

Измерения в EdTech:
чего не хватает для перехода к
доказательному образованию

Дмитрий Аббакумов

*доктор наук об образовании (PhD KU Leuven in Educational Sciences),
руководитель Центра вычислительных наук об образовании НИУ ВШЭ*



THIS IS A TRUE STORY

“Это реальные истории. Исследования, о которых рассказывается в этой лекции, опубликованы в международных научных журналах в 2010-2019 годах.”

THIS IS A TRUE STORY





- *Doctor of Educational Sciences* from KU Leuven (2019)
 - Specialization: Statistical Modeling and Methodology of Educational Sciences
 - Thesis: **Psychometrics of MOOCs: How to Measure Proficiency?**
- Руководитель Центра вычислительных наук об образовании НИУ ВШЭ

EdTech

LANGUAGE LEARNING

MANAGEMENT PLATFORMS

EXAMS & TEST PREP

WORKFORCE & UPSKILLING

ONLINE HIGHER EDUCATION

CURRICULUM & LEARNING PLATFORM

TALENT ACQUISITION

TUTORING & LEARNING SUPPORT

ROBOTICS, STEAM & CODING

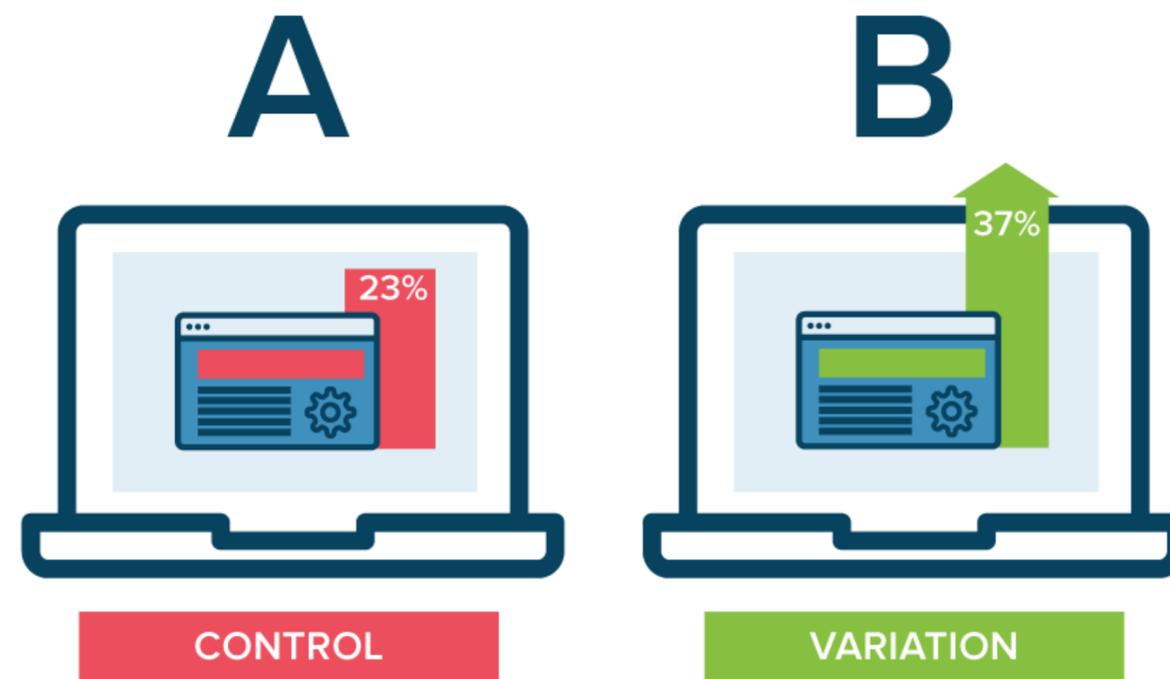
XR, GAMES & SIMULATION

"**Measurement** is the first step that leads to control and eventually to improvement. If you can't measure something, you can't **understand** it. If you can't understand it, you can't control it. If you can't control it, you can't **improve** it."

-H. James Harrington

A/B-тесты

- bucket testing, split(-run) testing
- (маркетинг.) методология исследования пользовательского опыта



A/B-тесты

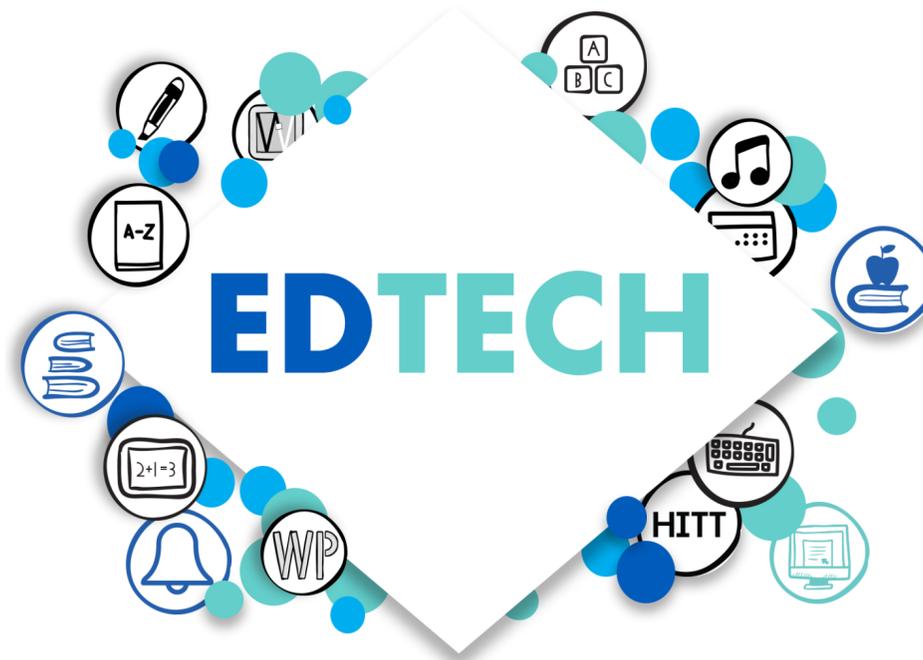
- bucket testing, split(-run) testing
- (*маркетинг.*) методология исследования пользовательского опыта

A/B-тесты

- bucket testing, split(-run) testing
- (*маркетинг.*) методология исследования пользовательского опыта
- форма и содержание сайта ~ количество регистраций на курс, количество и объем пожертвований в эндаумент



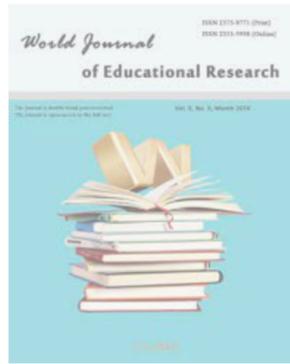
A/B-тесты



- форма и содержание учебных материалов ~ учебные результаты и прогресс
- двигатель экспериментальной педагогики



Educational Research Re...
journals.elsevier.com



World Journal of Educatio...
scholink.org



European Journal of Edu...
eu-jeer.com



Review of Educational R...
aera.net



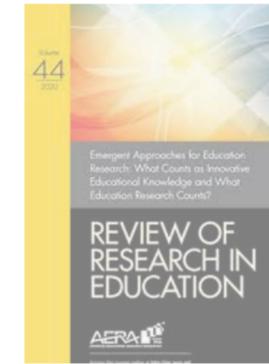
International Journal of E...
ijer.penpublishing.net



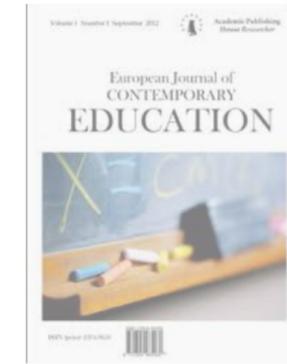
Journal of Educational a...
mcser.org



International Journal of Ed...
journals.elsevier.com



Educational Researcher...
journals.sagepub.com



European Journal of Con...
ejournal1.com



American Educational ...
en.wikipedia.org



British Educational Res...
bera-journals.onlinelibrary...



CEEOL - Journal Detail
ceol.com



International Journal of ...
ajol.info



Home | Journal of Educati...
ccsenet.org



Journal of Education: SAG...
journals.sagepub.com



International Journal of R...
ijрте.penpublishing.net



Journal of International Edu...
clutejournals.com



PDF) International Journ...
researchgate.net



Journal of Applied Research in...
emeraldgroupublishing.com



International Journal of Educat...
journals.elsevier.com



European Journal of Engineer...
sefi.be



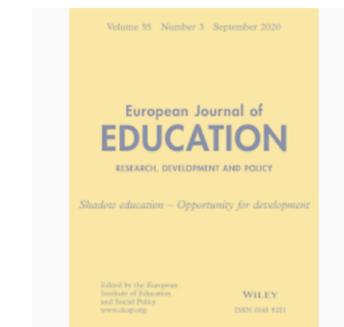
Educational research papers
teachingworks.org



Academic Journals and the Work of Editorial Boards: A...
cie.uprrp.edu



Journal of Education | Wheelock ...
bu.edu



European Journal of Educatio...
onlinelibrary.wiley.com

Три уровня проблем

- Индикаторы
- Концептуальная модель
- Психометрическая модель

Индикаторы

В отличие от А/В-тестов в маркетинге, в образовании мы имеем дело с **латентными** целевыми показателями

Индикаторы

В отличие от А/В-тестов в маркетинге, в образовании мы имеем дело с **латентными** целевыми показателями



Индикаторы

- Случайные

Индикаторы

- Случайные
 - *“Мы взяли 5 тестовых заданий.”*

Индикаторы

- Случайные
 - *“Мы взяли 5 тестовых заданий.”*
- Невалидные

Индикаторы

- Случайные
 - *“Мы взяли 5 тестовых заданий.”*
- Невалидные
 - *“Мы взяли два разных теста, но их разрабатывал один человек, и они оба на одной теме.” [Поэтому мы уверены, что эти тесты равнозначны.]*

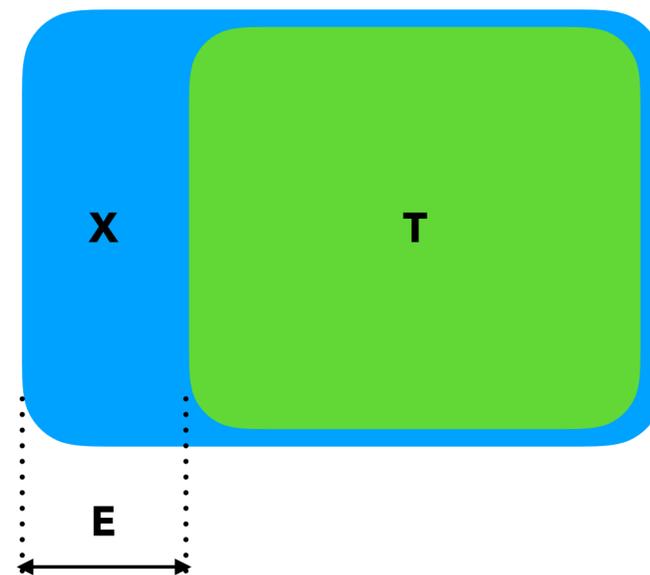
Индикаторы

- Ненадежные

Индикаторы

- Ненадежные

Насколько тесно целевой показатель (T) связан с индикатором (X)?



Индикаторы

- Ненадежные

A



B

Индикаторы

- Ненадежные

	Результат (из 20)
A	?
B	?

Индикаторы

- Ненадежные

	Результат (из 20)
A	13
B	17

Индикаторы

- Ненадежные

	Результат (из 20)
A	13
B 	17

Индикаторы

- Ненадежные

	Результат (из 20)	Доверит. интервал
A	13	10-16
B	17	14-20

Индикаторы

- Ненадежные

	Результат (из 20)	Доверит. интервал
A	13	
B	17	

Индикаторы

- Ненадежные

	Результат (из 20)	Доверит. интервал
A	12	10-16
B	17	14-20

The table contains a shaded gray area covering the 'A' row and the 'B' result cell. Inside this shaded area is a sad face icon.

Концептуальная модель

- Некорректная концепция измерения



Концептуальная модель: Пример 1

- Эффекты на **Mastering** (освоение)

Концептуальная модель: Пример 1

- Эффекты на **Mastering** (освоение)
 - использовался тест по последней попытке

Концептуальная модель: Пример 1

- Эффекты на **Mastering** (освоение)
 - использовался тест по последней попытке
 - не учтено количество попыток

Концептуальная модель: Пример 1

- Эффекты на **Mastering** (освоение)
 - использовался тест по последней попытке
 - не учтено количество попыток
 - не учтены эффекты перебора

Концептуальная модель: Пример 2

- Эффекты на **Прогресс** (прирост)

Концептуальная модель: Пример 2

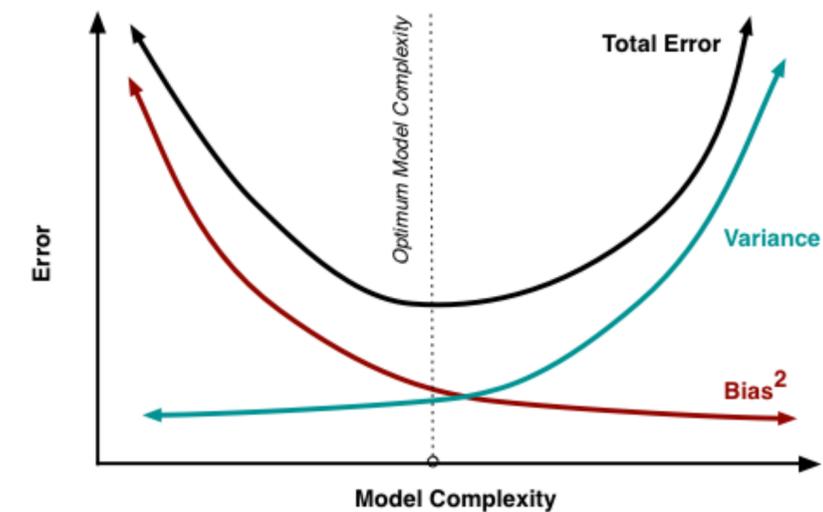
- Эффекты на **Прогресс** (прирост)
 - использовались два предъявлениям одного теста

Концептуальная модель: Пример 2

- Эффекты на **Прогресс** (прирост)
 - использовались два предъявлениям одного теста
 - не учтены эффекты запоминания заданий

Психометрическая модель

- Не отражает связь между индикаторами и целевым показателем
 - Не учтены внешние факторы
 - Неверная форма функциональной зависимости
 - Неверные допущения модели

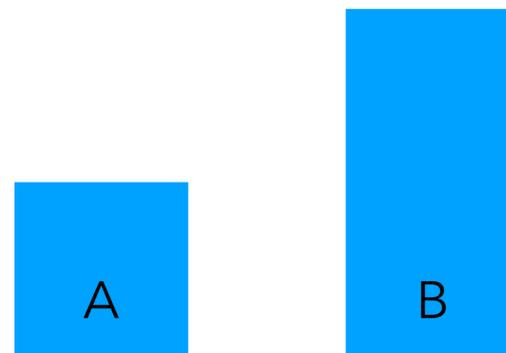


Неучтенные эффекты

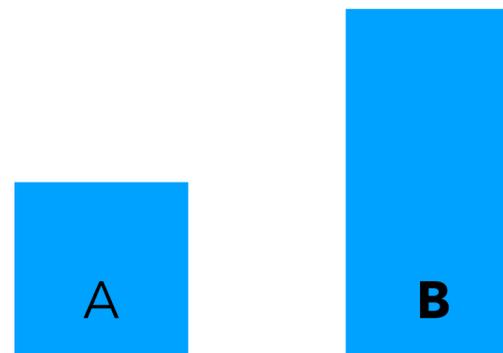
A

B

Неучтенные эффекты

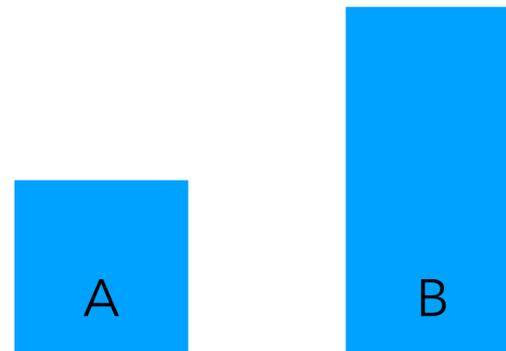


Неучтенные эффекты



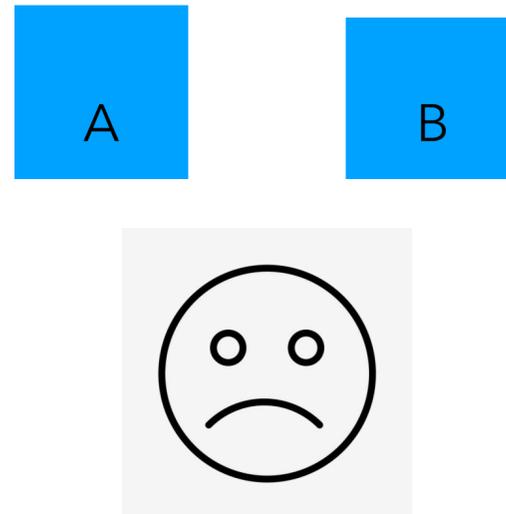
Неучтенные эффекты

Оценка = Знания + Трудность + Строгость



Неучтенные эффекты

Оценка = Знания + Трудность + **Строгость**



Другие примеры

- Внешняя многомерность (например, интерес)

Abbakumov, D., Desmet, P., & Van den Noortgate, W. (2020). Rasch model extensions for enhanced formative assessments in MOOCs. *Applied Measurement in Education*, 33(2), 113-123.

Другие примеры

- Внешняя многомерность (например, интерес)
- Внутренняя многомерность (например, два скилла)

Другие примеры

- Внешняя многомерность (например, интерес)
- Внутренняя многомерность (например, два скилла)
- Ложные допущения (например, бимодальное распределение)

Другие примеры

- Внешняя многомерность (например, интерес)
- Внутренняя многомерность (например, два скилла)
- Ложные допущения (например, бимодальное распределение)
- Ложные обобщения (например, response time как общий эффект)



*Н. Т.
Чертышевский*

ЧТО ДЕЛАТЬ?





Psychometrics is a scientific discipline concerned with the construction of assessment tools, measurement instruments, and formalized models that may serve to connect observable phenomena to theoretical attributes.

prof. dr. Denny Borsboom, UvA



James McKeen Cattell (1860-1944) и его психометрическая группа в Кембридже (1888/89)

SCIENCE

Vol. 85

FRIDAY, MARCH 5, 1937

No. 2201

<i>Psychology as a Quantitative Rational Science: Professor L. L. THURSTONE</i>	227
<i>Obituary: Edward Curtis Franklin: Professor C. A. KRAV'S. Recent Deaths and Memorials</i>	232
<i>Scientific Events: The Columbia University School of Medicine; The Electronics Institute at the University of Michigan; The Wisconsin Alumni Research Foundation; Meeting of Texas Geologists; The Foster Lecture Foundation at the University of Buffalo; The Installation of a Chapter of Sigma Xi at the George Washington University</i>	234
<i>Scientific Notes and News</i>	237
<i>Discussion: Dissemination of Scientific Literature by Means of Microfilms: DR. ATHERTON SEIDELL. Micro-photography or Photo-micrography: PROFESSOR W. J. LUYTEN. A Vasopressor Local Anesthetic: PROFESSOR CHAUNCEY D. LEAKE. The Action of Eserine and its Analogues on Skeletal Muscle: Z. M. BACQ and G. L. BROWN</i>	240
<i>Scientific Books: Recent Physics: PROFESSOR HENRY CREW. Nuclear Physics: PROFESSOR ALLAN C. G. MITCHELL</i>	243

<i>Special Articles: Vitamin B₁ and the Growth of Excised Tomato Roots: DR. WILLIAM J. ROBBINS and MARY A. BARTLEY. The Significance of the Adrenals for Adaptation: PROFESSOR HANS SELYE. The Synthesis of Substances Related to Lysergic Acid: DR. WALTER A. JACOBS and DR. R. GORDON GOULD, JR.</i>	246
<i>Scientific Apparatus and Laboratory Methods: An Automatic Dehydrating Device: JEAN PENNINGTON and DR. C. P. HICKMAN. The Isolation of Muscle Nuclei: GERMAIN CROSSMON</i>	249
<i>Science News</i>	14

SCIENCE: A Weekly Journal devoted to the Advancement of Science, edited by J. MCKEEN CATTELL and published every Friday by

THE SCIENCE PRESS

New York City: Grand Central Terminal
Lancaster, Pa. Garrison, N. Y.
Annual Subscription, \$6.00 Single Copies, 15 Cts.
SCIENCE is the official organ of the American Association for the Advancement of Science. Information regarding membership in the Association may be secured from the office of the permanent secretary, in the Smithsonian Institution Building, Washington, D. C.

PSYCHOLOGY AS A QUANTITATIVE RATIONAL SCIENCE¹

By Professor L. L. THURSTONE
THE UNIVERSITY OF CHICAGO

THE purposes of this society are not new, but they represent an emphasis and direction which have not hitherto received major consideration in psychological science. It seems proper that we should devote some share of our first program meeting to a consideration of our main objectives.

Our main purpose is briefly stated in the subtitle of the new journal, *Psychometrika*, namely, to encourage the development of psychology as a quantitative rational science. More briefly, this may be called mathematical psychology. We should justify our emphasis upon quantification and upon rationalization in science, as well as our conception of the fundamental nature of science.

I assume that we are in complete agreement that we can not suddenly quantify our comprehension of psy-

¹ Abstract of address by the retiring president of the Psychometric Society at Hanover, N. H., September 4, 1936.

chological phenomena over their entire range. As psychologists, we are as interested as ever in making exploratory studies of new psychological effects and in discovering hitherto unknown effects. At present, the range of psychological phenomena that can be profitably reduced to mathematical formulation is limited, and it is likely that every man who works on a problem of mathematical psychology will also concern himself with exploratory studies of other problems that are as yet too new for detailed rationalization.

After the discovery of a psychological effect, we naturally turn to the second phase of scientific inquiry, namely, to relate the new effect in a simple descriptive manner to what is already known. In this stage theories are devised to explain the experimentally known effects, and we try, of course, to make psychological theories less complicated than the effects that are to be explained. In this phase the descriptions of psycho-



Louis L. Thurstone (1887-1955)

Системное решение

Системное решение



Системное решение



Проектируйте измерения от пед. целей, а не от контента.

Системное решение



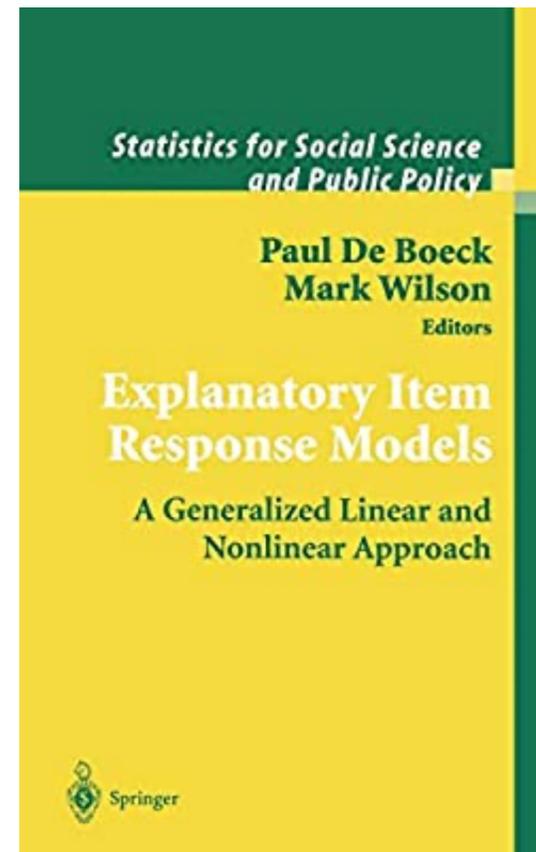
Обозначьте возможные внешние эффекты и опишите контекст.

Системное решение



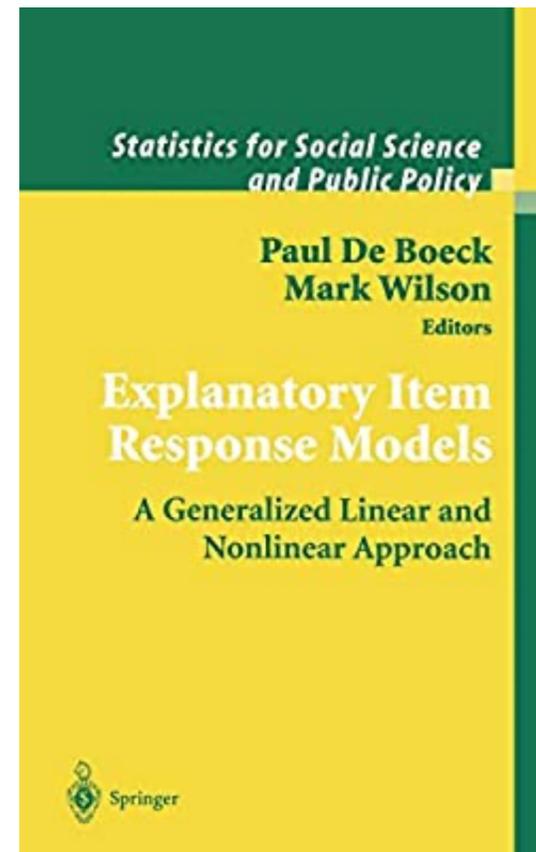
**Обозначьте возможные внешние эффекты и опишите контекст.
Мозговой штурм работает.**

Системное решение



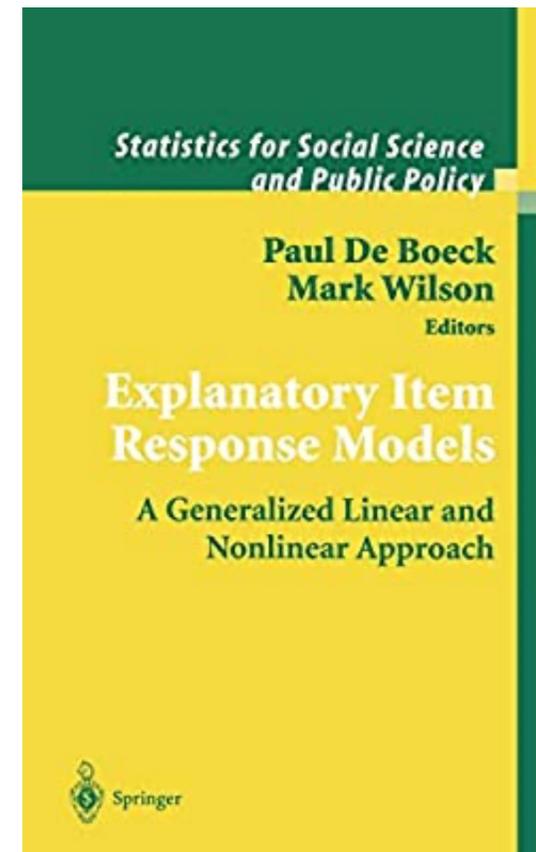
Используйте гибкие фреймворки для моделирования.

Системное решение



Используйте гибкие фреймворки для моделирования.
Пробуйте разные модели.

Системное решение



Используйте гибкие фреймворки для моделирования.
Пробуйте разные модели. **Делайте кросс-валидизацию.**

Нужен воркшоп по
этой теме?

Литература

- Углонова И. Л., Брун И. В., Васин Г. М. Методология Evidence-Centered Design для измерения комплексных психологических конструктов // Современная зарубежная психология. 2018. Т. 7. № 3. С. 18-27
- Abbakumov, D., Desmet, P., & Van den Noortgate, W. (2020). Psychometrics of MOOCs: Measuring learners' proficiency. *Psychologica Belgica*, 60(1), 115-131
- Van den Noortgate, W., De Boeck, P., & Meulders, M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics*, 28, 369-386.



Спасибо!

www.abbakumov.com