

Ethics and Tactics of Professional Crowdwork

Paid crowd workers are not just an API call—but all too often, they are treated like one.

By M. Six Silberman, Lilly Irani, and Joel Ross

DOI: 10.1145/1869086.1869100

Faster, cheaper, smarter, and more efficient. These words might bring to mind the latest Intel ad, Moore’s law, or hopes for cell phone processors—silicon, copper, and computation. These circuits, however, are not only embodied in semiconductors. Increasingly, masses of people sit at their keyboards computing the answers to questions artificial intelligence cannot. Programmers access these computing crowds through APIs or a GUI on a service such as Amazon Mechanical Turk (AMT). Working for a couple of dollars an hour, these anonymous computing workers may never meet the programmers who use them as part of their research and engineering efforts. Who are these mysterious workers? What kind of relationship do they have with the engineers who use them as human computation? How should we, as computing researchers, conceptualize the role of these people who we ask to power our computing?

In this article we discuss findings which suggest that these questions are increasingly important for those of us building the collection of technologies, practices, and concepts called human computation. We hope however that it will be understood as not only about human computation. Rather we hope to link the thus-far mainly technical conversations in human computation to discussions of engineering ethics that have gone on for at least forty years (see, e.g., Florman’s *Existential Pleasures of Engineering* and

Papanek’s *Design for the Real World*). Here we offer insight into the practical problems crowdworkers face to ground these discussions in the current conditions of human computation.

Our research has focused on Mechanical Turk, a web platform that allows people (“requesters”) to post information tasks called Human Intelligence Tasks (“HITs”) for completion by other people (“workers” or “Turkers”), usually for a fee between one cent and a few dollars. Many businesses with large amounts of data use Me-

chanical Turk to create metadata and remove duplicate entries from their databases. Audio transcription and moderation of user-generated content on “Web 2.0” sites are other popular applications (see **Figures 1 and 2**).

After a worker submits a HIT, the requester can decide to “accept” the work and pay the worker, or “reject” it and keep the work for free. The site keeps track of how often workers’ submissions are accepted and rejected, and requesters use these rates to screen workers. When requesters re-

ject work, they hurt workers' ability to get more work (especially more highly paid work) in the future. The frequencies with which requesters accept and reject work, however, are not made available to workers. This information asymmetry underlies many of the difficulties we discuss in this article.

With few exceptions, human computation research has focused on problems facing the requesters of human computation, and most investigations of workers have aimed to motivate better, cheaper, and faster worker performance. This makes sense sociologically: most researchers are requesters. Put simply, the requester's problem is to get good data from workers, quickly and without paying much. Workers, however, also have interesting and difficult practical problems.

Our last year studying Mechanical Turk from a worker point of view [1, 2] offers insights into opportunities for human computation researchers to think more broadly about the people who are crucial to the systems they build. We summarize the results of demographic studies of workers in Mechanical Turk and describe some of the problems faced by Turkers, as some workers call themselves. We present several projects, including one we built, that approach some of these problems. Finally, we explore open questions of interest to workers, requesters, and researchers.

THE CROWD AND ITS PROBLEMS

"I don't care about the penny I didn't earn for knowing the difference between an apple and a giraffe. I'm angry that AMT will take requesters' money but not manage, oversee, or mediate the problems and injustices on their site."

—An anonymous worker

Abstraction hides detail. The very abstraction that lets human computation researchers access thousands of workers in a click also renders invisible the practical problems faced by people in the crowdworking workforce. A number of surveys and active web forums offer glimpses behind the curtain where "artificial artificial intelligence" is made.

WORKING HARD FOR THE MONEY

The Mechanical Turk labor pool hosts a growing international population

Figure 1: The Mechanical Turk interface for HIT selection lets workers choose tasks to complete, ranging from audio transcription to content moderation.

Search for article using specific keyword		View a HIT in this group	
Requester: Roy Sencio	HIT Expiration Date: Aug 9, 2009 (5 hours 19 minutes)	Reward: \$0.05	
	Time Allotted: 60 seconds	HITs Available: 1	

Find an image for this MIT person		View a HIT in this group	
Requester: Timothy Kay	HIT Expiration Date: Aug 9, 2009 (1 day 2 hours)	Reward: \$0.02	
	Time Allotted: 60 seconds	HITs Available: 642	

Value of Human Life		View a HIT in this group	
Requester: Craig W French	HIT Expiration Date: Sep 12, 2009 (4 weeks 6 days)	Reward: \$0.01	
	Time Allotted: 60 seconds	HITs Available: 1	

Figure 2: Once a worker has selected a HIT, she is given instructions and a task to complete within a time limit.

Write an article in 5 different languages and with minimum 800 words			
Requester: James Wander	Reward: \$2.50 per HIT	HITs Available: 1	Duration: 2 weeks 4 days
Qualifications Required: Location is US			

Write an article in 5 different languages and 800 words long

Follow the instructions:

- Write an unique article about the influence of Einsteins Theory of Relativity on Mozambique's energy spendings
- Write this article in the languages: English, Jewish, Arabique, Chinese and Russian
- The articles have to be unique and not found elsewhere on the internet. I will check that
- It is NOT allowed to write the article in english and translate with Google translate or another tool to the other languages. I will check the articles being well written all languages!
- Each of the 5 articles must be minnum 800 words. Will also be checked.

Copy and paste the 5 articles in the box below

earning less than \$10,000 per year, some of whom rely on Turking income to make basic ends meet. Ross et al. [2], extending work by Ipeirotis [3], present longitudinal demographic data on Mechanical Turk workers.

While Indian residents made up only 5 percent of respondents to a November 2008 survey, they comprised 36 percent of respondents to a November 2009 survey and 46 percent in February 2010, at which point American Turkers, formerly the majority, comprised only 39 percent of survey respondents. Many of these new Indian Turkers are young men earning less than \$10,000 a year. Almost a third of Indian Turkers surveyed reported that they always or sometimes relied on their Turking income to "make basic ends meet." Between May 2009 and February 2010, the fraction of U.S. Turkers surveyed reporting reliance held steady at 13±1 percent.

Many Turkers see themselves as laborers doing work to earn money. In

survey data collected in February 2009 ($n=878$), the most commonly reported motivation for doing HITs was payment: 91 percent of respondents mentioned a desire to make money. Turking to pass the time, in contrast, was mentioned by only 42 percent of respondents. February 2010 data ($n=1,000$) from Ipeirotis confirms the importance of money compared to other motivations, with most respondents reporting they do not do HITs for fun or to kill time. 25 percent of Indian respondents and 13 percent of U.S. respondents reported that Mechanical Turk is their primary source of income.

OCCUPATIONAL HAZARDS

What challenges face these professional crowdworkers? Several researchers have engaged workers by posting open-ended questions to Mechanical Turk—a sort of online interview to access a generally invisible population and see the world from their perspective. We have also conducted interviews of

workers through Skype and participated in the forums where they share tips, talk about work, and virtually meet their coworkers. Turkers often advise one another on the occupational hazards of human computing:

Employers who don't pay: When workers submit work to employers through Mechanical Turk, they have no guarantee of receiving payment for their work. The site terms state that employers "pay only when [they're] satisfied with the results."

While this makes Mechanical Turk highly attractive to employers it leaves workers vulnerable to the whims of employers—or, just as likely, employers' evaluation software—judging the merit of their work. The amount of work often makes it impractical for employers to evaluate manually. Because employers hire hundreds or more workers at a time, they puzzle rejected workers with generic messages giving reasons for rejection, if they explain their decision at all. At worst, ill-intentioned employers post large batches of tasks with high pay, receive the work, and reject it as a way of obtaining free work. Such rejected work leaves workers feeling vulnerable, reduces their effective wage, and lowers their work acceptance rate.

Staying safe online: Mechanical Turk workers have to learn to identify illegitimate tasks to stay safe online. Administrator spamgirl on Turker Nation, a forum for workers, outlines tasks to avoid:

Do not do any HITs that involve: filling in CAPTCHAs; secret shopping; test our web page; test zip code; free trial; click my link; surveys or quizzes (unless the requester is listed with a smiley in the Hall of Fame/Shame); anything that involves sending a text message; or basically anything that asks for any personal information at all—even your zip code. If you feel in your gut it's not on the level, IT'S NOT. Why? Because they are scams...

The discussion that ensued identified malware, sale of personal information and wage theft as risks workers face choosing among jobs.

"Why is there no control?" Hit by several of the problems described above, a4x401 offered a newcomer's frustrated perspective with the worker side of human computation:

Being a newbie and having relatively decent PC skills, I have been checking all this stuff out and am somewhat upset [about] the things that I have discovered! It's no wonder that people don't trust the requesters, yes I did some of those HITs that one should not do and found myself having to repair my PC and remove some pop-ups. After having done that I really got into checking out the program and realized that it's too easy to manipulate it due to the fact that work can be rejected after it's finished but the work is still done. All [a requester] has to say is "not to our satisfaction"!!!! The other way is to just leave the HITs open; you still collect your work but don't have to pay! My favorite part is HITs that are way too complicated to complete in the time frame allowed! Why is there no control on any of this stuff? [Edited for punctuation and spelling.]

He echoes experiences many report on worker forums and in research surveys. Workers report trying to contact Amazon staff but receiving little response.

Costs of requester and administrator errors are often borne by workers: When a requester posts a buggy task or a task with inadequate instructions, they often don't get the responses they want from workers and reject the work. One worker wrote:

I would like to see the ability to return a HIT as defective so it dings the requester's reputation and not mine. Let's face it, if I'm supposed to find an item for sale on Amazon but they show me a child's crayon drawing...there really needs to be a way to handle that without it altering my numbers.

Similarly, occasionally requesters will post a task with a prohibitively short time limit, and the task expires before workers can complete it. This lowers workers' effective wage and affects the worker's reputation statistics rather than the requester's.

At present, largely owing to request-

er and administrator unresponsiveness, workers can do little to improve the conditions of their tasks. Unsurprisingly, some have expressed interest in a more relationship-oriented approach to distributing work. One Turker wrote:

We the Turks, in a world that requires productivity in working together, will work honestly and diligently to perform the best work we can. The Requestors, in turn, will provide useful work and will pay us fairly and quickly, providing bonuses for especially good work. The goal is to create a working environment that benefits us all and will allow us the dignity and motivation to continue working together.

These are some of the more prevalent phenomena we've encountered. For more, see turkernation.com and mturkforum.com, especially their "Suggestions" boards.

APPROACHES TO WORKER PROBLEMS

Software tools exist, some built by Turkers, that attempt to help Turkers manage these problems. Many are client-side scripts that add functionality to the Mechanical Turk interface. At least one platform aims to compete with Mechanical Turk.

Augmenting Mechanical Turk from the outside: Workers and requesters have made a number of Turking tools, including a list of all requesters, a script for recording your own worker history (not preserved by Mechanical Turk, but useful for tax purposes), and a client-side script to hide HITs posted by particular requesters.

Motivated by the problems above, we built Turkopticon (turkopticon.differenceengines.com) in 2008, a database-backed Firefox add-on that augments Mechanical Turk's HIT listing. The extension adds worker-written reviews of requesters to the interface (see **Figures 3, 4, and 5**); the next version will compute effective wage data for HITs and requesters. Some Turkers have been enthusiastic about Turkopticon. One early adopter posted to Turker Nation, "if you do not have this, please get it!!!! it does work and is worth it!!!"

This was a proud moment for us,

Figure 3: Turkopticon augments the MTurk interface with Turker feedback using a Firefox extension.



and we have attempted to respond to feature requests and provide support for new users. Turkopticon users have contributed over 7000 reviews of over 3000 requesters, but the user base has remained very small, especially compared to the total number of Turkers. Relatedly, some Turkers have pointed out that a third-party review database is no substitute for a robust, built-in requester reputation system.

Building alternative human computation platforms: CloudCrowd, launched in September 2009, aims to provide a “worker-friendly” alternative to Mechanical Turk. In a post to mTurk Forum, CEO Alex Edelstein writes that CloudCrowd will offer “a more efficient [worker] interface,” payment through PayPal (allowing workers to collect currencies other than USD and INR, the only choices for Turkers), and “credibility” ratings (in place of acceptance rates as in Mechanical Turk) as the measure of worker quality.

Kochhar et al., in a paper at HCOMP 2010 [4], documented the success of a relationship-oriented approach to distributing work in the design of a “closed” large-scale human computation platform.

Offering workers legal protections: Alek Felstiner has raised the question of legal protections for crowdworkers, asking, “what [legal] responsibilities, if any, attach to the companies that develop, market, and run online crowdsourcing venues?” In his working paper [5] he explores the difficulties that arise in the application of traditional employment and labor law to human computation markets.

Figure 4: Turkopticon is a feedback platform for Turkers. Turkers can rate requesters on communicativity, generosity, fairness and promptness.

Channel Intelligence A2YC4ECK3MFH4R HIT Group »	COMM: 4.09 / 5 FAIR: 4.57 / 5 PAY: 4.17 / 5 FAST: 4.83 / 5
Amazon Requester Inc. (Product Ads) AALZYJEWCVBZB HIT Group »	COMM: 2.00 / 5 FAIR: 2.05 / 5 PAY: 3.20 / 5 FAST: 3.14 / 5
kelsie greear A1R5Y8Z987CERT HIT Group »	COMM: 3.36 / 5 FAIR: 4.27 / 5 PAY: 3.86 / 5 FAST: 4.23 / 5
mtlabel A3VP3B2IX7IZPH HIT Group »	COMM: 4.33 / 5 FAIR: 4.29 / 5 PAY: 3.33 / 5 FAST: 4.43 / 5
Product Search A28JEQTWH76JDT HIT Group »	COMM: 1.92 / 5 FAIR: 2.48 / 5 PAY: 2.43 / 5 FAST: 2.38 / 5
Raphael Mudge A29XDCTJMAE5RU HIT Group »	COMM: 3.30 / 5 FAIR: 3.70 / 5 PAY: 4.10 / 5 FAST: 4.39 / 5

Figure 5: Turkers can provide free text feedback to explain their ratings.

Dolores Labs A2IR7ETVOIULZU HIT Group »	I usually really like their HITs, but I think .01 for this HIT is ridiculous. We should not do their HITs that are this low paying so that the improve the prices. Be fair, and this is not! POSTED BY <NABO...@G...>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 3 / 5 COMM: 5 / 5
Jason Gurwin A8LIB2HP9GRSE HIT Group »	Random rejects. Lots of random rejects. POSTED BY <XJU...@M...>	FAIR: 1 / 5 FAST: 1 / 5 PAY: 1 / 5 COMM: 1 / 5
Jason Gurwin A8LIB2HP9GRSE HIT Group »	Had about 25% of these hits rejected because I "Did not follow directions". The rest were approved and paid and I did nothing different with the rejected ones. ::Shrugs shoulders:: POSTED BY <SMI...@C...>	FAIR: 1 / 5 FAST: 1 / 5 PAY: 1 / 5 COMM: 1 / 5

OPEN QUESTIONS

The projects listed above are tentative steps toward addressing the problems facing Turkers and developing a richer understanding of the structure and dynamics of human computation markets. Many questions remain, including: How does database, interface, and interaction design influence individual outcomes and market equilibria?

For example, how would the worker experience on Mechanical Turk be different if workers knew requesters' rejection rates, or the effective wages of HITs? This has been explored in online auctions, especially eBay, but only tentatively in human computation (e.g., [6], which examines task search).

Another question is: What are the economics of fraudulent tasks (scamming and spamming)?

That is, how do scammers and spammers make money on Mechanical Turk, and how much money do they make? Work in this thread might draw on existing research on the economics of internet fraud (e.g., [7]) and could yield insights to help make human computation markets less hospitable to fraudsters.

A third question is: What decision logics are used by buyers and sellers in human computation markets?

We might expect workers to minimize time spent securing payment on each task, even if this means providing work they know is of low quality. Some workers do behave this way. We have found, however, that workers seem more concerned with what is "fair" and "reasonable" than with maximizing personal earnings at requester expense. The selfish optimizers that populate the models of economic decision-making may not well describe these "honest" workers, although as noted in [8] they can perhaps be extended to do so. So how do differently motivated actors in human computation markets shape market outcomes, and how can this knowledge shape design?

Finally, we can ask: What's fair in paid crowdsourcing?

Economists Akerlof and Shiller, in their 2009 book *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*, argue that "considerations of fairness are a major motivator in many economic

decisions" that has been overlooked in neoclassical explanations that assume people act rationally: "while...there is a considerable literature on what is fair or unfair, there is also a tradition that such considerations should take second place in the explanation of economic events" (pp. 20, 25).

At public events we have heard Mechanical Turk requesters and administrators say tasks should be priced "fairly," but fairness is difficult to define and thus to operationalize. The concept of a reservation wage—the lowest wage a worker will take for a given task—as discussed in [9] is useful but not definitive: the global reach of human computation platforms complicates the social and cultural interpretation of the reservation wage.

The question of fairness links interface design to market outcomes. If considerations of fairness are key to explaining economic decision making, but fairness is constructed and interpreted through social interaction, then to understand economic outcomes in human computation systems we need an understanding of these systems as social environments. Can systems with sparse social cues motivate fair interactions? Human computation and Computer Supported Cooperative Work may have much to learn from one another on these topics.

LOOKING FORWARD

This review of workers' problems should not be mistaken as an argument that workers would be better off without Mechanical Turk. An exchange in late 2009 on the Turker Nation forum makes the point concisely:

xeroblade: *I am worried that Amazon might just shut the service down because it's becoming full of spammers.*

jml: *Please don't say that :(*

With Mechanical Turk, Amazon has created work in a time of economic uncertainty for many. Our aim here is not to criticize the endeavor as a whole but to foreground complexities and articulate desiderata that have thus far been overlooked. Basic economic analysis tells us that if two parties transact they do so because it makes them both better off. But it tells us nothing about the conditions of the transaction. How did the parties come to a situation in

which such a transaction was an improvement? When transactions are conditioned by the intentional design of systems, we have the opportunity to examine those conditions.

Human computation has brought Taylorism—the "scientific management" of labor—to information work. If it continues to develop and grow, many of us "information workers" may become human computation workers. This is a selfish reason to examine design practices and workers' experiences in these systems. But the underlying question is simple: are we, as designers and administrators, creating contexts in which people will treat each other as human beings in a social relation? Or are we creating contexts in which they will be seduced by the economically convenient fiction alluded to by Mechanical Turk's tagline, "artificial artificial intelligence"—that is, that these people are machines and should be treated as such?

Biographies

M. Six Silberman is a field interpreter at the Bureau of Economic Interpretation. He studies the relation between environmental sustainability and human-computer interaction. His website is wtf.tw.

Lilly Irani is a PhD candidate in the Informatics department at University of California-Irvine. She works at the intersection of anthropology, science and technology studies, and computer supported cooperative work.

Joel Ross is a PhD candidate in the Informatics department at University of California-Irvine. He is currently designing games to encourage environmentally sustainable behavior.

References

1. Silberman, M. S., et al. Sellers' problems in human computation markets. In *Proceedings of HCOMP 2010*.
2. Ross, J., et al. Who are the crowdworkers? shifting demographics in Mechanical Turk. In *Proceedings of CHI 2010*: 2863-2872.
3. Ipeirotis, P. Mechanical Turk: the demographics. <http://behind-the-enemy-lines.blogspot.com/2008/03/mechanical-turk-demographics.html>.
4. Kochhar, S., et al. The anatomy of a large-scale human computation engine. In *Proceedings of HCOMP 2010*.
5. Felstiner, A. Working the crowd: employment and labor law in the crowdsourcing industry. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1593853.
6. Chilton, L., et al. Task search in a human computation market. In *Proceedings of HCOMP 2010*.
7. Franklin, J., et al. An inquiry into the nature and causes of the wealth of internet miscreants. In *Proceedings of CCS '07*: 375-388.
8. Jain, S. and D. Parkes. The role of game theory in human computation systems. In *Proceedings of HCOMP 2009*.
9. Horton, J. J. and L. Chilton. The labor economics of paid crowdsourcing. arXiv:1001.0627v1 [cs:HC], 2010.