

# Experimenting in Security Studies

22.4. 2022

# Outline of the Lecture

- Experimenting in Political Science and Security Studies
- Epistemology of Experiments
- Practical Considerations
- Security Studies Experiments
- Example of SS Experiment: „CyberTrain“
- Assignment Paper Options

# Experiments in Political Science/Security Studies

- The first political science experiment: Harold F. Gosnell. 1926. An Experiment in the simulation of voting. *American Political Science Review*, 20(4), 869-874.
- With few exceptions, no progress has been made for a long time.
- Exception: 1970s Experimental Study of Politics
- Breakthrough in the 1980s: Kinder and Iyengar (News that Matter)
- Progress since the 1990s"" , now becoming mainstream- 2010: Cambridge Handbook of Experimental Political Science, 2014: Journal of Experimental Political Science, NYU Annual Conference,
- Becoming more common in some areas of research, fuelled by methodological debates, experimenting becoming mainstream method

# Security Studies

- Increase since 9/11, still underrepresented as a method
- „Policies for managing dynamically evolving threats“, „Politics of Competence“
- Threat appraisal x coping appraisal (individual level), appraisal processes x state countermeasures (homeland security)
- „Libertarian Paternalism“ (Thaler-Sundstein 2008)
- „Experimental State“
- Fuelled by problem-solving applied research designs, still not a mainstream methodology
- Beale-Thompson 2017, Chytilek et al. 2022

# Areas of research: use of experiments in SS

Security taken to the individual level

- Threats, Risks (Terrorism, War- Risk Appraisal)
- Cybersecurity
- Security Measures (Coping Appraisal)

# Why experiment?

Exploring **CAUSALITY**

*„influence by which one event, process, state, or object (a cause) contributes to the production of another event, process, state, or object (an „effect“)*

- **Creates** and **compares** different states of the world
- The world in which the intervention took place and the world in which the intervention did not take place (or a different interventions took place)
- Intervention into the „data generating process“

# Key elements of experiments in the SS

Intervention into the DGP by the researcher

Manipulated variable

Agreed good practice of reporting

Random assignment of subjects (harder to meet in field experiments)

(experimental- and control groups, pre- and post measurements **no more** key elements)

# When we manipulate the „variable“

- ,we (researchers) decide when, how and at which values this variable will happen
- Example: We are interested how negative information affects support for particular policy
- Variable values: „1.received no information, 2.received negative information“ through newspaper article



# When we make random assignment

- we randomly assign our experimental subjects to either „no information“ or „negative information“ variable value.
- Some of them are negatively informed about the policy, while some receive no information, none of them is in both states of the world.
- Each subject has **the same chance** to be assigned to any variable value.
- **Random assignment** does not equal **random sampling!**

# When to experiment

- Experiments: costly, difficult to administer and get the results published (some potential in applied research however).
- It is thus nearly always better **not to experiment in the SS** except for instances when the elements of interest are:
  - too diffused/hard to identify to be measured in observational studies (too few observations)
  - hard to isolate/distinguish from other elements present in observational studies (too much complexity)
- Unfortunately both situations not uncommon 😊
- Analogy: states as cases, case selection with regard to the values of independent variables and control, Lijphart: „quasi-experimental method“.

# Types of experiments

	LABORATORY	ONLINE/SURVEY	FIELD
<b>Location:</b>	Lab	Online Survey	Real world
<b>Artificiality</b>	High	Medium to Low	Low
<b>Control</b>	High	Medium	Low
<b>Typical Use</b>	Wide Range of Topics	Framing Experiments, Manipulated Variable: Information	Policy Experiments
<b>Conduct</b>	Difficult	Easy	Difficult
<b>Ethical Issues</b>	Controlled	Sometimes problematic	Often problematic

# Let me demonstrate this idea

Suppose you are HRC, you are running for President and you would like to hit DJT with a following ad:

<https://www.youtube.com/watch?v=RaxNEzA3jRs>

At the same time, you would ask:

**„Does really such negative advertising (independent variable) increase chances of winning the election (dependent variable) for the person who produced it?“**

# The problem with observational research

Imagine we made the ad and started airing it. Then we would look for people in the population who have seen it and who haven't and ask them how likely they are to vote for HRC. We would then attribute the difference between the two groups to the clip. What do you see as the problem(s) with this „research design“?

# Experiments and „Four Causal Hurdles“

- First (mechanism): it is no easier or harder than in non-experimental research
- Second (sequencing): the dependent cannot cause the independent variable, timesequencing problem, and the independent variable is "caused" by us randomly.
- Third (correlation): we calculate the correlation between the independent variable and the dependent variable in a similar way to observational studies
- Fourth (third variables): the main advantage of experimental studies. We say that in experiments we **control the influence of third variables**".

# The „Fourth Causal Hurdle“ in Experiments

We want our effect (the difference between groups with different treatments) is not caused by the effect of/by some third variable (e.g., education or different fatigue of the subjects).

In a good experiment, **all third variables** are **equally distributed** across groups **due to random assignment** (groups are equally educated, equally tired, equally left-wing, equally mentally healthy, equally immortal and equally brilliant, have the same everything we can and couldn't think of).

This doesn't mean, of course, that third variables don't exist, it's just that they operate the same manner in both groups!

# When establishing causality, we are interested in the data in

- **Average treatment effect:** the difference in average support for HRC between the groups
- **The heterogeneity measures:** the variability of support within- and between groups



# Causality: Limitations of experiments

It seems that experiments are a great tool to determine causality because they overcome the four causal hurdles and we can say that independent variable causes the dependent variable. That's true, but but there are two significant limitations:

- It's **probabilistic causation**.
- It has **uncertain contiguity**.

# Limitations to establishing causal relationships in experiments

- **Uncertain contiguity:** usually we don't know when **exactly** the cause will have an effect. Does, e.g., the HRC ad exposure should make impact immediately or with some delay?
- **Probabilistic causation:** causes influence (raise, decrease) the probabilities of their effects, all else being equal, not „determine“ them („deterministic“ causation)

# Benchmarks for experimental good practice

- Internal Validity: inner quality of our research
- External Validity: generalizability of conclusions
- Realism (Ecological and Experimental)

# Internal Validity: Components

- **Statistical** (enough subjects to make statistical inferences, depends of number of experimental conditions, lower burden 200 subjects)
- **Causal** (ability to establish causal relations)
- **Construct** (experimental manipulations must correspond to the teoretical concepts of interest)

# External Validity

- Generalizability **beyond** the study time/space/population
- The theories are generally valid. They are also generally valid the results of the experiment?
- - Do the exp. subjects correspond to the real world? Does the treatment match the stimuli in the real world? Does the experimental context match the context in the real world?
- Achilles heel of experimental research? Political science is obsessed with external validity(Rose McDermott 2010).

# Establishing External Validity

- External validity must be established on the basis of empirical evidence
- EV is a matter of **SCIENTIFIC REPLICATION** to testing the same theoretical constructs
- On a new population
- On a new sample
- Using a different setting for specification of environmental conditions
- Stress test: change in experimental protocol without having a theoretical prediction of what it will mean. If the assumption of the original the theory doesn't hold, the theory is not robust.
- „Replication Crisis“ in SS

# „Realism“ in/of Experiments

- **Ecological Realism** (Connected to both External and Internal Validity):
  - Does the experimental environment match the real-world environment?
- **Experimental Realism** (Connected to Internal Validity):
  - Can the experiment induce in the subjects those processes that the theory predicts?
  - Are the subjects sufficiently motivated?
  - Is the intended manipulation strong enough to elicit what we want?

# Pretreatment

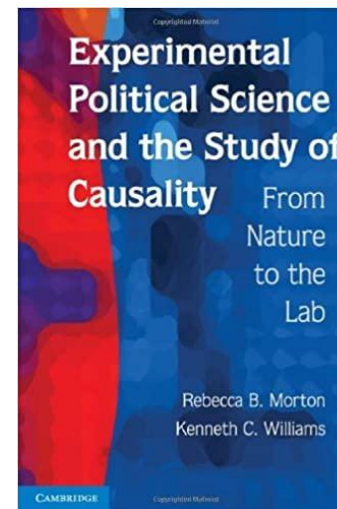
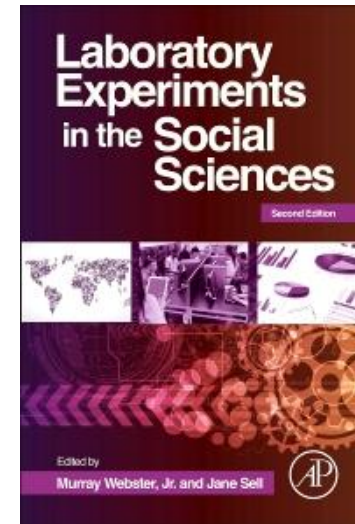
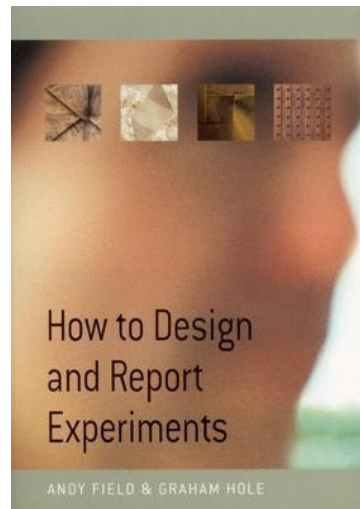
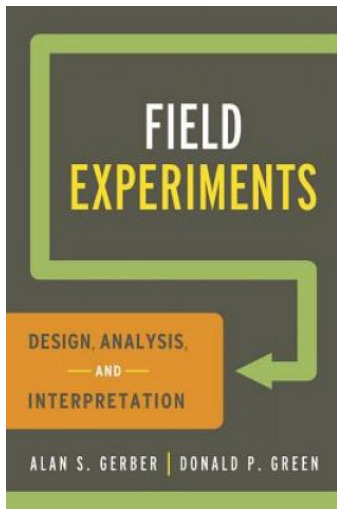
- Vexing problem for security experiments!
- Some/all subjects **receive manipulation before our experiment**, even those for which we **intend no manipulation**
- *Imagine we want to measure support for some restrictive policy and the manipulated variable are keywords that are supposed to activate "libertarian paternalism". The problem is that these words may have already been heard by the participants in the control group experiment in real world, libertarian paternalism has already been activated for them, and we will not detect a difference between the groups. At the same time, our stimulus does indeed activate libertarian paternalism, it just had already activated it in all people before our experiment.*



# Practical Issues: Conducting Experiments

- Conduct of Experiments
- Ethics, Deception

# Suggested Manuals for Conducting Experiments



# Experimental Protocol

- Script for designing and reporting experiments
  - Theories and hypotheses
  - Instrumentation
  - Population, sample, assignment to experimental conditions
  - Implementation
  - Analysis

# Theories and Hypotheses

- Precise specification of the dependent and independent variables (including what values the IP can/in the experiment will have)
- References to previous research in the field
- Overview of hypotheses based on the theory
- Reflection on how the theory and hypotheses have influenced what is in the experiment

# Instrumentation

- operationalisation of theoretical concepts
- experimental instructions
- method of measuring the values of the dependent variable,
- the decision of which third variables to measure
- pretest, pilot testing
- deciding what medium to use to expose the subjects to the experimental stimulus and through which through which medium the dependent variable will be measured

# Pretest, Pilot Study

- **Pretest:** the subjects act as informants/ assessors of the quality of our research procedures.
- In small-sample we check whether subjects interpret the situation in such a way that our manipulation is consistent with the theory to which we refer.
- Example: we want to verify whether the support of a repressive policy is somehow affected by how drastically the "threat" (NP) is presented. Subjects in the pretest evaluate the experimental material (photographs) and whether the "drasticness of the threat" (its level) is well operationalized.
  
- **Pilot Study:** A complete experimental session, followed by a thorough debriefing with participants, focusing on the progress of the experiment
- Helps to uncover problems that are not encountered in pre-tests because these often only focus on a subset of the experiment
- Examples: subject fatigue, emotions involved related to the instructions, missing instructions, poor logical consistency of the tasks

# Sample, Population, Random Assignment

- It is necessary to consider who the population is and how will the sample from it be recruited
- Typically, inviting selected individuals from the database preregistered subjects (report number of invited, enrolled, those who currently arrived)
- Need to think about the mechanism of random assignment (randomizer!?)

# Conduct of an Experiment (I.)

- When, where and (above all) how?
- Key question: timing of procedures/parts of the experiment.
- Usual scenario: welcoming information, „warm-up questions“, then the main tasks and things that might potentially embarrass the subjects (e.g. measuring political knowledge) at the end.
- In economic experiments, "practice rounds".
- We should have a plan ahead of time if things (in the lab) "start to go wrong" (e.g., someone starts to interfere)



# Conduct of an Experiment (II.)

- Data collection
- We need to verify and ensure that all data that is relevant to subjects' choices, often including latent variables (e.g. thinking time to make choices) are collected. It is usually already part of the pretest.
- Pre-designed environments (Z-tree, DPTE, Inquisit) usually record. More complicated if subjects, for example, work with printed materials, cards they hand in, audio is recorded, etc.

# Ethics

3 major principles:

- Voluntary participation
  - Informed consent
  - Anonymity
- 
- Biomedical research and ethical implications issues, often carboncopied into SS.
  - Morton and Williams (2010): social scientists **must actively create their own (ethical) norms**

# Golden Rules of Experimental Ethics

**Minimal risk:** probability and magnitude of harm or discomfort to be expected in the research must be no greater than the harm and discomfort in everyday life during the conduct of routine physical and psychological tests or tasks. The probability of the magnitude of the risk has to be considered.

- The **expected benefits** must exceed the **expected costs**
- **Expected benefits** = probability weighted by the magnitude of the benefit
- **Expected cost** = probability weighted by the size of the cost

## **Benefits**

- Social
- Therapeutic
- Research
- Collateral: Financial (but to what extent are rewards ethical?),
- Altruistic, educational

## **Costs/Risks**

- Physical
- Psychological
- Social
- Economic

# Deception

**We do not tell subjects the truth about the parameters of the research.**

- Common in psychological experiments
- **APA guidelines:**
- Deception as a last resort
- Justified by scientific or educational value
- Higher efficiency than non-deceptive techniques
- Must not threaten to do harm to subjects
- Deception must be explained to the subjects as soon as possible
- Data retractability
- **We are not telling the truth (intentionally)**
- On the identity of the experimenters
- The identity of the confederates
- The purpose of the experiment or the realism of the materials

# Challenge to Ethical Issues: Field Experiments

- No informed consent
- People don't know they are part of the study
- Can change social reality (behavior, voting results)
  
- „Montana 2014“: Bonic, Rodden, Kyle 2014.
- 100 000 letters in Montana
- Nonpartisan elections (judges), information on the ideological alignment of candidates on the liberalism-conservatism axis (Obama-Romney)
- They made nonpartisan elections partisan
- They used the official seal of Montana without permission.
- Violation of election law, Complaints, They didn't go through all the ethics committees, Apology letter to voters

# Two experimental traditions in SS

- No home-grown methodology
- Psychology vs. Economics

3 major differences, influences conduct of an experiments

- Stylization
- Incentives
- Deception

# Two Traditions Compared

## Psychology

- Focuses on psychological processes, attitudes, behaviors, emotions
- Resembles real-life situations
- Financial rewards for participation, never for actual choices
- Frequent deception

## Economics

- Focuses on decision making
- Highly abstract, stylized environment, neutral language („Player I“, „Game“, „Rules“)
- Financial rewards for choices made
- No deception (clear rules, strictly obeyed)



# Security Experiments I

- Shibata et al. 2021: **Effect of implementing security measures on fear of crime**
- <https://doi.org/10.1080/1068316X.2022.2061485>
- **Abstract:** Several studies have explored the relationship between the implementation of security measures and people's fear of crime. However, most were cross-sectional studies, and the causality between two variables is not clear. This study examined the relevant causal relationships using scenario-based experiments. In Study 1, Japanese university students ( $N=180$ ) read one of the two messages, and the fear of crime was found to be decreased in participants who received a message about guards patrolling. This result was inconsistent with previous research reported that the perception of security measures positively correlates with fear of crime, and inspired Study 2. In Study 2, Japanese participants ( $N=105$ ) read one of three scenarios that involved security measures against serious crime (the serious condition), security measures against minor crime (the minor condition), or no security measures (the control condition). The results showed that fear of crime was increased more in participants in the serious condition than in the other two conditions; fear of crime in the minor condition remained around the same level with that in the control condition. These results support the hypothesis that implementing security measures increases fear of crime only when security measures are suggestive of serious crimes.

# Security Experiments II

- Tomz, Weeks (2019): Human Rights and Public Support for War
- <https://doi.org/10.1086/705741>
- **Abstract:** One of the most important themes in international relations is the relationship between domestic politics and interstate conflict. In this article, we use experiments to study how the human rights practices of foreign adversaries affect domestic public support for war. Our experiments, embedded in surveys in the United States and the United Kingdom, reveal several important findings. First, citizens are much less willing to attack a country that respects human rights than a country that violates them, even when the dispute concerns military security rather than humanitarian intervention. Second, human rights affect support for war primarily by changing perceptions about threat and morality. Citizens are more likely to view human rights violators as threatening and have fewer moral qualms about fighting such countries.
- **Procedure:** Survey Experiment. Subjects were told that the U.S. might come into conflict with a country developing nuclear weapons within a few years, they had a range of information about the country, the manipulated variable was the level of human rights in that country. On this basis, they were to declare support for various US countermeasures.

# Security Experiments III

- Vance et al. 2011: **Enhancing Password Security through Interactive Fear Appeals: A Web-Based Field Experiment**
- <https://ieeexplore.ieee.org/abstract/document/6480204>
- Abstract: Passwords remain the dominant authentication mechanism for information security. Unfortunately, research has shown that most passwords are highly insecure. Given the risks of using weak passwords, there is a need to effectively motivate users to select strong passwords. In this study we examine the influence of interactivity, as well as static and interactive fear appeals, on motivating users to increase the strength of their passwords. We developed a field experiment involving the account registration process of a website in use in which we observed the strength of passwords chosen by users. Data were collected from 354 users in 65 countries. We found that while the interactive password strength meter and static fear appeal treatments were not effective, the interactive fear appeal treatment resulted in significantly stronger passwords. Our findings suggest that interactive fear appeals are a promising means of encouraging a range of secure behaviors in end users.

# Assignment: Option One

- 2019-2022 **literature review** of experimental research in either **cybersecurity experiments** or **experiments on emotional responses to terrorism** (main topics, research questions, manipulated variables, conclusions, types of experiments).
- Length: 18.000 characters
- Completion Due: June 15

# Assignment: Option 2

- Propose **your own experimental design** for a cybersecurity or terrorism/fear research problem. The structure of the proposal should conform to the standards of the „experimental protocol“, the proposal must be theoretically grounded, and include a description of the practical implementation of the experiment (even if it will not be practically implemented yet).
- Length: 18.000 characters
- Competition due: June 15