Mechanical Turk: Current Issues and Solutions

JOSEPH K. GOODMAN

THE OHIO STATE UNIVERSITY
FISHER COLLEGE OF BUSINESS

Special Thanks to my Coauthors

GABRIELE PAOLACCI
ERASMUS UNIVERSITY

SCOTT WRIGHT
PROVIDENCE COLLEGE
Agenda

Basics of Mechanical Turk: Issue or Non-issue?
What about editors and reviewers?

Solution to Many Problems: TurkPrime
- √ Screen based on demographics or any other criteria that you choose
- √ Provide real random completion codes and automatically approve
- √ Block repeat participants
- √ Calculate Completion & Bounce rates
- √ MicroBatch (or HyperBatch)
- √ Block duplicate IPs
- √ Anonymize worker IDs
MTurk Myth or Truth?

1. Not real people
2. Low quality data
3. Uneducated
4. Poor/low income
5. Internet geeks
6. The lab is still more reliable
7. They don’t pay attention
8. Mostly Indian
9. Mostly non-native English speakers
10. Editors/reviewers won’t accept papers with only MTurk data
   ◦ Surveyed marketing researchers at the top 150 schools (CB and non-CB)
   ◦ Asked about their use and opinions of MTurk
What Do Researchers Think about MTurk?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Neither agree nor disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers should not have all their data from MTurk</td>
<td>4.68</td>
<td>4.68</td>
<td>4.72</td>
</tr>
<tr>
<td>Mturk studies should be rerun using a non-Mturk sample.</td>
<td>3.64</td>
<td>3.15</td>
<td>3.76</td>
</tr>
<tr>
<td>Mturk samples are better than student samples.</td>
<td>4.00</td>
<td>3.96</td>
<td>4.04</td>
</tr>
</tbody>
</table>

Mechanical Turk: Current Issues and Solutions, SCP Pre-Conference 2018

Joseph K. Goodman, PhD
AE’s & Editors are Not All Convinced

"Papers should not have all their data from Mturk"

- Strongly Disagree
- Neither agree nor disagree
- Strongly agree

Distribution

<table>
<thead>
<tr>
<th></th>
<th>All P's (including quant and CCT)</th>
<th>ERB</th>
<th>AEs/Editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.11</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>0.14</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>0.06</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>5</td>
<td>0.23</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>6</td>
<td>0.18</td>
<td>0.08</td>
<td>0.15</td>
</tr>
<tr>
<td>7</td>
<td>0.40</td>
<td>0.30</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Solution 1: Stop Doing MTurk Studies
Solution 2: Get a job at Ohio State

(OR OTHER SCHOOL WITH 8000 SUBJECT HOURS/YEAR)
Solution 3: Measure Real Behavior on MTurk

Example$_1$: Tonietto, Malkoc, & Nowlis (forthcoming *JCR*)
- P’s entered their availability/calendar on Qualtrics
- IV: Generate times bounded/not bounded by tasks (btw SS)
- DV: Choice of MTurk study
  - 30 minute brief study for $2.50
  - 45 minute full study for $5.00
  - Neither

Example$_2$: Coin flip game to measure cheating
(Balasubramanian, Bennett, and Pierce 2017)
Solution 3: Measure Real Behavior on MTurk

Example 3: Take picture of artwork and upload or email it

Example 4: Costello and Malkoc (working paper)
- DV: Donate $ or donate time (micro-donations: counting/classifying animals, algae, weather patterns)
Solution 4: Go Beyond Surveys

Longitudinal studies
- 60% retention rates (but need to pay more)

Group interactions

Webcams for facial expression analysis and analyzing emotions (FaceReader)

Physiological measures
- Measure heart rate (Muender et al. 2016)

RA Work
- Classification of content
- web searches
- registering for websites
- viewing ads
- edit transcripts
- test software
- skilled tasks: logo & graphic design, ad copy
- See Fiverr, Upwork, Guru
Issue: Pre-screening & Non-naiveté

Workers are usually honest, but will lie to be eligible for participation
- 48/106 heterosexual workers claimed LGB to be eligible for study
- 8% women, 17% men changed gender to be eligible for study

Lying increases with compensation
Lie more about owning a dog than religion or politics

Non-naiveté
- We are sampling the same people across studies

Chandler & Paolacci 2017, Wessling, Huber, & Netzer 2017
A Solution to Many Problems: TurkPrime.com

- Pre-Screen based on demographics or other criteria (eg, liberals vs. conservatives)
- Non-naiveté: Block repeat participants
- Calculate Completion & Bounce rates
- Provide real random completion codes and automatically approve
- MicroBatch & HyperBatch
- Block duplicate IPs
- Anonymize worker IDs
Pre-screening on TurkPrime: Two Ways

Pay to use TurkPrime’s criteria
- Expensive, but lots of options and fast

Create your own database (see supplemental slides)
- Stage 1: Collect important info (e.g., gender, age, political affiliation)
- Stage 2: Launch HIT limited to target group
Guidelines for Researchers

Use qualifications and appropriate pre-screening
- > 95% approval
- Do not expect workers to screen themselves (use TurkPrime or other solution)

Be Ethical – pay a fair wage
- Min wage of $7.25/hr in US ($0.121/min)
- Paying more ≠ better data

Collect Data and Track your participants
- Collect basic demographics
- Prevents non-naiveté & can do anonymously via TurkPrime

Guidelines for Academic Requesters
http://wiki.wearedynamo.org/index.php/Guidelines_for_Academic_Requesters
- Clearly identify yourself
- Provide reasonable time estimates
- Approve work as soon as possible
- Maintain worker privacy
- Do not block workers
- Maintain a responsive line of communication

Know your reputation. Check out 3rd party websites:
Supplemental Slides: Instructions on How to Create Your Own Database of Workers
Many Issues are Solved by TurkPrime

See our paper "TurkPrime.com: A versatile crowdsourcing data ..." in Behavior Research Methods

Recruit Online Participants

Simple • Affordable • Powerful
Crowd|Research

Toolkit for Amazon Mechanical Turk
- Basic Toolkit Free for Academic Users
- MTurk Panels on Your Own Requester Account

Lab Service: Mechanical Turk Panels
- Target MTurk Workers by Ethnicity, Age and Custom Traits

Lab Service: Prime Panels
- Recruit from 20 Million Participants by any Demographic

Over 5000 Labs and Researchers Worldwide Use TurkPrime

Quinnipiac University
Tulane University
University of Waterloo
Université de Montréal
Nanyang Technological University

Joseph K. Goodman, PhD
Create a Study

Run a study on your own MTurk account

- MTurk account required.
- Gives you the flexibility to set payment amounts and other custom settings.
- Directly communicate with workers and monitor your own study progress.

Run a fully managed MTurk study

- No MTurk account needed.
- TurkPrime manages worker communication and study maintenance.
- Reach your desired population with our TurkPrime guarantee.

Need to run a complex study? Click here for a price quote.
**MTurk Toolkit**

Worker Payments and MTurk Fees funded with your MTurk Requester account. TurkPrime Lab Fees are funded through your TurkPrime Lab account.

1. **Panel Options (Optional)**
2. **Basic Info**
3. **Describe HIT**
4. **Setup HIT and Payment**
5. **Worker Requirements**
6. **MicroBatch (Pro Feature)**
7. **Pro Features**

---

**Total Number of Survey Participants**

The number of participants you want to take your HIT

<table>
<thead>
<tr>
<th>Category</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td></td>
<td>Less than $10,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$150,000 or more</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Over 70</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>All selected</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td>All selected</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td>All selected</td>
</tr>
<tr>
<td>Household Income</td>
<td></td>
<td>All selected</td>
</tr>
</tbody>
</table>

**Feasibility**

- Your study will most likely complete.

Note: The panel fee is a combination of Pro Features fees and is funded through your TurkPrime Lab account.

- **Panel Cost per Worker**: $0.00
- **Range**: $0.15 - $0.75/Worker
- **Number of Workers**: 1
- **Total Panel Cost**: $0.00 (0.00 x 1 Workers)

---

Joseph K. Goodman, PhD
<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic Disability</td>
<td>Do you have a disability or chronic illness of any type (either physical or mental)?</td>
</tr>
<tr>
<td>Migraines</td>
<td>How often do you experience migraines, if at all?</td>
</tr>
<tr>
<td>Future Exercise</td>
<td>Do you intend to engage in regular exercise in the next six weeks?</td>
</tr>
<tr>
<td>Dieting</td>
<td>How often are you dieting?</td>
</tr>
<tr>
<td>Fitness Tracker</td>
<td>Do you regularly use a pedometer or any other type of fitness tracker?</td>
</tr>
<tr>
<td>Chronic Pain</td>
<td>Do you have a diagnosis of chronic pain (pain lasting more than 12 weeks)?</td>
</tr>
<tr>
<td>Occupation</td>
<td>Which best describes your current occupation/job/career; If student, career goal?</td>
</tr>
<tr>
<td>Current Status</td>
<td>Which of the following best describes your current status?</td>
</tr>
<tr>
<td>Industry</td>
<td>Please select the industry that best describes your work and/or field of training;</td>
</tr>
<tr>
<td>Current College Student</td>
<td>Are you currently?</td>
</tr>
<tr>
<td>Supervisor</td>
<td>Are you employed full time in a supervisory position or higher where people report directly to ye</td>
</tr>
<tr>
<td>Hours Employed</td>
<td>How many hours per week do you work?</td>
</tr>
</tbody>
</table>
**Title**
Short Survey

Title of the survey that will be displayed to workers

**Description**
Answer a short survey. Takes about 1 minute!

A description of your survey

**Custom Instructions**
Survey takes about 1 minute.

*Note: This HIT is an academic study being conducted by university researchers. Your answers are completely anonymous and we are just interested in your opinions. It is not being conducted for commercial purposes or to make profit, and it won’t be sold or used by a company or marketers. You don’t have to search the internet on this HIT. It’s just opinion questions, so please take your time to answer the questions honestly.

**Requirements:**
- Must be 18 or older

(Optional) It is best to include instructions on your survey landing page. Nevertheless, you may include custom instructions to display to workers.

Standard instructions that workers should follow the hyperlink and submit a secret code (if set) are always included and will follow the custom instructions, if specified.

**Keywords**
- survey, study, academic, preferences

Keywords associated with your HIT

- This project may contain potentially explicit or offensive content, for example, nudity
### Worker Payment Per Survey

| $ | 0.1 |

Amount you will pay each worker who completes your HIT.

### Expected time to complete HIT

| 1 | Minutes |

The estimated amount of time for Worker to complete HIT. This will be appended to your HIT description.

### Time allotted per assignment

| 10 | Minutes |

Maximum time allowed for a Worker to complete the HIT.

### HIT expires in

| 7 | Days |

The amount of time you want your HIT to be active for.

### Survey Hyperlink

[https://osu.r1.qualtrics.com/jfe/form/SV_3xa1WCRuL9L5x](https://osu.r1.qualtrics.com/jfe/form/SV_3xa1WCRuL9L5x)

The URL of your survey.

### Query String Parameters

TurkServer will add the following query string parameters to your URL so you can take advantage of them in your data collection. The parameters are: `workerId`, `assignmentId`, and `hitId`.

- `workerId`: The MTurk Worker ID
- `assignmentId`: A unique ID that Amazon generates for a worker taking your HIT
- `hitId`: The ID of your HIT

For example, if your Survey Link is https://www.example.com, we will add the above parameters to the URL, so the final URL might look like https://www.example.com?workerId=ABCD143567hitId=35ghfsbda&assignmentId=xywjeblwujf. You can then take advantage of these parameters by setting up Qualtrics Embedded Data.

### Scheduled Launch Time (UTC)

Please note that you may leave the Scheduled Launch Time blank if you want to launch your HIT as soon as possible. This feature is useful if you want to automatically launch your Survey at a later time.

- **Scheduled Launch Time (UTC)**: 

  - **Pick an optional launch-off date...**

  Please note the time needs to be in UTC format so if you want to launch on Feb 15, 2018 at 02:00 PM Eastern Time, the time should be 02/15/2018 07:00 PM.

- **UTC Time Zone**: 02/14/2018 04:52 AM

- **Automatic Worker Bonus**: $ 0.0
The amount of time you want your HIT to be active for

<table>
<thead>
<tr>
<th>Survey Hyperlink</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="https://osu.az1.qualtrics.com/jfe/form/SV_3xa1WCRuJL9LSxH" alt="Survey Hyperlink" /></td>
<td>Preview link in new tab</td>
</tr>
</tbody>
</table>

The URL of your survey

<table>
<thead>
<tr>
<th>Query String Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>TurkPrime will add the following query string parameters to your URL so you can take advantage of them in your data collection. The parameters are: <code>workerId</code>, <code>assignmentId</code>, and <code>hitId</code>.</td>
</tr>
<tr>
<td><strong>workerId</strong> - The MTurk Worker ID</td>
</tr>
<tr>
<td><strong>assignmentId</strong> - A unique ID that Amazon generates for a worker taking your HIT</td>
</tr>
<tr>
<td><strong>hitId</strong> - The ID of your HIT</td>
</tr>
<tr>
<td>For example, if your Survey Link is <code>https://www.example.com</code>, we will add the above parameters to the URL, so the final URL might look like <code>https://www.example.com?workerId=ABCD1435&amp;hitId=35hjfsbdue&amp;assign...</code></td>
</tr>
<tr>
<td>take advantage of these parameters by setting up Qualtrics Embedded Data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scheduled Launch Time (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please note that you may leave the Scheduled Launch Time blank if you want to launch your survey right away. This feature is useful if you want to automatically launch your Survey at a later time.</td>
</tr>
<tr>
<td><strong>Scheduled Launch Time (UTC)</strong></td>
</tr>
<tr>
<td>Required Format UTC Time Zone</td>
</tr>
<tr>
<td>UTC Time Now: <strong>02/14/2018 04:52 AM</strong></td>
</tr>
<tr>
<td>Please note the time needs to be in UTC format so if you want to launch on Feb 15, 2018 Eastern Time, the time should be <strong>02/15/2018 07:00 PM</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Automatic Worker Bonus</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="https://osu.az1.qualtrics.com/jfe/form/SV_3xa1WCRuJL9LSxH" alt="Automatic Worker Bonus" /></td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>
Now go to your Qualtrics Survey ➔ “Survey Flow” ➔ “Set Embedded Data” ➔ enter variable names (case sensitive!)

Survey Flow  MTurk Demographics Survey

Show Block: Default Question Block (13 Questions)

Set Embedded Data:
- hitid: Value will be set from Panel or URL. Set a Value Now
- assignmentId: Value will be set from Panel or URL. Set a Value Now
- workerId: Value will be set from Panel or URL. Set a Value Now

Add a New Field

+ Add a New Element Here
Dynamic Completion Code For Qualtrics

Each worker will have a unique secret code. Your HIT can either be auto-approved or manually approved.

Setup for Qualtrics
1. Create or Edit a Survey
2. Click on "Survey Options"
3. In the Survey Termination Section: Click on Custom end of survey message...
   - Default end of survey message.
   - Custom end of survey message...
   - Thank you

4. If you have a Secret Key option that you previously created Select it and you are done/Otherwise, Click New Message
5. Enter Secret Key in the Description field. (Next time you can select this message.)
6. Click the Source button
7. Paste the following text exactly into the window as shown

```
```

Create a New Message

<table>
<thead>
<tr>
<th>Category</th>
<th>End of Survey Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Secret Key</td>
</tr>
</tbody>
</table>

8. Click Save

Joseph K. Goodman, PhD
Auto Pay Workers
(approve only if completion code is correct)
Blocks workers that participated in your other relevant studies

Blocks workers that participated in your other relevant studies
Other Qualifications...

Worker Qualifications

Worker Level
- Any Level
- Master

Location
- is one of
- Select Locations (up to 30)
- UNITED STATES

Note Regarding HIT Approval Rate
When you set HIT Approval Rate (e.g., 80-100%) TurkPrime automatically sets Number of HITs Approved to be at least 100 since MTurk assigns an Approval Rate of 100% to Workers who have completed less than 100 HITs (MTurk API Reference).

HIT Approval Rate (%) for all Requesters' HITs
- 95-100%

Number of HITs Approved
- 100
- 1,000,000
- Worker needs to have completed between 100-1,000,000 HITS

Group Requirements
Create a new Worker Group here
- Select Group Requirements (up to 5)
- New: Nalvete (Exclude most active Workers) (more details)
- Nalvete is free until April 1, 2015.

Next
MicroBatch

MicroBatch HITs
Some issues to be aware of with MicroBatch

- Please be advised that MicroBatch HITs take longer to complete. The average study takes 1 hour per 7 workers, but times may vary.
- MicroBatch will not be enabled if the Time allotted for your HIT is more than 90 minutes
- Additionally, MicroBatch will not be enabled if you include Workers

MicroBatch
The survey will be batched into small HITs
(Duration for survey must be 90 minutes or less and you cannot include any Workers for this to be enabled)
Pro Features

- **Block Duplicate IP Addresses**
  - Will prevent multiple responses from the same IP address
  - (more details)

- **HyperBatch**
  - The survey will be batched into small HITs and will complete faster than Microlatch
  - (will disable Microlatch)

- **Anonymize Worker IDs**
  - (more details)

Verify Worker Country and State Location

- (more details)

Survey Group

- No Survey Group

  - There are currently no surveys in the selected group

  - Create New Survey Group

US Regions

- Northeast
  - New England (States include: CT, MA, ME, NH, RI, VT)
  - Mid-Atlantic (States include: NJ, NY, PA)