A Community Rather Than A Union: Understanding Self-Organization Phenomenon on MTurk and How It Impacts Turkers and Requesters

Xinyi Wang
University of Minnesota
Minneapolis, MN 55455, USA
wang4831@umn.edu

Yu Cui
University of Minnesota
Minneapolis, MN 55455, USA
cuixx322@umn.edu

Haiyi Zhu
University of Minnesota
Minneapolis, MN 55455, USA
haiyi@cs.umn.edu

Joseph Konstan
University of Minnesota
Minneapolis, MN 55455, USA
konstan@umn.edu

Yangyun Li
University of Minnesota
Minneapolis, MN 55455, USA
lixx3524@umn.edu

Abstract

This paper aims to understand the self-organization phe-
nomenon among the workers of Amazon Mechanical Turk
(MTurk), a well-known crowdsourcing platform. Specifically,
we explored 1) why MTurk workers self-organize into on-
line communities (Turker communities), and 2) how the
workers' self-organization impacts the requesters and 3)
the workers. In the first study, we conducted a field experi-
ment by advertising the same survey tasks on both MTurk
and on Turker communities. In the second study, we inter-
viewed two founders of the Turker communities. The results
suggest that 1) workers' main motivation to participate in
communities is to "find good HITs". 2) For requesters, there
is no indication of differences in work quality between the
tasks posted on MTurk and the ones advertised on Turker
communities. 3) For workers, participation in Turker com-
munities is associated with higher income, controlled for
working hours. We also learned from the interviews about
why founders built the communities and their future plans.

Author Keywords

Crowdsourcing; Amazon Mechanical Turk; self-organization;
online community; work quality; working condition

ACM Classification Keywords

H.5.3 [Group and Organization Interfaces]: Computer-
supported Cooperative Work

Introduction

Getting work done has never been easier than right now with crowdsourcing. In 2006 Jeff Howe described crowdsourcing in *The Rise of Crowdsourcing* as "the new pool of cheap labor: everyday people using their spare cycles to create content, solve problems, even do corporate R & D" [7]. Nowadays, Crowdsourcing is no longer viewed as simply hiring a group of cheap workers to complete assembly-like piecework; it is an online, distributed problem-solving and production model [4] that enlists a multitude of humans [5] to undertake a task to satisfy a type of need [6].

Many websites, systems and applications fall under the broad definition of Crowdsourcing, such as Wikipedia, Linux, Yahoo! Answers [5], Threadless [3], Quora [14], and Amazon Mechanical Turk [9]. However, these Crowdsourcing instances differ by how the crowd workers work together. In systems like Wikipedia and Linux, workers work collaboratively: they communicate with each other directly about how to get the tasks completed. In contrast, in systems like Mechanical Turk (MTurk), workers are assumed to work on their own tasks by themselves. The independent working paradigm allows requesters to use algorithmic approaches such as majority voting to manage quality [10] and run user studies and experiments [11].

However, a growing trend of self-organization among MTurk workers (Turkers) [8] challenges the "independent worker" assumption [15]. To give a sense about the size of the self-organization, Table 1 shows four communities we observed: Turker Nation, MTurk Forum, MTurkGrind, and MTurk Crowd, each of which has members ranging from 2,000 to 58,000, contributing to over 3,700,000 messages.

This paper intends to explore the self-organization phenomenon among workers of Mechanical Turk, and how it could impact the requesters and the workers. Specifically,

Turker Community	Member	Discussion	Message
Turker Nation	18,022	24,102	485,996
MTurk Crowd	2,821	1,163	606,058
MTurkGrind	12,502	29,390	1,088,576
MTurk Forum	58,527	11,775	1,597,448

Table 1: The summary statistics of Turker communities (www.turkernation.com, mturkforum.com, mturkcrowd.com, mturkgrind.com) as of January 4, 2017.

we try to answer the following three research questions:

RQ1: Why do Turkers self-organize into groups?

Do Turkers self-organize because they want to "unionize" to demand better pay and benefits [1] , or socialize with others, or learn new skills from others, or something else?

RQ2: How does this phenomenon impact researchers who use MTurk as a source for research?

Particularly, will the rise of communities change the demographics of participants and their quality of work?

RQ3: How does this phenomenon impact the Turkers?

How much do Turkers participate in the Turker communities? How does participation impact their working conditions such as working hours and income?

Study Overview

Two studies were conducted to answer above questions. Study 1 was a field experiment with 3 conditions that advertised surveys on MTurk and Turker communities. To see a more complete picture of the phenomenon, we continued with Study 2: an interview with the communities founders.

Study 1: Field Experiment

Experiment Design
We posted three Human Intelligence Tasks (HITs) on MTurk. The content of the HITs were identical, but the channels where the HITs were posted were different:

C1 was a standard HIT posted on MTurk (as our baseline).

C2 was posted by a *requester* in the "Requester Support" section of Turker Nation and MTurk Crowd with a description message: *"Hello all! You are invited to participate in a research that intends to understand experience on MTurk."*

C3 was posted in the "Daily Thread" section (where Turkers recommend HITs) in Turker Nation and MTurk Crowd. The message was similar to C2, but posted by a real *community member*: *"Hi fellow Turkers! I've just completed a HIT and really want to recommend it to you! It was great and I liked it. Please click here if you are interested to do it on MTurk!"*

Task content: survey
The three HITs contained identical survey questions summarized into five categories: 1) the source where the Turkers saw this HIT. 2) Experience on MTurk, including work hours, income, trust, and motivation. 3) Experience in Turker communities, including the level of participation, trust, motivation, and hours spent in the communities. 4) Demographics (age, gender, and nationality). 5) An ideation task where we asked Turkers to "*come up with 15 different ways of using a 5-minute break*". We used this task to assess the Turkers' work quality, following Zhu et. al.'s method [16].

Experiment Procedure
C1, C2, and C3 were posted for 50 HITs each and closed at the end of the week. All participants were compensated for $.50 (market rate $5/h [2]) for 10 min of their time. 116

Table 2. To check for experimental manipulation, 82% of the workers in C1 were recruited from intended source (MTurk); 79.8% in C2 and 95.2% in C3 were from intended sources (MTurkCrowd and TurkerNation). Data from unintended source were then excluded, which left 101 altogether (42, 19 and 40 in C1, C2, and C3 respectively) for further analysis.

remained after excluding 3 incompletes, and 101 remained after the manipulation check (Table 2).

Measurements
Work quality is measured by how many answers one gave for the ideation task and how creative each answer was. The creativity of an idea was approximated by how frequent that category of ideas appeared. The more answers one gave, the higher the work quality score; the more less-frequently-appeared ideas one had, the higher the score. Overall, the most frequent idea (183 times) fell into the "*exercise*" category, and the second was "*social media*" (149 times); examples of the least frequent ideas include "*make paper airplanes*" or "*run the flight of stairs up and down*".

Participation is the self-reported level of participation in the communities: none, viewer (low), and member (high).

Study 1: Result

	Source	C1	C2	C3
MTurk	MTurk	82.0%	18.5%	0%
	Turker Nation	0%	22.2%	71.4%
Communities	MTurk Crowd	6.0%	55.6%	23.8%
	MTurk Forum	0%	3.7%	4.8%
	MTurkGrind	4.0%	0%	0%
	Other	8.0%	0%	0%

Table 2: Manipulation Check

RQ1: Why do Turkers self-organize into communities?

Based on surveys results, participants' top two *initial* motivations (i.e., why did they join the communities in the first place) were the same as the top two *current* motivations (i.e., why did they participate now): to "find good HITs" (73.5% and 77.5% respectively) and "learn skills to be a good Turker" (68.9% and 55.0%). For the third motivation,

47.0% joined initially to "ask technical questions" and now 24.5% "just socialize". Interestingly, 6.6% and 7.3% participants chose "unionize together to demand better pay" as initial and current motivation to join the communities.

RQ2: How does self-organization impact requesters?

Condition	(n)	C1(42)	C2(19)	C3(40)	All(101)
Gender	M	66%	54.2%	31%	49.1%
	F	34%	45.8%	64.3%	47.3%
Nationality	US	84%	74.1%	97.6%	84.0%
Age	18-29	42%	25.9%	21.5%	36.1%
	30-39	30%	14.8%	26.2%	28.4%
	40-49	16%	22.2%	28.6%	18.3%
	> 49	12%	25.9%	23.8%	15.4%
	Med	30-39	40-49	40-49	30-39

Table 3: The break down of demographics by condition.

Demographic diversity of the sample
To examine how posting the HITs to Turker communities impacts the demographics of the sample, Table 3 compared participants of three conditions on gender, nationality, and age, suggesting that posting to communities might yield a more gender-balanced, more non-US and older sample.

Work quality
We originally expected that tasks advertised in Turker communities (C2 and C3) might be completed in higher quality than posted on MTurk (C1). However, ANOVA on the task quality suggests no statistically significant difference among three conditions ($p = 0.51$). Note that this also suggests no difference on quality ($p = 0.70$) between tasks advertised by requester (C2) and recommended by Turker (C3).

RQ3: How does self-organization impact Turkers?

Participation
ANOVA on the participation of the Turker communities ($F(2, 98) = 7.94$, $p < 0.001$) confirms our assumption that tasks advertised in communities (C2 and C3) would attract more members than tasks only posted on MTurk (Figure 1). It is also interesting to see that even tasks only posted on MTurk (C1) still attract 77% of Turker community members, indicating the prevalence of the self-organization of Turkers.
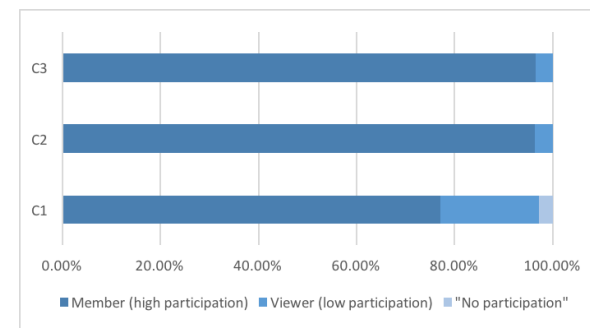

Figure 1: 77.06% of participants in C1 are members of Turker communities, while 96.43% and 96.61% are in C2 and C3.

Working hours per week
We used participation in Turker communities (*none, low, high*) to predict the working hours using a linear regression model. The result suggests that the more a worker participates in Turker communities, the longer hours s/he works on MTurk (coef. = 3.67, $p < 0.001$).

Income per week
The linear regression model to predict income shows positive association between participation and income while controlling for work hours (coef. = 0.17, $p < 0.001$). This suggests that participating in Turker communities might

be associated with better working conditions of Turkers. One possibility is that the workers can find better HITs in the communities and thus earn more even when they work same amount of time. However, this correlational analysis does not rule out alternative explanations such as better workers choose to participate in Turker communities more.

Study 2: Interview

Study 1 lets us understand the phenomenon from a normal Turker's view, but not from a community founder's perceptive. To see a more complete picture, we designed Study 2 - interviews with founders of the Turker communities. The 30 minute semi-structured Skype interviews were conducted with founders of Turker Nation and MTurk Crowd.

	Community	Seniority	Commitment
P1	Turker Nation	10 years	Full time
P2	MTurk Crowd	9 years	Full time

Table 4: The summary statistics of interviewees in study 2.

Study 2: Result
Founders: not a union, but a platform to communicate

P1 created Turker Nation because Amazon didn't allow Turkers to talk to each other and it kept them down.*"... not a union for Turkers but more like a connection and association. People could get together and exchange experience."* As for MTurk Crowd, P2 regarded it as *"a really open place for Turkers and researchers to communicate without any constraints."* P2 thought that Turkers self-organize, not to fight for better pay, but instead to help each other with a better experience on MTurk.

MTurk communities were designed to protect the quality of requesters' data

"In forums, people are not allowed to talk about the content of HITs." –P2. In other words, Turkers can only comment on the requester or the HITs themselves. Most Turkers there, especially "super Turkers" with years of experience on MTurk, *"take the work seriously." –P1.* Both founders told us they knew the importance of the authenticity of the data and tried to protect it.

For the future, two founders have different plans

P1 worried about the problem workers face, such as underpayment and job competition with robots, and was running a website named We Are Dynamo *"for Turkers to fight back resistance."* On the other hand, P2 thought that MTurk Crowd achieved her initial goal. *"In Mturk Crowd, there are new Turkers and experienced Turkers and they work together well. And I am really happy to see requesters come to my forum to get help and post ads." –P2*

Discussion

The two studies lead to several preliminary but interesting findings. First, the main motivation behind the self-organization was to "find good HITs", instead of unionizing to get better paid. Second, there is no sign of difference in the quality of work between the tasks posted on MTurk and the ones advertised in Turker communities, although the demographics of the samples differ slightly. Third, for workers, participation is positively associated with higher income even controlled for work hours, which suggests that participation in Turker communities might lead to better working conditions.

As we learned from the interviews with Turker community founders, these communities were designed to protect the data quality. The founders shared concerns Turkers have but have different plans for the future. The detailed takeaways for Turkers and requesters are summarized below.

For MTurk requesters

Not a niche community
Compared to standard Turker sample (34% female, 16% non-US, median age 30-39), sample recruited by posting the HIT to Turker community resulted in a more gender-balanced (46% female), slightly older (median age 40-49), and more of non-US Turkers (26%) group of participants.

Our preliminary study shows no evidence that adviertising tasks in Turker communities can increase quality
From this preliminary result, surprisingly, we do not find re-posting leading to increased performance despite additional channel of exposure, possibly due to the effect hard to be detected by limited sample size. On the other hand, we do find Turkers trust HITs posted by members more than by requesters, which we will explore more in future studies.

For MTurk workers

Higher the participation, higher the income
Turkers vary greatly. The general trend is that higher participation is associated with higher income controlling for work hours. In other words, participating in the communities might increase the working condition of a Turker.

Join community to find good HITs, learn skills, and socialize
Confirmed by the interviews, Turkers self-organized into communities to find good HITs, learn skills to be good Turkers, and socialize with each other. It's interesting to see that less than 10% of participants identified "unionize together" as their primary motivation. Because underpayment has been discovered to be one of the biggest issue in MTurk [12, 13], we hypothesized that Turkers may join to negotiate their wages with requesters and MTurk site owners. But this hypothesis was not confirmed by the survey responses. Admittedly, the framing of the survey question might have shaped the responses to a certain degree as the understanding of "union" may vary across the broad cultural swatch of Turkers. Interestingly, although Turkers did not explicitly choose "unionize" as their primary motivation, participation was indeed associated with higher income.

Limitations
In the paper, the measurement we used - the number of answers to the ideation question and how frequently each answer appeared - was a rough proxy of task quality. In future work, we plan to use other metrics (e.g. expert rating and peer rating) to capture work quality more accurately.

Also, the correlation nature of this study prohibits us from drawing causal conclusions. However, it still suggests some evidence that participation in Turker communities may be beneficial for the workers.

Future Directions
In the future we intend to further understand the self-organization phenomenon by using a larger sample of participants, which will help reducing the skew caused by "super Turkers". We would also benefit from conducting more interviews with Turker with a variety of positions (founders, ordinary community members, the non-community members, and those who opposed the communities). Interviewing with Turkers from various positions might help us to better understand their opinions on the self-organization phenomenon. Furthermore, we plan to follow up with the participants who chose "unionize together" as their motivation to join the Turker communities. Last but not least, we also want to learn more about how exactly Turkers use communities to find good HITs, such as using automatic scripts. This might lead to some design implications for crowdsourcing platforms in general to improve on.

References

[1] Benjamin B Bederson and Alexander J Quinn. 2011. Web workers unite! addressing challenges of online laborers. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, 97–106.

[2] Adam J Berinsky, Gregory A Huber, and Gabriel S Lenz. 2012. Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis* 20, 3 (2012), 351–368.

[3] Daren C Brabham. 2010. Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society* 13, 8 (2010), 1122–1145.

[4] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science* 6, 1 (2011), 3–5.

[5] Anhai Doan, Raghu Ramakrishnan, and Alon Y Halevy. 2011. Crowdsourcing systems on the worldwide web. *Commun. ACM* 54, 4 (2011), 86–96.

[6] Enrique Estellés-Arolas and Fernando González-Ladrón-De-Guevara. 2012. Towards an integrated crowdsourcing definition. *Journal of Information science* 38, 2 (2012), 189–200.

[7] Jeff Howe. 2006. The rise of crowdsourcing. *Wired magazine* 14, 6 (2006), 1–4.

[8] Lilly C Irani and M Silberman. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 611–620.

[9] Aniket Kittur, Ed H Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 453–456.

[10] John Le, Andy Edmonds, Vaughn Hester, and Lukas Biewald. 2010. Ensuring quality in crowdsourced search relevance evaluation: The effects of training question distribution. In *SIGIR 2010 workshop on crowdsourcing for search evaluation*. 21–26.

[11] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. *Judgment and Decision making* 5, 5 (2010), 411–419.

[12] M Silberman, Joel Ross, Lilly Irani, and Bill Tomlinson. 2010b. Sellers' problems in human computation markets. In *Proceedings of the ACM SIGKDD workshop on Human Computation*. ACM, 18–21.

[13] M Six Silberman, Lilly Irani, and R Ross. 2010a. Ethics and tactics of professional crowdwork. XRDS 17, 2 (December 2010), 39-43. (2010).

[14] Gang Wang, Konark Gill, Manish Mohanlal, Haitao Zheng, and Ben Y Zhao. 2013. Wisdom in the social crowd: an analysis of quora. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 1341–1352.

[15] Ming Yin, Mary L Gray, Siddharth Suri, and Jennifer Wortman Vaughan. 2016. The communication network within the crowd. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1293–1303.

[16] Haiyi Zhu, Steven P Dow, Robert E Kraut, and Aniket Kittur. 2014. Reviewing versus doing: Learning and performance in crowd assessment. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1445–1455.