

Udacity A/B Testing Lesson 2 Notes

Jung Ah Shin

Feb 2021

1 Policy and Ethics for Experiments

1.1 Introduction

Examples of Cases where participants may not have been protected

- Tuskegee syphilis experiment Approximately 600 rural African-American men in Tuskegee (1932-1972) thought they were receiving free healthcare from the government when they were being cited regarding the natural progression of untreated syphilis. These men were not informed of the possible treatment and not treated.
- Milgram experiment Early 1960s measured the willingness of studied participants to obey the authority figure who instructed them to administer electric shocks to a third person. Participants may have been damaged psychologically.
- Facebook experiment Studied how emotions can be spread on social media. Risk to the participants is substantially lower than the aforementioned ones and there is no discussion about the potential benefits that might result from such a study.

1.2 Four Main Principles

1. **Risk** What risk is the participant being exposed to?
2. **Benefit** What benefits might be the outcome of the study?
3. **Choice** What other choices do participants have?
4. **Privacy** What expectation of privacy and confidentiality do participants have?

1.3 Risk

Main threshold: Whether the risk exceeds "minimal risk"

Minimal risk: Probability and magnitude of a harm that a participant would encounter in normal daily life. Harm considered encompasses physical, psychological and emotional, social, and economic concerns.

If risk exceeds minimal risk, then informed consent is required.

Assessing Risk

Table 1: Is risk greater than minimal risk?

Situation	Risk
Search engine tests change adding price info to results page	
Health app lets users know possible consequence of their diets	✓
Online encyclopedia tests changing the location of the search bar	

- First point: The search engine is simply providing additional information that was already publicly available
- Second point: Potentially more than minimal risk as it's providing medical advice tailored to individuals. If advice is incorrect, could have health risks. Depending on the information given, the app may be subject to additional FDA regulation
- Third point: Harmless change

1.4 Benefits

What benefits might result from the study? It is important to be able to state what the benefit would be from completing the study.

1.5 Alternatives

What other choices do participants have? The main issue is that the fewer alternatives that participants have, the more issue that there is around coercion and whether participants really have a choice in whether to participate or not. In online experiments, the issues to consider are the other alternative services, switching costs in terms of time, money, information, etc.

1.6 Data Sensitivity

What data is being collected, and what is the expectation of privacy and confidentiality?

- Do participants understand what data is being collected about them?
- What harm would befall them should that data be made public?
- Would they expect that data to be considered private and confidential?

If new data is being gathered, then:

- What data is being gathered? How sensitive is it? Does it include financial and health data?
- Can the data being gathered be tied to the individual, i.e. is it considered personally identifiable?
- How is the data being handled? Security measures? What level of confidentiality can participants expect?
- What harm would befall the individual should the data become public, where the harm would encompass health, psychological/emotional, social, and financial concerns?

Three main issues with data collection with regards to experiments:

- For new data being collected and stored, how sensitive is the data and what are the internal safeguards for handling that data?
- For that data, how will it be used and how will participants' data be protected? How are participants guaranteed that their data, which was collected for use in the study, will not be used for some other purpose? This becomes more important as the sensitivity of the data increases.
- What data may be published more broadly, and does that introduce any additional risk to the participants?

Difference between pseudonyms and anonymous data

- **Identified:** Data is stored and collected with personally identifiable information. E.g. Names, IDs. HIPAA is a common standard, and that standard has 18 identifiers that it considers personally identifiable. e.g. device ID
- **Anonymous:** Data is stored and collected without any personally identifiable information. This data can be considered **pseudonymous** if stored in a randomly generated id such as a cookie that gets assigned on some event, such as the first time that a user goes to an app and does not have such an id stored. In most cases, anonymous still has time-stamps (one of HIPAA 18 identifiers).
- **Anonymized data:** Identified or anonymous data that has been looked at and guaranteed in some way that the re-identification risk is low to non-existent.

Assessing data sensitivity

Table 2: Should the following data be considered sensitive?

Situation	Sensitive
Existing census data with race, age, and gender by zipcode	
Number of users visiting specific sites each day (e.g. HitWise to gather counts)	
User-entered glucose levels with timestamp and anonymous id	✓
Achievements, kills, and level advances in an online game by anonymous id	
Number of purchases and total money spent by a zipcode	
Credit card information	✓

- First point: Too granular for re-identification (i.e. aggregated)
- Second point: Can't identify individuals. Only total numbers
- Third point: Timestamps could identify. Glucose levels are health data. Regulation such as HIPAA
- Fourth point: Game data not sensitive
- Fifth item: Totals are stored for zipcode. Too granular for re-identification. Ensure enough data points to guarantee that data is sufficiently anonymized
- Sixth item: Very sensitive

1.7 Summary

For online data being gathered:

- Are users being informed?
- What user identifiers are tied to the data?
- What type of data is being collected?
- What is the level of confidentiality and security?

What about informed consent?

Participants are told the risk they'd face, what benefits they'd get, what alternatives are available, what data is being gathered, and how the data is being handled.

Which tests need further review?

Table 3: Which tests need further review or possibly obtain user consent?

Situation	Need further review
Survey about net worth on financial news site. Test survey at bottom of article vs. after first page, not allowing user to advance until they complete the study	
Different menu layouts for meal delivery company	
Test different ways of presenting user heart rate data	✓

- First point: Depends on what information is being gathered in the survey. If the survey collects personal info, then info is sensitive. If no personal info is gathered, then survey should be fine to run. **Users can voluntarily choose what to do.** Ethical considerations also increase if user can only read the article only in exchange for their survey info.
- Second point: No new sensitive being collected. Minimal risk
- Third point: Need further review. Physical fitness is important to good health, presentations that confuser users or encourage users to exercise too much or too little could have risk. Data being collected is sensitive, so needs to be carefully secured and anonymized. Depending on the data, may not be possible to anonymize while keeping same level of granularity, therefore need to aggregate the data

Provided information

Table 4: Which information is ethically necessary to provide to users?

Information	Need to be provided to users
A Terms of Service (TOS) or a Privacy Policy	✓
History of funding	
Navigation bar	
List of experiments you are planning on running	
Search bar	

- First point: A clear statement of what data you'll collect and how you will use it is ethically necessary
- Second point: Not necessary
- Third point: Not ethical requirement
- Fourth point: Users could be fully informed, but for some experiments, giving out this information could skew the results
- Fifth point: Not ethically necessary

Internal Training

Table 5: Which information is ethically necessary for anyone who runs A/B tests to know?

Information	Need to know
Which questions to consider when evaluating ethics	✓
History of A/B testing	
History of IRBs	
Data policy detailing acceptable data uses	✓
Principles to uphold	✓