amazon mechanical turk™ Artificial Artificial Intelligence

About – Methodological Issues – Ethical Issues Basics and Setup – Creating Studies – Getting Results Tips and Tricks

> Jessie Taft / jgt43@cornell.edu October 6th, 2015

What is it?

- "AMT" or "MTurk"
- Popular online crowdsourcing platform "artificial artificial intelligence"
- Requesters pay workers ("Turkers") small amounts of money to complete small tasks ("HITs")



What is it useful for?

- Easy to use for:
 - Survey-based research
 - Transcription
 - Categorization and tagging
 - Some types of coding
 - User testing

- More difficult to use for:
 - Experiments with complex manipulations
 - Longitudinal studies
 - Group studies

Who are workers?

- About 500,000 workers in 190 countries, mostly US and India
- More diverse than student populations, but...
- Less diverse than a nationally representative sample (younger, better educated, more female, and poorer than US average)
- "super turkers" 10% of workers complete about 40% of HITs
 - Some even make a living doing MTurk HITs. These people can complete thousands of HITs per year, the majority of them academic surveys.
- Data quality and effect sizes are comparable to student samples; however, survey completion rates are lower

Methodological Issues

- Turkers can do the same HIT more than once -> duplicate responses
 - Scripts or screening tools can set a maximum number of survey repeats (eg. http://uniqueturker.myleott.com/)
- Demographics and other factors are difficult to control for
 - Can set preferences for country and state, everything else relies on honesty
- Super-users see the same scales and manipulations often
 - Check questions "have you seen these questions before?"
- Turkers can be distracted or have little motivation to put thought into their answers

Methodological Issues: Attention & Satisficing

- Online-only format means lab-style attention management isn't possible; low motivation can cause satisficing behavior
- Solutions:
 - Reverse-coded questions
 - Attention checks Click the banner at the top of the page rather than select a response.
 - Trap questions While watching television, have you ever suffered a fatal heart attack?
 - Follow-up questions How distracted were you while completing this survey?
 - Speed checks require participants to spend a certain amount of time on each question or set of questions

Ethics and Best Practices

- Set fair pay rates, and pay promptly
- Set clear guidelines for approval/rejection, and follow them
- Provide reasonable time estimates
- Provide accurate contact information, respond to worker questions and comments

Remember – Workers can keep track of your reputation, just as you can keep track of theirs. Treating them as human workers gets you better results.



Setup: Creating an Account

https://requester.mturk.com



Amazon Payments account is needed Choose personal vs. business account Choose your Requester Name wisely

Creating Your HIT: Picking a HIT Type

Home	Create	Manage	Developer
New Project Ne	w Batch with an	Existing Project	
Start a New P	roject		
Categorization			
Data Collection		Example of §	Survey Link
Moderation of an	Image		
Sentiment		Instruct	lians
Survey		Instruct	ions
<u>Survey Link</u>			onducting an academ
Tagging of an Ima	ige		bout social networks. /ou will receive a code
Transcription from	AV	survey.	
Transcription from	an Image		you will return to this
Writing		Templa	te note for Requester
Other		each Wo	orker to enter a <u>unique</u> s on how to generate this

Useful HIT Types:

- Survey Create survey directly in Mturk (HTML form); useful for simple surveys
- Survey Link Link to Qualtrics or other external survey; useful for longer survey studies or experiments with manipulations
- Transcription Upload a link to an audio or video file for participants to transcribe
- Categorization have participants rate an image or piece of text on scales that you define

Creating Your HIT: External Surveys

Information about your HIT – Requirements or anything you'd like participants to know before beginning.

Remember to remove this template reminder!

Enter the link to your own survey here

Instructions

We are conducting an academic survey about social networks. We need to understand your opinion about social networks. Select the link below to complete the survey. At the end of the survey, you will receive a code to paste into the box below to receive credit for taking our survey.

Make sure to leave this window open as you complete the survey. When you are finished, you will return to this page to paste the code into the box.

Template note for Requesters - To verify that Workers actually complete your survey, require each Worker to enter a <u>unique</u> survey completion code to your HIT. Consult with your survey service provider on how to generate this code at the end of your survey.

Survey link:

http://www.linktomysurvey.com

Provide the survey code here:

here: e.g. 123456

You must ACCEPT the HIT before you can submit the results.

To ensure that participants complete a survey and to make matching MTurk results with survey results easier, have Qualtrics generate a random number and ask Turkers to enter it into your HIT.

Creating Your HIT: Connecting Qualtrics

To create a random number for Turkers to enter, add a random number generator to the survey flow. This is recorded in both Qualtrics and MTurk for easy comparison. Instructions: <u>http://brentcurdy.net/qualtrics-tutorials/link/</u>

Other tips:

- Ask Turkers to enter their MTurk ID instead of their name on your consent form (yes, you do need a consent form!).
- If you're testing both student and Turker populations, make sure survey information is correct for each – don't include mentions of class credit, etc. in surveys for Turkers.
- Make sure your contact information is prominently displayed on the last page.

Creating Your HIT: Uploading External Data

If you'd like each participant to see a different value of a variable or different piece of data, you don't have to create new HITs for each value!

Imagine the following scenario:

You are texting with your friend, and the two o the message and then answer the questions b

\${message}

Α	В
message	
Hey there	!
Hello!	
Ahoy, mat	ey!
	Hey there Hello!

1. Insert formatted variable names into the text of your HIT, make sure your CSV has the same names as column headers.

Home	Create	Manage	Dev
New Project	New Batch with a	n Existing Project	
Start a Nev	w Batch with ar	Existing Proje	ct
Start a Nev	w Batch with ar	Existing Proje	×
Publish Batch	w Batch with an		

2. Select your pre-made HIT, click "Publish Batch", and upload your CSV file.

Imagine the following scenario:

You are texting with your friend, and the two of the message and then answer the questions b

Hey there!

А	AB	AC	
Days.	Input.message	Answer	r.(
)/0)	Hey there!	Female	
)/0)	Hey there!	Female	
)/0)	Hey there!	Male	
)/0)	Hey there!	Male	
-			

3. Values are automatically entered into your HIT text and recorded in your results file.

Creating Your HIT: Details and Worker Criteria



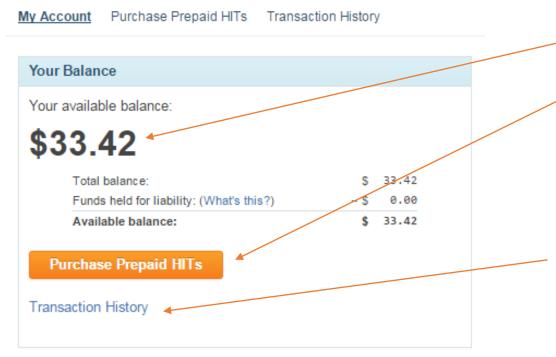
After creating your HIT, you need to enter HIT details and select the Turkers you'll use as participants.

Setting up your HIT			Ad	dvanced				
Reward per assignment	\$ 0.04	1	1				Worker requirement	s «
Number of assignments per HIT	Tip: Consider how long it will take a Worker to complete each task. A 30 second task that pays \$0.05 is a \$8.00 hourly wage.			Worker requirements:				
Number of assignments per first	80			Customize Worker Requirements		•		
Time allotted per assignment	How many unique Workers do you want to work on each HIT?			Specify ALL the qualifications Workers must meet to work on your HITs:				
	Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.			HIT Approval Rate (%) for a 🔹 greater than or equal to	۲	90	▼ remove	
HIT expires in	7 Days •			Number of HITs Approved	۲	50	▼ remove	
	Maximum time your HIT will be available to Workers on Mechanical Turk.			(+) Add another criterion (up to 5)				
Auto-approve and pay Workers in	8 Hours •							
	This is the amount of time you have to reject a Worker's assignment after they submit the assignment.		On	IIy Workers who qualify to do my HITs can preview my HITs.				
	Advanced »		0	Yes No				

On the Project Properties page, enter how much you'll pay per HIT, how many HITs you'd like, and how long each HIT will take. Specify worker type (Master/non-Master), approval rate, number of HITs approved, and worker location (country or US state).

Adding Funds to Your Account

Amazon charges a 20% fee on top of what's paid to workers. This increases to 40% for HITs done by 10 or more people. An additional 5% fee is charged for employing workers with Master qualifications.



- This is how much money is left in your account.

Click here to pre-pay for HITs. When you purchase HITs, you'll receive an estimate of how much your batch will cost, including fees, and how much will be left in your account.

Your transaction history will tell you how much you've spent on HITs in the past. Receipts from each past purchase can be downloaded and printed from here. Useful for reimbursements!

Monitoring Your HIT and Approving Results

You can monitor workers' progress as they complete your study.

Manage Batches

Click on the name of the batch to see more details

Metaphors - FB 13			Results Delete
Created:	July 17, 2015	Assignments Completed:	67 / 100
Time Elapsed:	3 days	Estimated Completion Time:	October 07, 2015 10:00 PM PDT (Wednesday)
Average Time per Assignment:	9 minutes 49 seconds	Effective Hourly Rate:	\$1.528
Batch Progress:	67% submitted	100% published	

Remember to approve results in a consistent and timely manner!

Review Results

Select the check boxes on the left to approve or reject results. You only pay for approved results. To evaluate results offline, select Download CSV.

Tips and Tricks: Group and Longitudinal Studies

Use MTurk for:

- Assembling participant pool
- Distributing informed consent information
- Collecting demographics and other survey data
- Paying participants (use bonuses!)
 - To award a bonus, go to Manage -> Workers, find and click on individual worker's ID -> Bonus Worker

Use other tools for:

• Collecting group data

Tips and Tricks: Preventing Duplicate Work

DON'T: Allow Turkers to do your hit multiple times, and then reject duplicate work.

DO: Be clear that duplicate work will not be accepted, or when a HIT has been re-posted.

DO: Use scripts that automatically block duplicate attempts, even across studies – <u>UniqueTurker</u>, <u>Turkitron</u>, <u>TurkCheck</u>

DO: Check results for copy-and-paste responses

References

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*, 351-368.

Buhrmester, M., Kwang, T., & Gosling, S. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3-5.

Chandler, C., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*, 112-130.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. Behavioral Research, 44, 1-23.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. Judgment and Decision Making, 5.

Useful Links

<u>http://wiki.wearedynamo.org/index.php/Guidelines_for_Academic_Requesters</u> - Guidelines and Best Practices for Researchers <u>https://experimentalturk.wordpress.com</u> – A blog un using Mturk for social science research

http://mechanicalturk.typepad.com/ - Official Mechanical Turk Blog, has a lot of tips, tricks, ideas, and critiques

http://www.mturk-tracker.com/#/general - Continuously updated data on Mturk activity, including demographics

<u>http://www.pbs.org/newshour/updates/inside-amazons-hidden-science-factory/</u> - A great article from PBS on people making a living through Mechanical Turk