RANDOMIZATION

PMAP 8521: Program Evaluation for Public Service

October 14, 2019

Fill out your reading report on iCollege!

PLAN FOR TODAY

The magic of randomization

The “Gold” Standard

Running and analyzing RCTs

THREATS TO VALIDITY

Internal validity

Omitted variable bias

Trends

Study calibration

Contamination

External validity

Construct validity

Statistical conclusion validity

INTERNAL VALIDITY

Omitted variable bias

- Selection
- Attrition

Trends

- Maturation
- Secular trends
- Seasonality
- Testing
- Regression

Study calibration

- Measurement error
- Time frame of study

Contamination

- Hawthorne
- John Henry
- Spillovers
- Intervening events

THE MAGIC OF RANDOMIZATION

WHY RANDOMIZE?

Fundamental problem of causal inference

$$\delta_i = Y_i^1 - Y_i^0$$

Individual-level effects are impossible to observe

WHY RANDOMIZE?

Average treatment effect

$$ATE = E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

$$\delta = (\bar{Y}|P = 1) - (\bar{Y}|P = 0)$$

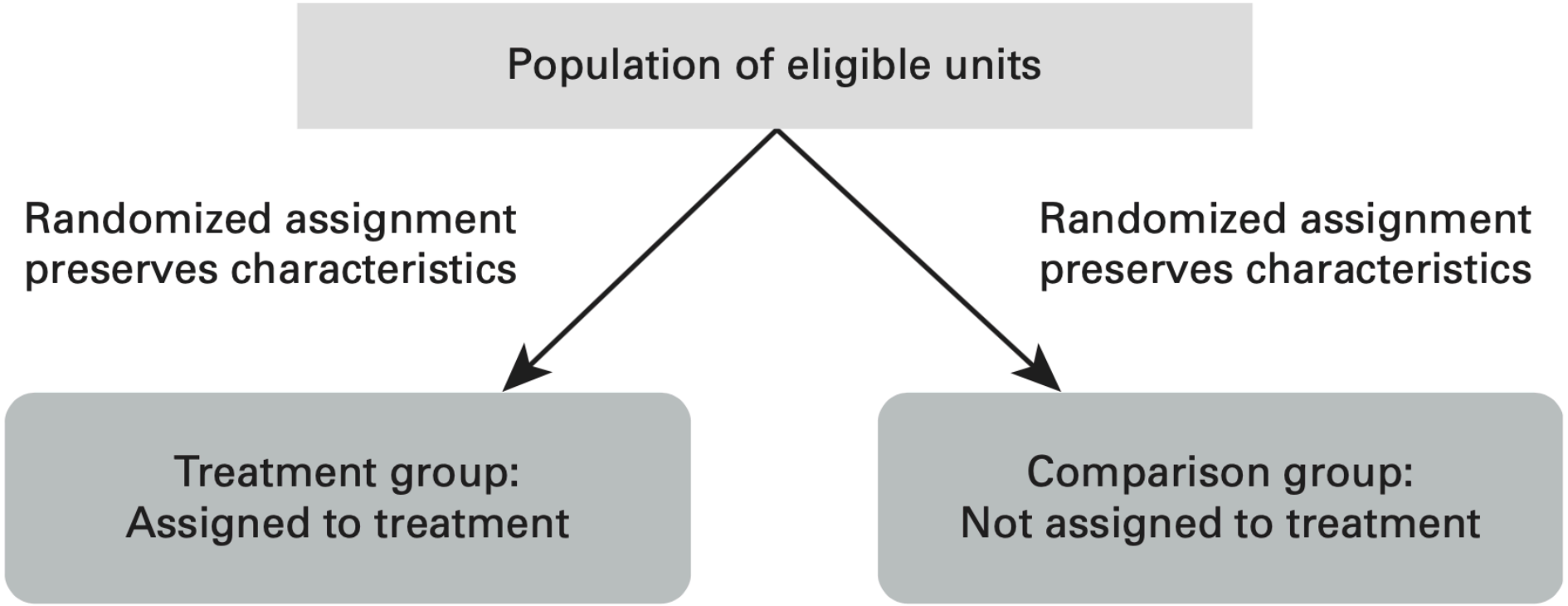
WHY RANDOMIZE?

$$\delta = (\bar{Y}|P = 1) - (\bar{Y}|P = 0)$$

This only works if subgroups that received/didn't receive treatment look the same

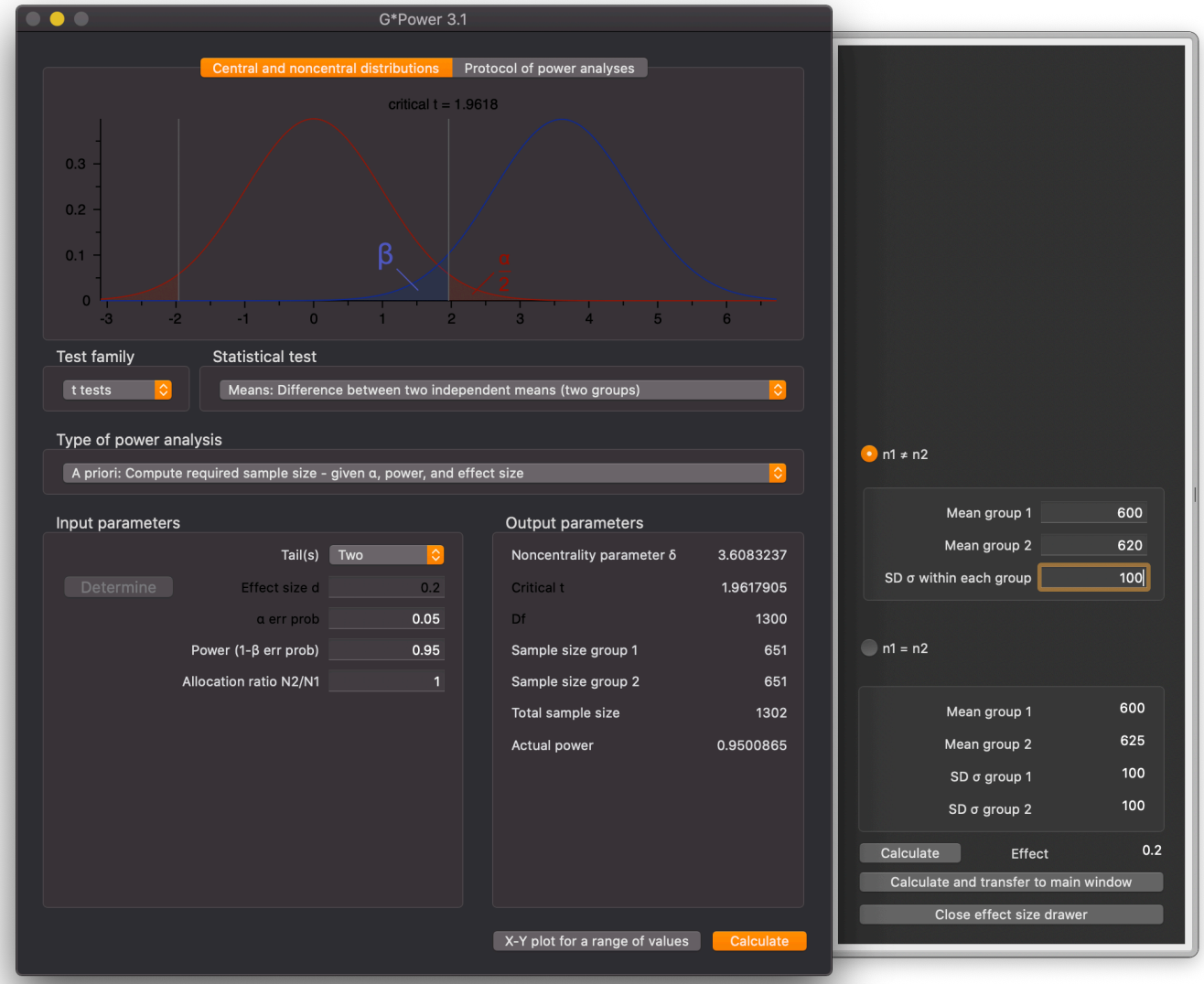
WHY RANDOMIZE?

With big enough numbers, the magic of randomization helps make comparison groups comparable



R example

How big of a sample?



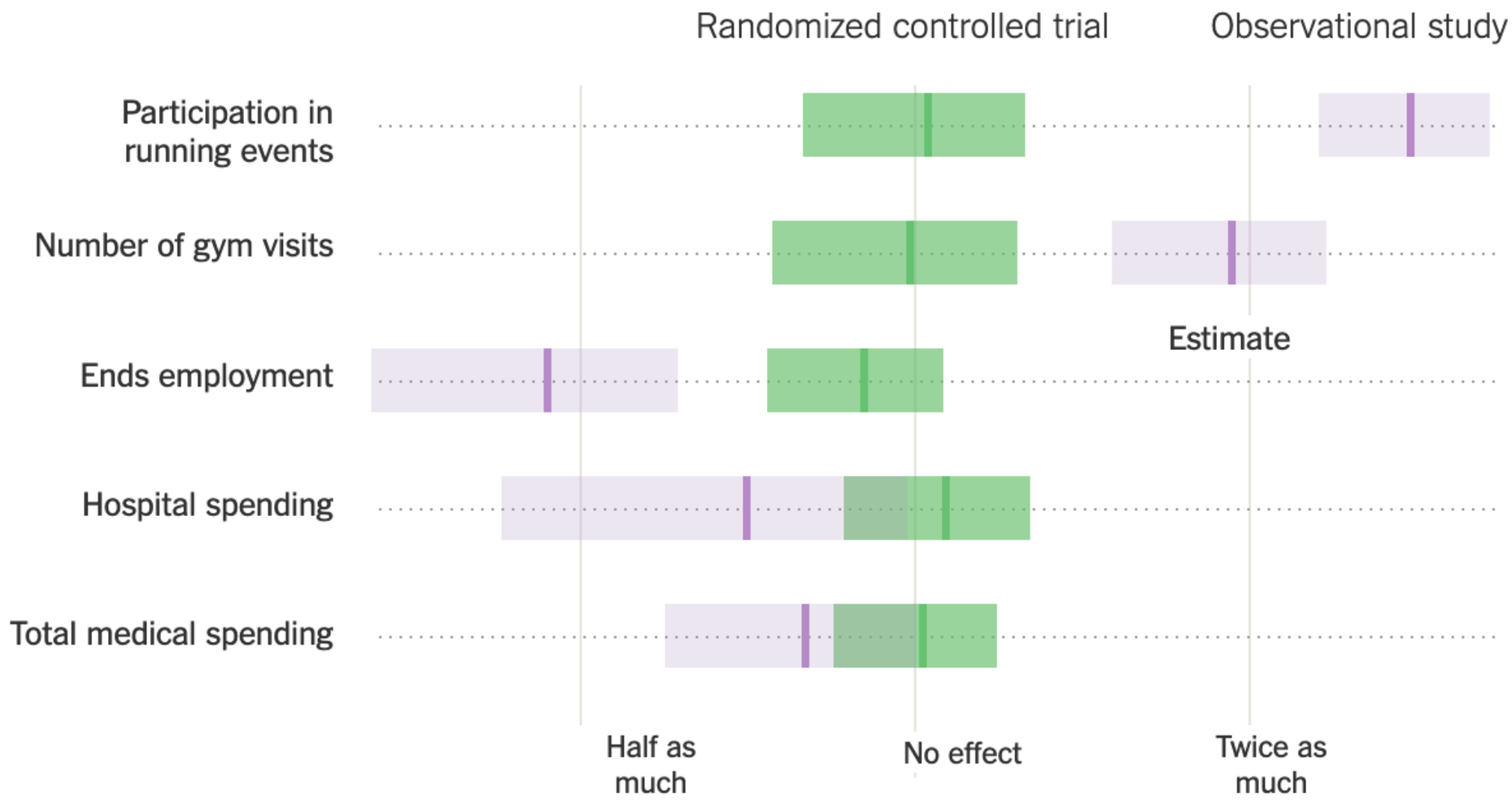
THE "GOLD" STANDARD

TYPES OF RESEARCH

Experimental studies vs. observational studies

Which is better?

How the Illinois Wellness Program Affected . . .



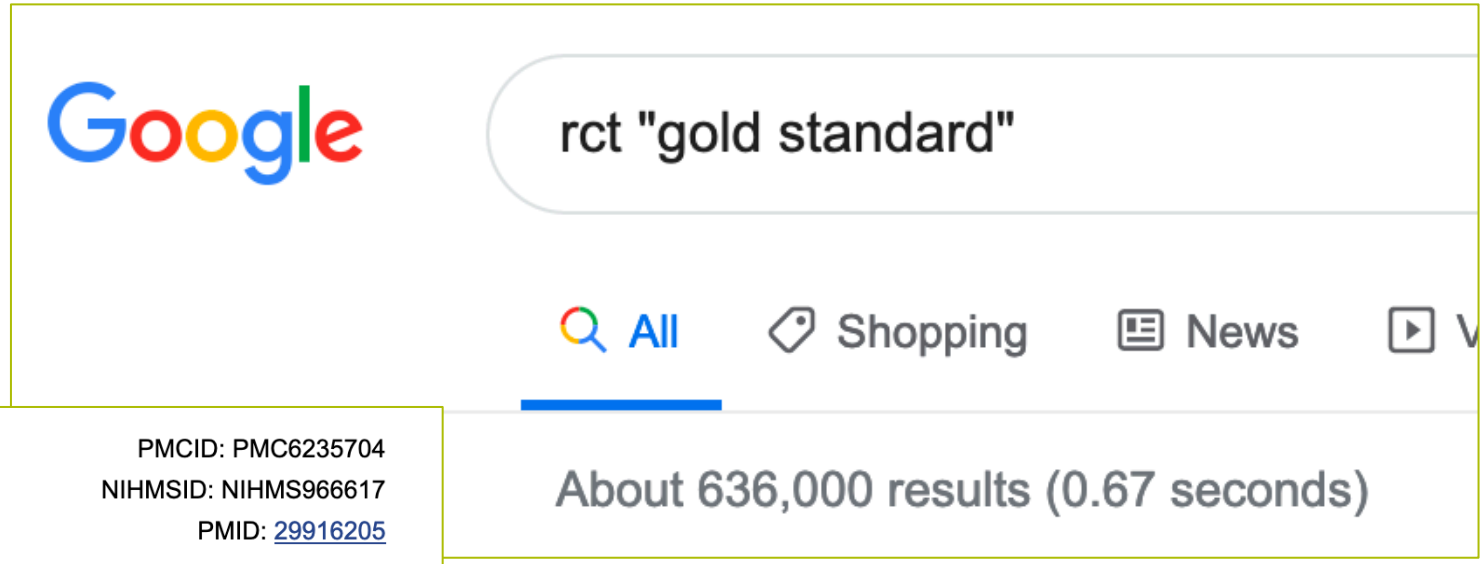
Source: What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study

TYPES OF RESEARCH

Experimental studies vs. observational studies

Medicine | Epidemiology | Social science

DAGs in RCTs?



BJOG. Author manuscript; available in PMC 2018 Dec 1.
Published in final edited form as:
BJOG. 2018 Dec; 125(13): 1716.
Published online 2018 Jun 19. doi: 10.1111/1471-0528.15199

PMCID: PMC6235704
NIHMSID: NIHMS966617
PMID: 29916205

Author Manuscript

Randomised controlled trials—the gold standard for effectiveness research

Eduardo Hariton, MD, MBA[1] and Joseph J. Locascio, PhD[2]

Randomized Assignment of Treatment

When a program is assigned at random—that is, using a lottery—over a large eligible population, we can generate a robust estimate of the counterfactual. *Randomized assignment* of treatment is considered the gold standard of impact evaluation. It uses a random process, or chance, to decide who is granted access to the program and who is not.[1] Under randomized assign-
ment, every eligible unit (for example, an individual, household, business,

RCTs are great!

Super impractical to do all the time though!

The Washington Post

Democracy Dies in Darkness

Business

3 share Nobel Prize in economics for ‘experimental approach’ to solving poverty

Esther Duflo, who at 46 is the award’s youngest winner, shares the hor
fellow MIT economist Abhijit Banerjee and Harvard’s Michael Kremer

J-PAL

ABDUL LATIF JAMEEL POVERTY ACTION LAB

Massachusetts Institute of Technology (MIT) @MIT · 5h

Professors Esther Duflo and Abhijit Banerjee, co-directors of MIT's @JPAL , receive congratulations on the big news this morning. They share in the #NobelPrize in economic sciences “for their experimental approach to alleviating global poverty.”

Photo: Bryce Vickmark

12 112 510

“Gold standard” implies that all causal inferences will be valid if you do the experiment right

We don’t care if studies are experimental or not

We care if our causal inferences are valid

RCTs are a helpful baseline/rubric for other methods

Moving to Opportunity

RCTS & VALIDITY

Randomization fixes a ton of internal validity issues

Selection

Treatment and control groups are comparable; people don't self-select

Trends

Maturation, secular trends, seasonality, regression to the mean all generally average out

RCTS & VALIDITY

RCTs don’t fix attrition!

Worst threat to internal validity in RCTs

If attrition is correlated with treatment, that’s bad

People might drop out because of the treatment, or because they got/didn’t get the control group

ADDRESSING ATTRITION

Recruit as effectively as possible

You don’t just want weird/WEIRD participants

Get people on board

Get participants invested in the experiment

Collect as much baseline information as possible

Check for randomization of attrition

RCTS & VALIDITY

Randomization failures

Check baseline pre-data

Noncompliance

Some people assigned to treatment won't take it; some people assigned to control will take it

Intent-to-treat (ITT) vs. Treatment-on-the treated (TTE)

OTHER LIMITATIONS

RCTs don’t magically fix construct validity and statistical conclusion validity

RCTs definitely don’t magically fix external validity

The Nobel Prize in economics goes to three groundbreaking antipoverty researchers

In the last 20 years, development economics has been transformed.
These researchers are the reason why.

By Kelsey Piper | Oct 14, 2019, 3:30pm EDT

Empiricism and development economics

The transformation of development economics into an intensely empirical field that leans heavily on randomized controlled trials hasn't been uncontroversial, and many of **the responses** to the Nobel Prize announcement acknowledge that controversy.

Critics have **complained that** randomization feels much more scientific than other approaches but doesn't necessarily answer our questions any more definitively. **Others worry** that the focus on small-scale questions — Do wristbands increase vaccination rates? Do textbooks improve school performance? — might distract us from addressing larger, structural contributors to poverty.

WHEN TO RANDOMLY ASSIGN

Demand for treatment exceeds supply

Treatment will be phased in over time

Treatment is in equipoise

Local culture open to randomization

When you're a nondemocratic monopolist

When people won't know (and it's ethical!)

When lotteries are going to happen anyway

WHEN TO NOT RANDOMLY ASSIGN

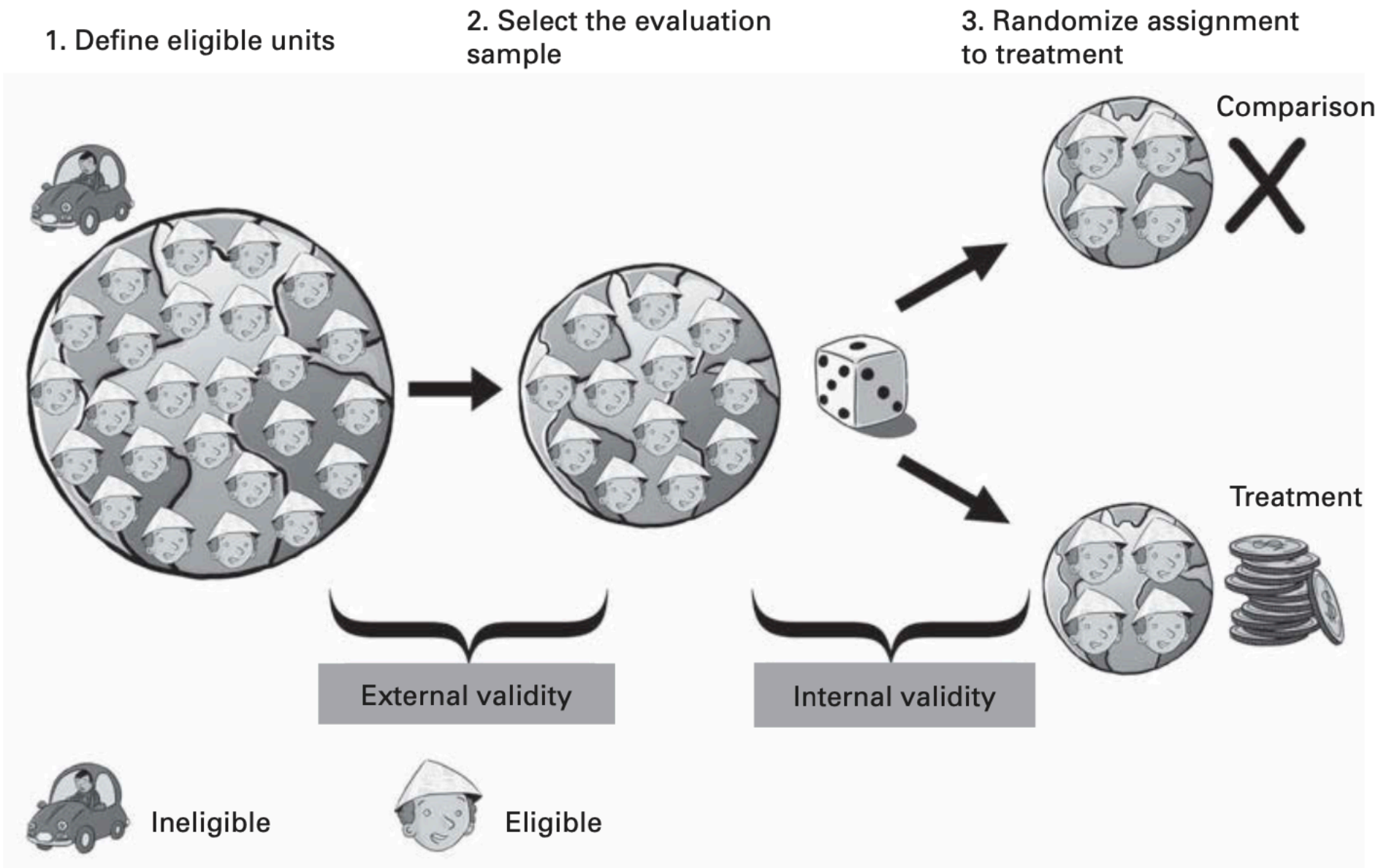
When you need immediate results

When it's unethical or illegal

When it's something that happened in the past

When it involves universal ongoing phenomena

RUNNING & ANALYZING RCTS


1. Define eligible units
2. Select the evaluation sample
3. Randomize assignment to treatment
Comparison
Treatment
External validity
Internal validity
Ineligible
Eligible

RANDOM ASSIGNMENT

Coins

Dice

Unbiased lottery

Random numbers + threshold

Atmospheric noise

random.org

R example

RCT with Qualtrics