Introduction to Experimental Research

William Bishop, Autumn Perkey, and Alauna Safarpour
University of Maryland, College Park
GVPT Methods Workshop
April 19, 2021

Outline

I. The Logic of Experiments
II. Types of Experiments
III. Analyzing Effects
IV. Performing Randomization checks
V. Manipulation Checks
VI. Attention Checks
VII. Select Advanced Issues: using covariates (CATEs); block randomization.
VIII. Practical Steps
    A. IRB
    B. Pre-registration
    C. Power-analyses
    D. GVPT Experimental Lab
    E. Survey Firms
    F. Other Fielding Options
    G. Funding Sources

The Logic of Experiments

Establish causal relationships using three key elements:

1. A Planned Intervention
    a. In contrast to observational research, experimental research involves the manipulation of some intervention (independent variable), rather than observing (and tracing) the independent variable and trying to use statistical controls to exclude competing explanations.
2. Random Assignment
    a. Crucial: In order to rule out other competing explanations, we randomly assign subjects to treatment or control conditions. This way, the only differences between groups is the treatment and chance variations (which on average, cancel out).
3. Careful and Precise Measurement

Types of Experiments

Between subjects design: Simplest method involving establishing a treatment effect by examining differences between treatment and control subjects after some experimental intervention. If randomization is successful, then any differences observed can be attributed to the treatment. Most experiments are between subjects designs.

Within subjects design: Pre-tests for all subjects enable a comparison to the same person at an earlier point in time as well as a comparison to control subjects conditions. If you can, it is usually better to take a within subjects design. This type allows for more precise estimation of treatment effects, and added CATEs (e.g. by baseline measures of your dependent variable).

Types of Experiments (cont.)

Quasi Experiment (sometimes called Natural Experiment): Typically lacks full experimental control over the random assignment because you are taking advantage of an as-if random assignment that occurs in the real-world. Typically used to study political phenomenon that it is impractical or unethical to study using a typical experimental design.

Lab experiments: Great design for internal validity due to maximum control, but faces problems of external validity and generalizability when studying convenience samples (usually college students). Typically used as a pilot study.

Survey Experiments: One of the most common type used in social science research. Usually involves presenting respondents with an experimental stimulus or control during an online survey (this can also be done by phone, in-person, and through mail surveys . Typically used to test hypotheses on candidate messaging and characteristics, news/information, and other attitudinal and behavioral outcomes.

Field: Experiments conducted in real-world settings. Help address concerns of external validity inherent in laboratory experiments. If you can, this is almost always your best option.

Additional Subtypes of Experiments: MANY subtypes of survey experiments most common are list and conjoint. List typically used to reduce social desirability (but see Blair, Coppock and Moor 2020 for whether it is even necessary). Conjoints vary multiple characteristics of a stimulus at once.

Analyzing Effects

Average Treatment Effect (ATE) = Difference in means:

Mean Outcome in Treatment– Mean Outcome in Control: $\bar{y}_{treatment} - \bar{y}_{control}$

Estimating ATE with regression:

Y=Constant+ Treatment: $y = \beta_0 + \beta_1 \text{ Treatment}$

where Treatment=1 if assigned to treatment, 0 if control.

ATE with Multiple Treatments:

Mean Outcome in Treatment A- Mean Outcome in Control: $\bar{y}_{treatmentB} - \bar{y}_{control}$

Mean Outcome in Treatment B- Mean Outcome in Control, etc: $\bar{y}_{treatmentA} - \bar{y}_{control}$

Even more complicated options are simple. E.g. If you used block randomization, you include dummies for the blocks to the regression. Same with covariates. Proper setup eliminates the need for complicated statistics. Did you cluster randomize by households? Cluster the SEs at the household level.

Conjoint analysis is slightly more complicated, sample code available upon request (analysis and setup available).

Performing Randomization Checks

Purpose: To ensure your randomization was successful.

Caveat: It is practically quite rare that randomization is unsuccessful, but finicky reviewers will sometimes want to see this. Also, certain journals will require it (e.g. JEPS).

How: For a basic experiment with two conditions (treatment/control) run a binary response regression (logit/probit) to predict assignment to the treatment condition using available demographic indicators (e.g. age, gender, education, race, income, etc). Report chi-squared and details on the indicators that fail to predict treatment assignment in a footnote.

Manipulation Checks

Purpose: To ensure that you actually manipulated your IV (ensure that subjects receive your treatment)

How: Asking factual questions about the treatment; Asking attitudinal questions to gauge the level at which your manipulation was successful.

Attention Checks

Purpose: To measure how much attention respondents are paying to your survey questions.

How: Give them instructions detailing how to respond to a survey question. You can tell whether they were attentive or not based on their response.

For more information see: Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2021). Using screeners to measure respondent attention on self-administered surveys: Which items and how many?. *Political Science Research and Methods*, *9*(2), 430-437.

Attention Check Example

Most modern theories of decision making recognize that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. To demonstrate that you've read this much, just go ahead and select both red and green among the alternatives below, no matter what your favorite color is. Yes, ignore the question below and select both of those responses.

What is your favorite color?

1. White
2. Black
3. Red
4. Pink
5. Green
6. Blue

Some advanced issues: Using Covariates

Purpose: 1) Increase the efficiency of your estimates by reducing variance 2) Compute conditional effects (CATEs).

Requirements:

They must be collected pre-treatment

Must be related to your dependent variable. Don't throw in all available demographic variables, ideally the ones you select will be either theoretically derived and based on prior empirical research. Remember, this is an experiment, not observational research so your use of covariates should be limited and purposeful.

Often Preferred:

Pre-registration

Sources: (Gerber and Green 2012; Mutz and Pemantle 2015).

Practical Steps

I. IRB
II. Pre-registration
III. Power analyses
IV. GVPT Experimental Lab
V. Survey Firms
VI. Other fielding options
VII. Funding Sources

Practical Steps I. IRB

- Ethical Review Required for All Human Subject Research
- CITI Training Required for Certification (online)
- Best to give yourself between 3 weeks–1 month before desired field date

https://research.umd.edu/irb

Practical Steps II. Pre-Registration

- Purpose: increase transparency, decrease fishing/ p-hacking, and assess publication bias.
- Growing in popularity and now required for some journals, e.g. JOP
- Involves brief statement on theory, hypotheses, methods, and expected analytical procedures
- Report filed (usually) prior to data collection

https://egap.org/registry/

For more information see: Blair, G., Cooper, J., Coppock, A., & Humphreys, M. (2019). Declaring and diagnosing research designs. *American Political Science Review*, *113*(3), 838-859.

Practical Steps III. Power Analysis

- Purpose:
  - Determine required sample size to detect statistically significant effects at some threshold of confidence to avoid committing a Type II error.
  - Can also be used to calculate the probability of detecting an effect given a certain sample size and effect size at a specified level of confidence. Can tell you whether your budget is large enough to make the experiment worthwhile.
- Software:
  - Stata has an easy interface for this (Statistics-> Power and sample size)
  - Packages in R: e.g. pwr
  - Also several online calculators.
  - More complex analyses may require simulation (e.g. interactive effects).

Practical Steps IV. GVPT Experimental Lab

I. Submit proposal to Antoine Banks (abanks12@umd.edu)

1. The rationale for the project and its contribution to the literature

2. The research question and specific hypotheses to be tested

3. A description of the experimental stimulus

4. A description of the analytical strategy (Method of data analysis; dependent variable; independent variable(s); specific links between hypotheses and analysis to be conducted)

I. Apply for IRB approval
II. Code survey into Qualtrics
III. Field survey if approved

Practical Steps V. Survey Firms

Firms: MTurk, Lucid, Qualtrics, SSI, YouGov, TESS

Procedures:

- Talk with faculty to get UMD contact person.
- ALWAYS note that you are a graduate student (can typically get a discount).
- Often helpful to note how much you have to spend, required sample size, and rough details on survey length, etc.
- Most firms you create the survey yourself in Qualtrics software, they field. Although some (e.g. YouGov) will program survey for you using their proprietary software.

Trade-offs:

- Cost
- Sample type and quality

Practical Steps VI. Other fielding Options

- Self sourced list of participants (e.g. emails of local elected officials/ journalists/ NGO leaders)
- Location Specific Participants (e.g. GOTV experiments, randomly selected city streets, houses on streets)
- Partnerships
  - States
  - NGOs
  - Civic organizations
  - Schools
- Critical Issues Poll (Shibley Telhami and Stella Rouse)

Practical Steps VII. Funding Sources

Internal (UMD) Options:

- Jean Elizabeth Spencer award (GVPT)
- Mabel S. Spencer award (Graduate School)
- Graduate School Summer Research Fellowship
- BSOS Dean's Research Initiative (pre and post candidacy options)
- GVPT Departmental Awards (Conley Dillon; Roger Davidson; Don Piper)
- Longest Award (Graduate School)
- Pelczar Award (Graduate School)
- 3MT (Graduate School)

External:

- APSA (NSF) Doctoral Dissertation Research Improvement Grant
- Cosmos Club Foundation
- Centennial Center Research Grant

Resources (Literature)

Druckman, J.N., Greene, D.P., Kuklinski, J. H., & Lupia, A. (Eds). (2011). *Cambridge Handbook of Experimental Political Science*. Cambridge University Press.

Field, A., & Hole, G. (2002). *How to design and report experiments*. Sage.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Kinder, D. R., & Palfrey, T. R. (Eds.). (1993). *Experimental foundations of political science*. University of Michigan Press.

Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The Sage encyclopedia of social science research methods*. Sage Publications.

McDermott, R. (2002). "Experimental Methods in Political Science." Annual Review of Political Science, 5(1), 31–61. https://www.annualreviews.org/doi/pdf/10.1146/annurev.polisci.5.091001.170657

Mutz, D. C. (2011). *Population-based survey experiments*. Princeton University Press.

Resources (Misc.)

Government and Politics Experimental Lab: Director Antoine Banks, current-
RA: William Bishop

APSA Experimental Research Organized Section (Section 42):
https://connect.apsanet.org/s42/

Journal of Experimental Political Science