**How to Use**

This template provides a simple and standardized format for reporting Amazon Mechanical Turk (AMT) results in the social sciences.[[1]](#footnote-1) This version is a minimal reporting template, including a recommended set of quantities to allow reviewers and readers to evaluate the general quality of the data, its applicability, and possible limitations or problems.

Because of the variety of possible uses and structures of AMT studies, a more extensive version is in preparation with a modular design providing recommendations for additional reporting and validation checks depending upon the purpose and design of the task.

Items in the first section should be included in all studies reporting AMT results. Items in the second section should be included whenever germane to the design of the study. Investigators are further encouraged to maintain a public repository with this documentation, copies of all instruments, and (when possible) an anonymized copy of the original output.[[2]](#footnote-2)

One such repository is provided at <https://github.com/ndporter/asamturk>. It also includes a suite of freely adaptable tools for Stata to help prepare raw MTurk output for analysis and public archival.

**Template for Reporting Social Scientific Data Collected using Amazon Mechanical Turk\***

|  |  |  |
| --- | --- | --- |
| **Recommended for all studies** | | |
| **Item** | **Description** | |
| Batch | Name or signifier of batch | |
| HITs | Number of HITs (unique cases in input file) | |
| Workers per HIT | Number of workers assigned to complete each HIT (e.g. provided identical input) | |
| Date(s) | The date(s) and time period during which the batch was collected | |
| Instrument(s)+ | HTML, complete description, or screen capture of instrument(s) for tasks exactly as implemented | |
| Source of input data | What defines cases in the input file and where the data are originally derived from | |
| Output variables | Descriptives or frequencies for output variables used in analysis (including missing patters and worker demography if applicable) | |
| Qualifications | List of requirements for workers to accept HITs (standard or custom) | |
| Rejection criteria | Description of how decision was made to approve or reject assignments | |
| Rejection rate | Proportion of assignments rejected and (when applicable) Type I and II error rates | |
| **Recommended whenever applicable** | | |
| **Item** | **Description** |
| Additional validation check(s) | |  | | --- | | Any additional procedures (other than qualifications or rejection criteria) used to verify data quality. These could include:   * Consistency between multiple workers on the same HIT (inter-coder reliability) | | * Accurate completion of items with known correct answers included in HIT | | * Worker attention checks (questions with obvious correct answers to ensure workers are reading questions and following directions) | | * Confirmation in later sequential HITs | | * Consistency with another method (e.g. machine coding or trained coders) | |
| Third-party tools | Name and version number (or date, if non-versioned) of any third party tools such as Qualtrix or SurveyMonkey used to administer HITs externally |
| Design features | Precise description of any contingency, experimental, or quasi-experimental design that is not clear from the instrument (often requires third-party tool) |
| Sampling methodology | Information on sampling cases for input (not workers), including the population being sampled, how cases were selected for inclusion, and whether with replacement |
| Weights | List of any weight or adjustment variables and their derivation |
| Panel attrition | Standard panel attrition statistics for longitudinal data collection |
| Repeat worker rate | For surveys, experiments, and other tasks collecting information about workers, what proportion of HITs were completed by workers who had already completed one or more HITs in the study? |
| Repeat worker consistency | For tasks collecting information about workers, what proportion of demographic responses was consistent between HITs by the same worker? |

\* Unless identical across batches, items should be reported for each batch of data collected using AMT  
+ We recommend starred items be included in reporting table as the URL of an online repository

1. We anticipate the template may be usable for other crowdsourcing platforms with only small modifications, but focus on AMT as the largest and most established platform for academic use. [↑](#footnote-ref-1)
2. Recommended locations for repositories are within online supplements to an article, open-access data archives, institutional repositories, or GitHub repositories. [↑](#footnote-ref-2)