Field studies, crowdsourced studies

Michelle Mazurek

Some slides adapted from Lorrie Cranor, Blase Ur, Marshini Chetty, Elaine Shi, Christine Trask, and Yu-Xiang Wang, Howard Seltman
Today

• Class presentation info
• Finish up experimental validity
• Crowdsourcing
• Field studies, more ecological validity
Principles for research design

• “The goal of any research design is to arrive at clear answers to questions of interest while expending a minimum of resources.” – Ramsey and Shafer

• Goal: Identify sources of experimental variation and try to minimize/control them
1. Internal validity

• Avoid confounds!
  – Avoid criticism about causality

• How?
  – Randomize condition/treatment assignment
  – Change only one variable at a time per condition
  – Use blinding
  – Use a control group
2. Construct validity

• What do our metrics measure?
  – Is it what we intended?
  – Discuss: How would you measure tech-savviness?
3. External validity

• What population does your sample represent?
  – Race, gender, age, nationality, education, others

• What environment does your sample represent?
  – Carefully controlled study vs. real world

• This is especially critical for security research
  – Security as secondary task
  – What you think you should do vs. what you actually do
Promote power

- Covariates: Measure possible confounds, include in analysis

- Blocking: Group similar subjects, distribute across conditions

- Within-subjects promotes power
  - But has other drawbacks, esp. in security research

- Is what you are measuring strong enough?
  - Do you have enough participants?
  - Potential tradeoff: Generalizability for power
CROWDSOURCED STUDIES
What is crowdsourcing?

• Merriam-Webster: “The process of obtaining needed services, ideas, or content by soliciting contributions from a large group of people, and especially from an online community, rather than from traditional employees or suppliers”

• Academic Daren Brabham: “online, distributed problem-solving and production model.”
In our context

- Finding study participants online
- Service handles details of recruitment, payment, etc.
Why crowdsourcing?

• Large numbers of participants
  – Without complicated logistics
  – From around the country, world

• Easily controlled conditions

• Relatively inexpensive
Why not crowdsourcing?

- No direct observation of participants
- Limited followups
- Some participants will enter garbage
- Specific demographics participate
  - Younger, more technical than general population
  - Better than recruiting all students!
  - Usually worse than, e.g., Craigslist recruiting
Participant problems

• Attempted repeaters
  – Especially if you pay too much
• Entering garbage / not paying attention
• Discussion in forums
  – What about deception?
• Terms of service may limit request types
Participant solutions

• Collect a lot of data
  – Noise distributed across conditions
• Use cookies, IP tracking, worker IDs
• Ensure there is no “shortcut”
• Use attention check questions, repeats
  – Carefully designed and placed
• Do NOT use well-known “trick” questions
• Monitor forums
Case study: Amazon MTurk

Requester → Task → Worker

Worker $ Worker $ Worker

Logistics: Recruiting

Participate in a Carnegie Mellon University research study.

You will be asked to fill out a 5-minute survey for a $0.55 payment. We will email you to come back in a few days to log in again and fill out another survey for a $0.70 bonus payment.

Do not submit this HIT multiple times. You will not be paid if you participate in this study more than once. If the study website tells you that you have participated before, please return the HIT.

(Note: This link will not become visible until you accept this HIT.)

When you have completed the study, enter your completion code here to submit your work for approval and payment.
Logistics: Consent

Research Consent Form

This task is part of a research study conducted by Lorrie Cranor at Carnegie Mellon University.

The purpose of the research is to survey users about their behavior visiting secure websites.

Procedures
Participants will be asked to answer survey questions, which should take about 5 minutes. They will be asked to return later to login and answer more survey questions, which should also take about 5 minutes.

Participant Requirements
Participation in this study is limited to individuals age 18 and older.

Risks
The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities.

I am age 18 or older. Yes ☐ No ☐

I have read and understand the information above. Yes ☐ No ☐

I want to participate in this research and continue with the task. Yes ☐ No ☐

Continue
Logistics: Infrastructure

• Directly within MTurk
  – Easiest, limited feature selection

• Redirect to survey software
  – UMD Qualtrics subscription
  – Well coordinated, not great for non-survey things

• Redirect to your own server
  – Best option for complicated studies
  – But requires design / management
Logistics: Payment

- MTurk account is prepaid
- You “approve” the work and Turker is paid
- MTurk takes a small percentage
Other useful features

- Screen and reject workers
  - Location, quality rating, etc.
- Send notifications (e.g. to come back for part 2)
- Prevent repeated workers in the same task
  - May need multiple tasks per study
- On average, 100 participants / day
  - Starts faster, slows down, repost
Kang et al., SOUPS 2014

• Survey on privacy attitudes and behavior
• Administered to:
  – Representative Pew phone sample
    • 775 Internet users
  – U.S. Turkers (182)
  – Indian Turkers (128)
Results: Demographics

- Turk younger, malar, more educated
  - Indian Turk even more so
Results: U.S. general vs. U.S. Turk

- Turkers more likely to seek anonymity
- Turkers more likely to hide content selectively
  - Except, general more likely to hide from hackers
- Younger, more educated say more data on them is available; take more steps to hide
- Turkers more concerned about privacy, more likely to say anonymity should be possible
Results: U.S. Turk vs. India Turk

• Indians say more personal data is online
• U.S. more likely to seek anonymity
  – Indians more likely to hide from boss/supervisor
• Indians less concerned about privacy, more satisfaction with gov’t protection
• Fewer Indians say anonymity should be possible
  – More comfortable with monitoring to prevent terrorism
Beyond Turk

• Crowdflower
• crowdsource.com
• Samasource
• Google consumer surveys
Resources

- https://experimentalturk.wordpress.com/
- http://www.behind-the-enemy-lines.com/
FIELD STUDIES
Why a field study?

• Better ecological validity
  – Validate a lab study result

• Because you can’t get the data any other way
Why not a field study?

• Logistically difficult

• Limited piloting / not easy to adjust
  – One shot at your participant pool

• Expensive (money and time)

Plan extremely carefully!
PhishGuru in the real world

• Anti-phishing training delivered when users follow a phishing link

• Training, phishing, legitimate emails delivered to 300 employees in a Portuguese company
  – Over several weeks

• Why a field study, is this necessary?

Kumaraguru et al, eCrime Summit 2008
Logistical problems

• Didn’t include a legitimate email before training to compare click rates
• Control and experimental not 100% parallel
• Participants talked to each other, sharing the training materials
• No one turned in the post-study questionnaire!

• How could these have been avoided?