Using MTurk for Research

JOSEPH K. GOODMAN

The Ohio State University
Fisher College of Business

Based in-part on:
Special Thanks to my Coauthors

GABRIELE PAOLACCI
ERASMUS UNIVERSITY

SCOTT WRIGHT
PROVIDENCE COLLEGE
Agenda

Myth or truth? Debunking a few myths
Editors and Reviewers: How do they feel?

Solutions and Tips
- CloudResearch (formerly TurkPrime)
- See also Positly.com as an alternative
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 hate it!
   ◦ Won’t accept papers with only Mturk data
   ◦ And other views about Mturk
11. Attention: Bots! Bots! Bots!
12. Is it MTurk, Mturk, or mturk?
13. Non-naivete
14. Drop outs/p’s pre-screening
15. Pre-screening
What Do Researchers Think about Mturk?

...Decrease in Quality Perceptions

How do you perceive the quality of data obtained from the following sources?

- Student Pool
- Non-student Pool
- Grad Student Pool
- Mturk
- Prolific
- Other Crowdsources
- Online Panel
- Field Data
- Panel Data
- Scanner Data

**February 2019 vs February 2020**

- Mturk: $p < .001$
- Other Crowdsources: $p = .029$

Using MTurk for Research
What Do Researchers Think about Mturk?

...Decrease in Quality Perceptions

Papers should not have all their data from Mturk

Mturk samples are better than student samples.

Using MTurk for Research
What Do Researchers Think about Mturk?

...Decrease in Quality Perceptions

- Trust seems to be decreasing
- Quality measures correlated with
  - Age
  - Rank
  - Usage
- Wisdom or Familiarity?
- Researchers expect Mturk use to grow
- What to do? Some solutions...

Perceptions of Mturk are Decreasing

all p’s < 0.01

- Mturk workers pay attention
- Changed the field for the better
- Papers should not have all their data from Mturk
- Question the Validity
Solution 1: Stop Doing MTurk Studies
Solution 2: Get a job at Ohio State (OR OTHER SCHOOL WITH 8000 SUBJECT HOURS/YEAR)
Solution 3: Other Platforms?

Prolific
◦ Requires > $6.50/hr
◦ Representative samples (extra $)
◦ Pre-screening (not extra $)
◦ Slightly slower
◦ Simple and more integrated
◦ Slightly less attention?

CrowdFlower

Positly and CloudResearch (will discuss)

Qualtrics Panels

Others?

Peer, Samat, Brandimarte, & Acquisti (2017) *JESP*

# of failed attention check questions between the platforms (% of P’s)
Solution 4: Measure Real Behavior on MTurk

Example₁: Tonietto, Malkoc, & Nowlis (2019)
- 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

Example₂: Coin flip game to measure cheating
(Balasubramanian, Bennett, and Pierce 2017)
Solution 4: 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)

◦ Do you have other examples?
Solution 5: 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: Data Quality and Bots

How bad is it?
- Different opinions, and changed over time
- Bot-gate: August 2018
  - Evidence that they are foreign workers with low English proficiency, posing as US workers
  - “Workers pay attention”
    - 2019: 4.33
    - 2020: 3.84

SOLUTIONS

Attention checks?
- More on this later...

Increase your sample size!
  - n = 100+ per cell (see datacolada.org)

Use CloudResearch or Positly
TaskMaster: cut and paste html into qualtrics

Use Qualtrics Bot Detection
(if your license has it)

- **Open Access.** Allow anyone to take this survey.
- **By Invitation Only.** Prevent people from taking the survey using an anonymous survey link.

- **Password Protection.** This password must be entered to take this survey:
  - Enter survey and display default message.
  - Enter survey and display custom message...
  - Redirect to a full URL, e.g., "https://www.qualtrics.com".
  - Continue survey and set an embedded data field.

- **Bot Detection.** Flag suspicious responses using reCAPTCHA.
- **Email Scan Roadblock.** Use ReCAPTCHA to prevent security scanners from starting surveys when testing your survey links.
- **RelevantID.** Flag suspicious panel respondents using RelevantID.
  - Reset RelevantID so respondents can retake survey after 6 Months.

- **HTTP Referer Verification.** The user must come from this URL to take the survey:
  - Prevent Indexing. A tag will be added to the survey to prevent search engines from indexing it.
  - Secure Participants' Files. Files uploaded as responses can only be viewed by users with permission to view responses.

- **Survey Expiration.** The survey will only be available for a specified date range.

Using MTurk for Research
Issue: Attention & Checks

Researchers are using them more:
- 2019: 74%
- 2020: 85%

Downsides
- “Attention” might mean not naïve or those with higher Need for Cognition or System 2 processing
- P-hacking: Pre-register your attention check

Mturk Masters?
- Not naïve, $$$
- 18% researchers use Masters

Solutions

CloudResearch
- 43% researchers some sort of blocks via CloudResearch (up from 31%)
- Now has
  - “Block Low Quality Participants”
  - “CloudResearch Approved Participants”

Be objective: Not arbitrarily decide what ‘bad data’ looks like post-hoc

Pre-register!
Issue: Pre-screening & Non-naiveté

Non-naiveté
- We may be sampling the same people across studies
- Sampling same 7,300 workers on average (Stewart et al. 2015)
- Consider sampling workers with less experience (e.g., < 500 HITs)

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
Pre-screening on CloudResearch: Two Ways

Pay to use CloudResearch’s criteria
- $$$, 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

Using MTurk for Research
Issue: Compensation

“You get what you pay for” does NOT apply to Mturk.
  ◦ Research shows: Paying More ≠ Better Data

What do others pay? Our surveys say
  ◦ $0.31 for a 5-minute study

Minimum wage?
  ◦ Min wage of $7.25/hr in US ($0.121/min)
Guidelines for Researchers

Use qualifications and pre-screening
- > 95% approval
- Do not let workers to screen themselves

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

Collect Data and Track your participants
- Collect basic demographics
- Prevents non-naiveté, anonymous via CloudResearch

Try other sources

Attention: Bigger samples or use objective checks and pre-register.

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

Monitor your reputation.
- Reddit forms
- Turkopticon (shut down)
- Others?

http://wiki.wearedynamo.org/index.php/Guidelines_for_Academic_Requesters
Other Questions?
How to Create Your Own Database of Workers
Many Issues are Solved by CloudResearch

Recruit Online Participants

**Simple • Affordable • Powerful CrowdResearch**

- **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
Create a Study

Mechanical Turk Studies

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.
Using MTurk for Research
**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
Using MTurk for Research

Worker Payment Per Survey

Expected time to complete HIT

Time allotted per assignment

HIT expires in

Survey Hyperlink

Query String Parameters

Scheduled Launch Time (UTC)

Please note that you may leave the Scheduled Launch Time blank if you want to launch your own survey away. This feature is used if you want to automatically launch your survey at a later time.

Required Format UTC Time Zone

UTC Time Now: 02/14/2018 04:52 AM

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

* Survey Hyperlink

https://osu.az1.qualtrics.com/jfe/form/SV_3xa1WCRuJL9LSXh

The URL of your survey

Query String Parameters
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: `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=ABCD1435&hitId=35hgfsbdlu&assignmentId`

Scheduled Launch Time (UTC)
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.

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 7:00 PM Eastern Time, the time should be 02/15/2018 07:00 PM

Automatic Worker Bonus 🍀

Using MTurk for Research
Now go to your Qualtrics Survey ➔ “Survey Flow” ➔ “Set Embedded Data” ➔ enter variable names (case sensitive!)
Using MTurk for Research
Using MTurk for Research
Using MTurk for Research

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...
   - Survey Termination: Custom end of survey message...
     - Secret Key
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

```
<iframe src='https://www.turkprime.com/TakeLaunchedSurvey/DynamicKey' width='100%' height='200'></iframe>
```

Create a New Message

Category: End of Survey Messages
Description: Secret Key

8. Click Save
Auto Pay Workers
(approve only if completion code is correct)
Using MTurk for Research

Exclude Workers Who Completed These Surveys (live studies are not included)

Allow only Workers Who Completed and Were Approved

of These Surveys (live studies are not included)

Exclude Workers with These Worker IDs (You can use either TurkPrime Worker IDs or Amazon Worker IDs). Enter each on a separate line (Maximum 2000)

Include Workers with These Worker IDs (You can use either TurkPrime Worker IDs or Amazon Worker IDs). Enter each on a separate line (Maximum 2000)

Blocks workers that participated in your other relevant studies

Only allows workers with IDs that you enter. Copy & Paste your list here!