Sep 28th, 10:00 AM - 10:20 AM

Program Evaluation: Diffusion from policy literature to improve assessment in information literacy instruction.

Seth Porter
Georgia Institute of Technology, seth.porter@library.gatech.edu

PRINCETON UNIVERSITY

What is Program Evaluation?

Outcomes not Outputs

CONTEXTUAL OVERVIEW

Program evaluation is, "a systematic application of scientific methods to design, implement, and improve the outcomes of programs".

Most importantly the systematic nature of program evaluation creates a framework for collection analyses of data that is used to measure the effectiveness and outcomes of a specific program, treatment, or service.

education

Public Policy

WHY

Program evaluation is used by government agencies, non-profit organization, and non-governmental agencies to help provide quality information to policy makers, practitioners, administrators, and other stakeholders to assist in decision making, improve processes and behaviors, improve failing programs, and to understand resource allocation and the **outcomes of their inputs.**

WHEN

Understand, verify or increase the impact of products or services on customers or clients.

Improve delivery mechanisms to be more efficient and less costly - Over time, product or service ends up to be an inefficient collection of activities that are less efficient and more costly than need be. **Evaluations can identify program strengths and weaknesses to improve the program.**

Program evaluation can facilitate management's thinking about what their program is all about, including its **goals,** how it meets it goals and how it will know if it has met its goals or not.

Produce data or verify results that can be used for public relations and promoting services in the community.

Produce valid comparisons between programs to decide which should be retained, e.g., in the face of pending budget cuts.

TOOLBOX

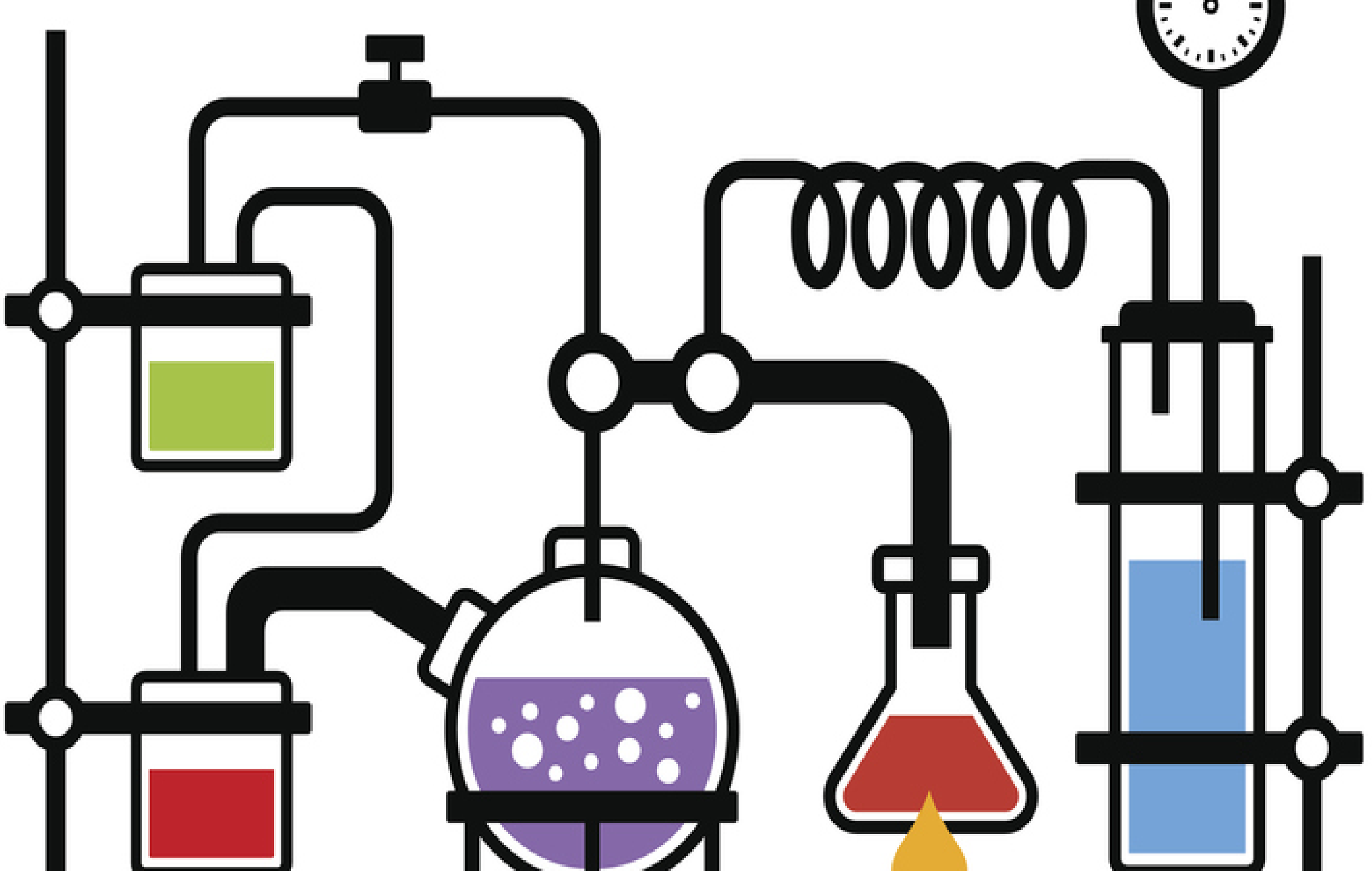
Experimental Design (RCT)

The Gold Standard

Quasi-Experimental Design

Cost-Effectiveness Analysis

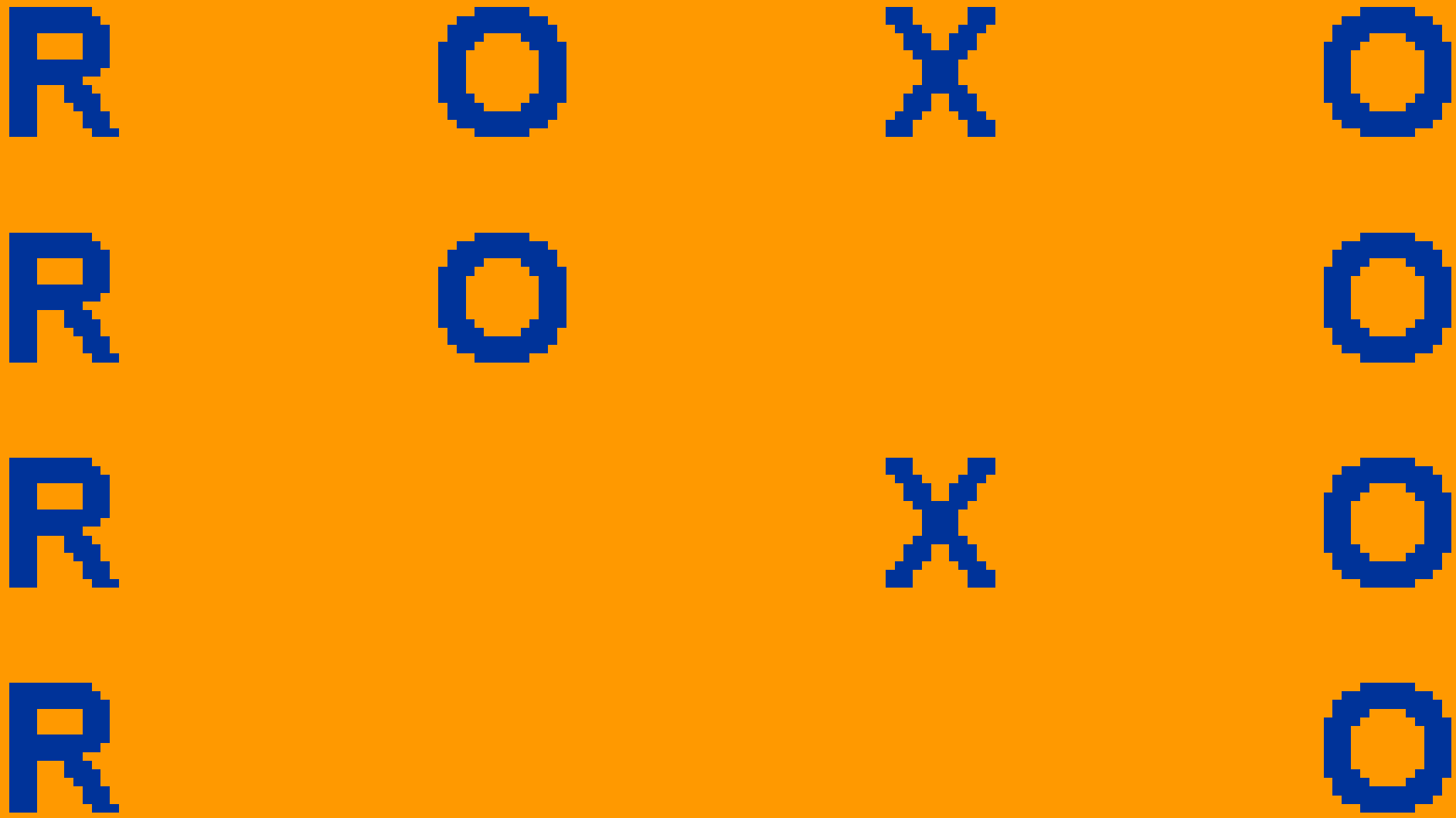
A/B Testing




14. OA ●B OC OD
15. ●A OB OC OD
16. OA ●B OC OD
17. OA OB OC
18. OA OB OC OD
19. OA OB OC OD

At the beginning of a student's freshman year they are randomly assigned to a treatment or control group. After random assignment, both groups of students will take a pre-test measuring their foundational knowledge of a certain topic, e.g., statistical literacy and decision making based on chosen peer-reviewed articles.

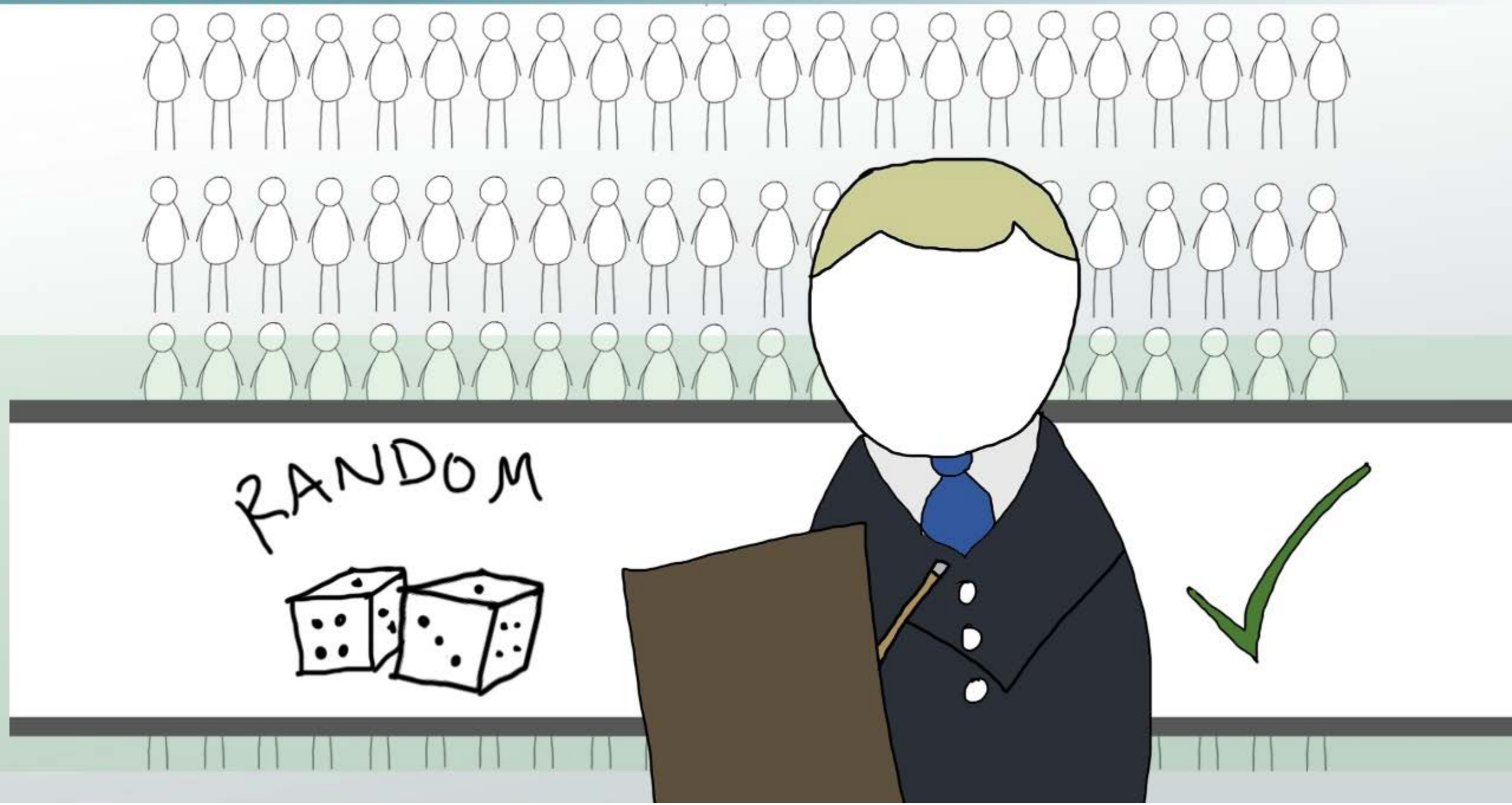
The treatment group would then go through either a one-shot instruction workshop teaching the best practices in analyzing data for statistical significance or a semester long course. At the end of the treatment both groups would again be tested on their knowledge of the topic. The two results would be compared.


R O X O
R O O
R X O
R O

Imagine four of the same courses that traditionally have academic librarians come into teach a two-hour long workshop on how to access, analyze, and implement research data into a semester long project.

Some of the classes will not get a pre-test, and some will not be taught this semester, but will have a pre-test. All the classes must be chosen at random as well as their classification of control or treatment.

ADVANTAGES AND DISADVANTAGES



If the pretest-post text randomized control trial was needed on a program underway, but two different sets of students groups were not chosen at random, the evaluators could still do a pretest-posttest and gain valuable information from the data, assuming the sample size was not invalid based on the artificial creation of the control and the treatment.

The comparison between the control and the treatment would still be a useful benchmarking for the outcome of the program; it just would not be as generalizable as the true experimental design.

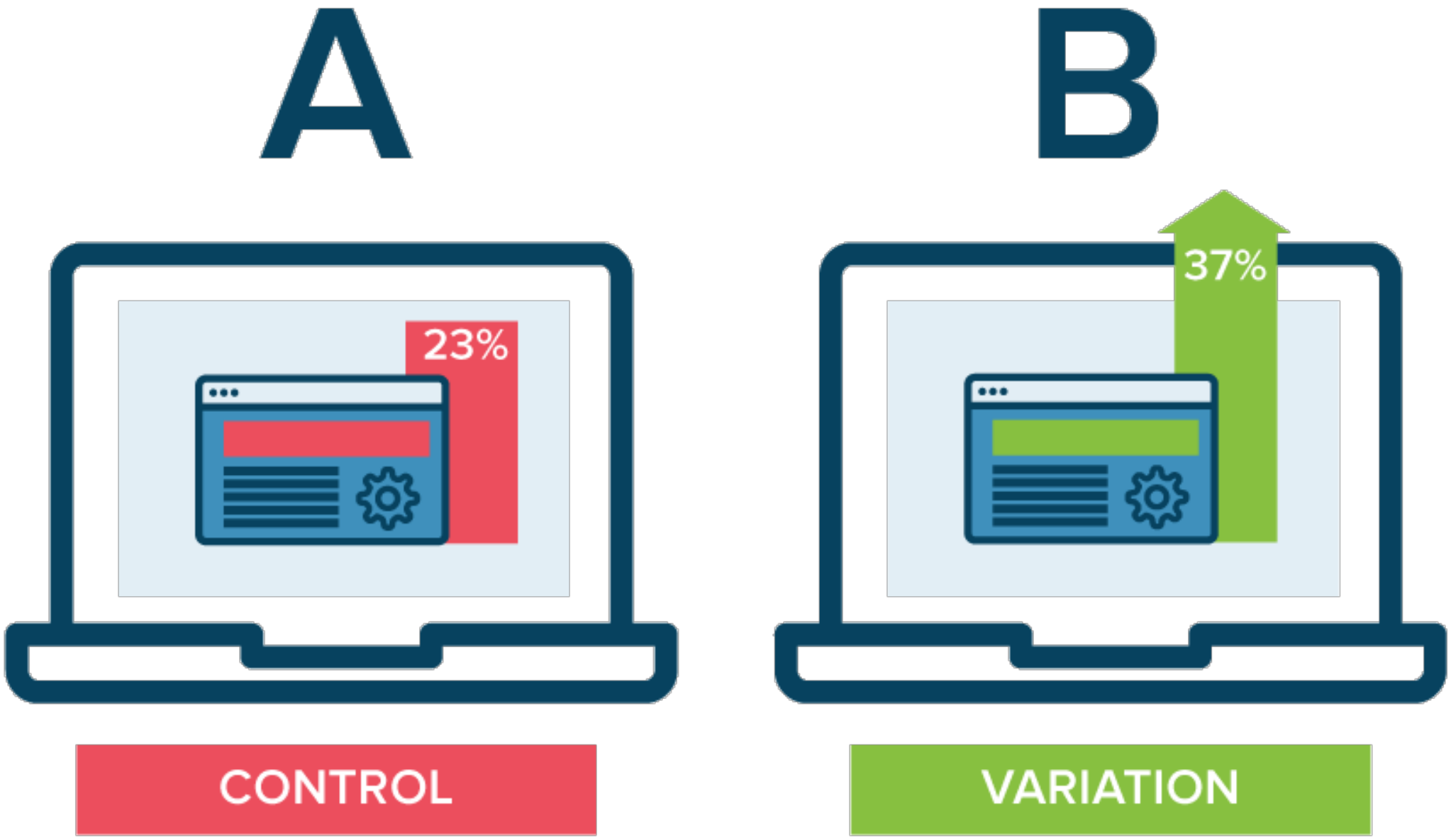
Attempt to replicate the Gold Standard with artificial groups


ERAL RESERVE NOTE
15155003 A
UNITED STATES · FEDERAL RESERVE SYSTEM
THIS NOTE IS LEGAL TENDER
ALL DEBTS, PUBLIC AND PRIVATE
Treasurer of the United States.
SERIES 2003
FRANKLIN
DF
H 12

Cost Effectiveness Analysis. A cost effectiveness analysis (CEA) is a natural fit within academic library assessment. It is a tool used to understand the resource allocation through the projected output and testing of a specific program (Bingham & Felbringer, 2002).

Essentially a cost effectiveness analysis analyzes the potential implementation of a program based on the comparison of potential needed resources and what the expected outcomes of the program will be. The value of a cost effectiveness analysis in an academic library is that you can use tools to measure holistic impacts and benefits that do not have an economic or financial cost (Metz, 2007).

Compare inputs to expected outputs and capital involved


A
B
23%
37%
CONTROL
VARIATION

A/B testing is very similar to an experimental or quasi-experimental design but it is used distinctly in virtual program evaluation and software assessment. Essentially, A/B testing or Bucket testing is an experiment where two or more webpages, digital learning objects, or apps, are tested against each other to evaluate the top performer (Optimizely 2017). This can be done with two completely different versions or small changes like a different font style, color, or content placement. The original version of the page, app, or digital learning object, would be the control and the updated version is the treatment. After running both versions the evaluator can collect the data through an analytics framework and compare the control and the treatment in the A/B test.

A library website that was newly designed and was on beta-test, but the organization was unsure if it was the best possible version and the most user friendly.

Create a few similar websites but with small differences, e.g., placement of the library catalog and discovery search, or naming the catalog something less technical like, “get books”. The designers would then run each version of the site for set amount of time. Following the test the stakeholders could analyze the differences in the use of the sites and if there was a drastic difference this could help guide the decision on the best version to use.

Use on Digital Learning Objects


EVALUATION

CURRENT USES

A/B testing similar classes taught by my librarians vs. similar classes not taught; and recording reference consultation statistics from each course.

CEA on outreach methods.

Solomon four-group design on course integrated instruction.

Assessing the content and methods of my librarians.

PRINCETON UNIVERSITY

Discussion

PREVIOUS WORK

- *Porter, S (2018). "Using Program Evaluation as a Proxy for Assessment: Diffusion from Policy Literature to Improve Academic Program Assessment".*
  - *Academic Libraries and the Academy: Strategies and Approaches to Demonstrate Your Value, Impact, and Return on Investment*, edited by Marwin Britto and Kirsten Kinsley.

- *Porter, Seth and Frizzell, Matthew (2018) "Assessing Instructional Initiatives and Services through Program Evaluation," Georgia Library Quarterly: Vol. 55 : Iss. 2 , Article 10.*

PRINCETON UNIVERSITY

Seth M. Porter, MA, MLIS

Head of Donald E. Stokes Library for Public and International Affairs & The Ansley J. Coale Population Research Collection,

Woodrow Wilson School Liaison

Donald E. Stokes Library

057 Wallace Hall

Princeton, NJ 08544

T 609.258.4782

http://bit.ly/2N1t72t

E smporter@princeton.edu


The End