Crowdsourcing and Human Subjects Research

Crowdsourcing and Human Computation
Lecture 21
Instructor: Chris Callison-Burch
TA: Ellie Pavlick

Website: crowdsourcing-class.org

Human Computation?

- So far we have focused on using crowdsourcing to solve problems that are beyond the reach of computers
- Alternately, we have used crowdsourcing to generate data for machine learning algorithms
- But many scientific disciplines are interested in studying people themselves

MTurk for social science research

- Many social science experiments require recruitment of a large number of subjects
- MTurk contains the major elements required to conduct research:
  - A participant compensation system
  - A large pool of potential participants
  - A streamlined process for study design, participant recruitment, and data collection

MTurk: A New Source of Inexpensive, Yet High-Quality, Data?

- Published in Perspectives on Psychological Science
- How Do MTurk Samples Compare With Other Samples?
- How Do Compensation Amount and Task Length Affect Participation Rates?
- How Does Compensation Amount Affect Data Quality?
- Do MTurk Data Meet Acceptable Psychometric Standards?

How Do MTurk Samples Compare With Other Samples?

- Field of psychology heavy relies on American college students for its samples
- The field started adopting Internet questionnaires only relatively recently
- Previously the field had a number of preconceptions that slowed adoption of the Internet for such surveys

Weird myths that people had about the Internet in 2004

Preconception	Finding
Internet samples are not demographically diverse	Mixed. More diverse than college students. Doesn't reflect whole population
Internet users are maladjusted, socially isolated, or depressed	Myth. Internet users do not differ from nonusers
Internet participants are unmotivated	Myth. You can provide means for motivating participants
Internet data do not generalize across different web sites	Myth. Findings replicate across sites (though samples differ)
Internet data are compromised by anonymity of participants	Fact. Repeat responders can be a problem. Be careful.
Internet-based findings differ from other methods	Myth. Internet-based findings are usually consistent

How Do MTurk Samples Compare With Other Samples?

- MTurk population is more diverse than college students (or non-students who reside in college towns)
- Good gender splits
- Good minority representation
- Large number of non-US participants

Other questions

- How Do Compensation Amount and Task Length Affect Participation Rates?
- How Does Compensation Amount Affect Data Quality?
- Do MTurk Data Meet Acceptable Psychometric Standards?

Compensation Amount / Participation Rates / Quality

- Similar to what we have discussed previously
- Participation is affected by compensation rate and task length, but participants can still be recruited rapidly and inexpensively
- Compensation rates do not appear to strongly affect data quality, but does appear to affect how quickly experiments are completed

Subject selection

- For psychology, cognitive science, behavioral economics, subjects can often be randomly drawn from the population at large
- For other types of experiments, researchers have to recruit subjects with specific traits (e.g. suffered from medical condition, part of a particular socio-economic group)

Using Mechanical Turk to Study Clinical Populations

- Participants with psychiatric symptoms, specific risk factors, or rare demographic characteristics are difficult to identify and recruit for participation in research
- But participants with these characteristics are crucial for research in the social, behavioral, and clinical sciences
- Can Mechanical Turk help?

Studying Clinical Populations

- People with some disorders are less likely to participate in clinical research and the mental health care system
  - Stigma associated with psychiatric illness
  - Direct consequence of their symptoms (e.g., social anxiety)
  - Online recruitment may alleviate anxiety
  - Recruiting from online communities may help find sufferers of rare disorders

Preconceptions about online clinical studies

- Adoption of online data collection methods has been slower in the clinical sciences than in other quantitative social and behavioral sciences
- Researchers have concerns about privacy and data quality
- They have administrative difficulties of managing online recruitment
- MTurk to the rescue!

Is MTurk viable for mental health data?

- What is the prevalence of clinical symptoms among Turkers?
- How feasibility is it to have Turkers to complete clinical surveys?
- Will they be honest when disclosing sensitive information online?

Is MTurk viable for mental health data?

- Shapiro et al (2013) tested Turkers for characteristics of depression and anxiety
- Asked about clinically relevant life events (e.g., trauma and drug and alcohol consumption)
- Attempted to assessed misrepresentation of basic demographic information and reporting of clinical symptoms

Study: part 1

- 500 US-based participants, paid $0.75 to complete a well-being study taking ~20 minutes
- Completed a mental health survey, which included a depression inventory, and anxiety index, a satisfaction with life scale
- Participants completed the “Infrequency-Psychopathology Scale” to assess the extent to which they report implausible psychological symptoms
- Participants also provided detailed demographic information

Study: part 2

- Followed up with same subjects 1 week later, paying $0.80. 400 returned.
- Repeated the depression inventory and demographic questionnaires, and added
  - A self-rating scale for mania,
  - Survey about 17 potentially traumatic events (e.g. fire or explosion)
  - Questions about recreational drug use
  - A screener for potential substance abuse

Reliability of participant reporting

- Participants were honest in reporting innocuous info – 97% consistently reported demographic info
- Did workers tended to fabricate psychiatric symptoms?
- *Infrequency-Psychopathology Scale* contains items that are rarely selected by healthy or clinical populations but are selected by those attempting to fake a psychiatric disorder
- 3% of Turkers were flagged as exaggerating

Psychiatric Demographics

Ever diagnosed w/ psychiatric/psychological condition	21%
Currently taking meds for condition	12%
Currently in talk therapy	6%
Screened positive for possible substance abuse	37%
Ever sought treatment for a substance abuse problem	4%
Ever diagnosed with chronic illness or physical disability	16%
Ever experienced a traumatic event	66%
Experienced ≥4 traumatic event	23%

Turkers v General population

- Equivalent levels of depression
- Are somewhat more prone to symptoms of social anxiety
- Have slightly lower satisfaction with life

Is MTurk viable for mental health data? Surprisingly, Yes.

- MTurk might actually be a useful resource for accessing and studying clinical populations
- Workers reported greater comfort disclosing clinical information in an online format than an a hypothetical in-person interview
- MTurk can be used to complete sophisticated research designs, including longitudinal studies, survey research, interviews

Human Subjects Research

Human Subjects Research

- In 1979 the federal government created the National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research
- It established ethical framework designed to protect human research subjects
- All human subjects research at universities is governed by this framework

Need for ethical principles and guidance

- Scientific research has produced substantial social benefits, but it poses troubling ethical questions
- In WW2, physicians and scientists conducted biomedical experiments on concentration camp prisoners
- U.S. Public Health Service conducted “Tuskegee Study of Untreated Syphilis in the Negro Male”
- Past social science studies carried risks of harm to psychological wellbeing
    - Milgram's Obedience to Authority Study
    - Zimbardo's Stanford Prison Experiment

Basic Ethical Principles

1. Respect for Persons – individuals should be treated as autonomous agents, and persons with diminished autonomy are entitled to protection
2. Beneficence – do not harm and maximize possible benefits and minimize possible harms
3. Justice – Who ought to receive the benefits of research and bear its burdens?

Respect for persons

- Autonomy – people must be empowered to make decisions concerning their own actions and wellbeing
- Mental Capacity – special care must be given for individuals with "diminished decision-making capacity" like children
- Voluntariness – people must be given choice to participate in or withdraw from research, and situations should be avoided where subjects feel pressure to participate

Informed consent

1. Disclose to potential research subjects information needed to make an informed decision
2. Help them the understand what has been disclosed
3. Emphasize the voluntariness of the decision about whether or not to participate in the research

Example: Impersonal Sex in Public Places

- In 1960s, a sociology grad student named Laud Humphreys studied gay men meeting for casual sexual encounters in public restrooms
- He gained the confidence of the men by pretending to be a participant and acting as a lookout
- Recorded license plate numbers of 100 regulars in order to contact them for later interviews
- Conducted in-home surveys in disguise and gathered data about their family relationships and religious background

Example: Impersonal Sex in Public Places

I already knew that many of my respondents were married and that all were in a highly discreditable position and fearful of discovery. How could I approach these covert deviants for interviews? By passing as deviant, I had observed their sexual behavior without disturbing it. Now, I was faced with interviewing these men (often in the presence of their wives) without destroying them.

Example: Harvard Facebook study

- Sociologists at Harvard University gleaned voluminous and detailed personal information from the Facebook profiles of an entire class of undergrads
- Created an extensive data set that included students' gender, home state, major, political affiliations, friend networks, photographs, and tastes in music, books, and film
- Made the data publicly available, removing students’ names.

Beneficence

- Most research in the social and behavioral sciences, education, and the humanities does not provide direct benefit to subjects, and risks of harm tends to be minimal
- Possible risk in psychological distress surrounding sensitive research topics and inadvertent disclosure of private information
- Studies of sexuality, mental health, interpersonal violence, and illegal activities expose subjects to embarrassment and reputational harm if private information becomes public

Anonymization

- Anonymize participants in order to protect their confidentiality
- Bits of information may be enough to allow for re-identification of subjects
- Store sensitive data securely
- Report only aggregate information

Is Mechanical Turk anonymous?

- Mechanical Turk already anonymizes subjects by exposing only their Worker IDs
  - I know that A23KO2TP7I4KK2 completed by task, but I don't know any other info (name, gender, location)
  - Great! Right?

https://www.amazon.com/gp/pdp/profile/A23KO2TP7I4KK2

amazon Try Prime

Recommendations Today's Deals Gift Cards Sell Help

Shop by Department

Search All Go

Hello. S
Your A

Your Amazon.com Your Browsing History Recommended For You Rate Items You Like Improve Your Recommend

Chris Callison-Burch

REAL NAME™

Helpful votes received on reviews: 50% (1

Reviews

Top Reviewer Ranking: 9,737,978 (Learn More) - Total Helpful Votes: 1 of

Pinzon Signature Diamond Matelasse Shower Curtain, White by Pinzon by Amazon.com

1 of 2 people found the following review helpf

HEAVY shower curtain, J

This shower curtain is very heavy. It p
the wall, and it seemed like it would bl
from my shower if I had managed to h
return it and just stick with a shower li

RSS feed · View Top Reviewers

Ethical experimentation

- Risks to subjects must be minimized using sound experimental design
- Risks to subjects must be reasonable in relation to anticipated benefits to subjects
- Research plan should make adequate to protect the privacy of subjects and to maintain the confidentiality of data
- How do you know if your experiments are ethical?

Human Subjects and Institutional Review Boards

- All research involving human subjects that is funded or regulated by the federal government must be approved by an IRB
- Definition of a **Human Subject**: a living individual about whom an investigator (whether professional or student) conducting research obtains:
  - Data through intervention or interaction with the individual, or
  - Identifiable private information

Private information

- Private Information includes:
- Information about behavior that occurs in a setting in which the individual can reasonably expect that no observation or recording is taking place
- Information that has been provided for specific purposes, other than research, where the individual can reasonably expect that it will not be made public (such as a medical record)

Authority of the IRB

- Approve research
- Disapprove research
- Modify research
- Conduct continuing reviews
- Observe / verify changes
- Suspend or terminate approval
- Observe the consent process and the research procedures

Composition of the IRB

- At least five members
- Member of both sexes
- Members that come from varied professions
- At least one member whose primary concerns are in nonscientific areas
- At least one member whose primary concerns are in scientific area
- At least one member who is not otherwise affiliated with the institution

IRB Requirements

- Risk / anticipated benefit analysis
- Informed consent and Assent process and documentation
- Selection of subjects
  - Equitable selection in terms of gender, race, ethnicity
  - Benefits are distributed fairly among the community's populations
  - Additional safeguards are provided for vulnerable populations susceptible to pressure to participate
- Safeguards and research plan for collection, storage, and analysis of data
- Information that demonstrates your research design / methods are scientifically valid and justify exposing subjects to research risks

IRB must also review

- The qualifications of the principal investigator (PI) and scientific collaborators
- A complete description of the proposed research
- Provisions for the adequate protection of rights and welfare of subjects
- Compliance with pertinent federal and state laws/regulations and institutional policy

PIs Must

- Protect the rights and welfare of human subjects who participate in research
- Understand the ethical standards and regulatory requirements governing research activities with human subjects
- Personally conduct or supervise the research
- Ensure that all staff, collaborators, and colleagues assisting in the conduct of the study are informed about the study, the regulations governing research, and the institutional policies.
- Ensure that all research activities have IRB approval and other approvals required by the institution before human subjects are involved.
- Obtain the informed consent of subjects before the subject is involved in the research and document consent as approved by the IRB

PIs Must

- Implement the research activity as it was approved by the IRB
- Maintain written records of IRB reviews and decisions and obtain and keep documented evidence of informed consent of the subjects or their legally authorized representatives
- Obtain IRB approval for any proposed change to the research protocol prior to its implementation
- Comply with the IRB requirements for timely reporting of unanticipated problems involving risks to subjects or others including adverse events, safety reports received from the sponsor, or data safety and monitoring summary reports
- Make provisions for the secured retention of complete research records and all research materials
- Ensure the confidentiality and security of all information obtained from and about human subjects

Failure to do so may result in

- Suspension of research project
- Suspension of all of a PI's research projects
- Inability to use data or publish results
- Notification of sponsors, regulatory agencies, and funding agencies of noncompliance
- Inability to receive funding from federal grants
- Termination of employment
- Loss of licenses
- Immediate shut-down of ALL research at an organization

Types of IRB approval

- Full/Convened Committee Review
- Expedited Review
- **Review for Exemption Status**

Research that is exempt

- Research conducted in established or commonly accepted educational settings, involving normal educational practices
- Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior
- Research involving the collection or study of freely available de-identified existing data, documents, records, pathological specimens, or diagnostic specimens
- Research and demonstration projects conducted by heads of government departments or agencies which are designed to evaluate public programs
- Taste and food quality evaluation and consumer acceptance studies

Additional tests

1. Does the research involve children?
2. Identifiability & Risks
    I. Will the human participants be identifiable, directly or through identifiers linked to the participants?
    II. Could any disclosure of the participants' responses place the participants at risk of criminal or civil liability or be damaging to the participants' financial standing, employability, or reputation?