Overview of research on crowdsourcing in evaluation

Miriam Jacobson
Claremont Graduate University

What is Crowdsourcing?

✓ Recruitment of many individuals for the completion of specific tasks[1]

✓ A task typically conducted by one individual is outsourced to a large group of people

✓ When done online, crowdsourcing can provide low-cost access to diverse, samples of people around the world

What is Crowdsourcing?

Complete a survey

Categorize information

Transcribe data

Translate text

Proofread documents

What is Crowdsourcing?

Complete a survey

Provide data for evaluation or research on evaluation

Assist evaluation tasks

Categorize information

Transcribe data

Translate text

Proofread documents

Amazon's Mechanical Turk (MTurk): A Crowdsourcing Service

amazon mechanical turk
beta
Artificial Artificial Intelligence

Your Account | **HITs** | **Qualifications**

Introduction | Dashboard | Status | Account Settings

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.
333,861 HITs available. **View them now.**

Make Money
by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. **Find HITs now.**

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → Work → Earn money

TASKS

Find HITs Now

or **learn more about being a Worker**

Get Results
from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. **Get Started.**

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account → Load your tasks → Get results
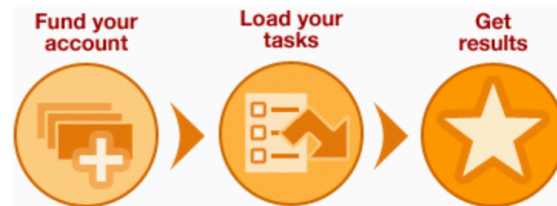
Get Started

Who is in the crowd?

People registered as workers on MTurk are:

- Located throughout the world, primarily in the United States (US) and India[2]

- The US sample is younger and more educated than the US public[2]

- Able to provide high quality data and completion of tasks[3,4]

Step 1: Create a HIT (Human Intelligence Task)

Start a New Project

Categorization

Data Collection

Moderation of an Image

Sentiment

Survey

Survey Link

Tagging of an Image

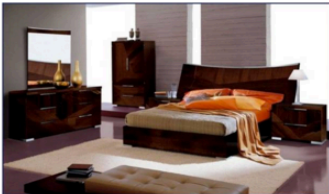Transcription from A/V

Transcription from an Image

Writing

Other

Example of Categorization

Create Project »

Choose the best category for this image

View Instructions↓

Select the room location in home for this picture. Seating areas outside are outside not living. Offices or dens are living not bedrooms. Bedrooms should contain a bed in the picture.

- ○ kitchen
- ○ living
- ○ bath
- ○ bed
- ○ outside

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

Step 1: Create a HIT (Human Intelligence Task)

Start a New Project

Categorization

Data Collection

Moderation of an Image

Sentiment

Survey

Survey Link

Tagging of an Image

Transcription from A/V

Transcription from an Image

Writing

Other

Example of Survey Link

Create Project »

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: [e.g. 123456]

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

Step 2: Customize HIT settings

Description of HIT
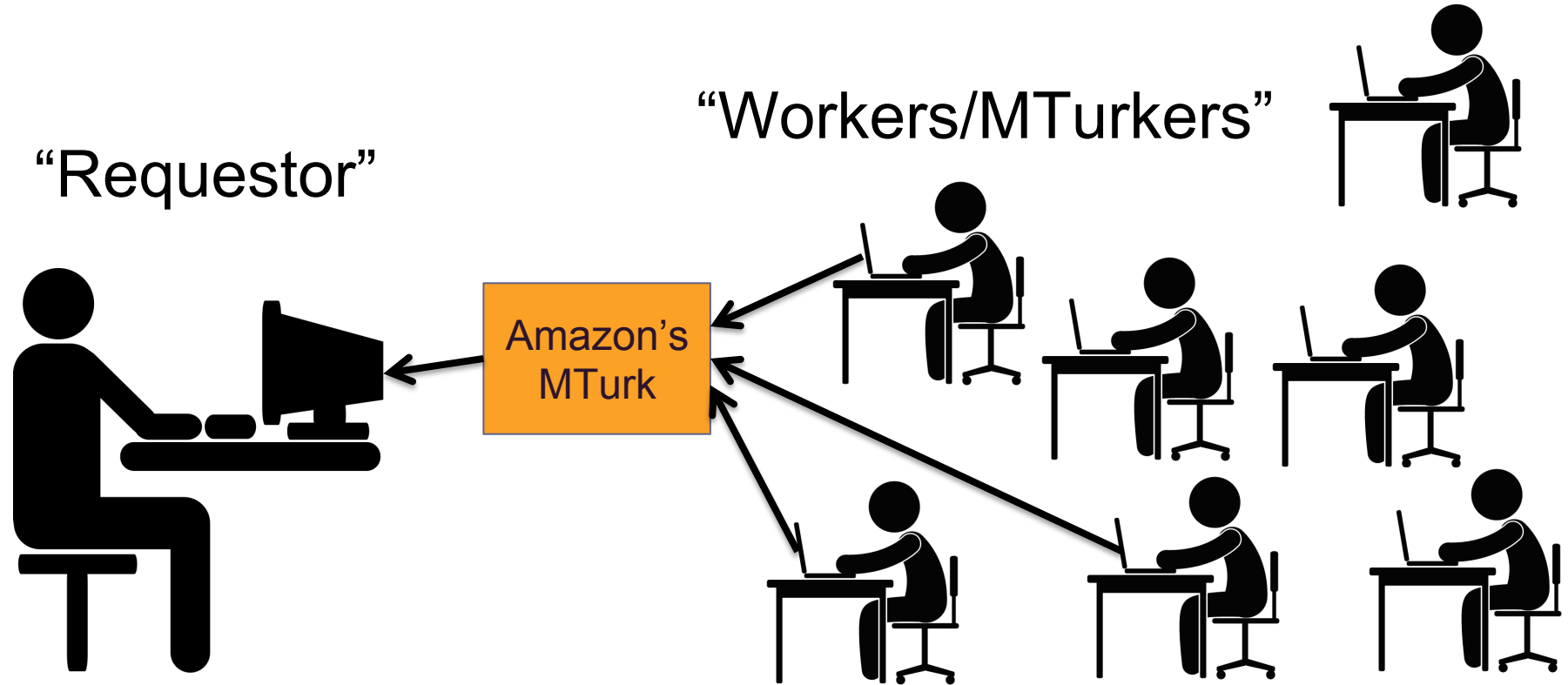
Time Allotted

Cost per HIT

Worker Requirements
- Example: Location, HIT approval rating

Step 3: Workers complete HITs

"Workers/MTurkers"
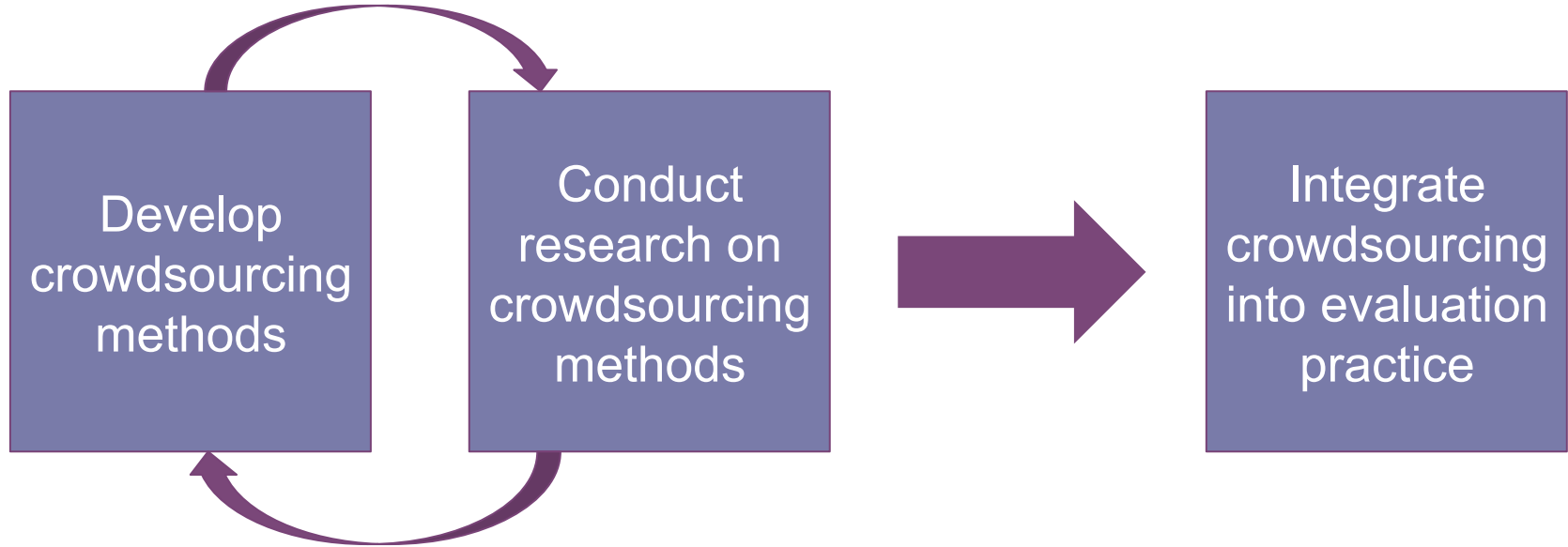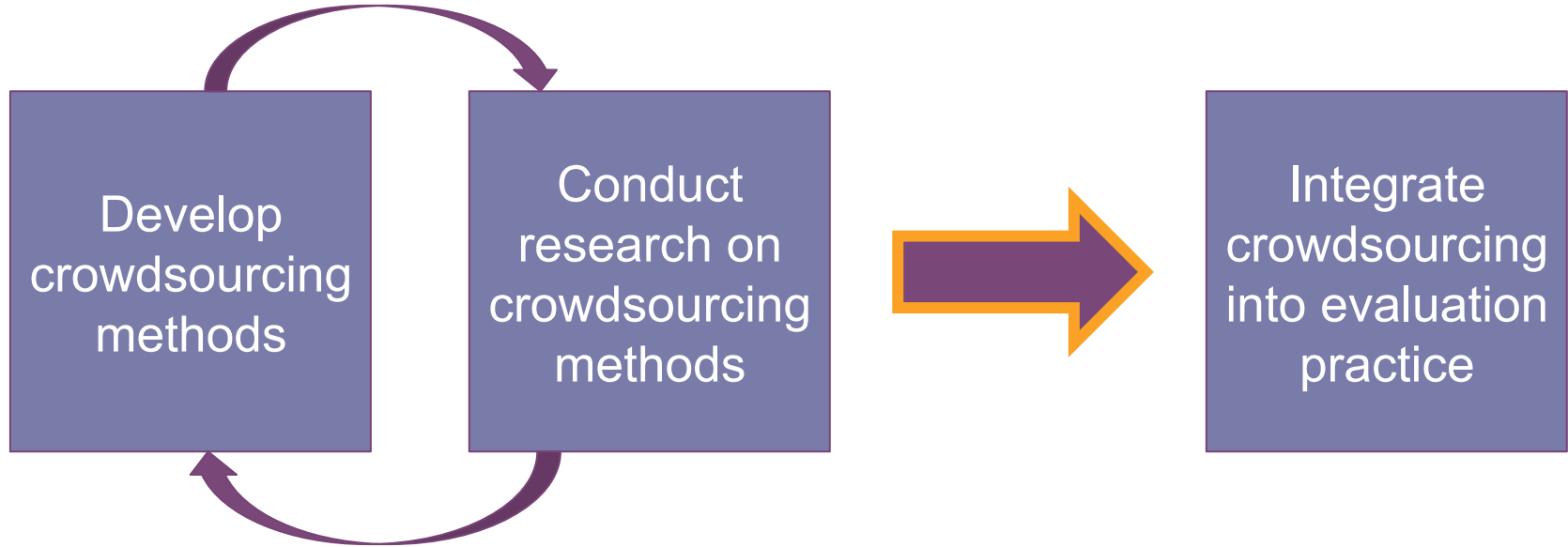
"Requestor"

Amazon's MTurk

Step 4: Approve work and pay workers

"Workers/MTurkers"

"Requestor"

Amazon's MTurk
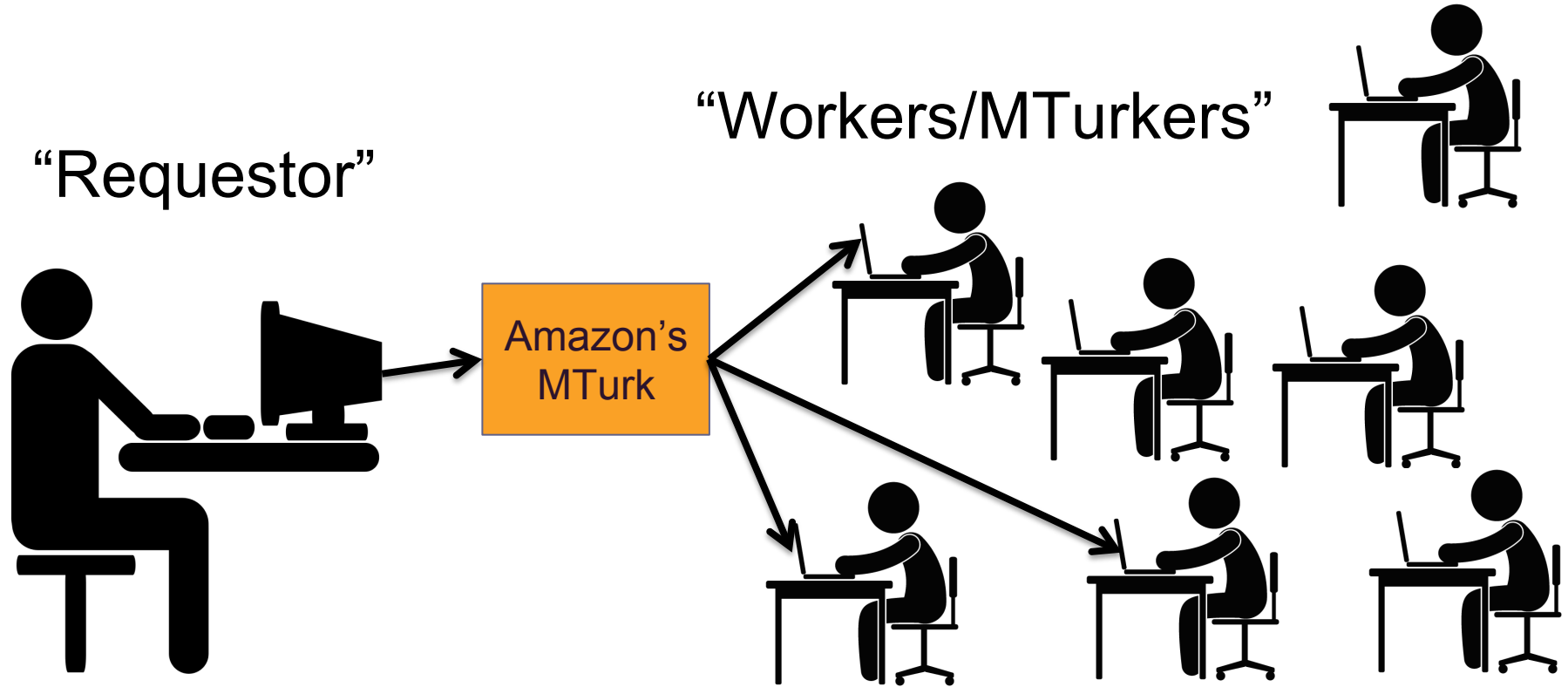
Research on Crowdsourcing

Research on Crowdsourcing

CDC Framework

Obtaining feedback on visualizations

Categorizing quotes from a survey

Developing evaluation criteria for a preschool program

Recruiting a sample for a health policy evaluation

Steps

Engage stakeholders

Describe the program

Focus the evaluation design

Gather credible evidence

Justify conclusions

Ensure use and share lessons learned

Standards

Utility
Feasibility
Propriety
Accuracy

References

1. Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5. doi: 10.1177/1745691610393980

2. Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23*(3), 184–188. doi:10.1177/0963721414531598

3. Azzam, T., & Jacobson, M. R. (2013). Finding a comparison group: Is online crowdsourcing a viable option? *American Journal of Evaluation, 34*(3), 372–384. doi:10.1177/1098214013490223

4. Simons, D. J., & Chabris, C. F. (2012). Common (mis)beliefs about memory: A replication and comparison of telephone and Mechanical Turk survey methods. *PLoS ONE, 7*(12), 1–5. doi:10.1371/journal.pone.0051876