

How to Conduct Valid Social Science Research Using MTurk – A checklist

The use of Amazon’s Mechanical Turk (MTurk) for social science research has increased exponentially in recent years. Although there is great excitement about the practical and logistical benefits, there is justifiable skepticism about the validity of research using data collected with MTurk. In this post, Herman Aguinis, Isabel Villamor, and Ravi S. Ramani provide 10 actionable best-practice recommendations and a checklist that can serve as a catalyst for more robust, reproducible, and trustworthy MTurk-based research.

Data collection is often the biggest logistical challenge faced by most social scientists. It is not surprising then that Amazon Mechanical Turk ([MTurk](#)) has quickly become ubiquitous. Why? Collecting data using MTurk is fast, inexpensive, and allows researchers to implement different types of research designs with sample participants from around the world. But, MTurk is no panacea and there are significant concerns about the validity of MTurk data and whether research results and conclusions based on those data can be trusted.

Specifically, in our article just published in [Journal of Management](#) (see [video abstract](#)) we describe 10 challenges to collecting data using MTurk: (1) inattention, (2) self-misrepresentation, (3) self-selection bias, (4) high attrition, (5) inconsistent English language fluency, (6) non-naïveté, (7) growth of MTurker communities, (8) vulnerability to web robots (or “bots”), (9) social desirability bias, and (10) perceived researcher unfairness. These are sufficiently serious that may render social science research flawed—and even misleading.

So, *what can researchers, journal reviewers and editors, and research consumers including funding agencies do to minimize these threats and improve the transparency and reproducibility of future MTurk-based research?* We provide 10 evidence-based best practice recommendations organized around the planning, implementation, and reporting of result stage of research. Here’s a brief summary.

Planning Stage

For trustworthy research, “[an ounce of prevention is worth a pound of cure.](#)” Given MTurk’s unique validity threats, careful consideration during this stage is even more essential. Recommended actions at this stage include:

1. Evaluate Appropriateness of MTurk to Develop or Test Theories. MTurk participants (MTurkers) can differ from more traditional samples. Rather than assuming comparability, researchers can:

- Evaluate alignment between desired target population and that of MTurkers
- Collect and report detailed sample characteristics

2. Decide Qualifications Used to Screen MTurkers. To counter threats due inconsistent MTurker English language fluency, self-misrepresentation, and non-naïveté, researchers can:

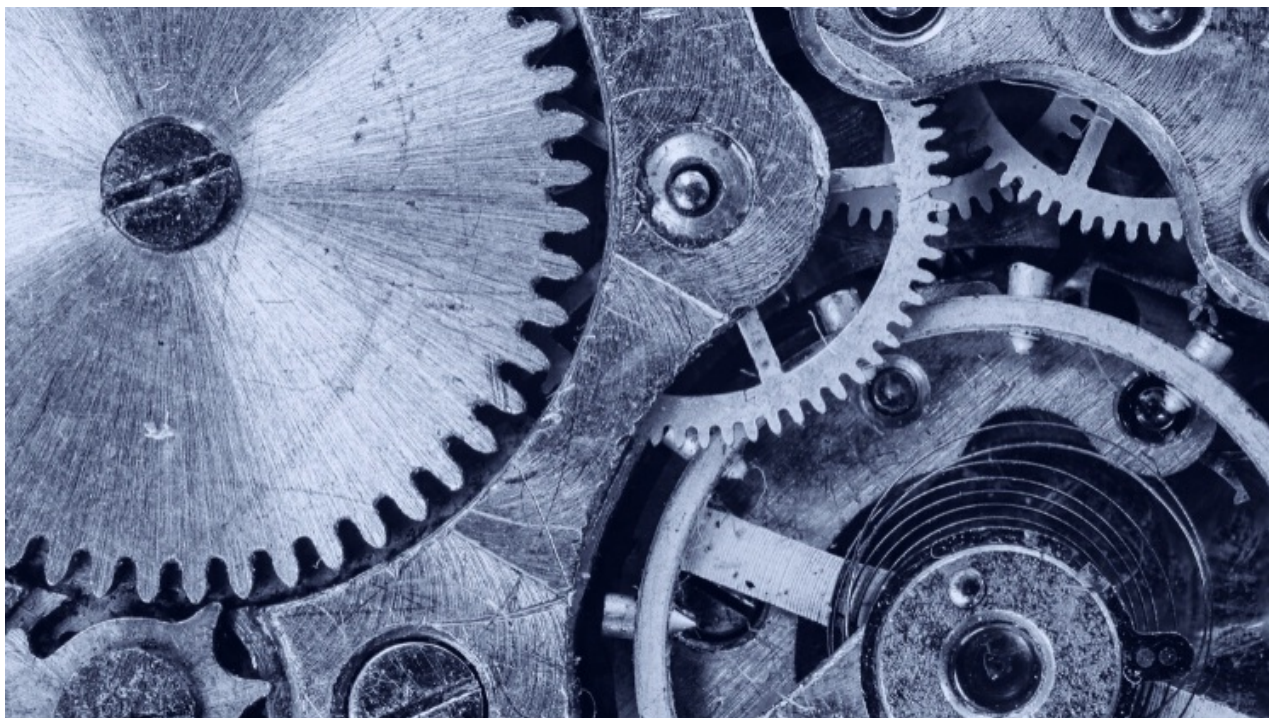
- Decide qualifications (e.g., demographics) relevant to the study
- Evaluate MTurkers using a screener study, and eliminate those who do not match desired criteria
- Determine whether to include only MTurkers from native-English-speaking countries (based on IP address), or whether measurement equivalence will be established
- Decide whether to use only highly qualified MTurkers (i.e., “Master Workers”), or to employ screening questions to gauge MTurker familiarity with research subject, stimuli, and, if applicable, manipulations

3. Establish Required Sample Size. Many responses are unusable due to high attrition rates and MTurker inattention. Therefore, in addition to the sample size determined through power analysis, researchers can:

- Collect data from at least an additional 15%-30% of MTurkers

4. Formulate Compensation Rules. Clear rules regarding compensation help address perceived researcher unfairness, while higher pay is linked to high-quality data. Therefore researchers can:

- Pay U.S. minimum wage or equivalent dependent on sample.
- Consider criteria (if any) used to refuse payment to MTurkers
- Use a consent form that includes details of compensation rules



5. Design Data-Collection Tool Used to Gather Responses. Well-designed tools can help researchers address threat due to web robots, self-misrepresentation, inattention, and perceived researcher unfairness. Thus, researchers can:

- Require MTurkers complete an informed consent form, including a “Captcha” verification
- Require MTurkers to provide MTurk ID and maintain database of past participants
- Use at least two attention checks
- Include an open-ended qualitative question
- Design a short study (approximately 5 minutes)
- Avoid using scales that only have “end” points labelled
- Include “quit study” and “contact researcher” option on each page of study

6. Craft the MTurk Task or HIT (i.e., “Human Intelligence Task”). A major MTurker complaint is that study directions are unclear. Thus, researchers can:

- Provide a detailed description of the study, accurate time commitment, describe what MTurkers will be asked to do, and specify compensation rules
- Avoid cues that might divulge the study’s aims or motivate MTurkers to engage in self-misrepresentation, or exhibit social desirability bias

Implementation Stage

Three specific actions can be taken at this stage.

7. Launch the Study, Monitor Responses, and Respond to Concerns. Researchers can:

- Conduct a pilot test with 10 to 30 participants that includes an open-ended question requesting feedback
- Monitor MTurker communities to gauge reactions to study
- Respond promptly to any questions or concerns raised by participants

8. Screen Data. Researchers can:

- Screen data using at least two or more tools to estimate unusable responses (e.g., MTurker self-reports of effort, answers to attention checks, response patterns and response times, statistical tools to evaluate consistency and identify outliers, IP addresses, and open qualitative questions)
- Adjust number of potential participants to achieve desired sample size

9. Approve or Deny Compensation for Completed Responses. Researchers can:

- Approve or deny compensation for responses within 24 to 48 hours of MTurker completing study
- Specify reason for rejecting compensation

Reporting Stage

10. Report Details to Ensure Transparency. Providing detailed information is key given a documented lack of transparency in MTurk-based studies. Therefore, researchers can:

- Report information regarding all procedures followed, decisions made, and results obtained during each stage of study?
- Provide data for future, secondary analyses (e.g., meta-analyses) of findings (e.g., demographic data, means, standard deviations, effect sizes)?
- Report details regarding HIT posting, qualifications used, and detailed sample characteristics
- Explain decisions regarding use of attention checks and screening techniques, including number of participants excluded for each, as well as decisions regarding sampling and non-naiveté
- Detail characteristics of study, including time commitment required and compensation provided?

Our consolidation of evidence-based best-practices provides actionable guidance for researchers considering MTurk. Journal editors and reviewers can use our checklist to evaluate the rigor and transparency of submitted manuscripts and provide developmental feedback, while practitioners can also use our recommendations to determine whether research using MTurk is sufficiently trustworthy.

*More detailed information about using MTurk for research can be found in the authors' paper, [MTurk Research: Review and Recommendations](#), published in the *Journal of Management*.*

Note: This article gives the views of the authors, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

Image Credit: [Pavlofox](#) via Pixabay.
