08. Surveys (continued), Qualitative Studies, Journalists

Blase Ur
April 25th, 2019
CMSC 23210 / 33210
Participants, ethics, and deception
Participants (1)

• Recruit people to do something remotely (e.g., online)
• Recruit people to come to your lab
• Recruit people to let you into their “context”
• Observe people (if possible, get consent! If not possible, consider necessity of design)
Participants (2)

• What recruitment mechanisms?
  – Craigslist, flyers, participant pools, representative sample, standing on street

• How do you compensate them?
  – Ethics of paying $0.00 vs. $10.00 vs. $100,000

• How do you get informed consent?

• What happens to their data?

• Prior knowledge / “what” are they?
Ethics

• How do we protect participants?
  – What risks do we introduce?
• Is there a less invasive method that would give equivalent insight?
• IRB is one arbiter of ethics; experimenters themselves are another crucial arbiter
• How do we make sure participation is voluntary throughout the experiment?
Deception

• Do we mind if participants know precisely what is being studied?
  – Sometimes, it’s crucial that we observe their organic responses in context

• What “deception” or “distraction” task can we introduce?

• How do we debrief people at the end?
Institutional Review Board (IRB)

- Is it research? Are there human subjects?
- Full review vs. expedited vs. exempt
- Fill out and submit protocol
  - Include all study materials (e.g., surveys)
  - Include recruitment text and/or poster
  - Leave plenty of time
What to submit to an IRB

• Full consent form (use UChicago model)
• All scripts, survey questions, instructions
• Recruitment plan
• Recruitment materials
  – Don’t emphasize compensation
• Information about how data will be handled
  – Password protection, encryption, etc.
  – Meetings to discuss
Survey design
Overall survey considerations

• How do we distribute it?
• How long should it be?
• One-time survey? Longitudinal survey?
• Will you use personalized data?
• What will participants learn?
  – What can we randomize to minimize this?
• Can we randomize the answer choices?
Are all answer options covered?

• With whom do you regularly share posts on social media?
  – Family
  – Friends

• Allow multiple answers?

• Include “other” option (write-in)?

• Do we care about previous use?
Are all answer options covered?

• I connect to Facebook over HTTPS
  – True
  – False

• What about “I don’t know”?
Are we biasing the answer?

• Strangers seeing your Facebook posts would cause you grave privacy concern.
  – Strongly agree
  – Agree
  – Neither agree nor disagree
  – Disagree
  – Strongly disagree
How will responses be distributed?

• For how long have you had Facebook?
  – Less than one day
  – Between one day and one week
  – More than one week
Should we force an answer?

• What gender are you? (* required)
  - Female  - Male

• What gender are you?
  - Female  - Male  - I prefer not to answer

• With what gender do you identify?
  - Female  - Male  - Non-binary  - I prefer to self-describe___  - I prefer not to answer
Likert-scale data?

- Respond to the following statement: Companies collect too much private data.
  - 7: Strongly agree
  - 6: Agree
  - 5: Somewhat agree
  - 4: Neutral
  - 3: Somewhat disagree
  - 2: Disagree
  - 1: Strongly disagree
Likert-scale data?

• I feel that companies collect too much private data.
  – 7: Strongly agree
  – 6: Agree
  – 5: Somewhat agree
  – 4: Neutral
  – 3: Somewhat disagree
  – 2: Disagree
  – 1: Strongly disagree
What demographics do we collect?

- Tech expertise, age, domain knowledge, gender, location, employment, etc.
- Don’t ask people to self-rate expertise
  - Ask questions with concrete answers
  - e.g., Have you earned a degree in, or held a job in, computer science, IT, or…
  - Include a knowledge test if you want to know about expertise
- Consider why you are collecting this info
Analyzing Qualitative Data
Qualitative Coding

• Many different approaches

• Key goal: capture themes in data

• Often, but not always, develop codebook containing themes observed

• For robustness, another person follows the codebook and independently codes data
  – Agreement metrics include Cohen’s Kappa
Anonymity and Journalists
Why is anonymity valuable? Or not?
Techniques for (some) anonymity

• Encrypt everything
• Use Tor to communicate
• Off-the-record (OTR) messaging
• Don’t use services that track you
Virtual Private Networks (VPNs)
Overview of Tor

• The Onion Router (Tor)
  – Onion routing introduced by U.S. Naval Research Labs ~ 20 years ago
  – Dinglydine, Matthewson, Syverson introduced Tor in a USENIX Security paper in ‘04
How Tor works (graphics from EFF)
How Tor works (graphics from EFF)

Step 2: Alice's Tor client picks a random path to destination server. **Green links are encrypted, red links are in the clear.**
How Tor works (graphics from EFF)

Step 3: If at a later time, the user visits another site, Alice’s tor client selects a second random path. Again, green links are encrypted, red links are in the clear.
How Tor works
What does Tor protect against?
What does Tor NOT protect against?
 Threats Against Tor

• Vulnerabilities in the protocol
• Vulnerabilities in the implementation
• Adversaries controlling large parts of the network and analyzing traffic/timing
• Vulnerabilities on the user’s end
  – E.g., old version of Firefox
• Human error on the part of the user
• Not enough users! (no hiding in the crowd)
Tor warnings

https://www.torproject.org/download/download#warning

Want Tor to really work?

You need to change some of your habits, as some things won’t work exactly as you are used to.

a. Use Tor Browser
   Tor does not protect all of your computer’s Internet traffic when you run it. Tor only protects your applications that are properly configured to send their Internet traffic through Tor. To avoid problems with Tor configuration, we strongly recommend you use the Tor Browser. It is pre-configured to protect your privacy and anonymity on the web as long as you’re browsing with Tor Browser itself. Almost any other web browser configuration is likely to be unsafe to use with Tor.

b. Don’t torrent over Tor
   Torrent file-sharing applications have been observed to ignore proxy settings and make direct connections even when they are told to use Tor. Even if your torrent application connects only through Tor, you will often send out your real IP address in the tracker GET request, because that’s how torrents work. Not only do you de-anonymize your torrent traffic and your other simultaneous Tor web traffic this way, you also slow down the entire Tor network for everyone else.

c. Don’t enable or install browser plugins
   Tor Browser will block browser plugins such as Flash, RealPlayer, Quicktime, and others: they can be manipulated into revealing your IP address. Similarly, we do not recommend installing additional addons or plugins into Tor Browser, as these may bypass Tor or otherwise harm your anonymity and privacy.

d. Use HTTPS versions of websites
   Tor will encrypt your traffic to and within the Tor network, but the encryption of your traffic to the final destination website depends upon on that website. To help ensure private encryption to websites, Tor Browser includes HTTPS Everywhere to force the use of HTTPS encryption with major websites that support it. However, you should still watch the browser URL bar to ensure that websites you provide sensitive information to display a blue or green URL bar button, include https:// in the URL, and display the proper expected name for the website. Also see EFF’s interactive page explaining how Tor and HTTPS relate.

e. Don’t open documents downloaded through Tor while online
   Tor Browser will warn you before automatically opening documents that are handled by external applications. DO NOT IGNORE THIS WARNING. You should be very careful when downloading
Making anonymity usable (example)

• Tor browser bundle
• TAILS (The Amnesic Incognito Live System)
• OTR (off-the-record) messaging tools
Why Johnny Can’t Blow the Whistle

• Identify stop-points in Tor Browser Bundle
• Highlight the security reason behind delays
• Combine Vidalia control window & browser
• Change icon
• Direct users to the right OS version
Guides to leaking and protesting

- https://ssd.eff.org/en/module/attending-protests-united-states
NY Times’s leak instructions

Got a confidential news tip?
The New York Times offers several ways to get in touch with and provide materials to our journalists. Learn more.
Got a confidential news tip?

Do you have the next big story? Want to share it with The New York Times? We offer several ways to get in touch with and provide materials to our journalists. No communication system is completely secure, but these tools can help protect your anonymity. We’ve outlined each below, but please review any app’s terms and instructions as well. Please do not send feedback, story ideas, pitches or press releases through these channels. For more general correspondence visit our contact page.
Huffpost’s leak instructions

GOT A TIP?

Do you have info to share with HuffPost reporters? Here’s how.
Need privacy?
If you’re concerned that being a source for a story poses a significant risk, take precautions:

- **Know your risks.** No form of communication is 100 percent safe from all observers. Make a plan about what you’ll do if the wrong person finds out you contacted us.

- **Do not contact us from your work computer or phone.** Your bosses can track your use of these devices. The same goes for your personal mobile phone, if you’ve ever installed apps from your employer — even if you later uninstalled them.

- Consider using **postal mail.** We’re at “HuffPost, PO Box 28154, Washington, DC 20038-8154.” Send from a public mailbox and don’t write a return address. Only we can read your message (unless a court provides a warrant).

- Use the same **encrypted email** service we do. Create a new [protonmail.com](http://protonmail.com) account — separate from your other email accounts — and use it to write us at [huffpostscoops@protonmail.com](mailto:huffpostscoops@protonmail.com). As long as you write to our Protonmail address from your Protonmail address, only we or someone who knows your password can read your message. Read more about Protonmail.

- **Use your browser’s “incognito” or “private browsing” mode.** Some sites (including, potentially, your employer’s) can access your browser history and see what websites you’ve visited. An incognito window masks this data. Open a new Incognito browser window to contact us, and close it immediately afterward. If