

Tips and tricks for online data collection

Sharon M. Noh

The University of Texas at Austin

Mturk data is publishable

Research Article

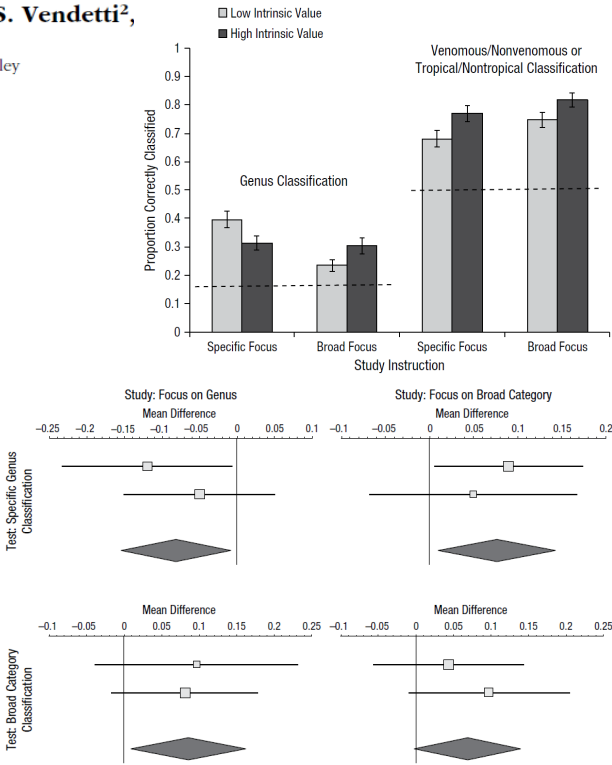
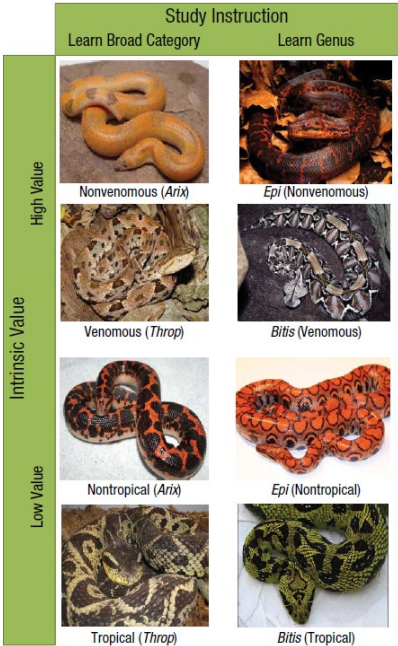
Multilevel Induction of Categories: Venomous Snakes Hijack the Learning of Lower Category Levels

Sharon M. Noh¹, Veronica X. Yan¹, Michael S. Vendetti², Alan D. Castel¹, and Robert A. Bjork¹

¹University of California, Los Angeles, and ²University of California, Berkeley



Psychological Science
2014, Vol. 25(8) 1592–1599
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614535938
pss.sagepub.com
SAGE



Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT



Original Articles

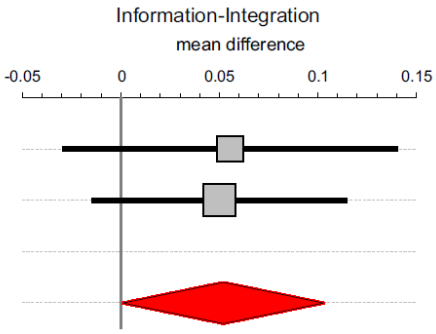
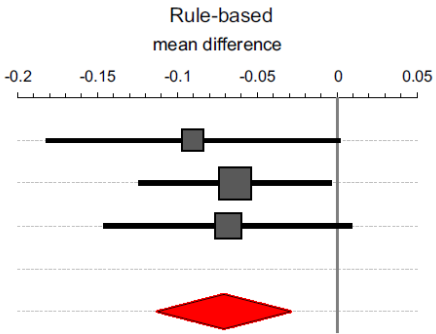
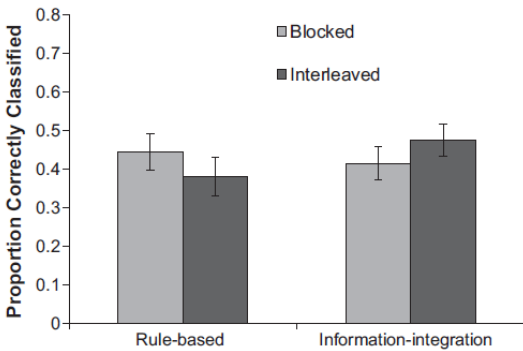
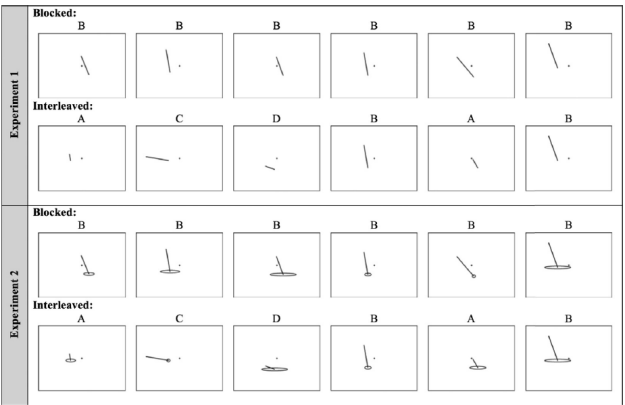
Optimal sequencing during category learning: Testing a dual-learning systems perspective



Sharon M. Noh^{a,*}, Veronica X. Yan^b, Robert A. Bjork^b, W. Todd Maddox^a

^a University of Texas, Austin, United States

^b University of California, Los Angeles, United States



Part 1: Setting up your experiment

Setting up your experiment: Study title and description

Describe your task to Workers

Title

Describe the task to Workers. Be as specific as possible, e.g. "answer a survey about movies", instead of "short survey", so Workers know what to expect.

Description

Give more detail about this task. This gives Workers a bit more information before they decide to view your task.

Keywords

Provide keywords that will help Workers search for your tasks.

“This is a psychology experiment conducted by the University of Texas at Austin. You will complete a brief 2-minute tutorial for \$0.50 and may be eligible for a longer study (25-30 minutes) for up to \$4 in bonus payments!”

Goal here is to get people to take your study, but also let them know what they’re signing up for

Important details to include:

- How long does your study take?
- How much will they get paid?
- What will they have to do?

You are space limited here, but you will have a chance to provide more info once they click on your study.

What participants see


























HIT Groups (1-20 of 865)

 Show Details

 Hide Details

Items Per Page:

20 

Requester	Title	HITs 	Reward 	Created 	Actions	
 James Billings	Market Research Survey	15,267	\$0.01	4h ago	Preview	Accept & Work
 Riku Lindholm	Find Mental Health Therapist Website and Email	9,711	\$0.14	2d ago	Preview	 Qualify
 Content Research	Clean Up How-To Questions	7,733	\$0.05	2d ago	Preview	 Qualify
 f8b64e4e-b7c8-47a8-9ee3-b161727f4c	Judge the reputation polarity of Article Clips	3,703	\$0.08	13d ago	Preview	 Qualify
 Klejda	Collect data from a Website: www.esources.co.uk, focusing on Home Supplies	1,576	\$0.01	6d ago	Preview	Accept & Work
Description		Time Allotted		Qualifications Required		Your Values
Review the given website and collect the information as requested.		180 Min		 HIT approval rate (%) is greater than 7		100
www.esources.co.uk, log in using: Username: AMZHL Password: email4manf, select Home supplies - filter on Manufactures and UK and EU countries		Expires		9/12/2020		
 Peter Birsinger	Find all advertisers in this text, if any	1,385	\$0.02	1h ago	Preview	Accept & Work
 Brain Bridge Lab	Identify objects in an image	1,381	\$0.01	5d ago	Preview	 Qualify
 Smaranda Muresan	Emotions in Art	1,108	\$0.05	5/18/2020	Preview	 Qualify
 Vision AI2	Recognizing Situations in a Video	985	\$0.75	1d ago	Preview	 Qualify
 str11223344	Tell us what this item is - General Contents - Batch ID #34554	919	\$0.08	8h ago	Preview	 Qualify
 Akari Asai	Wikipedia-based Open-domain Question answering (Aug 18)	880	\$1.50	6m ago	Preview	 Qualify
 ML Researcher	Make a question more concrete and smooth (Payment Adjusted)	830	\$0.52	7h ago	Preview	 Qualify

Setting up your experiment: Payment structure

Describe your task to Workers

Title

Describe the task to Workers. Be as specific as possible, e.g. "answer a survey about movies", instead of "short survey", so Workers know what to expect.

Description

Give more detail about this task. This gives Workers a bit more information before they decide to view your task.

Keywords

Provide keywords that will help Workers search for your tasks.

“This is a psychology experiment conducted by the University of Texas at Austin. You will complete a brief 2-minute tutorial for \$0.50 and may be eligible for a longer study (25-30 minutes) for up to \$4 in bonus payments!”

Setting up your task

Reward per assignment

This is how much a Worker will be paid for completing an assignment. Consider how long it will take a Worker to complete each assignment.

Number of assignments per task

How many unique Workers do you want to work on each task?

Time allotted per assignment

Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.

Task expires in

Maximum time your task will be available to Workers on Mechanical Turk.

Auto-approve and pay Workers in

This is the amount of time you have to reject a Worker's assignment after they submit the assignment.

Guaranteed amount participants will receive for completing your HIT

Batches under 10 are charged a 20% fee, otherwise the fee is 40%.

Free filters!

- **HIT Approval Rate (%) for all Requesters' HITs**
 - I set this to “greater than or equal to” 95
- **Location**
 - I include United States, and sometimes Canada (if I feel like it)
 - Can also filter by state
- **Number of HITs Approved**
 - Default “HIT Approval Rate” is 100% for people who sign up for the first time, which can be problematic
 - I usually set this to “greater than 100” for learning tasks, but you can set to 500 or 1000
- **Create your own qualification type**
 - Can be a little tricky to do; usually not necessary unless you are targeting a small pool

Worker requirements

Require that Workers be Masters to do your tasks [\(Who are Mechanical Turk Masters?\)](#)

☐ Yes ☒ No

Specify any additional qualifications Workers must meet to work on your tasks:

HIT Approval Rate (%) for all Requesters' HITs

greater than or equal to

95

[Remove](#)

Location

is one of

UGANDA (UG)
UKRAINE (UA)
UNITED ARAB EMIRATES (AE)
UNITED KINGDOM (GB)
UNITED STATES (US)

[Expand](#) | [Remove](#)

Number of HITs Approved

greater than

100

[Remove](#)

[\(+\)](#) Add another criterion (up to 2 more)

(Premium Qualifications incur additional fees, see [Pricing Details](#) to learn more)

Project contains adult content [\(See details\)](#)

☐ This project may contain potentially explicit or offensive content, for example, nudity.

Task Visibility [\(What is task visibility?\)](#)

☐ Public - All Workers can see and preview my tasks

☒ Private - All Workers can see my tasks, but only Workers that meet all Qualification requirements can preview my tasks

☐ Hidden - Only Workers that meet my Qualification requirements can see and preview my tasks

What participants see when they click on your HIT

Alien Learning Task (up to \$4.50 in total payments!)			
Requester: Preston Lab	Reward: \$0.50 per task	Tasks available: 0	Duration: 2 Hours
Qualifications Required: HIT Approval Rate (%) for all Requesters' HITs greater than or equal to 95 , Location is one of US , Number of HITs Approved greater than 100			

Instructions

For this study, you will play a game that involves learning about different aliens and their habitats! You will first complete a brief demo/tutorial on the task for \$0.50. If you are eligible for the full study, the experiment will continue after the tutorial and you will receive an additional \$4.00 for completing the full experiment. Please note that failure to follow instructions would result in the experiment terminating early, and you will not be eligible for the full bonus.

Please select the link below to participate. At the end of the experiment, you will receive a code to paste into the box below to receive credit for participating in our study. You will be paid \$0.50 for completing the 2-minute demo/tutorial, and \$4.00 in bonus payments for completing both portions.

NOTE: Please do NOT proceed with this experiment if you have already participated in this experiment before. Duplicate IDs will be flagged and will result in a rejection of the HIT. You can verify your eligibility by entering your Worker ID BEFORE accepting the HIT.

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

Experiment link:

<http://www.cogfog.com/SN/A2>

Provide the survey code here:

e.g. 123456

Submit

Part 2: Collecting good data

Demo 1: Instructions and demographics

Basic Information

Gender

☐ Male

☒ Female

☐ Other

Age

23

Birthday

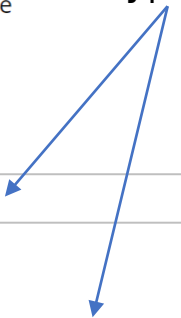
01/01/1988

Your age and birthday do not match up.
Are you sure you entered them correctly?

Education

Bachelors degree

Helpful for catching typos and mistakes



In every experiment I run:

Thank you for your participation. Please:

- Follow the instructions for each task and try your best to perform well
- Maximize your browser and focus completely on the task without any distractions
- DO NOT take notes during the experiment, because this interferes with our ability to accurately measure the learning process
- DO NOT participate if you feel you cannot fully commit to these requirements

When you have maximized your browser and are ready to begin, please click "Next" to continue to the task instructions.

The "Next" button will appear once you are in full screen mode.

Demo 2: Checking browser activity

- You can record how often the experiment browser is the active browser:
 - Clicking off the main browser window will start the recording timer
 - Pop-ups and alerts can also cause the experiment window to become inactive
- If you are recording responses during learning, you can alert people when the browser “loses focus”
- Note: this doesn’t catch everything...
 - Netflix open on a second monitor
 - Music playing in background
 - Participant walking away from computer
 - Other distractions

Demo 3: Tutorials

- Help participants understand how to do your task in the absence of an in-person experimenter and get used to timing, etc.
- Break down instructions into sections or small steps (especially when testing older adults), and cover stories are surprisingly helpful
- Make tutorials as interactive as possible
- Give participants a lot of feedback during this phase
- You can use data from the tutorial to exclude participants more objectively

Performance checks

- Include these throughout, starting with the tutorial
- Types of performance checks to consider:
 - Warnings based on unrealistic RTs
 - Probe or catch trials
 - Accuracy on an easy cover task
- Some strategies:
 - Warn people, but don't kick them out. Make the warnings annoying.
 - Warn people, then boot them (nicely) from the experiment.
 - People don't like getting booted, but they dislike HIT rejections more
 - You can warn people of this in the experiment description and instructions as well

Demo 4: Performance check example 1

- Include these throughout, starting with the tutorial
- Types of performance checks to consider:
 - Warnings based on unrealistic RTs
 - Probe or catch trials
 - Accuracy on an easy cover task
- Some strategies:
 - Warn people, but don't kick them out. Make the warnings annoying.
 - Warn people, then boot them (nicely) from the experiment.
 - People don't like getting booted, but they dislike HIT rejections more
 - You can warn people of this in the experiment description and instructions as well

Demo 5: Performance check example 2

- Include these throughout, starting with the tutorial
- Types of performance checks to consider:
 - Warnings based on unrealistic RTs
 - Probe or catch trials
 - Accuracy on an easy cover task
- Some strategies:
 - Warn people, but don't kick them out. Make the warnings annoying.
 - Warn people, then boot them (nicely) from the experiment.
 - People don't like getting booted, but they dislike HIT rejections more
 - You can warn people of this in the experiment description and instructions as well

Demo 6: Performance check example 3

- Include these throughout, starting with the tutorial
- Types of performance checks to consider:
 - Warnings based on unrealistic RTs
 - Probe or catch trials
 - Accuracy on an easy cover task
- Some strategies:
 - Warn people, but don't kick them out. Make the warnings annoying.
 - Warn people, then boot them (nicely) from the experiment.
 - People don't like getting booted, but they dislike HIT rejections more
 - You can warn people of this in the experiment description and instructions as well

Demo 7: Other considerations

- Progress bars give workers hope
 - People will sometimes drop out if an experiment seems too tedious or never-ending
- Breaks may be helpful for longer or more tedious studies
 - Keep in mind that you are paying people for this break, so don't make it too long.
- If you have a filler task between study and test phases, try to make it fun
 - People like getting paid to play Tetris

BEFORE collecting online data

- Run through the **entire** study (yes, on normal speed)
 - Good for catching lags (which tend to occur later) and other weird things that could happen
 - Things to test: different browsers, refreshing browsers, broken image, output
- Get RAs to run through the entire task while taking notes
 - Timing: important for determining payment
 - Are any of the instructions confusing?
 - Any bugs and/or general comments about the task (too easy, too difficult, etc.)
- Different browsers will display things differently and could cause problems
 - Recommend testing compatibility with Firefox, Chrome, and Safari
 - Safari tends to be a weird one...

Part 3: *After your study goes online...*

Once the study goes live...

- Monitor the email account (at least for the first round of participants)
- Approve participants reasonably quickly
- Canceling the batch can get messy...
- Recruit more participants than you would recruit for an in-lab study
 - 40-80% of recruited participants make it to the end (many get filtered or booted out)
 - General rule of thumb: release twice the number of HITs as your target N

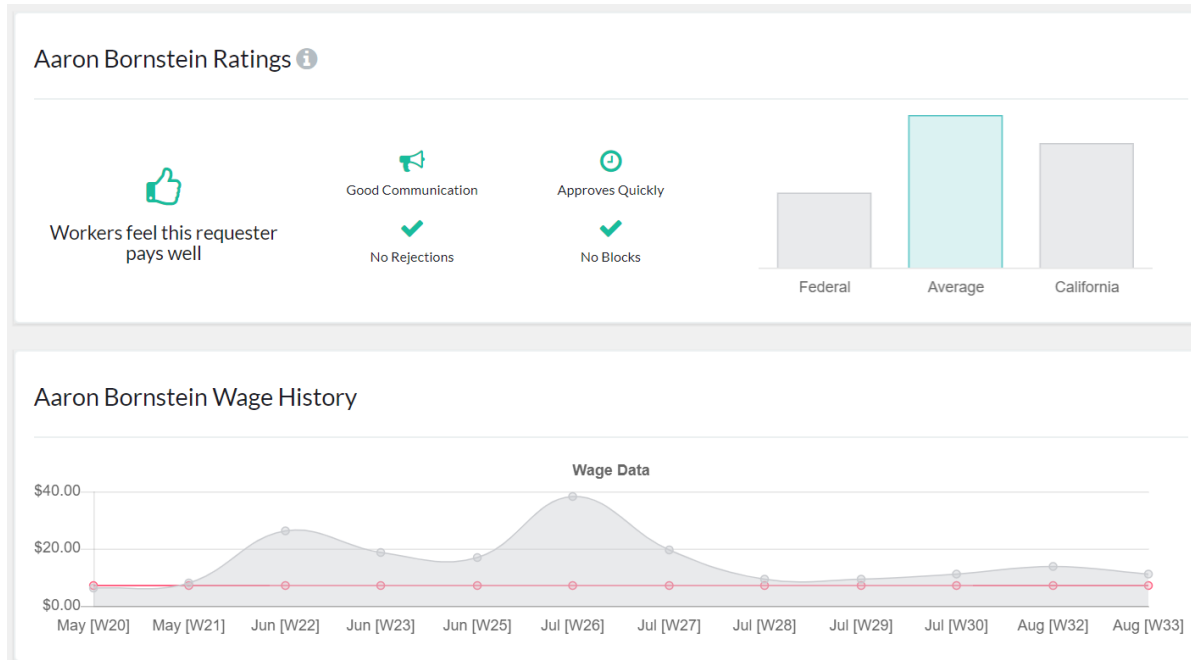
Feedback from Mturkers: Final Questions

- Did the experiment go smoothly or were there problems? Note: Your compensation will not depend on your answer below, so please be honest!
 - It went smoothly
 - There were minor problems
 - There were significant problems. I don't think my responses should be included in the data
- Which of the following problems did you have?
 - Sometimes a page wouldn't load and I would have to reload the page
 - Sometimes the page loaded but there was nothing there
 - During the experiment I experienced problems with my internet connection
 - Everything ran smoothly
- Were you doing anything else while participating in this study? PLEASE BE HONEST-- your compensation will **not** depend on your answer to this question.
- Is there anything you would like the experimenters to know? For instance, was the task too difficult, boring, etc?

Other feedback from MTurkers...

Aaron Bornstein A3GVHH5JG068XG Averages » HIT Group » Review Requester »	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: NO DATA	<p>I hate to say that this is the most tedious survey I've seen yet. It was supposed to take around 15 minutes but took around 45. That being said, the pay was very fair. The base pay was 50 cents, but I received the promised bonus of \$9.50 just a few hours later. As hard as this was, I would do it again for the pay.</p> <p>Jul 11 2020 Speechless flag comment flags, 1 comments »</p>
Aaron Bornstein A3GVHH5JG068XG Averages » HIT Group » Review Requester »	FAIR: NO DATA FAST: NO DATA PAY: 5 / 5 COMM: NO DATA	<p>Memory and Decision-Making Experiment(~ 10 minutes) \$2</p> <p>Took the whole hours to complete. Altogether I made \$13!</p> <p>Jun 27 2020 Sandooch flag comment flags, comments »</p>

Turkopticon
(turkopticon.ucsd.edu)

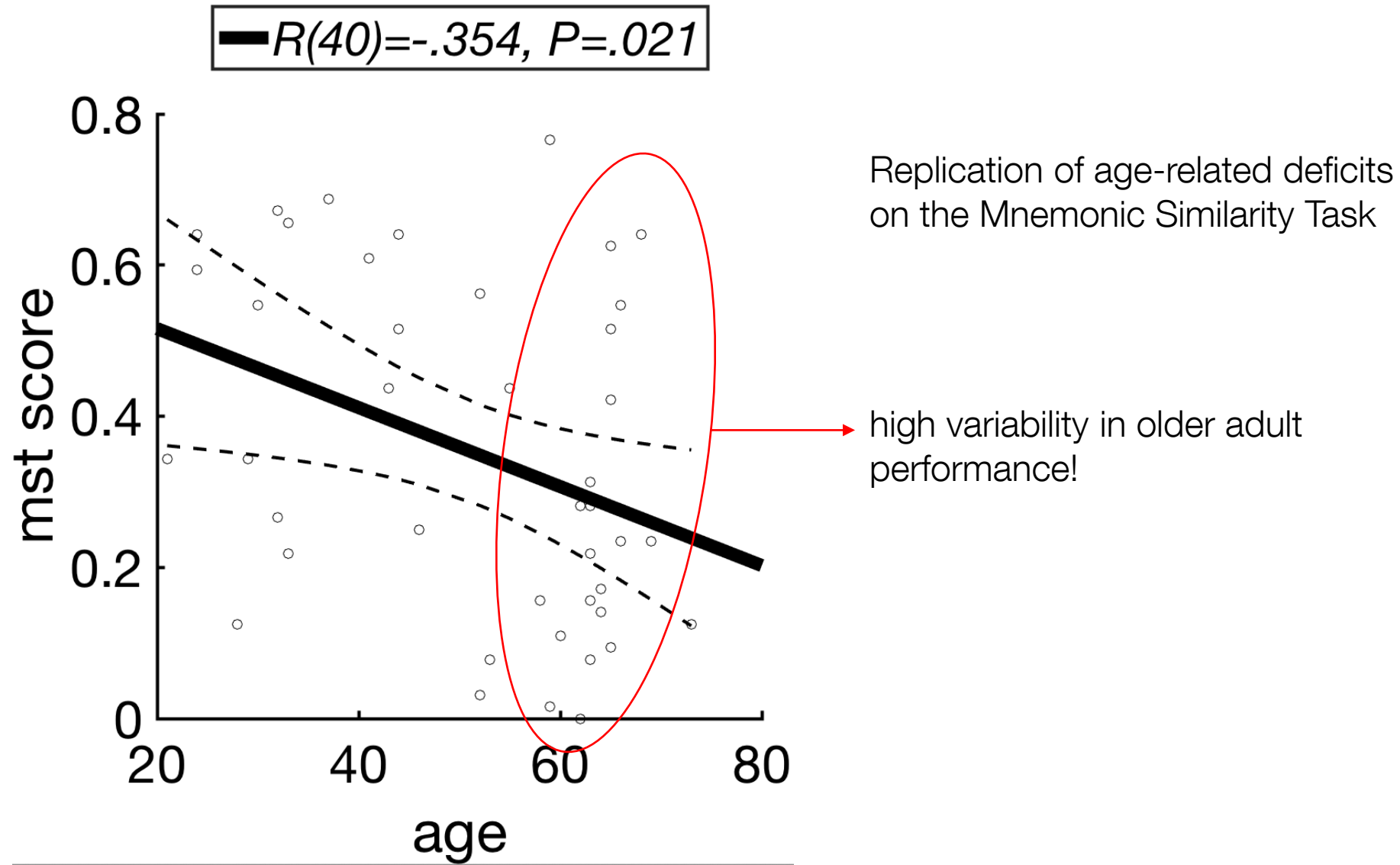


TurkerView (turkerview.com)

Miscellaneous things to expect

- Older adults will often perform better than younger adults on Mturk
 - Older adults may be less distracted and more motivated?
 - May need to consider additional measures to examine individual differences
- Younger adults are not that young (since COVID)
 - Expect people between 25-38 years
- Time of day seems to matter for recruitment
 - If you want more older adults, release HITs early in the morning
 - If you want younger adults, release HITs in the evening
- Many workers are relying on MTurk for income since COVID...
 - Try to be more generous with pay?
 - People are happy w/ \$10 hourly rate; \$12/hour is considered generous

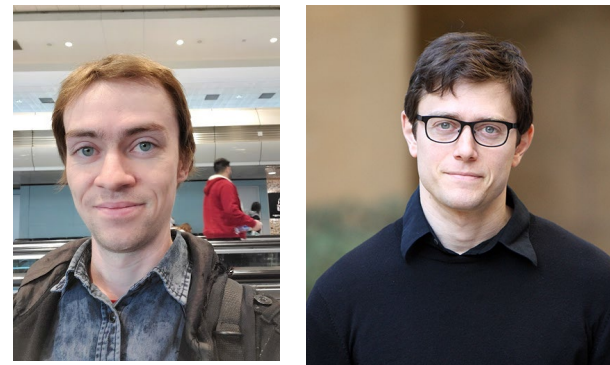
Mnemonic Similarity Task in MTurkers



Summary of Mturk Tips:

- The more info you can communicate to the participant, the better
- Use bonus payments to your advantage, and release HITs in small batches (<10)
- Take advantage of free filters; no need to use Master Workers
- Tutorials improve data quality and make it easier to exclude participants
- Put performance/data quality checks throughout the task
- Final questions can be important for getting feedback and excluding participants

Thank you!



<https://github.com/TysonKerr/Collector/tree/sn-demo>