Unlocking the Potential of the Crowd by Challenging its Assumptions

Amy Rechkemmer

Purdue University
arechke@purdue.edu

Motivation

Microtask crowdsourcing holds great potential as a flexible working environment for both requesters and workers. Requesters are able to quickly and easily put out an open call for many different types of tasks, while workers hold greater control over their schedules and the types of tasks that they can complete. With this flexibility, however, a set of assumptions also exist of the scope of work that can be performed, how much this work should be worth, and how long the arrangement between workers and a requester will last as a result of the original intentions of these platforms. For instance, as microtasks are considered to be simple tasks relying primarily on basic human skills to complete, often times more complex tasks are not considered suitable for distribution among the crowd, completing tasks is thought to be able to be performed by anyone who can comprehend them, and pay is very little and given on a task-by-task basis. Holding these assumptions of crowd work not only limits the types of work that requesters can utilize the crowd for, but also shows a lack of understanding of the worker experience both by not accounting for known usability challenges of crowd platforms and by devaluing workers' contributions. All in all, the full power of the crowd may be constrained by its traditions.

Background

A joint effort was made to identify key challenges of crowd work that stop it from reaching its true potential as a part of the rapidly expanding digital economy, and this effort culminated in Kittur et al. (2013). In this work, the question was posed, *"Can we foresee a future crowd workplace in which we would want our children to participate?"*. As an answer, it is argued that crowd work has the potential to become "complex, collaborative, and sustainable", but only through a series of worker and requester considerations that are not inherent to the traditional crowd platform structure.

From there, prior literature has proposed solutions that help make these suggestions a reality. For instance, numerous approaches have been explored to help facilitate the use of microtask crowd workers with the purpose of completing complex tasks, including decomposing complex tasks into simpler tasks (Bernstein et al. 2010) and creating teams of workers to tackle complex tasks (Retelny et al. 2014). Outside of the content of tasks themselves, the traditional structure of payment for crowd workers was bucked by Whiting, Hugh, and Bernstein (2019), providing a system that can simplify the process of providing workers with a $15/hr wage for their work, although use of this system is at the discretion of the requester.

Diverging from features of the work itself, previous literature has also been done to explore making crowd work a more usable employment option for the broader population. This has included examining the needs of sub-groups of workers who may be at a greater disadvantage in using the platform than the average worker, including workers with disability (Zyskowski et al. 2015) and elderly workers (Brewer, Morris, and Piper 2016). The findings from this body of work consistently showcase the challenges that workers in these sub-groups face, both in terms of general navigation of the studied crowd platforms and in completion of their tasks.

Description of Proposed Research

Following Kittur et al. (2013), I ask if there is greater potential that exists for microtask crowd platforms in which we can see a future that we would want our friends and family to join the crowd, altogether seeking to create a more attractive and sustainable form of work on crowd platforms. To that end, I propose examining the impact of both exploring new solutions for assumptions challenged in prior literature and challenging *new* assumptions that have not been previously studied. Testing these assumptions has the potential to provide greater satisfaction for workers, but also to broaden the scope of work that requesters can have completed by the wisdom of the crowd.

Key research questions include:

- Will challenging the assumptions of traditional microtask crowdsourcing platforms increase worker satisfaction?

- Will challenging the assumptions of traditional microtask crowdsourcing platforms be successful in expanding the scope of work that requesters can utilize the platform to accomplish?

Planned Methodology

To explore these questions, I plan to run experiments testing crowdsourcing tasks and tools that run counter to the traditional assumptions of microtasks. In running these experiments, I will explore both whether and how these expanded assumptions contribute to the satisfaction that recruited crowd workers have with the task at hand, and also whether these changes are effective for requesters to utilize, opening up new possibilities for tasks that can be distributively performed by the crowd. To measure worker satisfaction, self-reported worker perceptions will be used. To measure effectiveness of the new task parameters for requesters, task accuracy will be used. Worker retention and engagement will be measured to reflect both worker satisfaction and potential for requester benefit.

Proposed Experiments and Findings

As of now, there are four proposed directions of assumptions to be challenged that I believe have potential for exploration. Some of these directions consist of new solutions for previously challenged assumptions, and some are new assumptions not covered in prior literature. For each of the four directions, the benefit of challenging the notion for requesters and why the change may promote greater worker satisfaction will first be explained, followed by a discussion of current findings (if applicable) and proposed experiments in the given direction.

Greater Completion of Complex Tasks Under the assumption that crowd platforms geared towards microtasks can only contain simple tasks, requesters are limited in the work that they can receive for tasks that are more complicated to perform or require more sophisticated domain knowledge in order to complete. Therefore, by allowing for complex tasks to be more easily performed on these platforms, a greater variety of work can be requested. From the worker perspective, the capacity for complex tasks on crowdsourcing platforms can not only create for more interesting work content, but can also help promote learning and honing new knowledge and skills, two factors that may lead to greater workplace satisfaction. Here, I discuss a new solution for this previously challenged assumption.

I have explored training microtask crowd workers to complete a complex task in a three phase study conducted using Amazon Mechanical Turk (Rechkemmer and Yin 2020). To serve as a motivational strategy for increased learning and performance, the theory of goal setting was utilized to encourage recruited workers to have a goal when completing training for the complex task. From the requester standpoint, we found that goal setting was able to be an effective strategy for helping workers learn to complete complex tasks under specific conditions. Namely, matching a goal to a worker's goal orientation and having a worker decide to complete a high amount of practice were able to positively affect how much workers were found to have learned from the training materials, and even improve performance at a future date in certain cases. The results of this experiment provide further evidence that complex tasks are viable for microtask crowdsourcing platforms, although we find that extra care needs to be taken to receive greater benefits through goal setting.

From a worker satisfaction perspective, we found that our workers with a learning goal (a goal specifying the type of knowledge that they wished to gain while engaging with the training material) believed that they learned more and found the training material to be more helpful than workers with the other types of goals. Altogether regardless of their goal treatment, we found the retention of our workers to be quite high for the three phases of the experiment that we ran: 88% of our workers returned for Phase 2 two days after Phase 1, and 85% of our workers returned for Phase 3 one week after Phase 1. Suggested by the low churn of workers, there is the prospect for crowd workers to find greater enjoyment from completing complex tasks and the training that may be involved in equipping workers to complete such tasks.

Expanding the Crowd through Greater Usability Another existing assumption is that the only abilities required by workers to complete microtasks involve innate human knowledge. As many microtask crowdsourcing platforms currently exist, along with other usability issues that have been discovered, performing tasks and receiving payment can be a challenge for workers with a variety of disabilities. Although these issues do not affect all workers, bringing inclusivity to the use of crowd platforms helps crowd work become not only a more widespread opportunity for employment, but also helps provide requesters access to workers with a more diverse range of perspectives, experiences, and abilities. To that end, I further explored this previously challenged assumption by first examining its impact in order to identify solutions.

To begin our exploration towards greater usability, we conducted a study surveying Amazon Mechanical Turk workers self-identifying as having a disability about their experiences using the platform. Although we found that workers with disability are plagued with many of the same issues of the platform as workers without disability, their condition may disproportionately heighten the impact of these issues for them. For instance, many of our workers surveyed find that their condition causes them to take more time to complete tasks than the average worker, resulting in them being able to complete less tasks (and receive less payment) in the same time span, and even causing them to lose out on work altogether if its HIT timer expires.

A follow-up experiment that we would like to run based off of these findings would be the impact of a tool that can help requesters provide more appropriate durations for their tasks. This tool would keep track of how long workers are taking to perform a task that a requester has posted, sending reports to the requester with this information and also feedback from workers, including whether or not they believe that they typically need longer to complete a task than the average worker and allowing them the opportunity to disclose a condition or any other information. This information could help inform requesters of how long they should set time limits for future batches of that task, as well as help inform timing for future HITs that they may post to the platform.

Promoting Work-Life Balance for Workers Through our survey of workers with disability, we found that maintaining a work-life balance can be difficult to achieve for crowd workers, especially those who greatly depend on the money they earn through these means to pay for essential items. Even for those who do not have the same need for the money they earn from crowd platforms, stepping away can be difficult when workers are paid on a task-by-task basis and they utilize external queues that are always picking up newly available work. In line with this idea, an assumption of crowdsourcing platforms is that the actions taken by workers outside of tasks have no contribution to their work, creating the belief that there is no benefit to fully removing one's self from work. However, creating this balance and learning when to rest between tasks at appropriate times has the potential to combat fatigue and improve mental well-being for workers. I believe it may also improve long-term retention to the benefit of requesters, both in terms of how long workers will continue to work on a batch of tasks at one time and how long they will continue to use the crowd platform as a whole.

As such, I believe it would be worth exploring attributes of workers that are non-intrusive and can be used as predictors of when workers may need a break from tasks. Possible attributes could include the time spent working since their last break, the rate of mouse clicks, tabs that are currently open, and self-reports from workers. Compiling these attributes could be used to produce a model that can help workers take breaks at appropriate times that will be beneficial both for their own well-being and their productivity. Through an experiment comparing workers using a model for deciding breaks to randomly decided breaks (or no breaks), we can collect both self-reports of workers in terms of their state of mental exhaustion at various intervals and retention of these workers over different time periods.

Extending the Work Contract Finally, there is the assumption that workers should be hired by requesters solely on a task-by-task basis rather than making an informal contract to work together for a specified longer term. The promise of long-term work between a requester and a set of workers could not only help provide workers with a sense of job security, another factor that may contribute to greater workplace satisfaction, but also help guarantee for requesters that they have a set of workers that are experienced with the task at hand. Given the results of our goal setting experiment for complex tasks, training a set of workers has the potential to provide benefits in the long-run if those workers can be recruited again and again.

Challenging this assumption, I propose an experiment to explore whether informing workers up front about the potential of long-term, consistent work can lead to greater satisfaction with the work and retention of workers over this period of time. This can be studied by comparing the self-reported perceptions and retention of two groups of workers: one group that is provided work at consistent time intervals but no advance notice, and another group provided work at the same time intervals, but informed from the beginning of the given consistent opportunities.

Research Challenges

One of the challenges that I foresee is in being able to recruit enough crowd workers that we can get meaningful use and feedback from when it comes to using the tools that have been described here. It is no secret that recruiting a sizeable number of interview candidates can be tricky, and even in a long-form survey that we have already conducted in which we paid respondents a $15/hr wage for their time, we only ended up recruiting 31 respondents. An additional issue may be if the attributes that we attempt to use to predict when workers should take a break are not enough to produce a meaningful model. A similar study has been done and was tested on software developers at a large company, but many of the features attained are not feasible to collect from a widespread sample of workers and are considered intrusive (e.g., webcam data, heart rate, etc.). Aside from thoughts on these potential challenges, all around I would appreciate feedback on the proposed research itself, including its potential for impact and new potential directions. I also wonder if aligning this work with an existing framework of factors impacting satisfaction in the workplace would lend credibility to my proposed ideas.

References

Bernstein, M. S.; Little, G.; Miller, R. C.; Hartmann, B.; Ackerman, M. S.; Karger, D. R.; Crowell, D.; and Panovich, K. 2010. Soylent: a word processor with a crowd inside. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, 313–322. ACM.

Brewer, R.; Morris, M. R.; and Piper, A. M. 2016. "Why would anybody do this?": Older Adults' Understanding of and Experiences with Crowd Work. *Proceedings of CHI 2016* .

Kittur, A.; Nickerson, J. V.; Bernstein, M.; Gerber, E.; Shaw, A.; Zimmerman, J.; Lease, M.; and Horton, J. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 1301–1318.

Rechkemmer, A.; and Yin, M. 2020. Motivating Novice Crowd Workers through Goal Setting: An Investigation into the Effects on Complex Crowdsourcing Task Training. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 8, 122–131.

Retelny, D.; Robaszkiewicz, S.; To, A.; Lasecki, W. S.; Patel, J.; Rahmati, N.; Doshi, T.; Valentine, M.; and Bernstein, M. S. 2014. Expert crowdsourcing with flash teams. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 75–85. ACM.

Whiting, M. E.; Hugh, G.; and Bernstein, M. S. 2019. Fair work: Crowd work minimum wage with one line of code. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 7, 197–206.

Zyskowski, K.; Morris, M. R.; Bigham, J. P.; Gray, M. L.; and Kane, S. K. 2015. Accessible crowdwork? Understanding the value in and challenge of microtask employment for people with disabilities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1682–1693.