

**Title: ETHICAL CONCERNS ARISING FROM RECRUITING PARTICIPANTS
FROM AMAZON'S MECHANICAL TURK AS RESEARCH PARTICIPANTS**

Short Title: *ETHICAL CONCERNS*

Author information:

Ilka Helene Gleibs (ORCID ID: <https://orcid.org/0000-0002-9913-250X>)* & Nihan

Albayrak-Aydemir (ORCID ID: <https://orcid.org/0000-0003-3412-4311>)*

London School of Economics and Political Science

* These authors contributed equally to this work. Correspondence should be addressed to Ilka Helene Gleibs (Address: London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, United Kingdom; E-mail: i.h.gleibs@lse.ac.uk).

Keywords: data collection; ethics; Amazons Mechanical Turk; online platforms; participant recruitment; research practice

Author contributions:

Conceptualisation: Ilka Helene Gleibs & Nihan Albayrak-Aydemir.

Writing – the original draft: Ilka Helene Gleibs & Nihan Albayrak-Aydemir.

Writing – review & editing: Ilka Helene Gleibs & Nihan Albayrak-Aydemir.

Conflicts of interest:

The authors declare that there is no conflict of interest.

Abstract:

In this commentary, we respond to Burnette et al.'s (2021) paper, which gives significant practical recommendations to improve data quality and validity while gathering data via Amazon's Mechanical Turk (MTurk). We argue that it is also important to acknowledge and review the specific ethical issues that might arise when recruiting MTurk workers as participants. We particularly raise three main ethical concerns that need to be addressed when recruiting research participants from participant recruitment platforms: participants' economic vulnerability, participants' sensitivity, and power dynamics between participants and researchers. We elaborate on these issues by discussing the ways in which they may appear and be responded to. We conclude that considering the ethical aspects of data collection and the potential impacts of data collection on those involved would complement Burnette et al.'s recommendations. Consequently, data collection processes should be transparent as well, in addition to data screening processes.

Keywords: data collection; ethics; Amazons Mechanical Turk; online platforms; participant recruitment; research practice

ETHICAL CONCERNS ARISING FROM RECRUITING WORKERS FROM AMAZON'S MECHANICAL TURK AS RESEARCH PARTICIPANTS

Burnette et al.'s (2021) paper is timely and an important piece that raises central issues on the use samples recruited from Amazon's Mechanical Turk (MTurk), especially when dealing with more *vulnerable* populations. We appreciated the authors' efforts to openly share their experience in recruiting MTurk workers as research participants and to outline the steps they took before and after data collection to increase data validity. Based on their experience, they make a number of recommendations to improve data quality when collecting data from MTurk that are practical and should be widely used. However, we believe that there are also further ethical issues when recruiting MTurk workers as research participants that are important to consider here and go beyond the questions of data quality and validity. In particular, we have three main concerns.

Our first concern is about the economic vulnerability of some participants, which is often overlooked by scientists. Many MTurk participants are contingent workers; some of them are unemployed or underemployed. MTurk could be an important source of income for many of these workers. Although Burnette et al. (2021) paid above US federal minimum wage, many of their participants expressed financial pressures when a HIT was not approved. Issues of fair pay are important and therefore should also be included in the recommendations. This can include careful structures for who gets paid in our research and who is not. As scientists, we should act responsibly towards our participants and unequivocally state the conditions under which the payment can or cannot be made. It is also vital that every effort of potential participants in our work should be factored in in payment decisions. For instance, the authors share that the participants felt tricked or misled by two screening questions and that the survey ended for them if they did not pass the screening questions or if they failed the attention check. It is, however, unclear whether participants

received payment for the time they spent until the survey ended for them. If not being paid unless providing useful data, then participants must at least be informed about the criteria for payment (as well as the conditions under which there will be no payment) before starting the study. Importantly, researchers should strive to consider the norms and preferences of the microtask and crowdsourcing economy when designing compensation schemes. These expectations, and challenges have been articulated by worker communities themselves (wagepledge.org, 2021) and should be taken into account when using MTurk as a means to collect data.

A second concern we have is about potential vulnerability based on mental health problems among MTurk workers. Some studies point to the fact that mental health issues are higher amongst participants on MTurk compared to the general population (Arditte, Çek, Shaw, & Timpano, 2016). This situation, especially when dealing with additional sensitive topics such as eating disorders, increases their risk of being harmed by the research through reporting on sensitive topics. When researching sensitive topics in the lab or in person, researchers can directly react when they recognise that someone is distressed. In the remote situation of MTurk, this is not possible. So, not recognising the sensitive conditions of the participants could raise severe ethical concerns. Thus, we suggest that MTurk workers should be recognised as a sensitive population, especially when studying clinically relevant topics, such as eating disorders. Any pre- and post-data collection procedure (e.g., ethics application, interpreting findings) should therefore factor in this sensitivity and provide potential ways of minimising risks of re-traumatisation (for example, by giving information on mental health help-lines and access to mental-health care where possible).

The final concern, which is interrelated with the concerns discussed before, is about the power dynamics on participant recruitment platforms, which is commonly disregarded by scientists. Participants do not equal data and they should not merely be treated as such,

especially on participant recruitment platforms, such as MTurk. These platforms have control mechanisms that typically allow greater flexibility and agency for those who collect data than those who provide it. For example, it is common practice to select participants based on their HIT approval ratio and those whose rate is above 95%. Nonetheless, this ratio is not an objective measure as there may be various reasons for why participants' data is rejected. Outliers may, for instance, be seen as fakers and rejected, and this would decrease their approval rate. Similarly, participants might sometimes miss the numeric password to receive hit approval for payment and, therefore, try to open the same survey again. However, they are not allowed to do so because of using the same IP address. Eventually, the power researchers hold on these participant recruitment platforms would not only affect the research but the participants and potential participants as well. For this reason, better control mechanisms are needed on these platforms to ensure that both parties are ethically treated and seen as equally respectable in the data collection processes. Some examples of this could be providing an option for participants to withdraw consent while still being compensated through custom survey codes and following the code-of-conduct provided by the MTurk community, for example Turker Nation.

Although there might be other ethical concerns, we believe the points raised above require urgent attention from scientists. It is our responsibility not to exploit participants. Therefore, we should not only evaluate data collection processes taking place on these platforms considering data validity and quality but we also have to investigate the wider ethical aspects of data collection and their implications on everyone involved in the data collection processes. In that sense, we support the authors' suggestion that data screening processes should be transparent and reproducible. Yet, we also argue that this should be extended to data collection processes (before, during, and after data collection) as well.

Having transparency in data collection and screening processes would then, at least, help us to better evaluate research practices and improve upon them.

References

- Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological assessment, 28*(6), 684. doi: 10.1037/pas0000217
- Burnette, C. B., Luzier, J. L., Bennett, B. L., Weisenmuller, C. M., Kerr, P., Martin, S., Keener, J., & Calderwood, L. (2021). Concerns and recommendations for using Amazon MTurk for eating disorder research. *International Journal of Eating Disorders*. Advance online publication. <https://doi.org/10.1002/eat.23614>
- WagePledge.org. 2021. Crowdsourcing Wage Pledge - FAQ, version 30 August 2021. Accessed 30 August 2021. Available online at <https://wagepledge.org/faq.php>.