HRPP Guidance for Crowdsourcing Tools

Crowdsourcing tools provide users access to information, opinions, or work from a cohort of people sourced through social media, the internet, and smartphone applications. Uses of crowdsourcing tools in research vary widely amongst different disciplines but most Americans can see the effects of these tools daily, from navigation systems helping them avoid traffic slowdowns to looking up information on the crowdsourced tool Wikipedia. In the field of behavioral science these powerful tools can be utilized to query data from populations that the investigators may not have previous access to.

From the use of surveys to mobile application tracking and recording, any investigator seeking to capitalize on crowdsourcing tools must begin the study design process by further researching the advantages, disadvantages, best practices, and ethical concerns of each platform.

Platforms such as Amazon's MTurk, Prolific Academic, and CrowdFlower are examples of virtual labor markets in which enrolled "workers" complete online Human Intelligence Tasks (HITs) for compensation. These platforms appeal to researchers for their large and often diverse populations of motivated participants.

Advantages of these platforms include cost-effectiveness, speed of data collection, and offered built in processes for study design, participant recruitment, and integrated compensation systems.

Disadvantages include data quality issues, potential for respondents to not be whom they claim to be, and the necessity of multiple candidate screening questions to disqualify crowdsourcing cheaters, speeders, and bots (1,2).

Q: Are these services truly providing anonymous participants?

As these platforms develop over time, they improve providing protections for their participants. Unfortunately, there are no guarantees that any platform can truly provide complete anonymity for their participants.

For example, studies utilizing MTurk that derive from internal Amazon survey templates must consider the connection between MTurk Workers identifications (IDs) and the worker's account profile, Amazon account, and other personal services that may be connected through the users' digital wallet. When it is possible to link survey participants to their identifications their participation is not anonymous. Studies should be designed to implement controls to safeguard participant privacy, data confidentiality, and any sensitive intellectual property.

Researchers who engage with any crowdsourcing tools, whether through the platform and building templates directly within those systems or linking participants to third-party systems like Qualtrics, must remain vigilant on how participants access the platform and how participant data is linked to identifiers (even temporarily). Contact UIW HRPP staff to accurately report the collection of identifiers and develop sufficient plans to provide confidentiality protection.

Q: What are some ethical considerations using crowdsourcing tools?

Please consider the following ethical considerations while designing a research study with crowdsourcing tools:

1. **Country of origin of participants.** Be aware of international participation on these platforms and the privacy and confidentiality regulations that may apply for your cohort.
2. **Financial compensation of participants.** Participants depending on the task requested may spend time resources and provide labor that merit greater than minimal industry standards. Financial compensation amounts may influence participation rates and bias participant composition.
3. **Inclusion of vulnerable populations.** Be mindful that many crowdsourcing platforms have "workforces" in developing countries and increase the risk of inadvertently targeting economically disadvantaged persons.
4. **The principles of the Belmont Report.** Evaluation of the three principles, Respect for persons, Beneficence, and Justice, should initiate and guide any study design involving human subjects. Design and thoughtful implementation of controls are essential for the use of crowdsourcing tools.

Q: What are some best practices you recommend?

1. **Consider pilot studies.** Design is critical. Formative pilot studies can identify issues with design and provide insight on platform issues that may need to be addressed prior to a full study.
2. **Make data quality a priority.** Create data quality standards to include timeliness of task completion and response patterns.
3. **Utilize attention checks** to minimize speeders and bots for those utilizing the platforms for survey-based studies.
4. **Research the do's and don'ts of crowdsourcing tools within your respective field and professional associations and organizations.** The breadth and depth of these tools has increased exponentially over the past decade. Reach out to subject matter experts within your professional communities.
5. **Make every effort to ensure that participants are humans.** Be sure to include information technology industry standards such as reCAPTCHA.

Additional Resources:

1. Shmueli et. al. Beyond Fair Pay: Ethical Implications of NLP Crowdsourcing, 2021. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 3758-3769. Accessed online on September 11, 2023 at: https://aclanthology.org/2021.naacl-main.295.pdf
2. Use of Amazon Mechanical Turk ® in Iowa State University Research. Created 6/2/2020. Website resource accessed online September 11, 2023 at: https://compliance.iastate.edu/wp-content/uploads/sites/4/pdf/MTurk-guidance.pdf

3. Cobanoglu et. al. A beginner's guide and best practices for using crowdsourcing platforms for survey research: The case of Amazon Mechanical Turk (MTurk), 2021. Journal of Global Business Insights, Vol 6, Issue 1: 92-97. Accessed online on September 13, 2023 at: https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=1177&context=globe

4. Resource from Lehigh University: Crowdsourcing platforms discussed with links to additional sources for best practices using MTurk and its competitor Prolific. Accessed online on September 13, 2023 at: https://research.cc.lehigh.edu/crowdsourcing Resource from the University of Pittsburgh: HRPO guidance to crowdsourcing. Accessed online on September 11, 2023 at: https://www.irb.pitt.edu/crowdsourcing-platforms

5. Resource from the University of Florida: IRB MTurk Guidance. Accessed online on September 11, 2023 at: https://irb.ufl.edu/wp-content/uploads/IRB-Mechanical-Turk-Guidance.pdf