

6-2015

50 Days an MTurk Worker: The Social and Motivational Context for Amazon Mechanical Turk Workers

Gordon B. Schmidt

Indiana University - Purdue University Fort Wayne, schmidtg@ipfw.edu

This research is a product of the [Organizational Leadership](#) faculty at [Indiana University-Purdue University Fort Wayne](#).

Follow this and additional works at: http://opus.ipfw.edu/ols_facpubs

 Part of the [Industrial and Organizational Psychology Commons](#), and the [Organizational Behavior and Theory Commons](#)

Opus Citation

Gordon B. Schmidt (2015). 50 Days an MTurk Worker: The Social and Motivational Context for Amazon Mechanical Turk Workers. *Industrial and Organizational Psychology: Perspectives on Science and Practice*.8 (2), 165-171. Cambridge Press.
http://opus.ipfw.edu/ols_facpubs/84

This Article is brought to you for free and open access by the Division of Organizational Leadership and Supervision at Opus: Research & Creativity at IPFW. It has been accepted for inclusion in Organizational Leadership and Supervision Faculty Publications by an authorized administrator of Opus: Research & Creativity at IPFW. For more information, please contact admin@lib.ipfw.edu.

Industrial and Organizational Psychology

<http://journals.cambridge.org/IOP>

Additional services for *Industrial and Organizational Psychology*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Fifty Days an MTurk Worker: The Social and Motivational Context for Amazon Mechanical Turk Workers

Gordon B. Schmidt

Industrial and Organizational Psychology / Volume 8 / Issue 02 / June 2015, pp 165 - 171
DOI: 10.1017/iop.2015.20, Published online: 28 July 2015

Link to this article: http://journals.cambridge.org/abstract_S1754942615000206

How to cite this article:

Gordon B. Schmidt (2015). Fifty Days an MTurk Worker: The Social and Motivational Context for Amazon Mechanical Turk Workers. *Industrial and Organizational Psychology*, 8, pp 165-171
doi:10.1017/iop.2015.20

Request Permissions : [Click here](#)

Commentaries

Fifty Days an MTurk Worker: The Social and Motivational Context for Amazon Mechanical Turk Workers

Gordon B. Schmidt

Indiana University Purdue University Fort Wayne

The focal article of Landers and Behrend (2015) persuasively argues that universally condemning potential convenience data sources outside of traditional industrial–organizational (I-O) samples such as college students and organization samples is misguided. This author agrees that instead we need to consider the context, strengths, and weaknesses of more recently recognized potential data sources. This commentary will focus on understanding the context of one particular potential data source, Amazon’s Mechanical Turk (MTurk; <https://www.mturk.com/>). While some existing research has looked at the demographic characteristics of Amazon MTurk workers and how those workers’ answers compared with more traditional samples (Casler, Bickel, & Hackett, 2013; Goodman, Cryder, & Cheema, 2013; Paolacci & Chandler, 2014), for this commentary I decided to take a primarily different tack. For the space of approximately 50 days, I acted as an MTurk worker on the site and participated in online communities at which MTurk workers congregate. The purposes of this were to experience the MTurk worker environment firsthand and observe how MTurk workers interact with each other and the site. This was done in the spirit of participant-observer research. Stanton and Rogelberg (2002) argue that online communities might be a particularly fruitful avenue for such participant-observer research within the field of I-O psychology. I am quick to note here that I don’t see my efforts here as anywhere near as extensive as much of the participant-observer work of the past, and I did my time on

Gordon B. Schmidt, Organizational Leadership & Supervision, Indiana University Purdue University Fort Wayne.

Correspondence concerning this article should be addressed to Gordon B. Schmidt, Organizational Leadership & Supervision, Indiana University Purdue University Fort Wayne, OLS Neff 288D, 2101 Coliseum Boulevard, Fort Wayne, IN 46805. E-mail: schmidtg@ipfw.edu

MTurk in the spirit of such work rather than as a match for their methodological and analytical rigor. The observations I make in this commentary will be couched in my own experiences as well as the existing literature base on Amazon MTurk.

My Time on MTurk

My time as an MTurk worker ran for a period of approximately 50 days. During that time I participated in 806 human intelligence tasks (HITs), with a HIT basically being any single task a person, called a “requester,” wants done. These tasks varied significantly in nature, and while some were academic surveys (approximately 148 or 18.9% of my HITs done), other HITs were a number of different tasks, such as transcribing store receipts, classifying content into categories, providing tags for images, finding company information, moderating comments/pictures for websites, and a whole host of others. The time for the individual HITs I did ranged from literally 5 or 10 seconds to over 1 half hour. The value paid for a HIT ranged from \$0.00 (for a brief screener study for future potential work) to \$5.00 for a survey.

In engaging in the task through MTurk, I tried to take on the mindset of a person just starting out on the site and trying to earn a bit of extra money. While 806 HITs probably sounds like a large number, the amount of time put in was not nearly as sizable as one might think. Generally I did HITs while eating breakfast in the morning and after work while watching television in the evenings. I didn’t keep an exact time log, but I would estimate my usage averaged about an hour to an hour and a half per day. Estimates on how much time the average user spends on the site varies. Some estimates suggest just an hour per week (Chandler, Mueller, & Paolacci, 2014), while others suggest most MTurk workers spend less than 1 day a week on the site and complete 20–100 HITs per week (Goodman et al., 2013; Ipeirotis, 2010). While the average user has a relatively low amount of time on site, there are MTurk workers who seem to look at the site as more of a full-time job (Chandler et al., 2014), with Paolacci and Chandler (2014) finding some evidence across their own requester history that 10% of workers were responsible for 41% of their tasks completed.

For my own search for HITs to do on MTurk, I did some searches on the site itself, which allows a worker to sort HITs available to them by compensation per hit, creation time, expiration time, time allowed, and alphabetical order on request title. You can also search by keywords in the title of HITs available on the site, so for example you could search for the word “Twitter” if you had interest in doing HITs related in some way to that site. I also looked in two communities unaffiliated with Amazon MTurk but used by MTurk workers to share information on HITs that were deemed to be of high quality, usually with regard to good compensation compared with needed time for

completion. The sites were MTurk Forum (<http://www.mturkforum.com/>), which had a thread each day with site members sharing HITs they felt were worthwhile, and the sub-Reddit “HITsWorthTurkingFor” (<http://www.reddit.com/r/HITsWorthTurkingFor/>), where site users posted MTurk HITs of interest, with a bot checking to see whether a HIT was still active and a site macro crossing out HITs as they became no longer available. I’ll talk more about the nature of these sites in a future section.

Range of Tasks on Mechanical Turk

Landers and Behrend (2015) discuss one concern journals and reviewers have about sites like MTurk, which is that they potentially have “professional survey-takers,” with the implication that this will have a negative impact on quality or accuracy of survey results. On this front there is mixed news. On the bad news front, yes some users, and especially heavy users, are taking a lot of surveys. As discussed previously, Paolacci and Chandler (2014) found just 10% of site members were responsible for 41% of their survey response. This certainly suggests samples aren’t completely independent in terms of participants, although similar concerns can be had in psychology subjects pools or even in organizational samples where the same company is used for multiple studies. From my own experience, I did 152 academic research studies during my time on MTurk. As a researcher myself, perhaps I was drawn more toward research studies, but that number was during only a 50 day period, so long term users of the site would be likely to have a much higher number of studies done. This certainly could be a concern, although as Landers and Behrend (2015) point out, we can’t assume that because a person has done many research surveys their answers are somehow less accurate or of less value.

In the area of positive news, it is important to note that academic research is really only a small part of the population of tasks on MTurk. The majority of tasks appear to be for businesses. As the site was originally created by Amazon to get people to label site items (Behrend, Sharek, Meade, & Wiebe, 2011), many such tasks still are offered on the site in labeling and describing products or online content. Many requesters look for MTurk workers to transcribe sales receipts, business cards, and even recipes. Market research seems to be a significant part of the task population, with requesters asking workers to rate potential item titles and packaging. Other tasks include audio transcription, rating the sentiment of online posts, and even confirming final election voting totals.

As people that do academic research, it makes sense that we place our focus on academic surveys on MTurk, but that is just a part of the MTurk ecosystem. Taking surveys is not the main profession of MTurk users—rather it is just one of many tasks people do on the site. In some ways we

could see this as similar to the idea that part of the “job” tasks of a psychology student is to participate in the psychology subject pool. The MTurk workers have the potential to take significantly more studies over time than most subject pool participants do, but the pool of studies on MTurk is also exponentially larger, and there are many nonacademic tasks on MTurk also competing for their time and attention.

Community Around MTurk

Although participating in MTurk generally does not require a worker to interact with other MTurk workers (other than perhaps the occasional real dyadic interaction study), unofficial communities have sprung up online. These communities can potentially fill a worker’s need for connection with other “coworkers” as well as provide relevant information to workers on both specific tasks worth doing and general guidelines for successfully engaging in MTurk work. There are a number of MTurk related communities, but I will focus on two that I interacted with during my time on MTurk: The MTurk Forum (<http://www.mturkforum.com/>) and the sub-Reddit “HITsWorthTurkingFor” (<http://www.reddit.com/r/HITsWorthTurkingFor/>). I’ll talk about each in turn.

The MTurk Forum (<http://www.mturkforum.com/>) is a message board that covers a wide range of topics related to being a worker on Amazon MTurk. There is a section on the site for people to give tips for new members, a section for users to post their MTurk goals, a section to discuss particular requesters from the site, and even a section where requesters can advertise the HITs they are running. The site also includes links to various browser add-ons that can be used with the MTurk website to help search for HITs on the site, inform a person when new HITs are posted, and display user ratings of MTurk requesters so that a worker can know to steer clear of requesters who other workers have had problems with in the past. These applications are all third party and thus not endorsed by MTurk, meaning how well they function and interact with the site correctly can vary over time.

One of the most frequently used parts of the site is a forum called “Great HITs” where a new thread is started each day for site users to share HITs currently running that are seen as desirable. Posters will include a link to the actual study, the compensation, any listed qualifications, and often the amount of time the poster took to complete that hit. A good compensation rate for the time needed seems to be one of the primary consideration taken into account by those who post HITs in the daily thread. This fits with empirical research looking at MTurk workers by Chandler et al. (2014), who found participants said the topics they talk most about on message boards related to a HIT is how much a HIT pays and how long it takes.

Another criterion can be the number of HITs available for a single worker to take, which I saw talked about for multiple HITs posted. Workers will look for batches that they can do all in a row, potentially due to the value of not needing to search for a next HIT to do and the potential for quicker completion due to learning and familiarity to the task. This criterion is not very applicable to researchers, as we tend to want multiple participants to do a survey once, but for businesses with tasks like receipt transcription, having one user with good transcription skills do many HITs is beneficial.

Although the “Great HITs” thread is primarily a place to post such HITs, workers will use it for other functions as well. People will engage in informal conversations, discuss how they are doing during the week, and offer words of encouragement to each other. Thus the thread is also a way for MTurk workers to feel connected to these “coworkers” of sorts.

The sub-Reddit “HITsWorthTurkingFor” (<http://www.reddit.com/r/HITsWorthTurkingFor/>) has a more singular focus than the MTurk Forum does, focusing exclusively on new HITs that are seen as appealing. The site has a running list of threads that give links to individual studies. Generally each post has the name of the HIT, its compensation, approximately how long it takes, and a link to the actual HIT on MTurk. The site also has an automated bot that checks to see whether a HIT is still active, crossing it out if it is full or is expired. The site does not have the same type of community that the MTurk Forum has, but users will post on individual threads about any problems or issues they had with particular HITs. Other sites on Reddit fill the role of places where members can talk about other MTurk elements (<http://www.reddit.com/r/mturk>) and the general topic of making money online (<http://www.reddit.com/r/WorkOnline>).

As can be seen in the examples from both sites, a major focus is placed on helping site visitors to find and ultimately do the “good” HITs that are available on the MTurk. From a research perspective this can offer a dilemma, as offering higher compensation and having a good reputation with MTurk workers will help your studies to fill quicker due to this sharing of HITs on third-party sites but also can make your sample more likely to consist of these more heavy use MTurk workers rather than more casual or potentially naïve users. Whether this is actually problematic from a data quality or generalizability perspective is an open question (Goodman et al., 2013). In this regard I agree with Landers and Behrend that this is probably an empirical question rather than a question we can judge on its face.

The sites and browser extensions used by MTurk workers are likely to change over time, but it is important to note that these types of third-party resources are very likely to continue to exist. To some degree this suggests that doing research data collection on MTurk will always be somewhat of a “snowball” method, as workers who have taken a HIT can tell other workers

about studies that offer desirable compensation. Thus, in MTurk research data sets, it can be expected that at least a subset of participants will have at a minimum weak social ties with each other.

Conclusion and Next Steps

In this commentary I drew from my own experiences as an MTurk worker to discuss the context of Amazon's MTurk as to its range of tasks, the motivation of workers, and the community of MTurk workers. Landers and Behrend focus on the idea that understanding the context is important for understanding the relevance of a data source for particular research questions. From my experiences I now draw some tentative conclusions and suggestions for future needed research.

My time on MTurk does highlight for me the fact that when we consider the participant base from the site we do need to consider that at least a subset of participants are likely to have taken a large number of previous studies, with experience previously taking hundreds or thousands of studies not out of the question. For some I-O psychologists I'm sure this will be seen as a significant issue for use, perhaps a deal breaker, although I do agree with Landers and Behrend (2015) that experience participating in research doesn't necessarily mean that MTurk workers will provide inaccurate or poor data. It might suggest however that for studies that use deception in some way that MTurk workers are more likely to have seen it before and thus less likely to believe such manipulations. Empirical research could look at this deception manipulation concern by comparing MTurk samples with in-person samples. For example, Goodman et al. (2013) looked at how questions with factual answers were answered differently in an MTurk sample versus an in-person sample.

The second area that my time on MTurk highlights as important is the need to understand the community that exists around the site. MTurk workers are not necessarily autonomous individuals unconnected to other workers. Sites like the MTurk Forum offer workers places for social ties as well as discussions of what HITs are worth doing. This community influences the behaviors of workers. This affects the composition of samples, most likely in terms of well-paying studies by well-regarded requesters getting more participants who are the heavy MTurk users who frequently visit such sites. This could potentially impact research results. It also is a phenomenon that on its own could be interesting for I-O psychologists to study. MTurk and the community around it offer data for researchers on a potentially growing area of employment—freelance, fully online work.

With regard to Amazon MTurk the next needed step is for researchers to empirically examine more closely the impact this unique population has on sample data and for researchers to use that data to judge when MTurk

workers are appropriate or inappropriate samples for research questions relevant to the field of I-O psychology. MTurk is a potentially valuable source of data for I-O psychologists, and as such, we need to examine and explore it more fully.

References

- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43, 800–813.
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29, 2156–2160.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, 46, 112–130.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26, 213–224.
- Ipeirotis, P. (2010). *Demographics of Mechanical Turk* (CeDER Working Paper-10-01). New York: New York University. Retrieved from <http://hdl.handle.net/2451/29585>.
- Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology: Perspectives on Science and Practice*.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23, 184–188.
- Stanton, J. M., & Rogelberg, S. G. (2002). Beyond online surveys: Internet research opportunities for industrial–organizational psychology. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology*. Oxford, United Kingdom: Blackwell.

Amazon Mechanical Turk for Industrial and Organizational Psychology: Advantages, Challenges, and Practical Recommendations

Sang Eun Woo, Melissa Keith, and Meghan A. Thornton
Purdue University

We are in almost full agreement with Landers and Behrend's (2015) thoughtful and balanced critiques of various convenience sampling strategies focusing on the four most frequently used data sources in our field. In this

Sang Eun Woo, Melissa Keith, and Meghan A. Thornton, Department of Psychological Sciences, Purdue University.

Correspondence concerning this article should be addressed to Sang Eun Woo, 703 Third Street, West Lafayette, IN 47907. E-mail: sewoo@purdue.edu