CS 5306
INFO 5306:
Crowdsourcing and Human Computation

Lecture 23

11/21/17

Haym Hirsh

- Final exam: Tuesday, December 12, 9:00- 11:00am OLH155

- Project 2: Wednesday, December 6, 4:30pm
  (late projects accepted, no penalty through Dec 12)

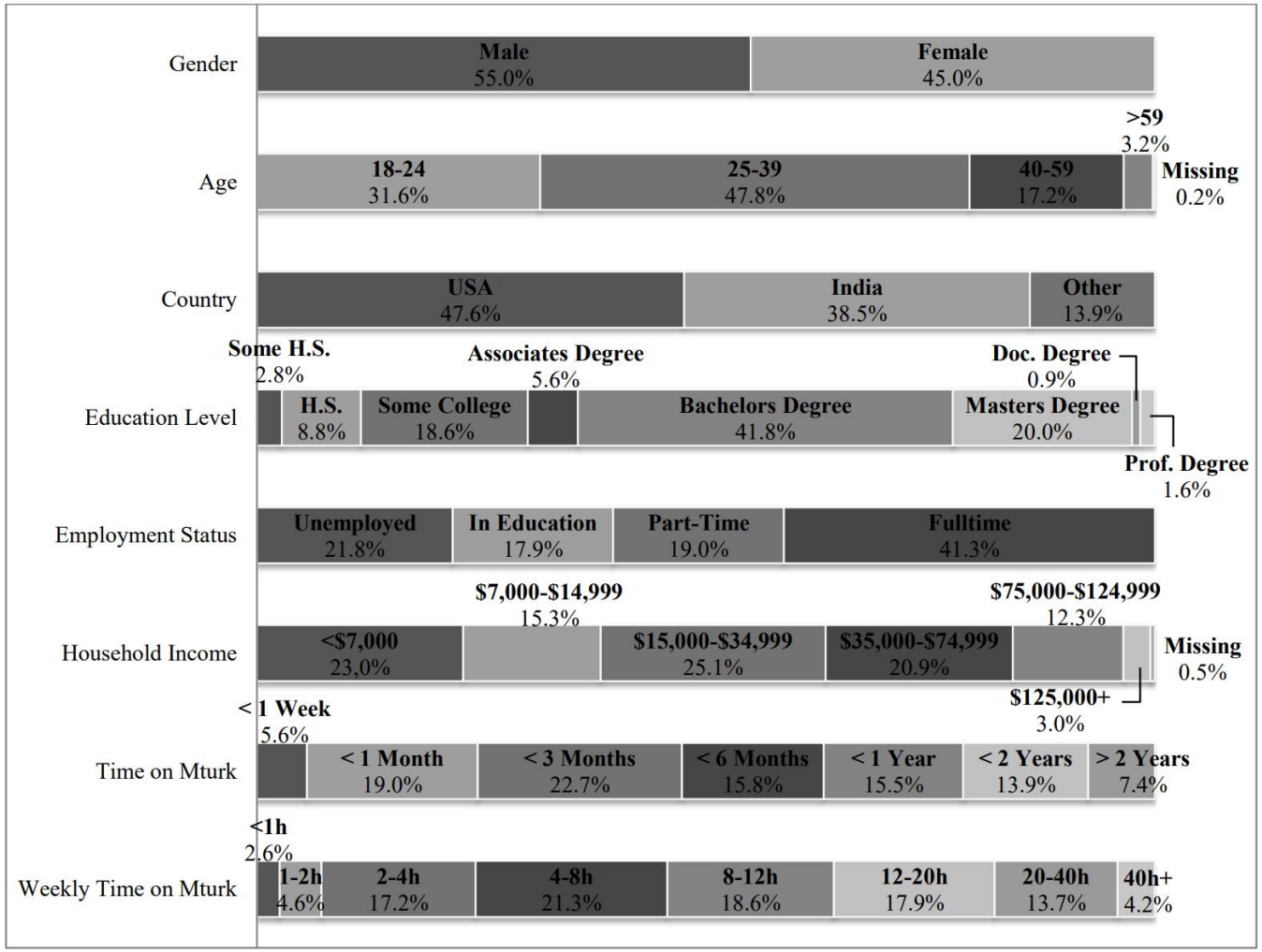
Today

What do we know about the crowdworker experience?

Demographics

- Ipeirotis, P. Demographics of Mechanical Turk. CeDER Working Papers No. CeDER-10-01. (2010) http://hdl.handle.net/2451/29585
- Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: shifting demographics in mechanical turk. In Proc. of CHI EA '10, ACM (2010), 2863–2872.
- Fort, K., Adda, G. and Cohen, K.B., 2011. Amazon mechanical turk: Gold mine or coal mine?. *Computational Linguistics*, *37*(2), pp.413-420.
  - Between 15,059 and 42,912 active workers
  - 80% of the tasks are carried out by the 20% most active (3,011–8,582) Turkers
- Kaufmann, N., Schulze, T. and Veit, D., 2011, August. More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk. In *AMCIS* (Vol. 11, No. 2011, pp. 1-11).

“More than fun and money.
Worker Motivation in Crowdsourcing-A Study on Mechanical Turk”
Kaufmann, N., Schulze, T. and Veit, D.
In *AMCIS* Vol. 11pp. 1-11, 2011



Demographics

demographics.mturk-tracker.com

Motivations

“More than fun and money.
Worker Motivation in Crowdsourcing-A Study on Mechanical Turk”
Kaufmann, N., Schulze, T. and Veit, D.
In *AMCIS* Vol. 11pp. 1-11, 2011

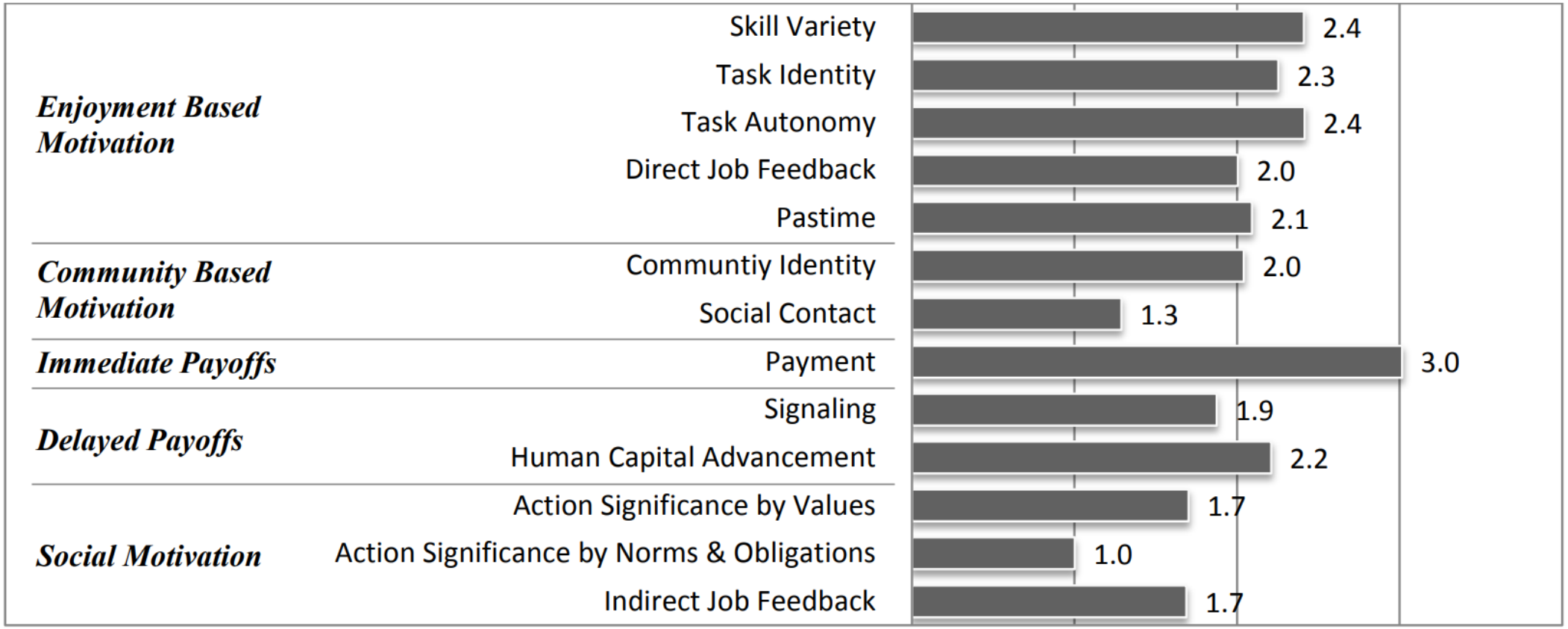
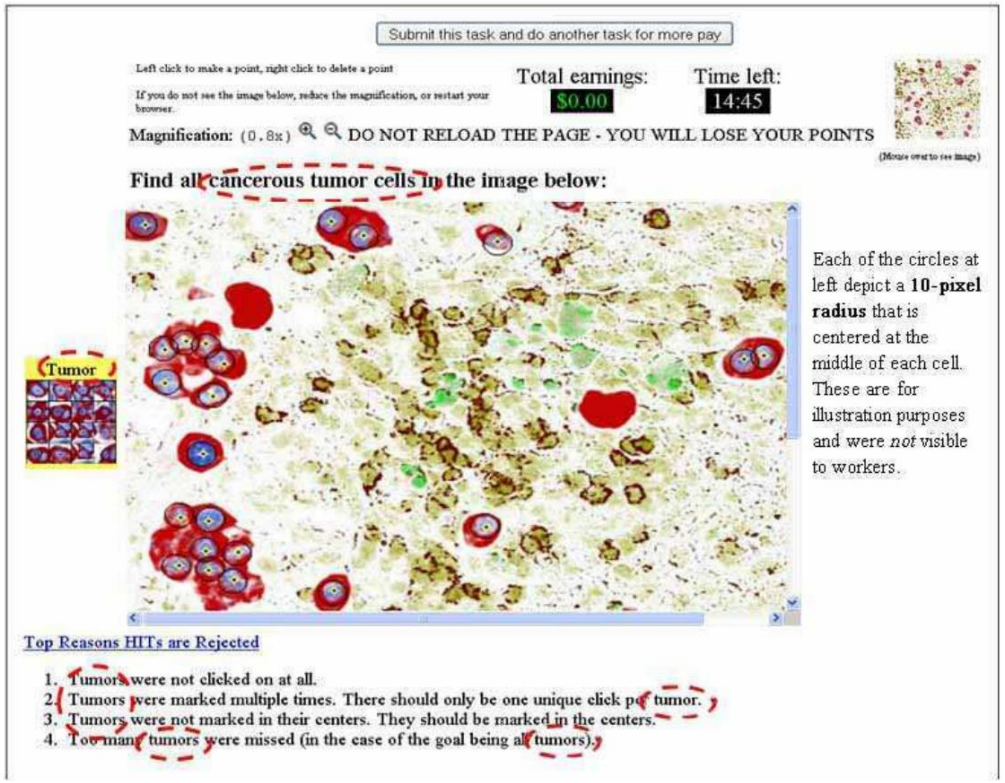


Figure 3: Mean construct scores after standardization

“Breaking monotony with meaning: Motivation in crowdsourcing markets”

Chandler, D. and Kapelner, A.

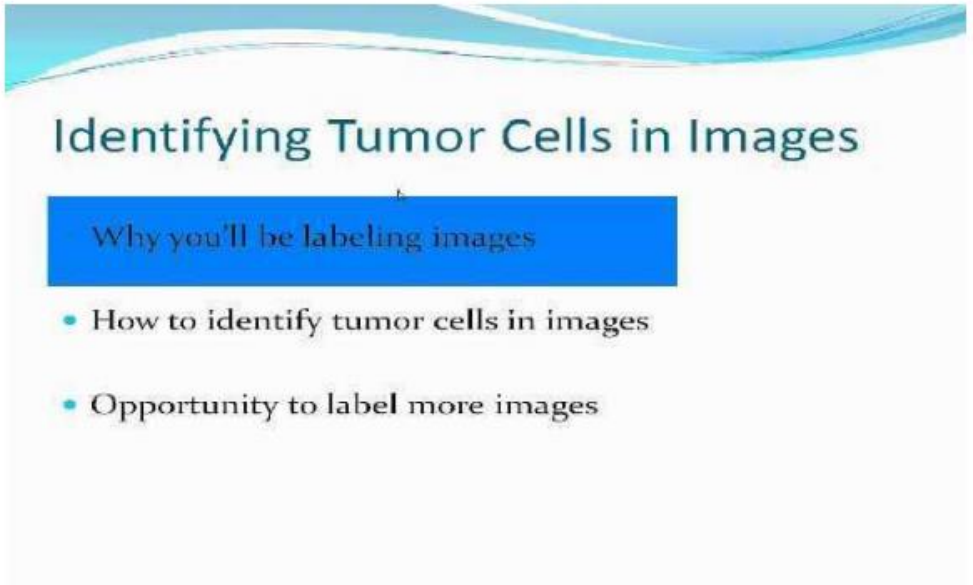
Journal of Economic Behavior & Organization, *90*, 2013



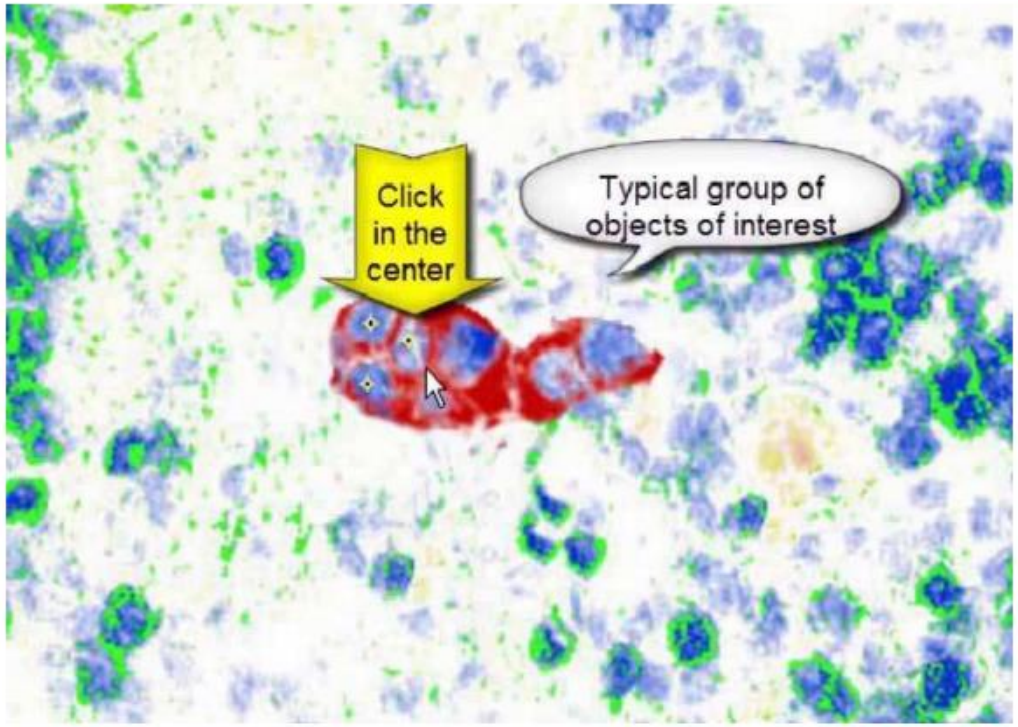
“Breaking monotony with meaning: Motivation in crowdsourcing markets”

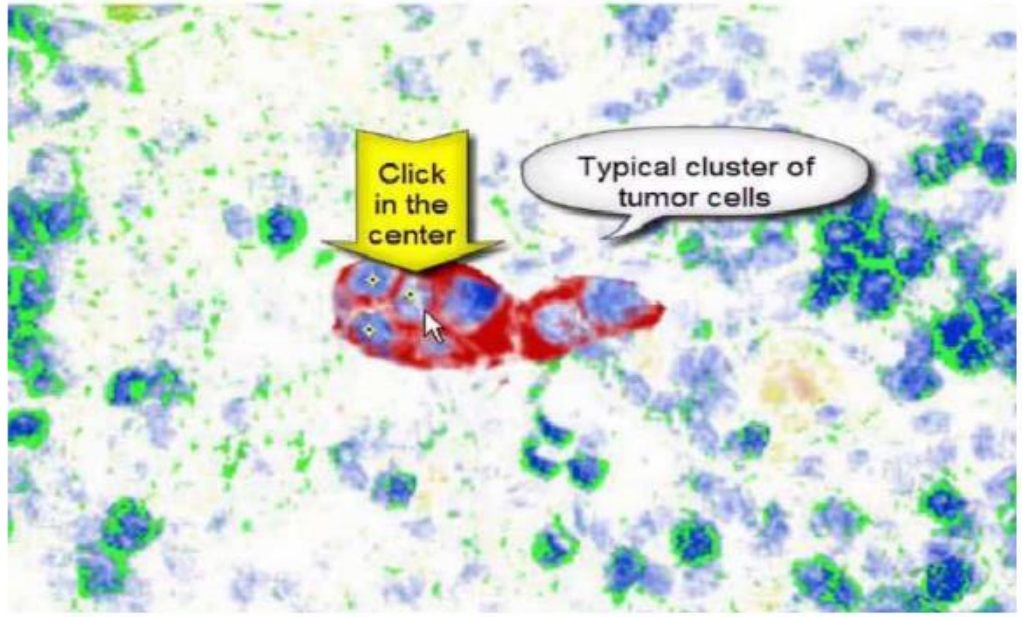
Chandler, D. and Kapelner, A.

Journal of Economic Behavior & Organization, *90*, 2013



“Breaking monotony with meaning:
Motivation in crowdsourcing markets”
Chandler, D. and Kapelner, A.
Journal of Economic Behavior & Organization, *90*, 2013





"Breaking monotony with meaning:
Motivation in crowdsourcing markets"
Chandler, D. and Kapelner, A.
Journal of Economic Behavior & Organization, *90*, 2013

	Induced to work	Did ≥ 5 labelings	Fine Quality	Average Hourly Wage
Meaningful	↑ 4.6%*	↑ 8.5%***	↑ 0.7%	↓ 4.5%
Meaningful (US)	↑ 5.1%*	↑ 8.9%**	↑ 3.9%	↓ 7.7%
Meaningful (India)	↓ 2.3%	↑ 7.0%*	↓ 3.1%	↑ 0.5%
Shredded	↓ 4.0%	↓ 2.8%	↓ 7.2%***	↑ 5.6%
Shredded (US)	↓ 2.3%	↓ 5.0%	↓ 6.1%*	↑ 9.5%
Shredded (India)	↓ 6.8%	↓ 1.6%	↓ 8.7%**	↓ 1.4%

* $p < .05$, ** $p < .01$, *** $p < .001$

“Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India”
Antin, J. and Shaw, A., *SIGCHI* 2012

“Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India”
Antin, J. and Shaw, A., *SIGCHI* 2012

"Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India"
Antin, J. and Shaw, A., *SIGCHI* 2012

Social desirability bias:

- Wikipedia: "tendency of survey respondents to answer questions in a manner that will be viewed favorably by others"

List experiment:

- Provide a list of choices, have respondent say *how many*
- Leave one choice (rotating through each of the choices) out of each survey

“Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India” Antin, J. and Shaw, A., *SIGCHI* 2012

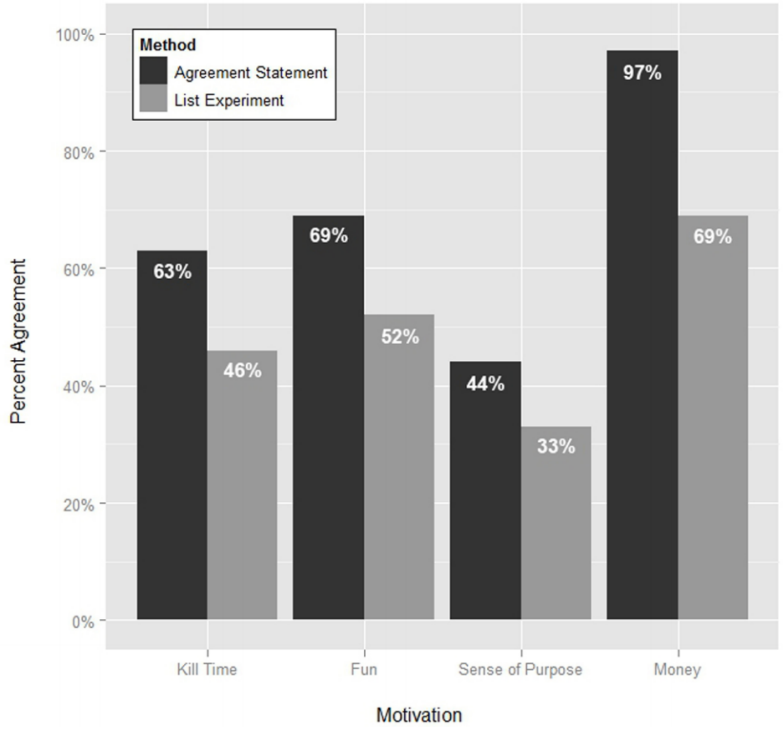


Figure 1. The proportion of US participants who selected each motivation using agreement statement-style questions or the list experiment. N = 1132.

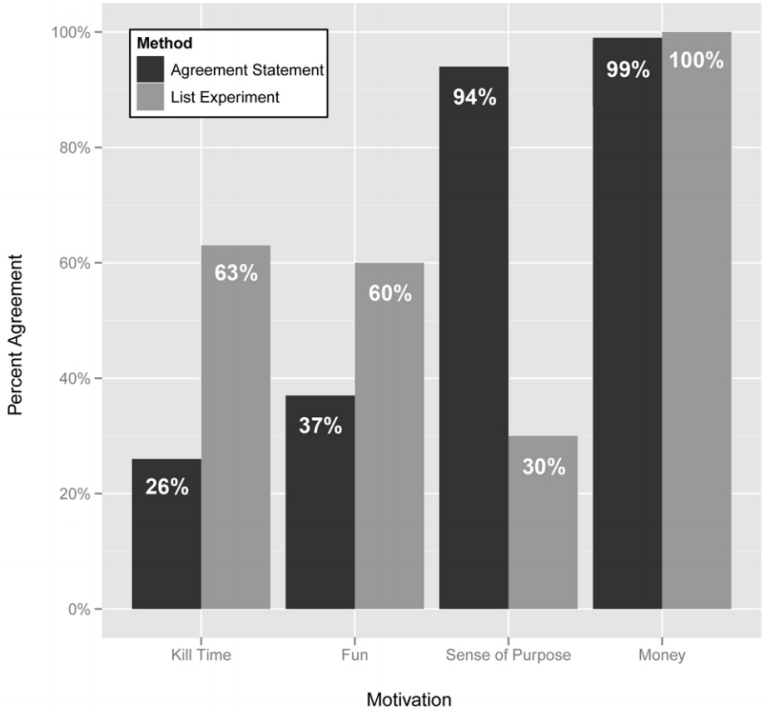
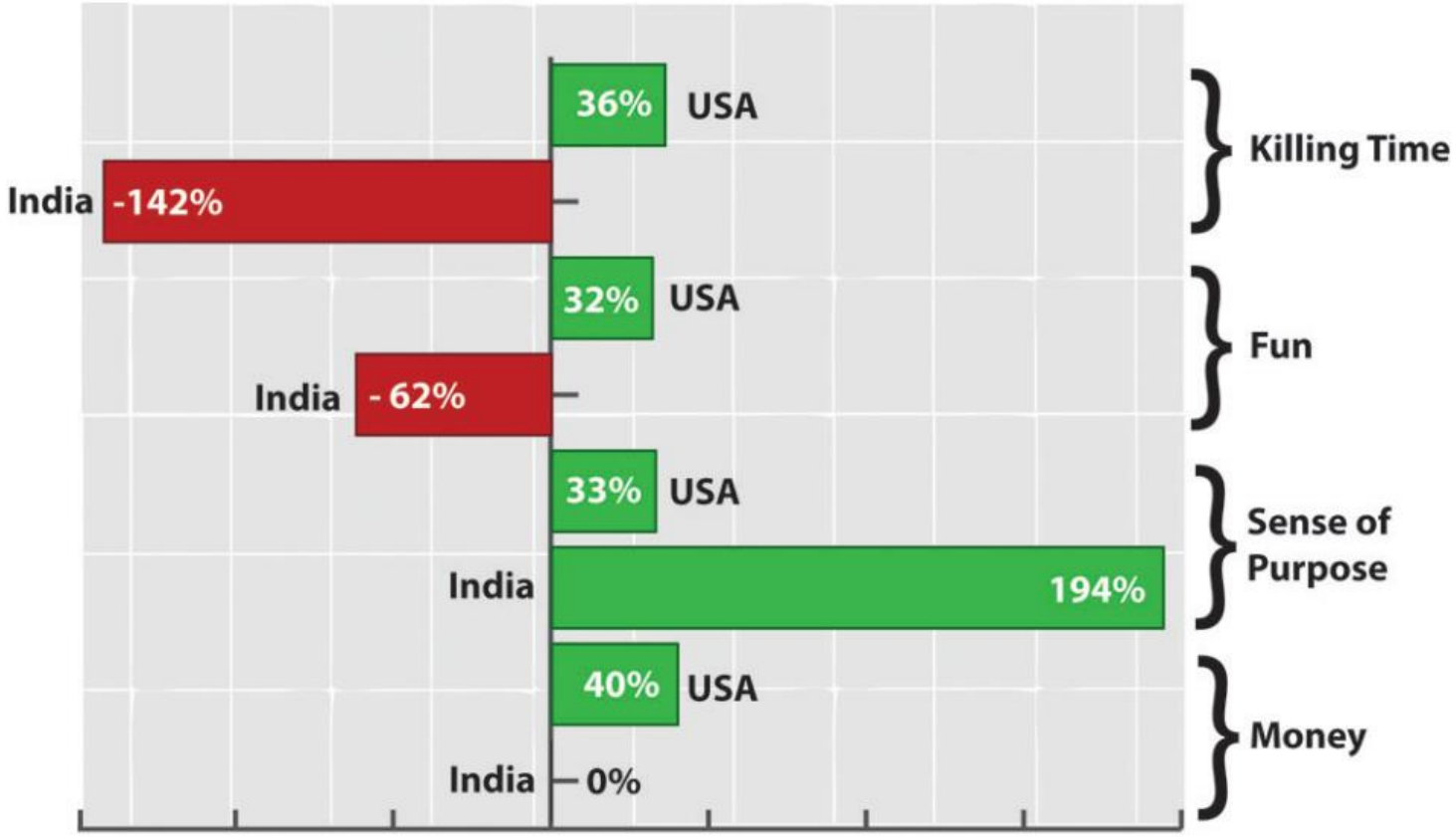


Figure 2. The proportion of India participants who selected each motivation using agreement statement-style questions or the list experiment. N = 898.

“Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India”
Antin, J. and Shaw, A., *SIGCHI* 2012



The Worker Experience

“Ethics and tactics of professional crowdwork”
Silberman, M., Irani, L. and Ross, J.
XRDS: Crossroads, The ACM Magazine for Students 17(2) 2010

- “Occupational hazards”:
  - Requesters not paying
  - “Staying safe” – inappropriate tasks
  - Little support from Amazon
  - Bad tasks (ambiguous, not enough time, …)
- Solution approaches:
  - From the “outside” (Turkopticon, Turker Nation, …)
  - New platforms
  - Legal protections

“Amazon mechanical turk: Gold mine or coal mine?” Fort, K., Adda, G. and Cohen, K.B., *Computational Linguistics*, *37*(2), pp.413-420l, 2011

- AMT:
  - Between 15,059 and 42,912 active workers
  - 80% of the tasks are carried out by the 20% most active (3,011–8,582) Turkers
- Concerns:
  - Low pay
  - Legal status: Employees?

“Web workers unite! Addressing challenges of online laborers” Bederson, B. B. and Quinn, A. J., CHI 2011

- Discussion of ethical and labor issues for workers
- Requester system design guidelines
  - Pay guidelines
  - Objective metrics
  - Feedback guidelines
  - Grievance process
  - Provide task context
  - Limit anonymity
  - …

“Turkopticon: Interrupting worker invisibility in amazon mechanical turk”, Irani, L.C. and Silberman, M., CHI 2013

- Problem: Human computation relies on/exploits worker invisibility
- Turkopticon: Platform for workers to discuss (and rate) requesters
- Discusses many issues for the future of crowd work

“The future of crowd work”
Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M. and Horton, J., CSCW 2013

“would we want our children to be crowdworkers?”

(discussion of research directions for the future)

“Being a turker”
Martin, D., Hanrahan, B.V., O'Neill, J. and Gupta, N., In *Proceedings CSCW,* 2014

- Ethnomethodological study of Turker Nation forum
- Many results/observations
- Key issues for workers:
  - AMT is a labor marketplace where pay is the most important consideration
  - Information quality for selecting tasks
  - Relationship with requesters
  - How to act collectively

“We are dynamo:Overcoming stalling and friction in collective action for crowd workers”

Salehi, N., Irani, L.C., Bernstein, M.S., Alkhatib, A., Ogbe, E. and Milland, K., CHI 2015

- Dynamo: Platform for organizing and collective action for AMT workers

"Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk" McInnis, B., Cosley, D., Nam, C. and Leshed, G.

- Main issues:
  - Risk and trust
  - Anonymity
  - Policy asymmetry => Risk, mistrust
    - Workers use outside tools/platforms to address the asymmetry
- Risk factors:
  - Task and interface design
  - Unclear evaluation criteria
  - Unresponsive, arbitrary resolution of rejections
    - “Amazon has no policy about Requesters’ responsibilities in communicating with Turkers and resolving disputes”
  - Lack of information on requesters
  - Inexperienced and unfamiliar requesters
  - Tasks with poor return
  - Prioritizing efficiency over quality

"Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk" McInnis, B., Cosley, D., Nam, C. and Leshed, G.

- Possible solutions (not requiring AMT action):
  - “Broken task” alarm button
  - Integrate existing “mutual aid” systems
  - Fast fail – provide feedback
  - Repairing rejected work
  - Tools for collective interaction with requesters
  - “Turker task design collective” – get turker help on task design

“The crowd is a collaborative network. “ Gray, M.L., Suri, S., Ali, S.S. and Kulkarni, D. CSCW 2016

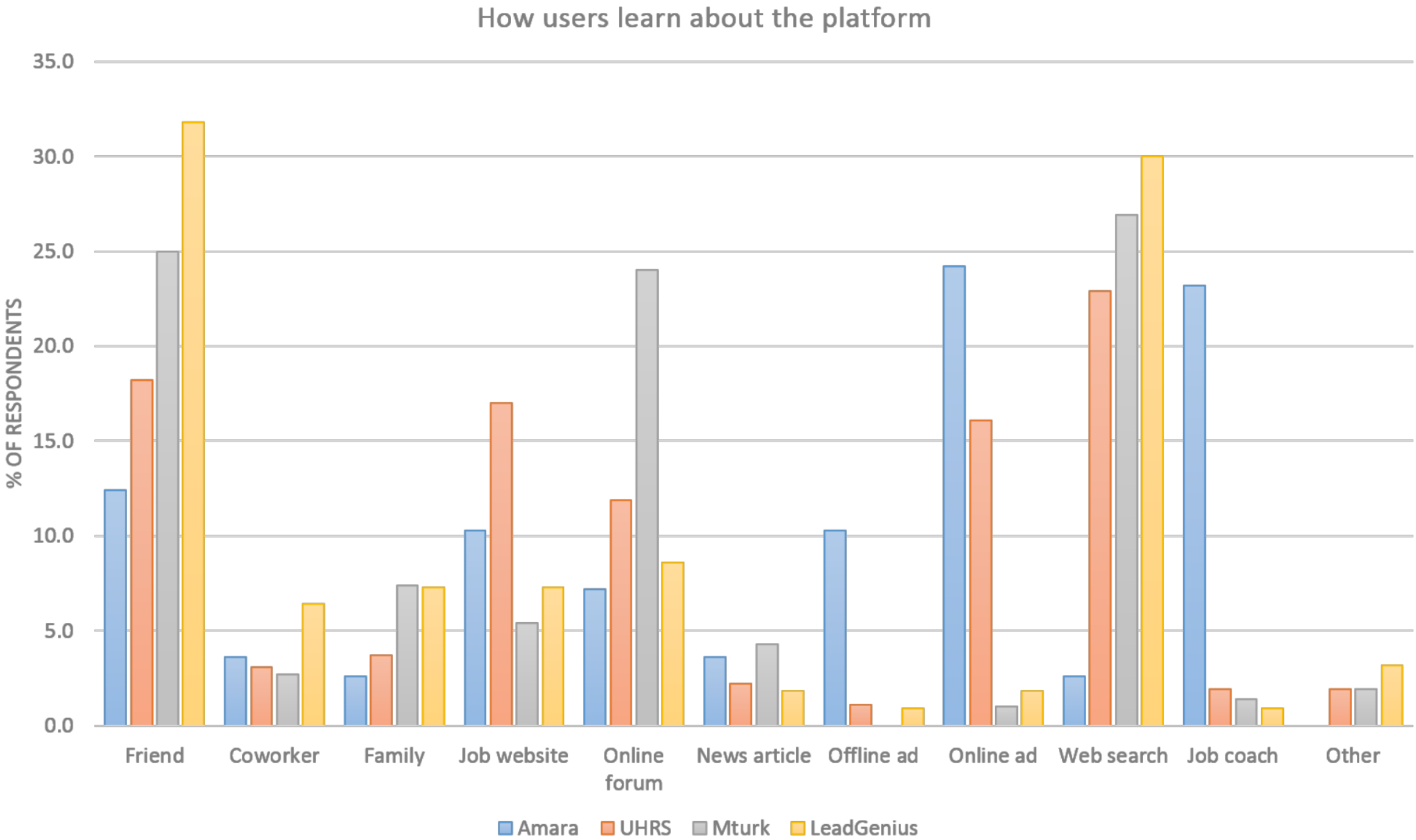
- Data:
  - AMT
  - Microsoft’s Universal Human Relevance System (UHRS)
  - LeadGenius
    - Provides sales “leads” via AI + human computation
    - Minimum wage based on each country’s cost of living
  - Amara.org
    - Video subtitles

“The crowd is a collaborative network. “ Gray, M.L., Suri, S., Ali, S.S. and Kulkarni, D. CSCW 2016

- Findings
  - Forms of collaboration
    - Reducing administrative overhead

“The crowd is a collaborative network. “ Gray, M.L., Suri, S., Ali, S.S. and Kulkarni, D. CSCW 2016

- Findings
  - Forms of collaboration
    - Reducing administrative overhead
    - Sharing tasks as employment opportunities


How users learn about the platform
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
% OF RESPONDENTS
Friend
Coworker
Family
Job website
Online forum
News article
Offline ad
Online ad
Web search
Job coach
Other
Amara
UHRS
Mturk
LeadGenius

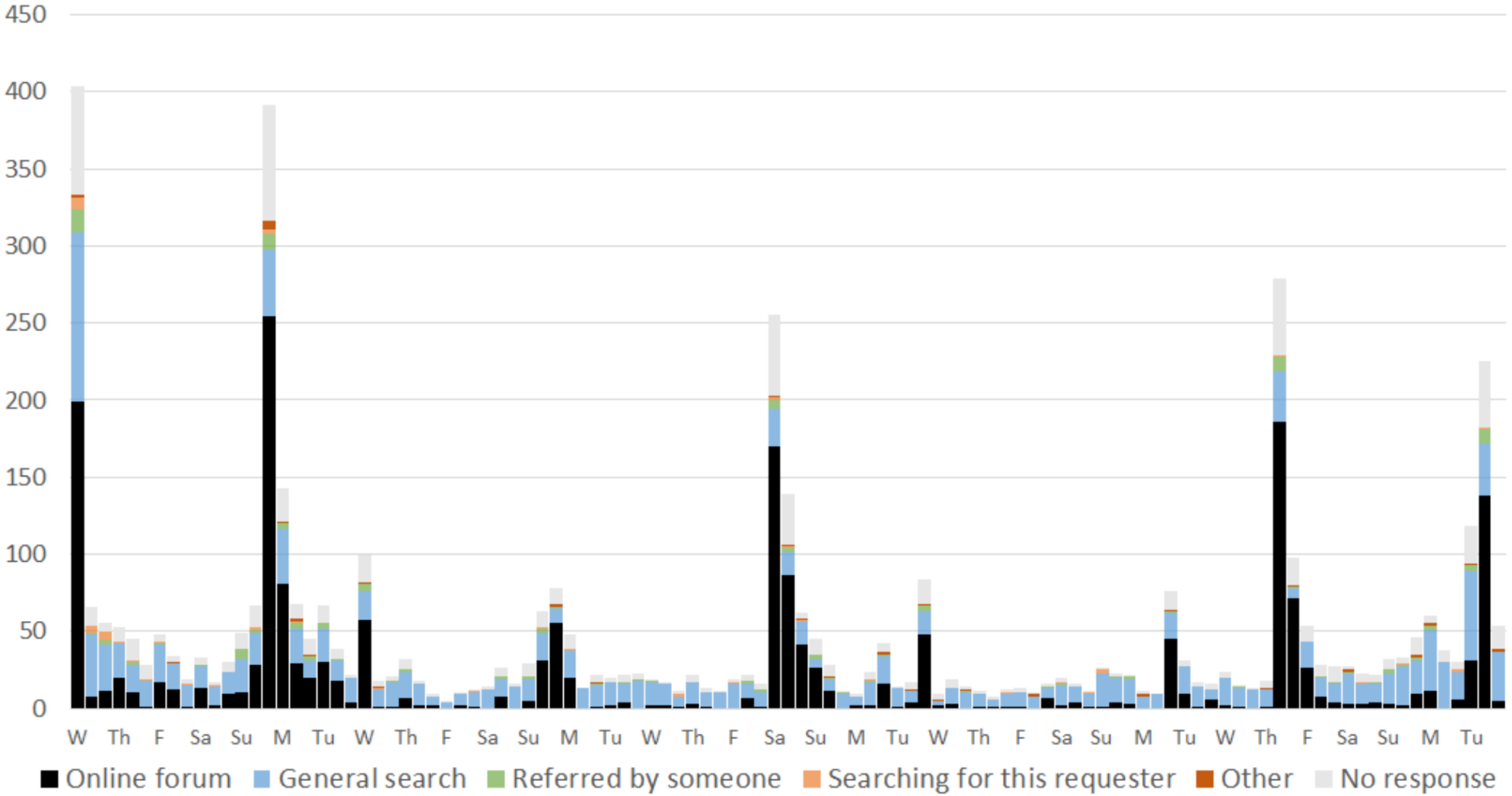
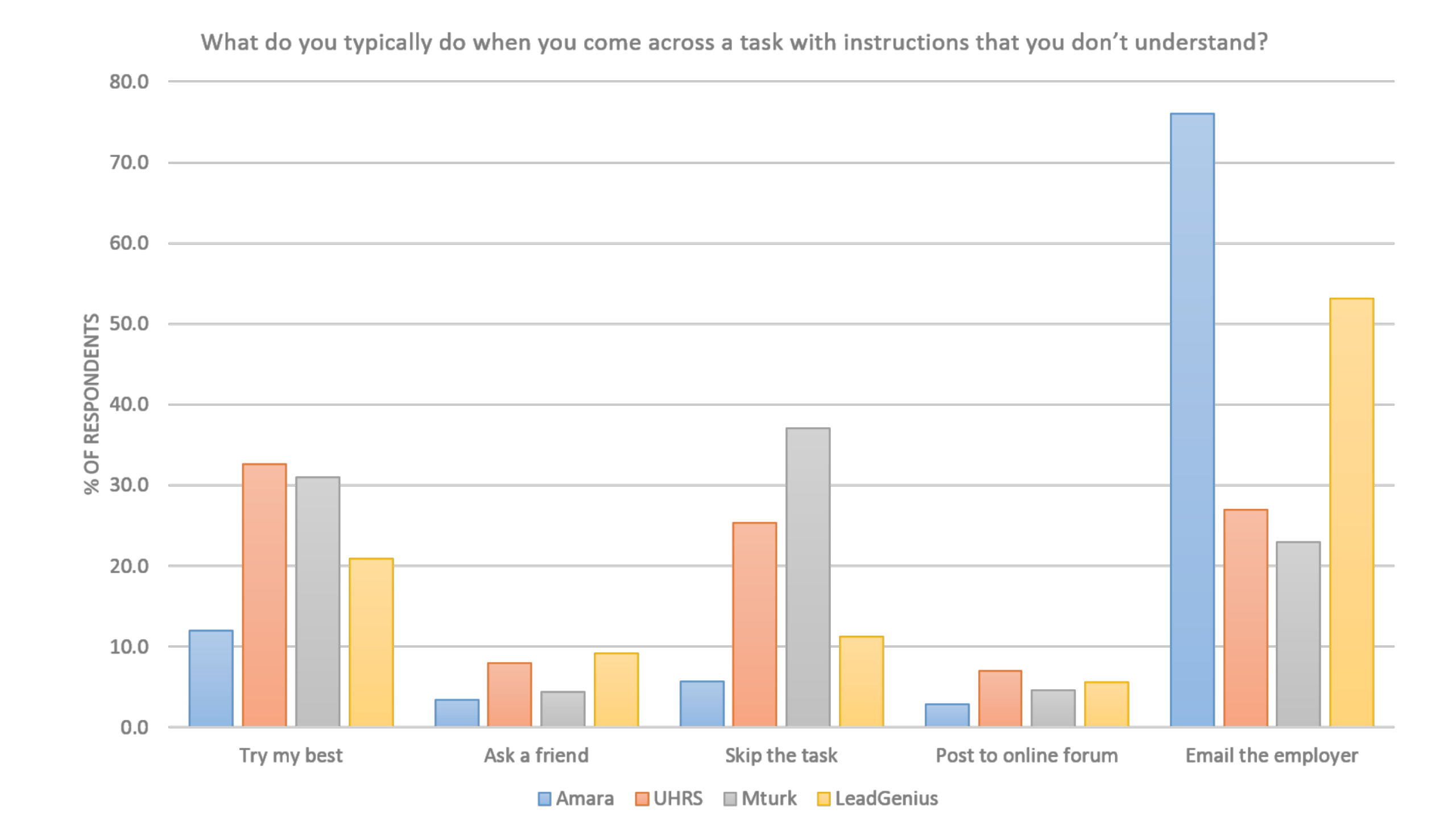


Figure 2. How workers found our mapping HIT (n=4,856) which ran from April 23 - May 28, 2014

“The crowd is a collaborative network. “ Gray, M.L., Suri, S., Ali, S.S. and Kulkarni, D. CSCW 2016

- Findings
  - Forms of collaboration
    - Reducing administrative overhead
    - Sharing tasks as employment opportunities
    - Helping others


What do you typically do when you come across a task with instructions that you don’t understand?
% OF RESPONDENTS
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Try my best
Ask a friend
Skip the task
Post to online forum
Email the employer
Amara
UHRS
Mturk
LeadGenius

Project 2

- Pick a cognitive bias that you conjecture might impact a worker's performance on AMT
- Evaluate your conjecture

Project 2

- Examples:
  - Does putting a “happy face” on a task page cause sentiment analysis tasks to be get more positive responses?
  - Does telling workers what previous workers have answered (akin to Asch’s study) bias their answers (possibly away from the correct one)?
  - Are workers more accurate if they are primed with a god-evoking cue?
  - Do workers favor answers that rhyme?

Project 2

- Sources for cognitive biases
  - Wikipedia's list of cognitive biases
  - The Cognitive Bias Cheat Sheet
    - (Depicted in the Cognitive Bias Codex)

Project 2

- Submission format:
  1. Define/explain the cognitive bias (independent of human computation)
  2. Explain the effect that you think it will have
  3. Describe the experiment that you ran to test it
  4. Document your outcomes

- Teams of size 1 or 2

- Due: Wednesday, December 6, 4:30pm
  (late projects accepted, no penalty, through Dec 12)