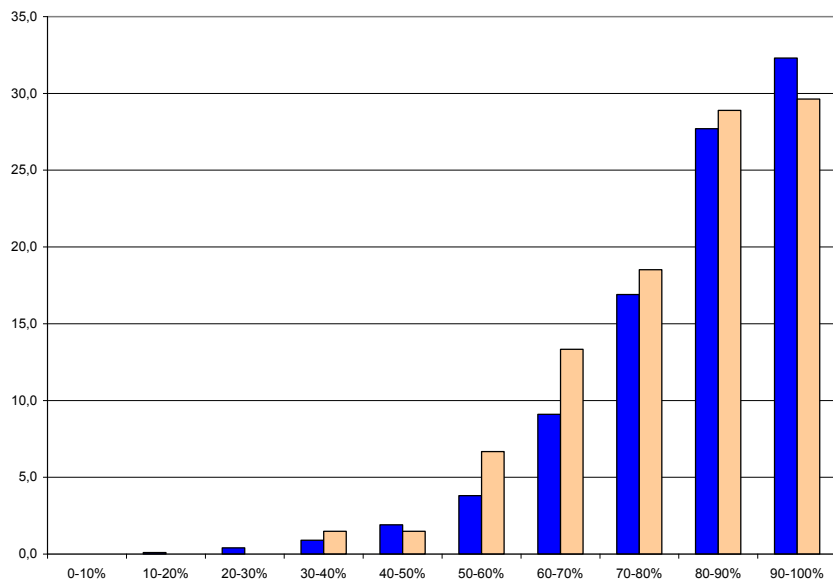
0,70 (0,72)	0,47 (0,59)	0,52 (0,60)	0,69 (0,70)	0,31 (0,28)	0,46 (0,48)
----------------	----------------	----------------	----------------	----------------	----------------

Люди реагируют на стимулы...

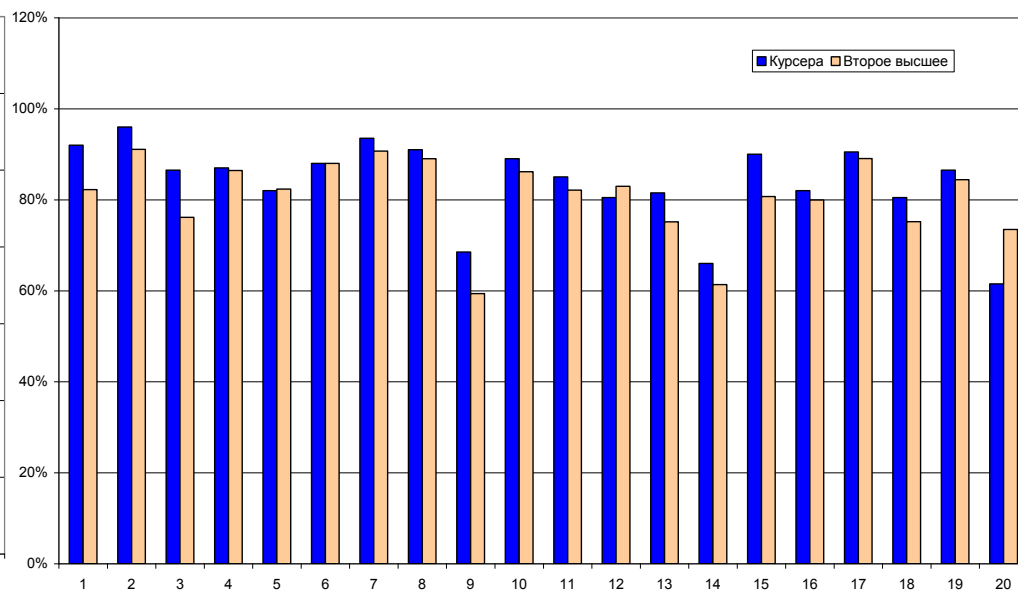
- Выгоды некорректного поведения
- Издержки некорректного поведения



Сравнение результатов контрольного тестирования двух разных групп обучающихся

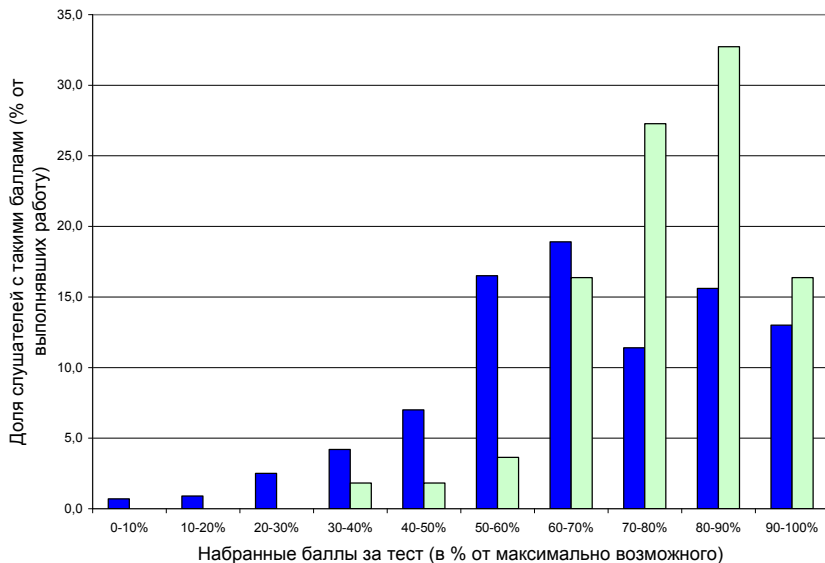


Общие результаты

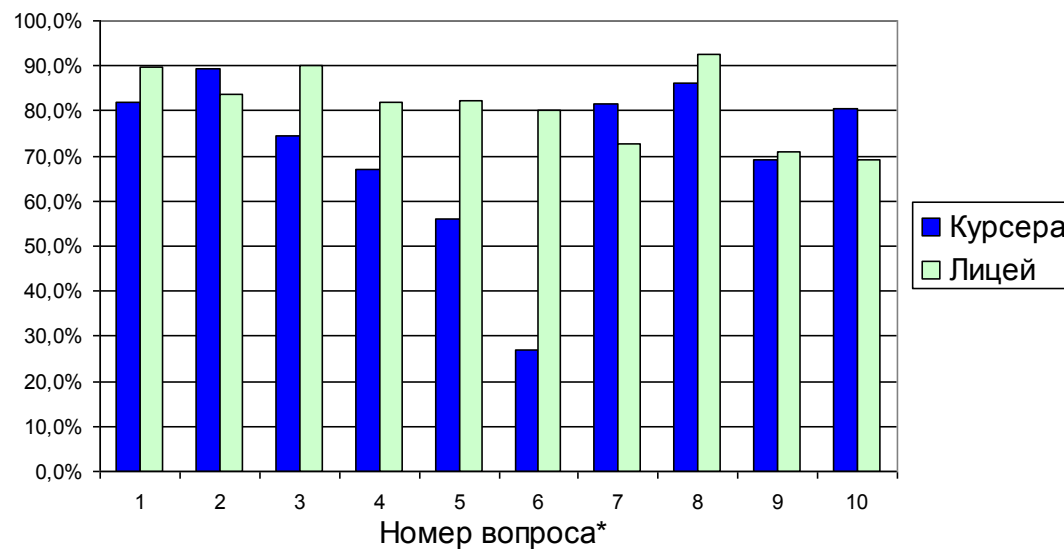


Статистика по отдельным
вопросам

Сравнение результатов контрольного тестирования двух разных групп обучающихся (2)



Общие результаты



Статистика по отдельным вопросам

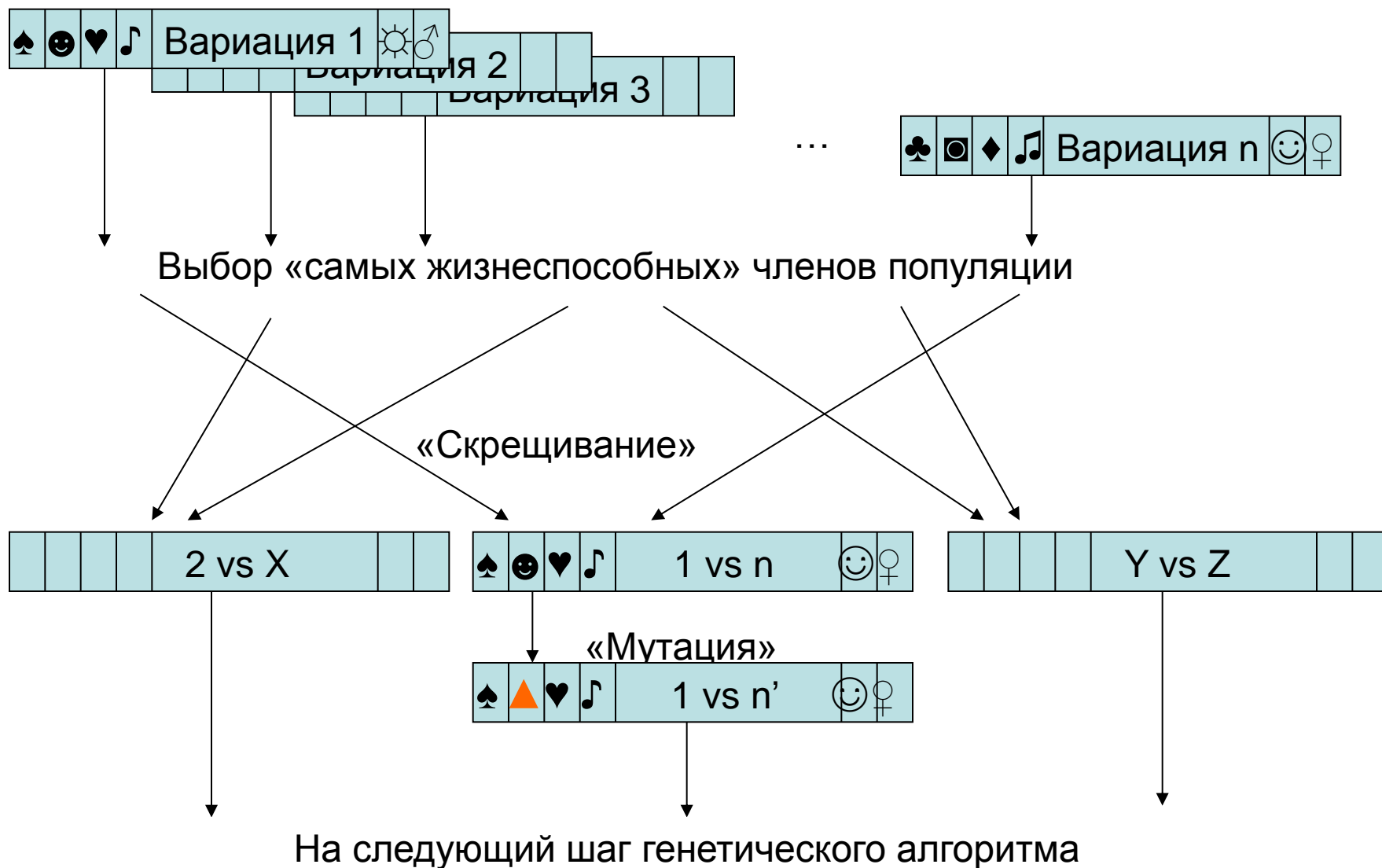
Как улучшить различающую способность тестов?

- Один и тот же вопрос теста/дистрактор вопроса для разных групп обучающихся (например, разных когорт на Курсере) может проявить себя по-разному.
- Вопросы не должны быть скучными и тривиальными, но при этом должны быть корректными и пониматься однозначно...
- ...но преподаватель не всегда может предсказать, «что поймут неправильно».
- Тесты должны регулярно обновляться.

Генетические алгоритмы

- Повторяют природный механизм естественного отбора.
- Уже используются в математике, информатике, конструировании, менеджменте.
- Предполагают создание некоторой исходной "популяции" вариантов тестовых заданий с разными комбинациями вопросов.
- По результатам нескольких поколений выделяются наилучшие варианты, которые «скрещиваются» между собой.
- Наихудшие члены популяции удаляются из нее без возможности скрещивания.
- Некоторые члены популяции подвергаются «мутации».

Один шаг генетического алгоритма



Направления дальнейшего исследования

- Исследование статистики по онлайн-тестам с прокторингом
- Экспериментальная проверка работоспособности генетических алгоритмов посторения базы данных тестов



Спасибо за внимание!