INTRODUCTION TO MTURK AND PROLIFIC

NERISA DOZO
ETHICS AND WELL-BEING HUB CO-ORDINATOR
CAVEAT

- I am by no means an expert on either of these platforms. Please always refer to online help guides and younger PhD students whose expertise and information I have stolen and used for today’s presentation.
- This is also a basic introduction to the two platforms along with a list of resources to get you started.
- Open discussion!
- Ask questions. Share your experiences.
This session was recorded and is available here - https://youtu.be/QbOBT8hza0g

All slides are on the EWB Hub website.
## TODAY

### What is MTurk and Prolific Academic?

### Money!

### Introduction to MTurk
- Creating an MTurk account
- Setting up your study
- Launching your study, watch the wave, and payment
- Communication and etiquette
- Bots and bad faith participants

### Introduction to Prolific
- How to set up a study
- What the platform looks like to participants
- Viewing completed progress
- Communication and payment
- Etiquette

### Complicated and longitudinal study design tips

### So which one should you use?

### Resources
WHAT IS MTURK AND PROLIFIC?

- Mechanical Turk (MTurk) and Prolific are crowdsourcing marketplaces making it easier for researchers to connect with participants worldwide.

- MTurk – is a crowdsourcing platform where you (as a Requester) can publish and coordinate a wide set of Human Intelligence Tasks (HITs), such as your online study or survey, as well as transcriptions and classifications. Participants are known as Workers and scroll through MTurk’s board to choose tasks they wish to complete. Workers earn a small amount depending on how much you choose to pay them.

- Prolific is very similar to MTurk except it was created with research studies in mind. In the same format, you as a researcher upload your study to Prolific and participants can scroll through to find studies they wish to participate in. No other terminology 😊
MTurk – US Dollars

- 1 AUD = 0.613 US Dollars
- Have to purchase pre-paid HITs (loading money onto MTurk) before you can run studies
- Can only load $2,000 AUD per day via a credit card
- Speak to Pras about using the school credit card to do this
- Cannot refund / return any unused money loaded to MTurk

Prolific – British Pounds

- 1 AUD = 0.496 Pounds
- Have to top up your account before you can run studies
- This can be done via credit card, bank transfer, or invoice
- Invoicing is the preferred method. Made out to the School and your lab / hub / person.
- Finance only pays invoices at the end of the month
- Can refund / return unused funds
YES, THIS MEANS YOUR BUDGET WILL BE SLASHED IN HALF (OR A LITTLE LESS) THANKS TO THE CURRENT EXCHANGE RATE.
MTurk is based in the USA with participant pools worldwide – Australia, India, France, Germany, and others.

However, the majority of participants are from the USA. This might be due to payment, as US workers can get direct bank payments of their study earnings. Those outside of the USA can only transfer their study earnings into Amazon gift cards.
You can set up an MTurk account with any email but we recommend using something that is shared and controlled by your Hub or Lab.

All emails from participants will come through to this email address so if you are sharing an account in a Hub or Lab, everyone needs to be able to access it.

If you have any issues with MTurk (e.g. getting a refund from a bad participant), you have to use the email associated with the account to contact MTurk.

You will need to share your log in details with Pras in order to get payments uploaded and also for Pras to check in on the budget every now and again.

Recommend changing your password at the end of every year.
CREATE A NEW PROJECT
The project name is not displayed to Worker. Establish a naming convention in your hub or lab so it is easier to identify projects across different studies and researchers.

The title is displayed to Workers. It is generally the first thing they will see when deciding whether or not to do the study.

Be as specific as you can without giving away important details about the study. We recommend including timing information as this will either attract or repel people to your study.

Workers will also use this description to decide whether to accept the task, so provide as much information as you can. It is also searchable within MTurk, so use relevant terms.

e.g. Read brief descriptions of social experiences (<50 words) and one possible interpretation of the situation. Rate your agreement with that interpretation.
Before you set up your HIT, have a good sense of how long your study takes to complete. Pilot, pilot, and pilot some more because you do not want to risk underpaying your participants.

Once you have a time estimate for your survey (which you will also need for your Plain Language Statement), calculate payment pro-rata of an hourly payment of US$7.50 per hour

Federal minimum is 7.25 USD however this can go as high as 13 USD in states like Washington DC and California.

At the very least, you should match federal minimum wage but there have been lots of recent discussions about small payments and the over-use of MTurk participants so you really should consider paying them more.

Workers tend to prioritize higher-paying tasks and we have to acknowledge we are getting away with paying pretty small amounts for these expert workers.

- A 30 minute survey should offer a minimum of US$3.70
- A 15 minute survey should be paid a minimum of US$1.80.

https://www.dol.gov/agencies/whd/minimum-wage/state
Usually this would be the number of participants you want but MTurk fees have increased so you will need a workaround (next slide).

Payment

This is NOT how long you think your study will take but rather how long a worker can have it open for to work on. Keep it longer than you normally would but not too long. We recommend 2-3 hours.

How long is your study active for? This really won’t matter as most studies are completed within a couple of hours and once you reach your number of respondents, your study is completed.

Workers can see this and tend to choose tasks with short auto-approve timeframes. We recommend only a few days.
MTURKS 20% FEE PRICE GOUGE

- Long story short, MTurk now charges a 20% fee for the first 10 participants and then 20% for anything over – so a 40% fee in total.
- Their reasoning = ??????????
- In order to avoid this fee gouge, lovely researchers have come up with a workaround where you run your study in batches of 10 or less participants. Ultimately, refreshing your study every 10 participants until you reach your maximum (sometimes slightly over or under)
- Follow their instructions and you will avoid this. Just do not forget to do it!
How to Avoid MTurk’s Extra 20% Price Gouge: A Click-by-Click Guide for Academic Requesters

Compiled by Charles Foster of Stanford's Laboratory for Social Research (lsr.stanford.edu)
ESTIMATE YOUR TOTAL COST BEFORE LAUNCHING

- Time taken
- Cost per participant
- Number of participants
- 20% fee of the total

- A 20% fee is also added to any bonuses you award participants.
WORKER REQUIREMENTS

Require that Workers be Masters to do your HITS

- NO! Masters are MTurk experts - completing over 5000 studies and with a good rating. However they cost more and there does not appear to be a lot of them so your data collection will be slow.
- If you are not collecting data quickly enough (when launched), you may have accidentally clicked yes to this. Cancel your study and re-start.

Qualifications:
Qualifications restrict who can take your surveys. To maximise high quality responding it is recommended you set the following:

- HIT approval rate greater than or equal to 90%
- Number of HITS approved greater than 100
- You may also want to set the location if you want all your participants to be in the US.

You can also use qualifications alongside a screening survey to recruit specific samples. For example, if you only wanted women to participate in your study, you could run a quick paid screening survey on a bigger sample using a question about gender (among other questions so it is not obvious to participants what you are screening for). You could then assign all the women in the sample a qualification and then only invite those with that qualification to the actual study.
DESIGN AND LAUNCH

- Here you will design the look of your study posting and also include a few basic instructions for participants.
  - You will include the link to your study
  - They should open your study in a new tab
  - You may wish to generate a random code for participants to enter back into MTurk. This gives you an easy way to verify in MTurk that they have done your study.
  - Any restrictions in terms of device or browsers

- When everything looks good (Preview), you are ready to launch!
- A few things to remember before launching
  - If you are using US participants, are you launching at a good time?
  - Do you have ethics clearance for your study and does your lab / hub / supervisor know about you using these funds?
  - Are you 100% sure about how long your study takes?
### HIT Groups (1-20 of 1015)

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Title</th>
<th>HITs</th>
<th>Reward</th>
<th>Created</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>James Billings</td>
<td>Market Research Survey</td>
<td>9,636</td>
<td>$0.01</td>
<td>3m ago</td>
<td>Accept &amp; Work</td>
</tr>
<tr>
<td>Content Research</td>
<td>Clean Up How-To Questions</td>
<td>6,401</td>
<td>$0.05</td>
<td>11th ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Phil Maher</td>
<td>&quot;Archival Video Reviewer&quot; Create Title Description, Add 5 Keywords, and pick a Category by watching a 5 second video.</td>
<td>6,057</td>
<td>$0.01</td>
<td>4th ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Gerry Gorman</td>
<td>Find email addresses for attorneys from given information. New!</td>
<td>5,365</td>
<td>$0.03</td>
<td>2d ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Trent Woodbury</td>
<td>Select the best fitting category to this transaction</td>
<td>3,941</td>
<td>$0.10</td>
<td>11th ago</td>
<td>Preview</td>
</tr>
<tr>
<td>UniMeb NLP Group</td>
<td>Identify documents with breaks in narrative flow</td>
<td>1,974</td>
<td>$0.25</td>
<td>4d ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Panel</td>
<td>Extract purchased items from a shopping receipt (6-10 items)</td>
<td>1,407</td>
<td>$0.06</td>
<td>13s ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Panel</td>
<td>Extract purchased items from a shopping receipt</td>
<td>1,406</td>
<td>$0.08</td>
<td>40s ago</td>
<td>Preview</td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill</td>
<td>Inference from the Text</td>
<td>1,329</td>
<td>$0.13</td>
<td>27s ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Iljanzi</td>
<td>Paraphrase machine-generated questions</td>
<td>1,204</td>
<td>$0.10</td>
<td>8d ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Panel</td>
<td>Extract purchased items from a shopping receipt (3-5 items)</td>
<td>962</td>
<td>$0.03</td>
<td>29s ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Oliver Boyne</td>
<td>Label dog keypoints</td>
<td>948</td>
<td>$0.03</td>
<td>2/19/2020</td>
<td>Preview</td>
</tr>
<tr>
<td>AQUX</td>
<td>Evaluate Search Results - 15 mins - $0.20 bonus - NO MOBILE</td>
<td>884</td>
<td>$0.05</td>
<td>1d ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Anita</td>
<td>Collect season ticket prices from a website</td>
<td>877</td>
<td>$0.01</td>
<td>8th ago</td>
<td>Preview</td>
</tr>
<tr>
<td>ffb54e4e-b7c8-47a8-9ee3-3a1617274ca2</td>
<td>Judge the reputation polarity of Article Clips</td>
<td>801</td>
<td>$0.08</td>
<td>5th ago</td>
<td>Preview</td>
</tr>
<tr>
<td>Panel</td>
<td>Extract purchased items from a shopping receipt (1-2)</td>
<td>697</td>
<td>$0.01</td>
<td>38s ago</td>
<td>Preview</td>
</tr>
</tbody>
</table>

**What the participant sees**
WATCH THE DATA ROLL IN

Once your study is live, you will be able to watch completed responses come in under the Manage tab.

If your study takes 15 minutes, check in around the 20-30 minutes and see how many completed you have. If none, something has gone wrong.

With participants galore, your data collection should be pretty quick so 100 participants might be collected in a couple of hours.

You can stop any live studies if there is an error. Participants in your study will still be able to finish (or not if there is a major error) and you should pay everyone who had the study open.
If something has gone wrong, if participants were unable to enter their code, or they have any issues, you will receive an email to the email used to create the account.

You should be on your email when launching, for the first hour, and then soon after (especially if you are launching at bad hours in Aus) so you can respond to questions in a timely manner.

**EWB Hub recommended etiquette**

- Respond to all questions
- Approve your payments promptly (ie. less than a few days)
- Don’t screen people out without pay - use a small eligibility study
- MTurk workers are really starting to dislike large likert scale surveys. Pages and pages of “respond on a scale of 1-7” gets boring. If your research uses large multi choice or radio type responses, consider breaking up the pages and including things to make it more visually interesting. At the very least, warn participants at the beginning of the survey that there may be lots of these questions.

- Say thank you!
PAYING PARTICIPANTS

- Participants will like you more and your account may be better reviewed if you pay participants pretty quickly after completing the study.
- You can review your results whilst the study is live and approve payments immediately or wait until your study is completed (all participants collected) and approve all.
MTurk is, to an extent, experiencing a crisis with bots and bad faith participants producing low quality data. There is no fool-proof method for eliminating either of these types of participants but there are checks you can put in place to “have evidence” to reject these participants.

Advice from the EWB Hub is that participants can be rejected for payment only if they fail MULTIPLE attention checks or you have some other form of proof that they are not in the country you are sampling.

There is a fine balance between we get easy data for pretty cheap and by approving all payments, we are increasing the approval rate for bad eggs.

If you are going to reject participants, warn participants about this in your study advertisement (on MTurk) and first page of your study. Give them plenty of warning and a chance to choose not to do your study.
BOTS AND BAD FAITH PARTICIPANT CHECKS

Analogies are a good way to detect bots. old is to young as adult is to ____
- five is to three as pentagon is to ____
- sun is to moon as blue is to ___

You can create your own as the aim is to create new analogies that do not appear on the first page of Google.

Response Speed
Responses of faster than 1 second per item have been found to be inconsistent and usually are a result of careless participants.

Response Consistency
You may be able to reject participants who fail to consistently respond to the same question, asked in different ways, throughout the study.

If responses are wildly different (not just a possible mistake) you may consider rejecting payment.

Unusual Comments
Include an open-ended question asking whatever you like. Participants who respond with all capital letters, single and irrelevant words, or copied and pasted text from Google could be rejected.
Attention checks asking participants to respond “agree” to an item or record an item described in the previous sentence can still be used.

English proficiency questions
Provide participants with a word and ask them to name (or choose from a list) a synonym.

For example, ask “Which of the following words is most related to “moody”? Participants unfamiliar with basic English or who provide responses without reading are likely to fail the screener.

Cultural checks
Cultural checks have been implemented to ensure workers are Americans and not foreigners.

For example, asking participants to name the vegetable in a picture. If you have a picture of an eggplant, Americans will name it as such, British participants might say aubergine, and Indians would say Brinjal.

To further test for American participants only, you can have a picture of a date circled and ask participants to write it out. American English usually has the date written as MM/DD/YY, not DD/MM/YY.
Be sure to check the MTurk acceptable use policy for certain restrictions on studies on MTurk.

A few to note – you cannot collect any personal information (e.g. emails if you wanted to contact them later) and you cannot ask them to download a different program or register on another website to take part in what you advertise on MTurk.

https://www.mturk.com/acceptable-use-policy
QUESTIONS ABOUT MTURK?
Formerly known as Prolific Academic, Prolific was created by PhD students who wanted a cleaner and more functional way to getting participants in their studies.

Prolific shamelessly capitalized on MTurk’s issues to say it was the savior of academic research. This is debatable but the platform sure is nicer to use.
CREATING AN ACCOUNT WITH PROLIFIC

- Similar to MTurk, you can set up a Prolific account with any email but we recommend using something that is shared and controlled by your Hub or Lab.
- Unlike MTurk, all communication from participants will come through the internal chat system of Prolific.
- If you email Prolific for support related to a specific study, you will need to email them from the email registered to the account.
NEW STUDY

Study name as seen by participants

Study description – include descriptive details on what is involved in your study but no need to provide timing information

Internal name, not visible to participants
You enter your study link and also include any parameters you wish to record in your dataset. For example, the Prolific ID of the participant and your study ID if you have multiple studies running.

https://researcher-help.prolific.co/hc/en-gb/articles/360009220993-Recording-participant-IDs-in-your-study-survey
On Prolific, it is important to redirect participants back to the Completion URL found on the study creation page or provide them with a completion code.

The will look something like: https://app.prolific.co/submissions/complete?cc=I2PWSFRG where 'I2PWSFRG' is the completion code.

Participants will visit this URL, or enter the code, to prove that they have completed your study. Therefore, this code should only be given to the participants who submit a complete and valid response to your survey.

Please note that submissions marked with an incorrect completion code (such as NOCODE) are not necessarily invalid. It could be that their browser closed early or the webpage did not redirect. Always ensure you check the raw responses in your survey for the Participant ID to check if they should be paid.
Prolific has participants complete a screening questionnaire when first signing up (gender, age, location) and they verify a few of these things with ID, uni emails, etc. This allows you to pre-screen before launching by applying custom screening based on your factors.

The number of matching participants will change depending on your audience criteria.
- Prolific removes the moral ambiguity on how much you should pay and enforces a minimum £5.00 per hour
- Ensure your estimate of time is correct because Prolific was also set its own timer (greater than your estimate) which will automatically kick participants out if they have not completed
- Participants will complain and you may have to pay participants who are kicked out if you underestimate time to completion
PARTICIPANT PAYMENT

Manipulative psychology at work!

Prolific will also provide a total cost of your study, including all fees.

Prolific fees are 33% of your total.
What participants will see

Qualification for a larger task! Everyone who won’t match the requirements but will fill out the survey will be rejected
Hosted by Wow Making
£0.75 • 9 minutes • £5.80/hr • 170 places remaining

Market Research Survey
Hosted by Bryan Tyner
£1.32 • 11 minutes • £6.08/hr • 270 places remaining

Believability of Events
Hosted by Ben Silver
£2.47 • 30 minutes • £4.93/hr • 43 places remaining

Food perception Copy
Hosted by J G
£1.06 • 10 minutes • £6.54/hr • 306 places remaining

Retort food questionnaire
Hosted by Tae OR Kim
£0.25 • 2 minutes • £7.50/hr • 268 places remaining

Email Usage
Hosted by Durek DeBellis
£0.48 • 4 minutes • £7.22/hr • 13 places remaining

Our research team is interested in examining experiences of playing online games among participants in Prolific. In this survey, you will be doing a short quiz before you play an online game and answer some questions.

This research is very important to us. If you randomly click to answer, we cannot use the data and compensate you. So please pay attention and try to answer sincerely. Thank you very much for your kind and generous
LAUNCH YOUR STUDY AND WATCH THE DATA ROLL IN

- When everything looks good (Preview), you are ready to launch!

- A few things to remember before launching
  - If you are using UK participants, are you launching at a good time?
  - Do you have ethics clearance for your study and does your lab / hub / supervisor know about you using these funds?
  - Are you 100% sure about how long your study takes?
When your study was launched (your local date and time)

How much participants are earning per hour. This adjusts as more participants take part

Eligible participants of their total (used pre-screening)

How many participants are in progress / completed

From here you can also message and bonus all participants or search for specific IDs
You can also approve all or by participant as they come in.

You can also stop your study if there is an error and same as MTurk, pay anyone who is currently in your study
CAN YOU REJECT PARTICIPANTS?

- Prolific is not the cure to MTurk’s problems. You should include the same attention checks as you would in any study and provide warning to participants that you will be screening for this.

- Prolific also provides guidance on acceptable reasons to reject payment - https://researcher-help.prolific.co/hc/en-gb/articles/360009092394-Reviewing-submissions-How-do-I-decide-who-to-accept-reject-
  - The study was completed exceptionally fast
  - The participant missed crucial questions
  - The participant failed multiple, fair, attention checks
  - The participant did not fully engage with a task (e.g. wrote a few words when you asked for at least a couple of sentences)
COMMUNICATION AND ETIQUETTE

If something has gone wrong, error with the study, or they are timed out but did actually complete (it happens!), participants will contact you via the message system inside Prolific.

You should be on your email when launching, for the first hour, and then soon after (especially if you are launching at bad hours in Aus) so you can respond to questions in a timely manner.

**EWB Hub recommended etiquette**

- Respond to all questions
- Approve your payments promptly (ie. less than a few days)
- Do not reject Timed Out of NOCODE participants. Check to see if they have contacted you or check their data in your study (match by ID) and see if perhaps they did complete everything but something went wrong.
- Say thank you!
The Prolific Help Centre really does answer all your questions on how to use Prolific and if you have any complicated study designs – e.g. longitudinal studies, use of other software.

Prolific is compatible with most survey software (e.g. Qualtrics) and allows other programs to integrate with it. Prolific also allows you to ask participants to download software or other programs to run your study. Your external software or program should not collect any personal information and if you have questions, email Prolific.


Bots and Bad Faith Participants

At the moment, the verification and screening process of new participants makes some bots less likely but you should still include some attention checks to ensure your participants are invested.
Australia: everyone needs to work from home and school will be run online

The NBN:

QUESTIONS ABOUT PROLIFIC?
For longitudinal studies, both MTurk and Prolific allow you to create workers hits or a whitelist which only allows participants in your created list to take part.

For example, we ran a study for 200 participants at Day 1 and then only followed up with 150 for Days 2 through to 7. The whitelist meant the study was advertised to all but only those we pre-selected were eligible to take part.

Prolific has the advantage of messaging these participants when they study becomes active but does not have an automatic way to launch the study at the right time. MTurk does not message or launch studies automatically but there are R workarounds for both these features.

Prolific also provides some support for dyadic experiments. The ruse that participants are being paired with someone else is harder to pull off online.

https://researcher-help.prolific.co/hc/en-gb/articles/360009223593-Dyadic-experiments
LONGITUDINAL OR COMPLICATED STUDY DESIGNS

- **MTurkR-package: R Client for the MTurk Requester API**

- **CloudResearch – when you do not want to deal with MTurk yourself**
  - [https://www.cloudresearch.com/](https://www.cloudresearch.com/)
  - [https://www.cloudresearch.com/resources/blog/longitudinal-and-follow-up-surveys-on-mechanical-turk/](https://www.cloudresearch.com/resources/blog/longitudinal-and-follow-up-surveys-on-mechanical-turk/)

- **Longitudinal studies in Prolific**

- **Advanced study features in Prolific**
MTURK OR PROLIFIC?

- Prolific is not the saviour of MTurk’s issues and not everything is doom and gloom on MTurk.
- Prolific is more user friendly but currently the largest pool of participants is only in the UK and the total pool is smaller than MTurk
- MTurk can take some getting used to but has a large USA sample and most studies are completed within hours
- Advice provided here can help you with both platforms
- Whatever you choose, pilot test small batches first. Better that you make a mistake with 10 participants than 1000.
RESOURCES

- MSPS Advice for online experiments
  - https://docs.google.com/document/d/1UY-N18_wVQqHzHQG-hbI_XwMHGXSgbgcYJOkRXiJL0
- Prolific Academic Help Centre
  - https://researcher-help.prolific.co/hc/en-gb
- MTurk Help
  - https://www.mturk.com/worker/help
- TurkerNation
  - https://www.reddit.com/r/TurkerNation/
- CloudResearch
  - https://www.cloudresearch.com/
- Completion code for Qualtrics and MTurk
- Turkopticon where MTurk Requesters are reviewed
  - https://turkopticon.ucsd.edu/
- Reddit for both MTurk and Prolific
  - https://www.reddit.com/r/ProlificAc/
  - https://www.reddit.com/r/mturk/
THANKS

- Melanie McGrath
- Reb Rebele
- Joshua Rhee
- Many more students within the EWB Hub
- Elise Kalokerinos
- Sean Murphy
- Amy Perfors
- And everyone who has edited the MSPS guide for running online studies