Crowdsourcing principles and practices: MTurk and beyond

Brad Turner
PhD Candidate, MIT-Sloan Economic Sociology
Wednesday February 17, 2020
Crowdsourcing research fora and participant pools

Self-serve
- Amazon Mechanical Turk
- Cloud Research
- Prolific
- LUCID
- Qualtrics

Managed
- Dynata
- Audience

Lab
- MIT BRL
- Sona Systems
- NORC
Online crowdsourcing research: Why to use

• Huge pools (MTurk has ~7,500 FT and 500,000 PT workers)
• Easy, fast, cheap, and flexible
• More diverse, less WEIRD
• Quality is generally good, vs. alternatives, w/ protections
  – “Substantively identical results” on personality and political scales between MTurk and “benchmark national samples” (Clifford et al. 2015)
  – $0.75 /person MTurk outperforms $3.75/person Qualtrics (Kees et al. 2017)
  – MTurk “adequate and reliable” for interactive experiments (Arechar et al. 2018)
• Not always dominating: Lab and managed pools offer advantages (e.g., specialized panels, quality filters)
Online crowdsourcing research: Why to be wary
(Aguinis, Villamor, and Ramani 2020, JM)

• Inattention: 15% of MTurkers fail attention/compliance checks
• Self-misrepresentation to meet eligibility criteria (10-83%, most on income, education, age, family status, and gender).
• Self-selection bias ($, boredom, employment status, or location)
• Non-naivete (10% MTurkers do 40% of completed studies)
• MTurker communities (61% interact with other participants)
• Attrition rates (often >30%)
• Some journals, editors, reviewers have intermittently rejected manuscripts using MTurk
Why you should still use self-serve platforms

• You can and should still use MTurk
• Prolific is better (Peer et al. 2021)
• Other pools are not always better, including managed ones
  – Outsourcing quality control: Often weak service for high cost
  – Subject to similar manipulations (Kees et al. 2017 and reply)
  – Any quality filters trade off naivete, introduce selection
  – Know what you pay for
• Educate yourself to describe and defend a rigorous approach using the latest (!) research
  – E.g., Arechar and Rand 2020 “Turking in the Time of Covid”
My goals for this presentation: Principles to practices

My Goals for this Presentation

1. Frame principles for online research in terms of values
2. Propose ‘best practices’ for any platform
3. Compare and contrast platforms
   - Including qualification features
4. Highlight many other resources
Values for online research

• Data quality
• Cost
• Ease and efficiency (time / no headache)
• Fair / decent treatment of participants (Ps) (IRB, Ethics, Law)
Principles: Connecting values to best practices

Values: data quality, cost, ease, & decent treatment of Ps.

1. Understand your workers
2. Pay fair for (good) work
3. Encourage good work
4. Discourage bad work: Protect yourself from and punish bad behavior.
1. Understand your workers

Has anyone done HITs?
Take a walk in a Turker’s shoes... [trigger warning]

Dhruv Mehrotra, Gizmodo, January 2020

$500 for >1,000 responses asking workers about their experiences
Take a walk in a Turker’s shoes

Mehrotra, 2020, “Horror stories from inside Amazon’s Mechanical Turk,” Gizmodo

respondents wrote about bizarre experiences wherein they were asked for personal information for questionable purposes. For example, four respondents told Gizmodo that they had been asked to draw pictures of their genitals for reasons unknown, and more than a dozen respondents claimed to have been asked for photographs of their feet for an undisclosed purpose.
Take a walk in a Turker’s shoes

Mehrotra, 2020, “Horror stories from inside Amazon’s Mechanical Turk,” Gizmodo

Since the requestor allegedly never paid, never responded to the workers’ inquiries, but still received the images before ultimately rejecting the job, they concluded that the HIT was probably a scam. “But I always wonder what happened to all those pictures of feet that he collected.”
Take a walk in a Turker’s shoes

Mehrotra, 2020, “Horror stories from inside Amazon’s Mechanical Turk,” Gizmodo

- HITs underpaid (essentially all Ps)
- HITs baselessly rejected (14%)
- HITs rejected because no code was provided (20%)
- Asked for or shown personal info like social security #s (10%)
- Graphic image or video (11%)
  - E.g., training artificial intelligence to moderate offensive content
- Worst experience in an academic survey (12%)
Take a walk in a Turker’s shoes

Mehrotra, 2020, “Horror stories from inside Amazon’s Mechanical Turk,” Gizmodo

“I have had to look at the faces of people who shot their faces off trying to commit suicide, and they survived. They were unrecognizable as even human. ... I will never forget those faces.

worker said, “There was NO warning that I would have to look at images like this, and it paid A QUARTER.”
Take a walk in a Turker’s shoes

Mehrotra, 2020, “Horror stories from inside Amazon’s Mechanical Turk,” Gizmodo

While categorizing the graphic videos or images can be traumatic, many respondents said that completing academic surveys can be even worse.

“Someone paid me like 50 cents to recall the most painful memory of my life,” recalled another respondent, “it fucked me up for the entire day.”
1. Understand your workers: Who are they

1. Fairly representative, managed sample.
2. Often subject to abuse
3. Attuned to pay fairness
4. Primarily motivated by money.
   - Mostly supplemental income, but some main job.
   - Need for speed balanced by need for quality for reputation
5. Secondary motivations?
   - Interesting studies? Sense of accomplishment? Community?
6. Some Turkers are jerks: Farmers, scripters, grifters, slackers.
2. Pay fair for (good) work
   (my second principle)


2. Pay rates
   - $0.10/minute = $6/hr
   - $0.12/min ≈ $7.25/hr (fed min.)
   - $0.16/min ≈ $9.6/hr (BRL, Prolific)
   - $0.22/min ≈ $13.50 (MA min.)

3. Filling in bubbles vs. reading, thinking, writing, stress/risk

4. Give realistic completion times (pilot).

5. Only filter Ps out with the first few questions (in MTurk).
   - Attention/comprehension checks OR demographics
3. Encourage good work: Better pay = better data? (principle #3)

1. Depends who you ask:
   - Not really: Burmester et al. 2011, Litman et al. 2015
   - Maybe: Bartneck, Duenser, Moltchanova, and Zawieska 2015
   - Yes: Lovett et al. 2017, Casey et al. 2017

2. Based on my experience, I believe it does:
   - Workers with better records can command better pay, therefore many turn their noses up at peanuts.
   - If they feel bilked or bullshitted, even otherwise sincere workers can become rageful saboteurs.
Encourage good work: Setting expectations

1. Set and calibrate survey expectations
   1. Make it easy: Be simple, clear, concise, and colloquial
   2. Be honest + accurate (time, attention checks, tasks, content) (Lovett et al. 2017)
   3. Put key info on separate lines or pages
   4. Repeat key info
   5. Use *visible* countdown timers for longer reading.
   6. Tap into workplace logics (e.g., express gratitude for work well done)
   7. Use attestations (e.g., will you… yes/ no)

2. Calibrate your own expectations.
   1. Nobody reads consent forms.
   2. You get what you pay for.
   3. No matter what you do, some people will flout your will.
Set expectations: Simple survey

Consent form: Use headers. Nobody reads.

Consent

Your participation in this study is completely voluntary. You can choose to continue or not. If you choose to continue for now, you may subsequently withdraw at any time. All responses will be kept anonymous and confidential.

This study is conducted by Bradley Turner of the Massachusetts Institute of Technology. If you have any problems, please contact bradley1@mit.edu. You do not waive any legal claims, rights or remedies by participating in this research study.

By clicking the button below to proceed, you consent to participate in this study.

Use page breaks, separate lines/pages – Accentuate what’s most important.

Attention Questions

Please pay close attention to the events that are described on the next screens.

We will ask questions to test your comprehension; failure to read carefully may lead us to reject your HIT.

It may be helpful to write down some details as you go along.

Get logistics out of the way.

You must have audio enabled to complete this survey.
Set expectations: Two-part (complicated) survey

1. Platform setup

   **Title**
   - Two-part survey: Part 1

   **Description**
   - PLEASE take this 15-minute HIT ONLY if you can do Part 2 in a week.
   - A description of your survey

   **Custom Instructions**
   - [ ] [ ] [ ] [ ] [ ]

   This is the first part of a two-part study with two surveys. We will post a follow-up survey in a new HIT about one week from today. We will message you through MTurk when the HIT is posted.

   The first-part survey will take about 15 minutes and pay $1.

2. Survey setup for both parts: repeat info + attestation

   **Introduction**
   - This is the first part of a two-part study with two surveys. We will post a follow-up survey in a new HIT about one week from today. We will message you through MTurk when the HIT is posted.

   The first-part survey will take about 15 minutes and pay $1.

   The second-part survey will take about 15 minutes and pay $2.

   Please complete this first-part survey ONLY if you will be able to complete the second-part survey in about one week, Jan 20th (Tuesday).

   Will you be able to complete both parts of the survey?
   - [ ] Yes
   - [ ] No

3. Survey setup for part 1: Intro with logistics + attestation

   In this first-part survey today, you will be asked questions about a movie and about a restaurant.

   You must be able to focus on this survey, read carefully and take your time for 15 minutes.

   You will also need audio and video to watch a full movie trailer.

   Will you focus on this survey for 15 minutes; read carefully and take your time; and you have audio and video to watch a movie trailer?
   - [ ] Yes
   - [ ] No

...then when Part 2 became available: we sent platform message to all participants.

Result: 80% completion across both parts
4. Discourage bad behavior: Protect yourself from and punish (principle #4)

1. Insufficient effort responding (IER) (Huang, Liu, & Bowling, 2015)
2. Non-diligent Ps add noise and reduce statistical power (Oppenheimer et al 2009)
4. Can come from
   1. Sloppy experimental set-up
   2. Participants’ need for speed: reasonable rushing
   3. Poor comprehension
   4. Cognitive overload, distracting environment
   5. Bad behavior: Speed racers, cheats, slackers, bots/scripts, survey farmers
5. Calling cards: Gobbledygook, seconds on substantive questions, crazy text responses, and random or ‘straight-line’ responses.
Protect yourself and punish: Lines of defense

1. Filter the pool with platform settings
   - US Only
   - # of Approved HITs (100, 500, 1000)
   - HIT approval rate (95%-98%, ≥99%)
   - Platform-specific filters (I’ll show)
   - Requester account reputation

2. (dynamic) Survey completion code

3. Qualtrics filters
   - e.g., US IP address check, mobile device check

4. Run in the morning.

5. Survey questions
Protect yourself and punish: Why you should use survey questions

When you mount your defenses, you control your sample

1. You can describe and defend your approach
2. By setting high standards of recruitment you may artificially exclude people from analysis.
3. DIY requires effort, but saves money
4. Balance data quality and naivete
Protect yourself and punish: Survey question strategies

1. **Eject** (first question(s) only…not Prolific)
   - Why: Eliminate worst offenders, Ps who cannot speak English. Risk of sample bias.
   - Questions: Captchas. Super simple comprehension + response.

2. **Nudge**
   - Why: Improve engagement without more work, bias risk (Oppenheimer et al. 2009).
   - Questions: Simple read and response attention checks.

3. **Exclude** from primary analysis
   - Why: Use as a quality metric, robustness check. Takes time
   - Questions: Not-super-simple comprehension checks. E.g., analogies.

4. **Reject**
   - Why: Saves *some* money. Ding bad actors (tragedy of the commons). Takes time
   - Questions: Timing (record time to submission). Text entry.
# Protect yourself and punish: Survey question types

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Place in survey</th>
<th>Use / Don’t Use</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Captcha</td>
<td>Beginning</td>
<td>Use</td>
<td>Eject</td>
</tr>
<tr>
<td>Super simple comprehension / attention check</td>
<td>Beginning</td>
<td>Use</td>
<td>Eject</td>
</tr>
<tr>
<td>More complicated comprehension / attention check</td>
<td>Middle</td>
<td>Use</td>
<td>Nudge or Exclude</td>
</tr>
<tr>
<td>Timing (record)</td>
<td>Anywhere</td>
<td>Use</td>
<td>Reject</td>
</tr>
<tr>
<td>Timing (enable submit after)</td>
<td>Anywhere</td>
<td>Use</td>
<td>Do nothing</td>
</tr>
<tr>
<td>Text entry</td>
<td>End</td>
<td>Use</td>
<td>Reject or Exclude</td>
</tr>
<tr>
<td>Reverse scale trick questions</td>
<td>Middle to end</td>
<td>Use or Don’t Use</td>
<td>Exclude</td>
</tr>
<tr>
<td>Repeat questions</td>
<td>Middle to end</td>
<td>Use or Don’t Use</td>
<td>Reject</td>
</tr>
<tr>
<td>Javascript</td>
<td>Anywhere</td>
<td>Use or Don’t Use</td>
<td>Depends</td>
</tr>
</tbody>
</table>
## Protect yourself and punish: Survey question types (my favorites)

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Place in survey</th>
<th>Use / Don’t Use</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Captcha</td>
<td>Beginning</td>
<td>Use</td>
<td>Eject</td>
</tr>
<tr>
<td>Super simple comprehension / attention check</td>
<td>Beginning</td>
<td>Use</td>
<td>Eject</td>
</tr>
<tr>
<td>More complicated comprehension / attention check</td>
<td>Middle</td>
<td>Use</td>
<td>Nudge or Exclude</td>
</tr>
<tr>
<td>Timing (record)</td>
<td>Anywhere</td>
<td>Use</td>
<td>Reject</td>
</tr>
<tr>
<td>Timing (enable submit after)</td>
<td>Anywhere</td>
<td>Use</td>
<td>Do nothing</td>
</tr>
<tr>
<td>Text entry</td>
<td>End</td>
<td>Use</td>
<td>Reject or Exclude</td>
</tr>
<tr>
<td>Reverse scale trick questions</td>
<td>Middle to end</td>
<td>Use or Don’t Use</td>
<td>Exclude</td>
</tr>
<tr>
<td>Repeat questions</td>
<td>Middle to end</td>
<td>Use or Don’t Use</td>
<td>Reject</td>
</tr>
<tr>
<td>Javascript</td>
<td>Anywhere</td>
<td>Use or Don’t Use</td>
<td>Depends</td>
</tr>
</tbody>
</table>
Eject: Super simple comprehension check

Please help us catch people not paying attention. Please select "Neither agree nor disagree."

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
Eject or nudge: Super simple comprehension check - matrix table

Please **carefully** answer the questions below.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many wheels does a normal bicycle have?</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>How many wheels does a normal car have?</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>How many wheels does a normal unicycle have?</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
Nudge or exclude: Not nice comprehension/attention check trap

An attention check trap question tests if people are paying attention and nudges or disqualifies them if they fail. Such a question seems easy to skimmers, but includes instructions like "select orange for this answer so we know you're paying attention" in the middle of a text block. If the participant fails to answer orange, you can kick them out, reject their submission, or exclude their data from your primary analysis. Give it a shot!

What color is the night sky?

- Orange
- Blue
- Black
Nudge: Nicer comprehension/ attention check trap

You failed our attention check. Please read everything carefully. It is necessary for our research. Please try again.

An attention check trap question tests if people are paying attention and nudges or disqualifies them if they fail. Such a question seems easy to skimmers, but includes instructions like "select orange for this answer so we know you're paying attention" in the middle of a text block. If the participant fails to answer orange, you can kick them out, reject their submission, or exclude their data from your primary analysis. Or you can just tell them that they messed up and ask them in good faith to try harder. Give it a shot!

Based on the text above, what color is the night sky?

Orange

Blue
Nudge or exclude: Comprehension / attention check for survey info

Which of the following best describes the scenario you read?

- Jack bought an expensive new car and I had to rate how he'd perform at work.
- I evaluated a creative person for my boss.
- Michelle is applying for a job and I had to decide whether or not she would get the job.
- I evaluated a creative team for my boss.
- I didn't read a scenario.
- I don't know how to answer this question.
Nudge or reject: Timing countdown + time-to-submit

Scenario 1
Once upon a time and a very good time it was there was a moocow coming down along the road and this moocow that was coming down along the road met a nicens little boy named baby tuckoo....
His father told him that story: his father looked at him through a glass: he had a hairy face.
He was baby tuckoo.
   O, the wild rose blossoms
   On the little green place.
He sang that song. That was his song.
   O, the green wothe botheth.

You can submit after 30 seconds. Please read everything carefully to be able to answer our questions.

Visible to P (a separate “Timing” question type in Qualtrics)
Reject: Text entry questions

What do you believe is the purpose of this study?

Did you notice anything odd or suspicious about this study?

In a few words, please describe what you ate for breakfast today?  
(Please answer this seriously, it helps us identify those not paying attention)
### Text entry questions – see anything suspicious?

<table>
<thead>
<tr>
<th>What do you believe is the purpose of this study?</th>
<th>Did you notice anything odd or suspicious about this study?</th>
<th>What did you have for breakfast today?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our ability to apply risk assessment to various situations?</td>
<td>One situation was a wedding which doesn't impact others, and one was about driving, which affects everyone?</td>
<td>A breakfast sandwich and some ginger tea.</td>
</tr>
<tr>
<td>Not sure</td>
<td>No</td>
<td>Eggs (?)</td>
</tr>
<tr>
<td>10-minute academic study paying $1. Requires reading, thinking, writing purpose good.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Analyzing Scenarios Study (?)</td>
<td>Good study.</td>
<td>Pancakes (?)</td>
</tr>
<tr>
<td>I'm actually not really sure.</td>
<td>I didn't.</td>
<td>Just coffee so far.</td>
</tr>
<tr>
<td>Therefore, authority is a form of knowledge that we believe to be true because its ... Knowing how to do research will open many doors for you in your career.</td>
<td>If you see suspicious activity, report it to local law enforcement or a person of authority. ... Unusual items or situations: A vehicle is parked in an odd location, ...</td>
<td>Apr 5, 2010 - What Did You Eat For Breakfast Today? ... Depending on the morning I tend to eat a granola bar or some steel-cut oats for breakfast. Sometimes yogurt. And always strong black coffee, brewed in a French press.</td>
</tr>
<tr>
<td>Seeing if people understand reasons for common failures</td>
<td>No, I did not</td>
<td>Sausage McMuffin</td>
</tr>
<tr>
<td>What do you believe is the purpose of this study?</td>
<td>Did you notice anything odd or suspicious about this study?</td>
<td>What did you have for breakfast today?</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Our ability to apply risk assessment to various situations?</td>
<td>One situation was a wedding which doesn't impact others, and one was about driving, which affects everyone?</td>
<td>A breakfast sandwich and some ginger tea.</td>
</tr>
<tr>
<td>Not sure</td>
<td>No</td>
<td>Eggs (?)</td>
</tr>
<tr>
<td>10-minute academic study paying $1. Requires reading, thinking, writing purpose good.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Analyzing Scenarios Study (?)</td>
<td>Good study.</td>
<td>Pancakes (?)</td>
</tr>
<tr>
<td>I'm actually not really sure.</td>
<td>I didn't.</td>
<td>Just coffee so far.</td>
</tr>
<tr>
<td>Therefore, authority is a form of knowledge that we believe to be true because its ... Knowing how to do research will open many doors for you in your career.</td>
<td>If you see suspicious activity, report it to local law enforcement or a person of authority. ... Unusual items or situations: A vehicle is parked in an odd location, ...</td>
<td>Apr 5, 2010 - What Did You Eat For Breakfast Today? ... Depending on the morning I tend to eat a granola bar or some steel-cut oats for breakfast. Sometimes yogurt. And always strong black coffee, brewed in a French press.</td>
</tr>
<tr>
<td>Seeing if people understand reasons for common failures</td>
<td>No, I did not</td>
<td>Sausage McMuffin</td>
</tr>
</tbody>
</table>
Crowdsourcing research fora

Self-serve
- Amazon Mechanical Turk
- Prolific
- CloudResearch
- LUCID
- qualtrics

Managed panels
- Dynata
- Audience

Lab pool
- MIT BRL
- Sona Systems
- NORC
## Comparing self-serve crowdsourcing platforms

<table>
<thead>
<tr>
<th></th>
<th>MTurk</th>
<th>CloudResearch (TurkPrime)</th>
<th>Prolific</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pool size</strong></td>
<td>Huge</td>
<td>“”</td>
<td>Very Large</td>
</tr>
<tr>
<td><strong>HIT type(s)</strong></td>
<td>Everything</td>
<td>“”</td>
<td>Academic</td>
</tr>
<tr>
<td><strong>Ease of use</strong></td>
<td>Hard</td>
<td>Easy</td>
<td>Easy</td>
</tr>
<tr>
<td><strong>Data quality</strong></td>
<td>Adequate</td>
<td>Adequate</td>
<td>Good</td>
</tr>
<tr>
<td><strong>Platform cost</strong></td>
<td>20% P-cost (avoid +20% for HIT of &gt;9 Ps by microbatching)</td>
<td>30% of P-cost</td>
<td>33% of P-cost</td>
</tr>
<tr>
<td><strong>Free Qualifications</strong></td>
<td># HITs Approved, HIT Approval Rate, Country, Worker Groups</td>
<td>“” + Block 'Low Quality Workers', Block Multiple IP Address, Approved Participants List, Block/Allow Lists, Simultaneous Groups, U.S. Regions</td>
<td>“” + Expanded Demog., Block/Allow Lists</td>
</tr>
<tr>
<td><strong>Cost Qualifications</strong></td>
<td>Demog. ($0.05-$0.50 / worker), Master (+5% of cost)</td>
<td>Demog. ($0.13-$0.67 / worker), Naivete ($0.25 / worker)</td>
<td>-</td>
</tr>
<tr>
<td><strong>Add. Services</strong></td>
<td>-</td>
<td>Rep. sample, Managed panels</td>
<td>Representative sample</td>
</tr>
<tr>
<td><strong>More stuff</strong></td>
<td>API and MTurkR package</td>
<td>Hyperbatching, Easy groups, Easy communication, Auto bonus, Delay start, Dynamic termination code</td>
<td>Easy groups, Minimum pay $0.16 / min., No ejecting</td>
</tr>
</tbody>
</table>
Comparing self-serve crowdsourcing platforms

So what should you use? (take these with a pinch of salt)

1. If you’re poor but scrappy: MTurk
2. If you’re rich and lazy: Prolific OR TurkPrime
3. If you want the best data with the least hassle: Prolific
4. If you’re using demographic qualifications: Prolific (they’re free)
5. If you’re surveying a special population (e.g., Managers): Consider managed pools like Dyata or another provider – but know what you pay for.
6. If your tasks are very cognitively demanding: Consider MIT’s Behavioral Research Lab – student pool, in lab, or video conferencing.
7. If you want a different set/type of participant (and you’re scrappy): Lucid
More topics

1. **Lucid** Survey Sampling and Marketplace
   - A useful addition to MTurk and Prolific
   - Survey marketplace model. Different participants (many access through mobile, Xbox, or Playstation). Different incentives (e.g., in-app incentives). Lower quality. Pay goes less to Ps.

2. Longitudinal research
   - Easy with TurkPrime or Prolific.

3. Building personal, group, or lab panels
   - Multi-stage surveys: First survey is a short demographic panel to identify qualified participants. Subsequent surveys are the research. Issues: Self-misrepresentation.

4. Video conferencing and online research, see Li, Leider, Beil, and Duenyas 2020

5. **Microbatching in MTurk with UniqueTurker**

6. **Using the MTurkR package in R to access MTurk’s API**
Resource repositories?

Building shared resources for Sloan researchers through the BRL

1. Qualtrics surveys
   1. Survey questions for eject, nudge, reject, and exclude
   2. Generic Javascript questions
   3. Tips, tricks and embedded tutorials
More resources

1. If you read one thing: Lit review and best practices from Aguinis, Villamor, and Ramani. 2020 MTurk Research: Review and Recommendations.

2. Academic research (see next page)

3. CloudResearch blog and data quality filters plus:
   - Best Practices That Can Affect Data Quality on MTurk, from CloudResearch
   - “New Solutions Dramatically Improve Research Data Quality on MTurk”

4. Prolific blog, attention check explainer, new researcher community forum
   - Event: “Supercharging Your Research Friday March 5th, 16:00 - 17:30 GMT”

5. MIT Behavioral Lab knowledge base


Works cited (also see MIT BRL Knowledge Base):


- Peer, Eyal and Rothschild, David M. and Evernden, Zak and Gordon, Andrew and Damer, Ekaterina, MTurk, Prolific or panels? Choosing the right audience for online research (January 10, 2021). Available at SSRN.


Thanks!

Questions?

Comments?