Internet-Based Human Subjects Research

So what's the concern?

Ben Mooso

Associate Director, IRB Administration

UC Davis

CARE-Q

Consortium for Applied Research Ethics Quality

Human Subjects Research conducted via the internet has become more complex…

- Research Participants – children & international
- Information to be collected – mundane vs sensitive
- Type of interaction – Survey Monkey vs Zoom
- Informed Consent – Waivers vs Documented Consent
- Compensation – How much, method
- Privacy/Confidentiality – data storage, ownership, access, sharing

CARE-Q
Consortium for Applied Research Ethics Quality

We'll touch on just a few of these Considerations…

- Research Participants
  - Children
  - International
- Type of interaction with participants
  - Indirect
  - Direct
  - Quasi
- Type of Information Collected
  - Public
  - Private
  - Sensitive/Illicit
  - Potentially Illegal/Illegal
- Privacy/Confidentiality
  - Data Ownership
  - Access
  - Storage
  - Sharing

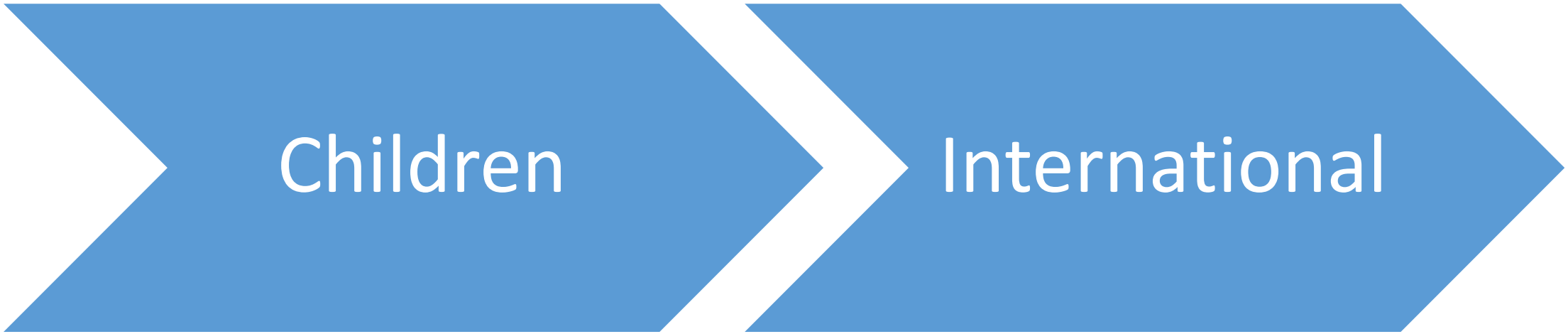
CARE-Q
Consortium for Applied Research Ethics Quality

Research Participants

PROS

- Can accommodate school schedules and extracurricular activities
- Can accommodate parental needs
- Often adept at using technology

- Able to reach a broader and more diverse subject pool
- No need to have personnel available at odd hours



CONS

- Need to comply with various Child/Internet protection laws (e.g. COPPA)
- Difficulty in excluding children
- Difficulty in verifying parental consent

- Countries in which participants will come from may be unknown at initial review
- Differing international laws regarding research

CARE-Q
Consortium for Applied Research Ethics Quality

Children as Research Participants

- Children's Online Privacy Protection Rule ("COPPA")
  - Applies to operators of websites and online services which are:
    - Directed at children under 13 years old
    - Directed at a general audience when there is “actual knowledge” that a child under 13 years old has provided personal information
  - Compliance
    - Provide privacy policy notice for information collected from children
    - Obtain verifiable parental consent
    - Provide parents access to collected information and allow deletion
    - Maintain confidentiality and retain data for minimum necessary time
- Ethical Implications
  - Need to obtain verifiable parental consent

CARE-Q
Consortium for Applied Research Ethics Quality

International Research Participants

- General Data Protection Regulation (GDPR) – EU/EEA
  - Applies to personal data collected from individuals in the EU/EEA and companies established within the EU/EEA
  - Requires lawful basis for processing of personal data
  - Researchers will also generally need explicit consent to process sensitive data
  - Grants certain rights to subjects including the right to withdraw data, correct data, view data, and “be forgotten”
- Protection of Personal Information Act (POPIA) – South Africa
- California Consumer Privacy Act (CCPA) – California, US
- Personal Information Protection and Electronic Documents Act (PIPEDA) – Canada
- Ethical Implications
  - May not know ahead of time what countries subjects will come from
  - May be encounter conflicting regulations depending on the countries involved
  - Differing levels of state control of internet traffic

CARE-Q
Consortium for Applied Research Ethics Quality

Example

- Dr. Smith wants to study linguistic development in children from around the world
- She will collect speech samples from 1-3 years olds in English, Mandarin, Hindi, Spanish, and Arabic
- While she won't target a specific country, she hopes to get participants from all over the world
- The study will collect personal data including race and ethnicity
- What issues does the IRB need to consider?

CARE-Q
Consortium for Applied Research Ethics Quality

Privacy/Confidentiality

- Participant Privacy
- Data Confidentiality
  - GDPR, CCPA, POPIA, PIPEDA
- Data Ownership and Third-Party Agreements
  - In-house systems
  - Third-party systems

CARE-Q
Consortium for Applied Research Ethics Quality

Privacy/Confidentiality

PROS

- Subjects can choose their own physical environment in which to participate
- Easy to invite others to be with subjects during interactions if desired

- Many security tools available to protect data
- Most institutions offer single-sign-on access and encryption
- Back-up files reduce risk of lost data in an emergency

- Many third-party services available for data collection and analysis
- More easily share data will colleagues for collaboration
- Out-of-the-box software

CONS

- Hard for researchers to ensure that any conversations are private on the subject's end
- Person claiming to participate may not be who they say they are

- Constant threats from outside forces
- Easier to lose storage devices or send files to wrong person
- Potential to identify subjects from coded or de-identified data

- Agreements may have unacceptable data collection, ownership, and sharing terms
- Third-party services may not be accessible unless agreements are signed

CARE-Q
Consortium for Applied Research Ethics Quality

Data Confidentiality and Ownership

- In-house systems
  - Usually secure
  - Vetted by IT
  - Prone to lagging behind in security standards if not consistently updated
  - Prone to cyber attacks if not consistently updated
  - PI retains data ownership
  - PI has only some or little control of security measures
- Third-party systems
  - Usually secure
  - Requires IT vetting/validation
  - Updated for security more often/regularly
  - Ownership of data determined by agreement with provider
  - May sell/release data externally depending on agreement
  - PI has little or no control of security measures
- Ethical Implications
  - Adequately informing subjects of the various risks associated with these systems

CARE-Q
Consortium for Applied Research Ethics Quality

Type of Interaction

- Indirect
  - Collection of existing information
  - Web-based Surveys/Questionnaire
    - Use of Qualtrics, Survey Monkey, REDCap
- Direct
  - Individual Interviews/Focus Groups
    - Use of Zoom, Google Chat, WebEx
- Quasi
  - Use of Avatars
  - Interaction in Dark Web

CARE-Q
Consortium for Applied Research Ethics Quality

Type of Interaction

PROS

- Data can be provided by subjects at times convenient to them
- Possibility of collecting data anonymously

- Enables research otherwise impracticable over long distances or during pandemic
- Eliminates need for a physical space to be available

- Provides an extra layer of protection in discussing taboo topics by facilitating anonymity of subjects while still allowing for real time communication

CONS

- Uncertainty of who is actually participating in the study (e.g. children)
- Potential for collateral risks to non-consented individuals

- Data security concerns when using third-party vendors/ services
- Issues with availability of resources (e.g. microphone) depending on study population

- Ethical implications if subject reveals illegal or harmful intent
- Data collected (e.g. demographics) may remove the benefits of anonymity

CARE-Q
Consortium for Applied Research Ethics Quality

Collateral Subjects and Risks

- More common in internet-based research
- Occurs when a participant interacts with a non-consented individual and the interaction becomes part of the data for analysis
- Ethical implications
  - New subject is not aware of their participation
  - No opportunity to object to participation
  - New subject is not made aware of possible risks
  - Selection bias (subject selects new subjects, not researchers)
  - Others?

CARE-Q
Consortium for Applied Research Ethics Quality

Example

- Dr. McCoy will give subjects who agree to participate a story to post
- The subjects won't know if the story is true or false
- Researchers will watch to see how many people interact with the post
- Researchers will also see how the public posting behavior of those who do interact with the post changes
- After 5 days, the subject will be provided a debriefing to post
- What issues does the IRB need to consider?

CARE-Q
Consortium for Applied Research Ethics Quality

Type of Information Collected

- Public
  - Social media posts, public websites, contact information
- Private
  - Browser history, email memberships, shopping activity
- Sensitive/Illicit
  - Family planning, pornography, online banking
- Potentially Illegal/Illegal
  - Dark web usage, drug use, weapons purchases

CARE-Q
Consortium for Applied Research Ethics Quality

Type of Information Collected

PROS

- Readily available
- Less personnel needed
- Able to collect tons of information relatively easily

- Can be collected at any time
- No need for subject visits
- Can be collected anonymously

- Allows for more representative sample
- Potential for more responses if not conducted in person

- Allows for anonymous responses
- Able to reach an audience where it may not be illegal

CONS

- Ethical issues of “data scraping”
- Blurred line between public and private online (e.g. Social Media)

- Data confidentiality concerns, especially when using a third-party website
- Data validity (e.g. internet trolls, Amazon Mturk survey takers, etc.)
- Handling of anonymous participant complaints

CARE-Q
Consortium for Applied Research Ethics Quality

Data Scraping and the Blurred Line of Social Media

- What is “data scraping”?
  - The use of a software tool to extract data from an electronic system
  - Commonly describes a situation in which a software program extracts various pieces of data from a website (e.g. email address, phone number, etc.)
- Ethical Implications
  - What about data which was originally shared privately but then re-shared publicly?
  - What about data which is shared with “friends of friends” but not “publicly”?
  - What about confidential data that has been leaked (e.g. WikiLeaks)?

CARE-Q
Consortium for Applied Research Ethics Quality

Example

- Dr. Bashir has designed a data scraping “bot”
- He wants to use it to get email addresses to recruit for his study
- He will have the bot search for users who made any political post on various social media platforms (Facebook, Twitter, Instagram, and TikTok)
- The bot will then go to their profile and collect their email address
- Dr. Bashir will then send a survey link to each email address
- What issues should the IRB consider?

CARE-Q
Consortium for Applied Research Ethics Quality

QUESTIONS?

CARE-Q
Consortium for Applied Research Ethics Quality