Experiments and Observational Studies
Thought Question 1:

When experimenters want to compare two treatments, such as an old and a new drug, they use *randomization* to assign the participants to the two conditions.

If you had 50 people participate in such a study, how would you go about randomizing them?

Why do you think randomization is necessary?

Why shouldn’t the experimenter decide which people should get which treatment?
5.1 Defining a Common Language

Explanatory variable is one that may explain or may cause differences in a response variable (or outcome variable).

Example:
Study found that overall left-handed people die at a younger age than right-handed people.

Explanatory \(=\) Handedness
Response \(=\) Age at death
A treatment is the change applied to the explanatory variable(s) assigned by the experimenter.

Example: Aspirin Experiment (Ch 1)

Explanatory = whether or not you received an aspirin

Two treatments: aspirin or placebo

Response = Heart attack
Randomized Experiment versus Observational Studies

• **Randomized experiment**: create differences in the explanatory variable and examine results (response variable).

• **Observational study**: observe differences in the explanatory variable and notice whether these are related to differences in the response variable.
Two reasons why we must sometimes use an observational study instead of an experiment …

1. It is **unethical** or **impossible to assign people** to receive a specific treatment.
2. Certain explanatory variables are **inherent traits** and cannot be randomly assigned.
Confounding Variables

A confounding variable is …

- related to the explanatory variable, and
- affects the response variable.

The effect of a confounding variable on the response variable cannot be separated from the effect of the explanatory variable on the response variable.

Confounding variables are a bigger problem in observational studies. Researchers try to measure possible confounding variables and see if related to the response variable.
Example: Confounding Variables

Study of the relationship between smoking during pregnancy and child’s subsequent IQ a few years after birth.

- **Explanatory variable**: whether or not the mother smoked during pregnancy
- **Response variable**: subsequent IQ of the child
- Women who smoke also have poor nutrition, lower levels of education, or lower income.
- **Possible Confounding Variables**: Mother’s nutrition, education, and income.
Interactions Between Variables

An interaction between explanatory variables occurs when the effect of one explanatory variable on the response variable depends on what’s happening with another explanatory variable.

Example:

If smoking during pregnancy reduces IQ when the mother does not exercise, but raises or does not influence IQ when the mother does exercise, then we would say smoking interacts with exercise to produce an effect on IQ.

*If two variables interact, results should be given separately for each combination.*
5.2 Designing a Good Experiment

- **Randomization**: reduces the effects of confounding
- **Control Groups**: handled identically to the treatment group
- **Placebos**: imitation of real drug
- **Blinding**: double blind and single-blind
- **Matched Pairs, Blocks, and Repeated Measures**
Matched Pairs, Blocks, and Repeated Measures

Matched-Pair Design
• Use two matched individuals or the same individual to receive each of two treatments.
• Randomization used to assign the order of the two treatments.

(Randomized) Block Design
• Extension of matched-pair to three or more treatments.

Repeated-Measures Design
• Block designs in which the same participants are measured repeatedly.
5.3 Difficulties and Disasters in Experiments

Potential Complications

1. Confounding variables
2. Interacting variables
3. Placebo, Hawthorne, and experimenter effects
4. Ecological validity and generalizability
Confounding Variables

**Problem:** variables connected with explanatory variable can distort results because they may be agent actually causing change in the response.

**Solution:** randomization => effects of confounding variables should apply equally to each treatment.

**Example 3: Nicotine Patch Therapy**

- Nicotine patch more effective when no other smokers in home.
- If first 120 volunteers assigned to placebo and last 120 to nicotine patch, and if those with no other smokers in home more eager to volunteer => treatment would have been confounded with whether there were other smokers at home.
- **Randomization** => impact would be similar across two groups.
Interacting Variables

**Problem:** second variable interacts with explanatory variable but results reported without noting it.

**Solution:** researchers should measure/report variables that may interact with explanatory variables.

**Example 4: Other Smokers at Home**

- Interaction between treatment and whether other smokers in home. Researchers measured and reported it.
- After 8 weeks: proportion of nicotine group quitting only 31% if other smokers at home, whereas 58% if not; proportions quitting were same whether other smokers or not for placebo.
- Misleading if only reported the 46% of nicotine group quit.
Placebo, Hawthorne, and Experimenter Effects

Problem: power of suggestion (placebo effect), just being included in a study (Hawthorne effect), and experimenter recording data erroneously, treating subjects differently – all these can bias results.

Solution: use double-blinding and control group, have data entered automatically in computer as collected.

Example 5: Dull Rats (Rosenthal and Fode, 1963)

- 12 experimenters each given 5 rats that had been taught to run a maze, all similar: six experimenters were told rats were bred to do well and other six told rats were not expected to do well.
- Experimenters told they had ‘maze bright’ rats reported much faster learning rates than those with ‘maze dull’ rats.
Ecological Validity and Generalizability

**Problem:** variables measured in labs or artificial setting, results do not accurately reflect impact in real world; results for volunteers may not extend to larger group.

**Solution:** try to design experiment that can be performed in natural setting with a random sample from the population of interest; measure other variables to see if related to the response or the explanatory variables.

**Example 6: Real Smokers with a Desire to Quit**

- Used a standard intervention that other physicians could follow.
- Used participants at three different locations around country with a wide range of ages (20 to 65).
- Recorded other variables and checked to be sure not related to the response variable or the patch treatment assignment.
5.4 Types of Observational Studies

Case-Control Studies

• ‘Cases’ who have particular attribute or condition are compared with ‘controls’ who do not.
• Efficient and reduces confounding

Retrospective or Prospective Studies

• *Retrospective*: participants are asked to recall past events.
• *Prospective*: participants followed into future, and events recorded. Better because people often do not remember past events accurately.
5.5 Difficulties and Disasters in Observational Studies

Potential Complications

1. Confounding variables and the implication of causation
2. Extending the results inappropriately
3. Using the past as a source of data
Confounding Variables and the Implications of Causation

**Problem:** no way to establish causation with an observational study – can’t separate out all potential confounding factors w/o randomization.

**Solution:** measure potential confounding variables; choosing controls as similar as possible to cases.

**Example 7: Smoking During Pregnancy**

- IQs lower for children of women who smoked.
- Difference as high as 9 points before accounting for confounding variables (diet and education); reduced to 4 points after accounting for those factors.
- Can’t conclude smoking *caused* lower IQs in children.
Extending the Results Inappropriately

Problem: many use convenience samples, not representative of any population.

Solution: researchers should use entire segment of population of interest.

Example 8: Baldness and Heart Attacks

- Observational study only used men who were hospitalized.
- Should consider whether results should be extended to all men.
Using the Past as a Source of Data

Problem: retrospective studies unreliable – ask people to recall past behavior; confounding variables in past not similar current ones.

Solution: use prospective studies if possible; else use authoritative sources versus memory.

Example 9: Do Left-Handers Die Young?

- Retrospective: sent letters to next of kin asking about handedness of deceased.
- Average age of death of LH was 66 versus 75 for RH.
- In early 20th century, many children forced to write RH. Many in study may have been influenced.
5.6 Random Sample versus Random Assignment

Extending Results to a Larger Population: *Random Sampling*

- Often impractical to obtain a random sample.
- Extent to which results extend depends on extent to which participants are representative of population.

Establishing Cause and Effect: *Random Assignment*

- Evens out confounding variables across treatments.
- Without it, naturally occurring confounding variables can result in an apparent relationship.
Chapter 6

Getting the Big Picture
How to Evaluate a Study

STEP 1: Determine if the research was a sample survey, an experiment, an observational study, a combination, or based on anecdotes.

STEP 2: Consider the Seven Critical Components in Chapter 2 to familiarize yourself with the details of the research.

STEP 3: Based on the answer in step 1, review the “difficulties and disasters” inherent in that type of research and determine if any of them apply.

STEP 4: Determine if information is complete. If necessary, find the original source of the report or contact the authors for missing information.
How to Evaluate a Study

STEP 5: Ask if the results make sense in the larger scope of things. If they are counter to previously accepted knowledge, see if you can get a possible explanation from the authors.

STEP 6: Ask yourself if there is any alternative explanation for the results.

STEP 7: Determine if the results are meaningful enough to encourage you to change your lifestyle, attitudes, or beliefs on the basis of the research.
Case Study 6.3: Drinking, Driving, and the Supreme Court

Craig v. Boren, 429 U.S. 190, 1976:

- Challenged Oklahoma state law prohibiting sale of 3.2% beer to males under 21 but allowed it to females under 21.
- Laws allowed to use gender-based differences if ‘serve important governmental objectives’ and ‘are substantially related to achievement of these objectives’.
- Defense argued traffic safety was an important governmental objective and data show young males more likely to have alcohol-related accidents than females.
- Supreme Court shown two sets of data.
- Review the data – do you think law should be upheld?

Source: Gastwirth, 1988, pp. 524-528.
### Case Study 6.3: Drinking, Driving, and the Supreme Court

**TABLE 6.1** Arrests by Age and Sex in Oklahoma, September–December 1973

<table>
<thead>
<tr>
<th></th>
<th>MALES</th>
<th>FEMALES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18–21</td>
<td>Over 21</td>
</tr>
<tr>
<td>Driving under influence</td>
<td>427</td>
<td>4,973</td>
</tr>
<tr>
<td>Drunkenness</td>
<td>966</td>
<td>13,747</td>
</tr>
<tr>
<td>Total</td>
<td>1,393</td>
<td>18,720</td>
</tr>
</tbody>
</table>

2% of males 18-21 0.18% of females 18-21

**TABLE 6.2** Random Roadside Survey of Driving and Drunkenness in Oklahoma City, August 1972 and August 1973

<table>
<thead>
<tr>
<th></th>
<th>MALES</th>
<th>FEMALES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Under 21</td>
<td>21 and Over</td>
</tr>
<tr>
<td>BAC* over .01</td>
<td>55</td>
<td>357</td>
</tr>
<tr>
<td>Total</td>
<td>481</td>
<td>1926</td>
</tr>
<tr>
<td>%BAC over .01</td>
<td><strong>11.4%</strong></td>
<td>18.5%</td>
</tr>
</tbody>
</table>

*BAC = Blood alcohol content
Case Study 6.3: Drinking, Driving, and the Supreme Court

STEP 1: Determine if research was sample survey, experiment, observational study, a combination, or based on anecdotes.
Table 6.1: observational, representing only a subset of those who committed the crimes. Table 6.2: sample survey, based on a convenience sample.

STEP 2: Consider the Seven Critical Components.
Not told how the ‘random roadside survey’ was conducted.
Case Study 6.3: *Drinking, Driving, and the Supreme Court*

**STEP 3:** Review “difficulties and disasters” inherent in that type of research and determine if any of them apply.  
Table 6.1: Males more likely to be stopped for other traffic violations. Table 6.2: drivers questioned at these certain locations may not represent *all* drivers.

**STEP 4:** Determine if information is complete. If necessary, find original source of report or contact authors for missing info.  
Relatively complete (except for Step 2).

**STEP 5:** Ask if results make sense in larger scope of things. If counter, see if a possible explanation from the authors.  
Nothing suspicious.
Case Study 6.3: Drinking, Driving, and the Supreme Court

STEP 6: Ask if any alternative explanation for the results.
Table 6.1: Males more likely to be stopped for other traffic violations. Table 6.2 shows almost 80% of drivers were male. Males more likely driving helps explain why Table 6.1 results so different. Table 6.2 results of 11.4% males versus 9.4% females were statistically indistinguishable.

STEP 7: Determine if results meaningful enough to change lifestyle, attitudes, or beliefs on the basis of the research.
Supreme Court overturned law concluding Table 6.2 data ‘provides little support for a gender line among teenagers’.
STUDY: SMOKING MAY LOWER KIDS’ IQS

ROCHESTER, N.Y. (AP)—Second-hand smoke has little impact on the intelligence scores of young children, researchers found. But women who light up while pregnant could be dooming their babies to lower IQs, according to a study released Thursday. Children ages 3 and 4 whose mothers smoked 10 or more cigarettes a day during pregnancy scored about 9 points lower on the intelligence tests than the offspring of nonsmokers, researchers at Cornell University and the University of Rochester reported in this month’s Pediatrics journal.

That gap narrowed to 4 points against children of nonsmokers when a wide range of interrelated factors were controlled. The study took into account secondhand smoke as well as diet, education, age, drug use, parents’ IQ, quality of parental care and duration of breast feeding.

“It is comparable to the effects that moderate levels of lead exposure have on children’s IQ scores,” said Charles Henderson, senior research associate at Cornell’s College of Human Ecology in Ithaca.

Case Study 6.4: Smoking During Pregnancy and Child’s IQ

STEP 1: Determine if research was sample survey, experiment, observational study, a combination, or based on anecdotes. Observational Study. Smoking behavior observed.

STEP 2: Consider the Seven Critical Components. News report too brief. Missing info provided in original report: Supporting grants from many sources not related to tobacco products; participant details (primiparous women, many teenagers, unmarried or poor, mostly white); two groups compared (smoked 10+ cigarettes per day average versus none); IQ measured at 12 months; many potential confounding variables measured and checked; not clear if study was single-blind.
Case Study 6.4: *Smoking During Pregnancy and Child’s IQ*

**STEP 3:** Review “difficulties and disasters” inherent in that type of research and determine if any of them apply.

Prospective study. Difference in IQ for two groups reduced from 9 to 4 points when potential measured confounding variables included – there could be others not included, such as maternal report of illegal drug/alcohol use or child’s exposure to side-stream smoke. Results may not extend to older mothers or less disadvantaged groups.

**STEP 4:** Determine if information is complete. If necessary, find original source of report or contact authors for missing info.

Original report fairly complete.
Case Study 6.4: Smoking During Pregnancy and Child’s IQ

STEP 5: Ask if results make sense in larger scope of things. If counter, see if a possible explanation from the authors. Authors do speculate on what might be cause.

STEP 6: Ask if any alternative explanation for the results. Confounding variables not measured; possible experimenter bias if those measuring IQ were not blind.

STEP 7: Determine if results meaningful enough to change lifestyle, attitudes, or beliefs on the basis of the research. If you were pregnant and concerned about your child’s IQ, results might lead you to quit smoking during pregnancy.
Case Study 6.5: Guns and Homicides at Home

Study Details:

• Multistate study of homicides => keeping gun at home nearly triples likelihood that someone in household will be slain there.

• Records of three populous counties (sample representative of the entire nation) were studied. Looked only at homicides that occurred in homes of victims—about 400 deaths.

• Members of households with guns were 2.7 times more likely to experience a homicide than those in households without guns.

• In 77% of cases, victims were killed by a relative or someone they knew. In 4% of cases were victims killed by a stranger. In remaining cases, the identity could not be determined.