Research Using Crowdsourcing Platforms

1.0 General Information

1.1 Crowdsourcing services are used for multiple reasons, but are becoming popular ways to recruit potential research study participants from the general public. These guidelines have been created to assist UWinnipeg researchers who are recruiting participants through a crowdsourcing service.

1.2 These guidelines have been reviewed by researchers in the field, UHREB members, as well the UWinnipeg Research Data Management Librarian. These guidelines are not policy. They are intended to support researchers. The guidelines are meant as an interpretation of existing UHREB policies (e.g., TCPS2 2018) in the context of crowdsourced research and are not meant to take precedence over them.

2.0 Amazon Mechanical Turk (MTurk)

2.1 The most common crowdsourcing platform used by UWinnipeg researchers is Amazon Mechanical Turk (MTurk)**.** Although this guidance was developed with MTurk in mind, elements of it may also be applicable to other crowdsourcing websites (e.g., Prolific Academic) or commercial survey panels (e.g., Qualtrics Panels).

2.2 Amazon describes MTurk as "a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually (www.mturk.com). This could include anything from conducting simple data validation and research to more subjective tasks like survey participation, content moderation, and more. MTurk enables companies to harness the collective intelligence, skills, and insights from a global workforce to streamline business processes, augment data collection and analysis, and accelerate machine learning development."

2.3 Academic researchers are using MTurk to access a large population of participants for research studies. Researchers generate a Human Intelligence Task (HIT) that provides a title and description of the online task (including average time required and compensation). MTurk was not designed for research studies with an eye to the experience of research participants and is subject to Amazon's policies, but has become a popular avenue of data collection. MTurk is primarily used for recruitment and then participants are rerouted to another site (e.g., Qualtrics) to engage in the research requirement. As the use of crowdsourcing websites has become more common in social science research, ethical issues have been raised about the shift to sampling from a population of 'gig-economy contract workers versus more traditional populations" (Auer et al., 2021). In particular, researchers should consider the vulnerability of participants to exploitation as 'workers,' and issues of compensation for time with potentially vulnerable populations.

3.0 Considerations & Recommendations

Academic researchers need to consider the following when submitting an ethics application for a study using crowdsourcing services for recruitment: risk level, recruitment, data collection and confidentiality, informed consent and data security, and debriefing.

3.1 Risk Level

Online crowdsourcing such as MTurk should generally be considered for minimal risk research. Research using MTurk or other crowdsourcing services for recruiting vulnerable populations, studying highly sensitive topics, or studies that require personal or health-related information would require strong justification to balance any risks to participants. Due to standards of informed consent, studies involving deception or partial disclosure should be justified and only engage in deception that would be unlikely to adversely affect the welfare of the participant.

3.2 Recruitment

When using MTurk or other crowdsourcing services, the title of the study and the description of the HIT act as a form of recruitment. Researchers should include the title and HIT description as part of the ethics application. In the description, researchers should include the time required (including screening), the nature of tasks, any additional software required, and compensation. As compensation must be provided even if the participant does not complete the tasks, wording in recruitment should be clear that participants may skip any and all tasks and still receive compensation when they reach the end of the survey. That is, barring unusual circumstances that are described in the application and part of the ethics review, researchers should not "reject" HITs for non-completion of a study.

3.3 Data Collection and Confidentiality

It is important to understand the risks associated with services like MTurk, and to take necessary precautions to protect participants' data.

3.3.1 Avoid using MTurk's internal HITs to collect response data. The data collected using MTurk's internal HITs are collected and stored by Amazon on Amazon servers (Mason & Suri, 2011). These data may become identifiable as response data can be linked with any other personal information collected by MTurk/Amazon, including IP addresses, name, age, address, etc. Additionally, with internal HITs, MTurk Worker IDs are automatically linked to survey responses. One study showed that MTurk Worker IDs can be linked to product reviews, ratings and any information provided on a personal Amazon Profile (Lease et al., 2013).

3.3.2 Researchers should avoid collecting personal information whenever possible, including MTurk Worker IDs. If it is necessary to collect Worker IDs, this should be disclosed on the consent form and researchers must take all precautionary steps to ensure that they are kept confidential and secure, and are not linked back to survey data.

3.3.3 Instead of MTurk's internal HITs, it is recommended that researchers use MTurk as a recruitment tool, by embedding a link in MTurk that redirects workers to an external online survey service such as Qualtrics. With this option, survey data are not collected or accessible to Amazon.

More information on MTurk's internal HITs vs external HITs here:
https://cphs.berkeley.edu/mechanicalturk.pdf

3.4 Informed Consent and Data Security

3.4.1 Researchers have the responsibility to ensure participants are informed about, and consent to, how their data is managed throughout and after the research study. When using

crowdsourcing services like MTurk, as with many research practices, complete confidentiality or anonymity cannot be guaranteed to participants. Crowdsourcing services like MTurk often collect and store personal information about their users, thereby posing potential privacy risks to research participants which should be included in informed consent.

3.4.2 Researchers must clearly indicate in the consent form the Terms of Use or Privacy Policy of the crowdsourcing service they are using so participants are informed about how their data is being collected and by whom. Further, researchers should indicate the steps that will be taken to minimize security risks to the data that are collected.

3.4.3 Researchers should be aware of any policies or procedures in place by MTurk and/or online survey software companies when a breach of data occurs.

3.4.4 The first page of the online survey should be the consent document. The online consent will have all of the elements of a regular consent, but can adapt the 'next' button to "I consent to participate" on the informed consent page.

3.4.5 Recommended Informed Consent Language (Researchers are strongly encouraged to carefully consider, select, and adapt the language provided to suit the specific needs of their research.)

- *The research team associated with this survey does not have access to any personal information obtained by MTurk/Amazon. The research team will have access to the information you provide in your responses, collected using Qualtrics, a separate research survey software. As with all research, there is a chance that confidentiality could be compromised; however, we are taking the following precautions to minimize this risk concerning the data we collect: [Explain data security measures to be taken, e.g., coding, encryption, password protection, storage, limited access to data, destruction of code key after data is collected etc.].*

- *The research team associated with this survey does not have access to your personal information obtained by MTurk/Amazon, only the information you provide in your responses, collected using Qualtrics, a separate research survey software. This survey asks respondents to provide their MTurk Worker ID for the purpose of paring responses collected at different points in time. The research team will protect your MTurk Worker ID by storing it securely and it will be accessed only by approved members of the research team. All personal identifying information collected about you, including your MTurk Worker ID, will be destroyed once it is no longer needed for the study. As with all research, there is a chance that confidentiality could be compromised; however, we are taking the following precautions to minimize this risk concerning the data we collect: [Explain data security measures to be taken, e.g., coding, encryption, password protection, storage, limited access to data, destruction of code key after data is collected etc.].*

3.5 Debrief

If the researchers are using deception or partial disclosure, then debriefing at the end of the survey is required. This debriefing form would be embedded into the last page of the survey and, before exiting, the participant would be forwarded to the debrief. The participant would be asked to answer a final question allowing researchers to use their data (or not use their data) now that they know the true purpose of the study.

4.0 References

Amazon Mechanical Turk, "Amazon Mechanical Turk." Retrieved on January 6, 2021 from https://www.mturk.com/

Auer, E.M., Behrend, T.S., Collmus, A.B., Landers, R.N., Miles, A.F. (2021). Pay for performance, satisfaction and retention in longitudinal crowdsourced research. *PLoS ONE, 16*(1). https://doi.org/109.1371/journal.pone.0245460

Lease, M., Hullman, J., Bigham, J., Berbstein, M., Kim, J., Lasecki, W., Bakhshi, S., Mitra, T., & Miller, R. (2013). Mechanical Turk is Not Anonymous. Available at SSRN: https://dx.doi.org/10.2139/ssrn.2228728

Mason, W. & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. Behavioral Research Methods, 44(1), 1-23. Behav Res. https://doi.org/10.3758/s13428-011-0124-6

Portage Network, "Sensitive Data Toolkit for Researchers Part 3: Research Data Management Language for Informed Consent." Retrieved on January 6, 2021 from http://doi.org/10.5281/zenodo.4060460

University of California, Berkeley, "Mechanical Turk (MTURK) for Online Research." Retrieved January 6, 2021. Retrieved January 6, 2021 from https://cphs.berkeley.edu/mechanicalturk.pdf

University of Massachusetts Amherst, "MTurk Guidance." Retrieved on January 6, 2021. Retrieved: https://www.umass.edu/research/guidance/mturk-guidance