Who are the Crowdworkers? Shifting Demographics in Mechanical Turk

Joel Ross

Department of Informatics
University of California, Irvine
Irvine, CA 92697-3440 USA
jwross@uci.edu

Lilly Irani

Department of Informatics
University of California, Irvine
Irvine, CA 92697-3440 USA
lirani@ics.uci.edu

M. Six Silberman

Department of Informatics
University of California, Irvine
Irvine, CA 92697-3440 USA
msilberm@uci.edu

Andrew Zaldivar

Department of Informatics
University of California, Irvine
Irvine, CA 92697-3440 USA
azaldiva@uci.com

Bill Tomlinson

Department of Informatics
University of California, Irvine
Irvine, CA 92697-3440 USA
wmt@uci.edu

Abstract

Amazon Mechanical Turk (MTurk) is a crowdsourcing system in which tasks are distributed to a population of thousands of anonymous workers for completion. This system is increasingly popular with researchers and developers. Here we extend previous studies of the demographics and usage behaviors of MTurk workers. We describe how the worker population has changed over time, shifting from a primarily moderate-income, U.S.-based workforce towards an increasingly international group with a significant population of young, well-educated Indian workers. This change in population points to how workers may treat Turking as a full-time job, which they rely on to make ends meet.

Keywords

Mechanical Turk, demographics, user surveys, crowdsourcing, human computation

ACM Classification Keywords

H5.3. Group and Organization Interfaces: Computer-supported cooperative work; H5.3 Group and Organization Interfaces: Web-based interaction.

General Terms

Human factors.

Introduction

Amazon Mechanical Turk[1] (MTurk) is an online crowdsourcing [9] system that allows users to distribute work to a large number of workers. This work is broken down into simple, one-time tasks that workers are paid to complete. Such tasks are frequently those that are difficult for computers and yet simple for humans (e.g., image labeling), creating a kind of "artificial artificial intelligence" [6]. Requesters create Human Intelligence Tasks, or HITs, specifying the amount paid for the job's completion—usually ranging from as little as $0.01[2] for a quick task up to a few dollars for more involved jobs, such as transcribing audio clips. Workers (or "Turkers") who log into the MTurk website are able to choose which tasks they perform (after previewing the HIT), creating a micro-task marketplace [13]. Launched in November 2005, Amazon reports that the system has now more than 400,000 workers registered, and there are about 50,000-100,000 HITs to work on at any given time.

The Mechanical Turk is a reference to Wolfgang von Kempelen's 18th century "automaton" that traveled through Europe, stunning audiences for 50 years with feats of mechanically calculated chess playing prowess. Underneath the table that supported the chessboard, the Turk's designer had hidden the secret to this computational wonder: a small grandmaster, pulling levers to control the Turk's mechanical hands. Hidden humans powered von Kempelen's wondrous machine; people, almost as invisibly, similarly power MTurk. In this paper, we describe the people who provide computational power that MTurk requesters rely on.

[1] Mechanical Turk can be accessed at http://www.mturk.com.

[2] All prices and monetary amounts are reported in U.S. dollars.

MTurk is increasingly popular as a tool for research, being used for performing user studies [8,13], image labeling [7,18], natural language processing [17], and relevance evaluation [4], as well as replicating classical economic and social science experiments [14]. As MTurk becomes a common research platform, it is important to understand the anonymous Turkers that enable these projects, in order to know how they may effect and be affected by the research performed. For example, obscuring worker identity may disassociate those users from requesters and potentially contribute to workers being exploited: because workers are decontextualized, requesters may be more likely to offer lower, unfair prices on HITs, or even refuse to pay for work performed [12]. Thus understanding Turkers is vital to the overall health of the system and its users, and may increase the effectiveness with which researchers can engage and interact with this population.

The demographics of MTurk workers have been previously reported informally by Ipeirotis [10,11]. This prior research describes a worker population based primarily in the U.S. and made up of workers who tended to be young, well-educated, and female, with moderately high incomes (more than 40% made more than $40,000 annually)—demographics that contrasted with the then-prevalent view of Turkers as low-income workers in developing countries. We update and extend this work with more recent data with additional factors.

Furthermore, by periodically repeating our surveys of the Turker population and combining the results with those presented by Ipeirotis, we can view the data longitudinally, looking at how the population has changed over almost two years (see Table 1)—nearly

Survey Date	Sample Size
Mar. 2008 [10]	n = 300
Nov. 2008 [11]	n = 1010
Feb. 2009	n = 878
May 2009	n = 512
Aug. 2009	n = 578
Nov. 2009	n = 733

Table 1. We consider 6 samplings of MTurk workers over a period of 20 months. Because of changes in sample selection method, the sample sizes are different for each survey. For example, the Ipeirotis surveys had a maximum number of respondents (gathered over as long as a month), as did the May 2009 survey (gathered over just a few days). On the other hand, the Feb. 2009 survey is a compilation of three consecutive weeklong surveys at increasing prices. The Aug. 2009 and Nov. 2009 surveys were each available for a full week before they were closed.

half the lifetime of the system. Thus we can begin to understand not only the make-up of Turkers, but also how that composition has changed as the system has matured. (Note that because we asked different questions than Ipeirotis, and because we refined and adjusted our questions between surveys, we do not have data for each demographic factor at each time sample. Nevertheless, we believe these data indicate trends in the changing worker population).

Based on this survey data, we argue that the MTurk workforce is shifting increasingly towards a more international population, of which a significant portion rely on the income earned through the site and treat Turking as a part- or full-time job. This change raises a number of ethical, analytical, and methods issues for HCI research involving such crowdsourcing systems.

Survey Method
Our surveys were delivered to users in the form of a HIT on the MTurk website—workers could choose to take the survey and be compensated through the system, as with any other task. Workers accepted the HIT and were redirected to a web-based survey, at the end of which they were given a confirmation code to enter back into MTurk. Note that because the external link to the survey was available in the HIT preview, workers were able to view (and even complete) the survey before actually accepting the HIT. For our surveys, workers were compensated $0.10, with median completion times of around 2 minutes.

Previous research [4,13,18] has described the potential unreliability of MTurk workers. While this literature has suggested the use of qualification pre-tests and/or explicitly verifiable questions, these were not

appropriate for the current study—qualification tests would exclude a portion of the Turker population, and there was no way of verifying user demographics. Indeed, although MTurk HITs default to being answerable only by users with at least a 95% approval rating (meaning that 95% of the worker's submitted HITs have been approved by the requester of the work), we removed all restrictions for accepting the surveys—allowing any worker to complete the HITs—in order to reach as broad a user population as possible.

Thus there is the possibility some respondents may have given purposefully false answers. Collecting responses through a HIT creates the possibility of demand characteristics (where subjects change behavior in response to being measured): workers may have shaped their responses based on what they believed we wanted to hear in order to be assured of being paid for their time. In addition, the method of delivering these surveys means that respondents were self-selecting—thus our results may be biased towards Turkers who enjoy taking surveys and are willing to provide information about themselves, rather than reflecting the worker population as a whole. Nevertheless, we believe these samples are still large and consistent enough to provide meaningful insight into the changing demographics of MTurk workers.

Results
Our combined surveys show that the Turker population is becoming increasingly international, particularly with workers from India making up a larger and larger percentage of workers (see Figure 1). While Ipeirotis reports that Indian workers were only a small minority even in Nov. 2008 (5%), our surveys suggest that a year later, Indian Turkers make up more than a third of

Figure 1. Nationality of MTurk workers over time. Countries with more than 1% of respondents include Canada, the U.K., and the Philippines. We do not have country data for Feb. 2009.
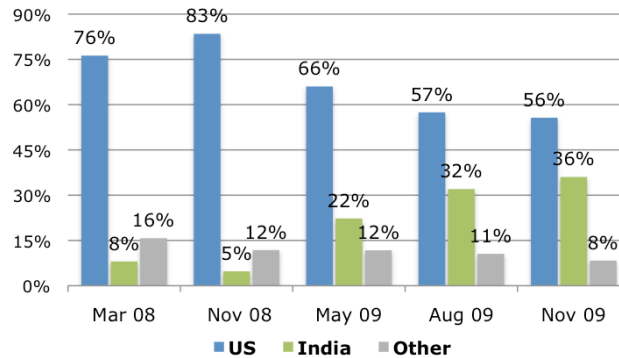

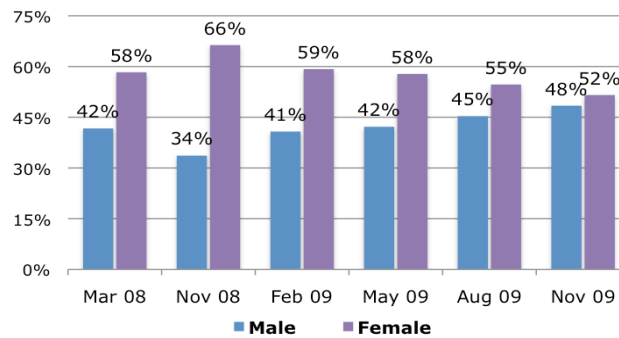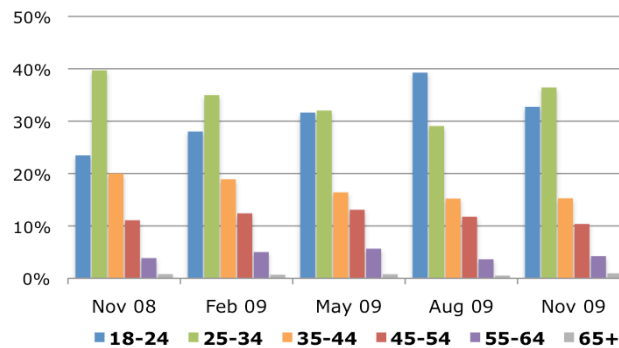**Figure 2.** Gender of MTurk workers over time.


Figure 3. Age of MTurk workers over time.

the worker population (36%). Similarly, our surveys show a fairly steady increase in the number of male workers (Figure 2)—a population changing from one that is 60% female to one that is more evenly split between the genders (52% female in Nov. 2009). Furthermore, the Turker population seems to be getting younger—the average age has dropped slightly from 32.9 in Nov. 2008 to 31.6 in Nov. 2009, particularly as the percentage of workers 18-24 increases (Figure 3).

Comparing these surveys also reveals a change in Turker annual incomes. Ipeirotis reports a median annual income in the $25-40,000 range, or even in the $40-60,000 range in Nov. 2008. He also reports a relatively small percentage of workers making less than $10,000 or $15,000 annually (16% and 10% in Mar. and Nov. 2008, respectively. See [10,11] for details). In these earlier surveys, Turkers on average had moderately-high incomes. However, in our more recent surveys, a larger and larger portion of MTurk workers report making less than $10,000 annually—indeed, nearly a third of respondents in Nov. 2009 (see Figure 4). Thus over time, the number of lower-income workers has increased, along with the number of young workers, male workers, and Indian workers.

These changing demographics all seem to stem from the increased global population of Turkers. Indian workers have consistently been younger than workers from the U.S., with an average age of around 26-28 years old (compared to 33-35 years old for U.S. Turkers; see Table 2). Indian workers are also substantially more likely to be male than U.S. Turkers—while around 2/3 of U.S. Turkers are female, 2/3 of Indian Turkers are male. Similarly, Indian workers have significantly lower annual incomes than U.S. workers

Figure 4. Above: reported annual household income of MTurk workers (in U.S. dollars). Note that due to survey adjustments, this question was not asked in the May 2009 survey.
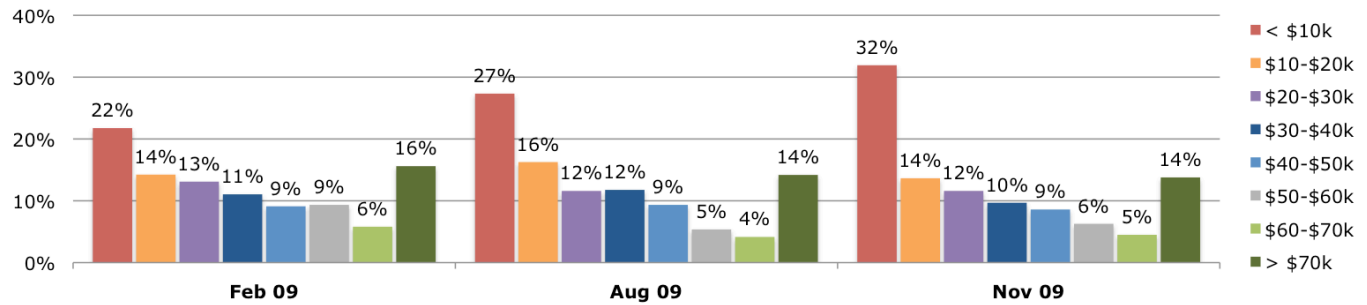
Below: reported annual household income by country, from Nov. 2009.
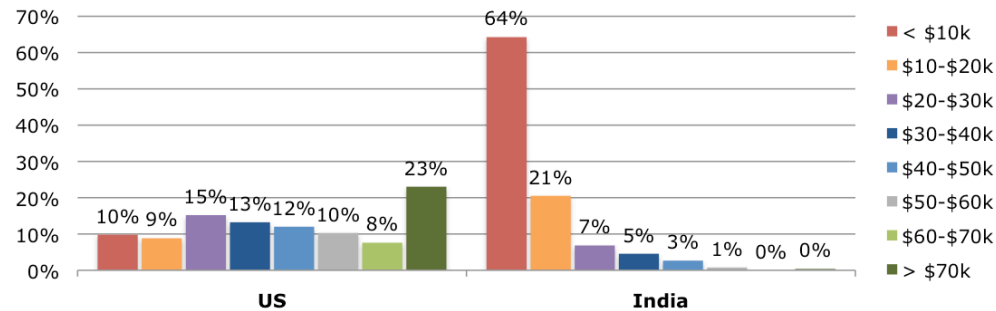
(Figure 4, lower), corresponding to the differences in the average wage and cost of living between the U.S. and India. Thus as Indian workers make up a larger percentage of the Turker population, the average values reported in demographics surveys are shifting. It is not that a larger number of young, male and/or low-income U.S. workers are now using the system, but that MTurk hosts a more and more international work-force.

There are other demographic differences between U.S. Turkers and Indian Turkers as well (also detailed in Table 2). Although younger on average, Indian workers report having higher levels of education. Indeed, the Turker population as a whole seems to be highly

educated—in Nov. 2009, overall 41% of Turkers reported having Bachelor degrees, and 18% report having Graduate degrees. Our later surveys have shown Indian workers are also slightly less likely to report themselves as being unemployed (around 25% unemployed vs. around 35% unemployed for U.S. workers), but further study is needed to see if this trend holds. It is possible that Indian workers may culturally be less willing to present themselves as unemployed, and/or consider their work on MTurk as part-time employment.

Indeed, Turkers from India are more likely to be reliant on the money they earn for doing work on MTurk (Figure 5). Around 27% of Indian workers report

Table 2. Comparison of select demographics between U.S. workers and Indian workers, over time.

		Nov 08	May 09	Aug 09	Nov 09
Average Age	*US*	33.6	34.3	33.2	35.4
	India	28.5	28.8	27.6	26.4
Gender	*US*	28% male, 72% female	34% male, 66% female	31% male, 69% female	37% male, 63% female
	India	75% male, 25% female	61% male, 39% female	69% male, 31% female	66% male, 34% female
Education	*US*	32% Bachelors, 11% Graduate	34% Bachelors, 14% Graduate	34% Bachelors, 19% Graduate	38% Bachelors, 17% Graduate
	India	69% Bachelors, 29% Graduate	56% Bachelors, 18% Graduate	56% Bachelors, 13% Graduate	45% Bachelors, 21% Graduate

themselves as sometimes or always requiring MTurk "to make basic ends meet", compared to around 14% of U.S. Turkers (overall, about 18% of Turkers sometimes or always rely on MTurk). Thus while only a minority of workers rely on the pay earned from completing HITs, they still make up a sizable percentage of the Turker population—a percentage that may be growing as MTurk gains an increasingly international user base and companies such as SamaSource (samasource.com) actively recruit low-income populations to MTurk.
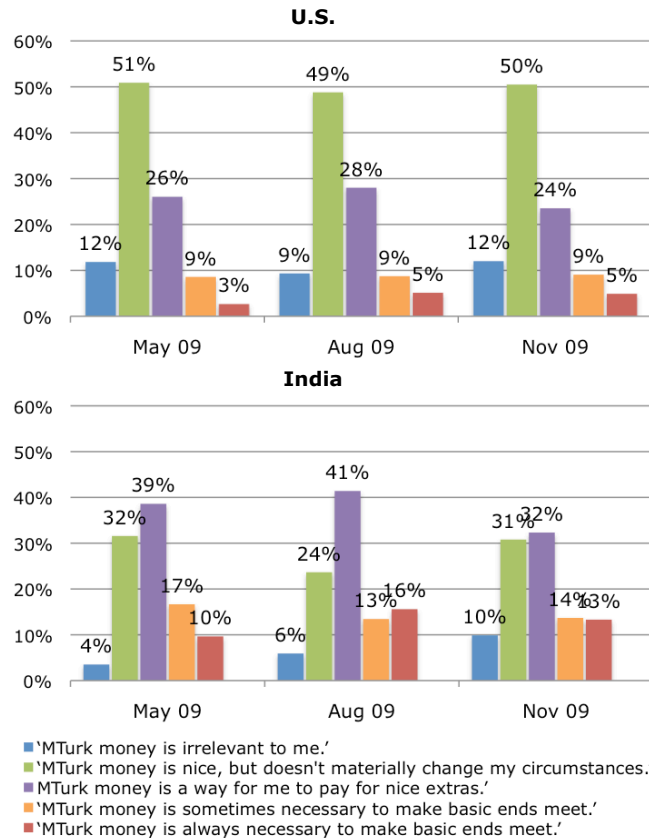
Finally, Figure 6 shows how much money workers report earning from MTurk and how much time they spend working in order to earn that money. On average, Turkers earn just under $2.00/hour, with Indian workers earning less than U.S. workers ($1.58/hour vs. $2.30/hour on average in Nov. 2009)— a discrepancy potentially due to the country restrictions on the best-paying HITs. So while the majority of workers spend less time Turking than many Americans do watching TV, a consistent 18% of workers spend more than 15 hours/week doing HITs. For these Turkers, MTurk functions as a part- or even full-time

job—a job that pays less than $2.00/hour. In fact, because Amazon's policies position workers as independent contractors, they are not guaranteed minimum wage for their work [1]. In this way, Amazon's platform supports broader trends in organizational management [5,16] of using freelance and part-time labor that can be hired and fired as company needs fluctuate.

Survey Limitations

There are a number of possible factors that may limit or bias the results of these surveys and their comparison. One such factor would be differences in method between the studies performed by Ipeirotis (the Mar. 2008 and Nov. 2008 datasets) and those we performed. For example, Ipeirotis would have presented the survey HITs somewhat differently—while we solicited Turkers to "Take a short research survey about Mechanical Turk workers", other surveys may have presented the task in a way that may have gathered different respondents. Furthermore, Ipeirotis appears to have paid a different reward for completion of the HIT than we did (he paid $0.01, we paid $0.10),

U.S.

India

- ■ 'MTurk money is irrelevant to me.'
- ■ 'MTurk money is nice, but doesn't materially change my circumstances.'
- ■ MTurk money is a way for me to pay for nice extras.'
- ■ 'MTurk money is sometimes necessary to make basic ends meet.'
- ■ 'MTurk money is always necessary to make basic ends meet.'

Figure 5. Reported reliance on the money earned from MTurk for U.S. and Indian workers, over time.

Figure 6. Overall hours spent doing HITs and amount earned from MTurk per week, over time. The *average hourly wage* (in bold) was calculated at the middle value of each bin for each respondent.

which again may have led to different populations taking the surveys. Nevertheless, we feel that that because these trends can also be seen—if perhaps less dramatically—in the latest surveys (May, Aug., and Nov. 2009), which followed a near identical method. Thus we feel that the fact that earlier surveys, designed slightly differently, provide results that are in line with our analysis, actually helps to validate our conclusions.
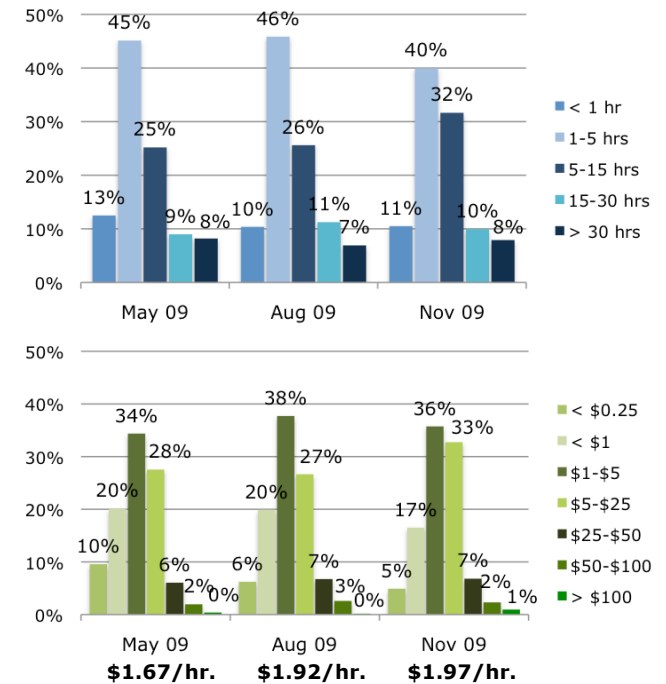
Still, our method of using HITs to gather demographic information may have lead to some level of self-selection—Turkers who often completed surveys may have been more likely to respond to ours. Indeed, when we asked Turkers what types of HITs they work on, 54.2% of the Nov. 2009 respondents completed surveys more often than other kinds of tasks.

However, a set of one-way ANOVAs comparing each

demographic factor of workers who mostly complete surveys and workers who mostly perform other tasks finds that this difference is not significant across most considered dimensions. Age, gender, education, and employment all have similar distributions, and reliance is only slight significant ($p < 0.05$). But Turkers who answer surveys are significantly more likely to be from the U.S. ($p < 0.005$) and to have higher reported annual income ($p < 0.005$). The differences in nationality may come from text-based surveys being more difficult for non-native speakers, and the income differences is likely a result of surveys often being less lucrative on MTurk—workers looking primarily to maximize their income are likely to find other HITs with a higher pay rate. Yet because our surveys paid more than reported average earnings ($3.00/hr. vs. about $2.00/hr.), we likely attracted a wider sample population. Thus we believe our results are not significantly driven by self-selection factors.

Discussion

Overall, this series of surveys suggests that the MTurk worker population has become significantly more international over the past 20 months, with many more workers from India. These Indian Turkers are mostly young, well-educated males—demographics that put them in contrast with some of the common personas of MTurk workers. Early surveys such as those by Ipeirotis presented a worker population inhabited by: stay-at-home moms who want to supplement the household income; office workers Turking during their coffee breaks; college students making a few dollars while playing a game; and recession-hit Turkers doing what they can to make ends meet.

Yet our surveys show an increase in young, highly-educated Indian males—a different persona that may change how we as researchers and developers view the crowd that is completing these HITs. For these workers, MTurk may increasingly function as a passable (and even greatly needed) part- or full-time job. Indeed, a significant minority of Turkers spend enough time doing HITs for MTurk to act as a form of employment, with an equal percentage relying on the income earned from MTurk to make ends meet. Nevertheless the average wage earned by Indian Turkers is still only about 75% of the Indian minimum wage, indicating potential concerns about fairness of compensation and payment for an activity firmly positioned as "work". Indeed, as workers come to spend more time performing this cognitive labor—and rely on the payment for completing that labor—a number of issues are raised for performing HCI research using these systems.

Labor and Livelihood in HCI

For example, the significant number of people relying on MTurk to always or sometimes "make basic ends meet" suggests a need for consideration of labor fairness in HCI design and development processes not prominent in literature on HCI, crowdsourcing, and human computation thus far. MTurk has been taken up in HCI primarily in two ways: as a means of evaluating technologies such as through usability testing (e.g., [8,13]), and as a part of the design of a computational system, responding to user inputs and contributing to outputs in real or semi-real time (e.g., [15]).

PAID EVALUATION AS AN OCCUPATION

While paying people little to nothing to participate in usability studies is nothing new, performing HCI research and evaluation with human subjects through MTurk creates the possibility of workers coming to rely

on such research participation to make a living. While a subject may participate in a few research and usability engagements each month with face-to-face usability and design research, MTurk workers can participate in a continuous stream of these evaluations as a source of income. This raises several questions: what is an ethical design research or evaluation engagement in such a setting? What issues should the HCI community consider? How does a subject's participation in a range of such studies affect the validity of experiment results?

HUMAN COMPUTATION IN HCI SYSTEMS
Human computing systems [2,19] involve using human thinking to drive computational systems. For example, Games With A Purpose [3] have used anonymous crowds of gamers on the web to label images, identify music, and perform other human intelligence tasks to support a variety of projects with low operational costs. Although scalability and viability of technological systems is frequently linked to these low costs, in human computing systems this goal becomes direct pressure to keep human labor costs down.

With the exception of participatory design, HCI has not developed a language or conceptual framework for considering questions of labor and livelihood in system design. Should HCI researchers rely on the low wages paid through the system or even seek to push them lower as part of making (human) computation more usable and acceptable? What are the effects of working as part of a human computation system over the short, medium, and long term? What kinds of work conditions for these human computers can the HCI community design? Design is unlikely to be a sufficient vocabulary for thinking about the conditions of human

computation. Likely, a growing HCI appreciation for labor and livelihoods research will be important as well.

Future Work
There is much further work that can be done to understand the anonymous Turkers working in the MTurk system. We are continuing our longitudinal repetition of surveys, looking at whether the trends identified here seem to be continuing. We have also begun looking at the motivations and reasons why workers complete HITs through MTurk, in order to see what effects this may have on how work is organized and performed. Nevertheless, more work along these lines continues to be necessary in order to understand the increasingly prevalent practice of crowdsourcing cognitive labor.

Conclusion
In this paper, we have presented a compilation of surveys of the demographics of MTurk workers over a 20-month period, showing how the Turker population appears to have become more international, with a growing number of young, male, Indian Turkers. If we take MTurk as a typical example of a system for crowdsourcing cognitive labor, then this globalization may be indicative of other crowdsourcing systems—particularly as they may take advantage of the willingness of people in lower-income countries to do these kinds of cognitive tasks for smaller payments. As with other aspects of the Internet, such tasks are not limited to a particular nationality or a particular socio-cultural context. Such a globalizing of the crowd workforce may affect the manner in which HITs are performed, as well as the social effects of offering this type of work in other crowdsourcing systems as well.

References

[1]   Amazon Mechanical Turk - Participation Agreement. https://www.mturk.com/mturk/conditionsofuse.

[2]   von Ahn, L. Human computation. *Pro. K-CAP*, ACM (2007), 5-6.

[3]   von Ahn, L. and Dabbish, L. Designing games with a purpose. *Commun. ACM 51*, 8 (2008), 58-67.

[4]   Alonso, O., Rose, D.E., and Stewart, B. Crowdsourcing for relevance evaluation. *SIGIR Forum 42*, 2 (2008), 9-15.

[5]   Barley, S.R. and Kunda, G. Contracting: A new form of professional practice. *Perspectives - Academy of Management 20*, 1 (2006), 45.

[6]   Barr, J. and Cabrera, L.F. AI gets a brain. *Queue 4*, 4 (2006), 24-29.

[7]   Deng, J., Dong, W., Socher, R., Li, L., Li, K., and Fei-Fei, L. ImageNet: A large-scale hierarchical image database. *Proc. CVPR 2009*, IEEE (2009), 248-255.

[8]   Evans, B.M. and Chi, E.H. Towards a model of understanding social search. *Proc. CSCW 2008*, ACM (2008), 485-494.

[9]   Howe, J. The rise of crowdsourcing. *Wired Magazine 14*, 6 (2006).

[10] Ipeirotis, P. Mechanical Turk: The Demographics. *A Computer Scientist in a Business School*, 2008. http://behind-the-enemy-lines.blogspot.com/2008/03/mechanical-turk-demographics.html.

[11] Ipeirotis, P. Turker Demographics vs. Internet Demographics. *A Computer Scientist in a Business School*, 2009. http://behind-the-enemy-lines.blogspot.com/2009/03/turker-demographics-vs-internet.html.

[12] Irani, L. and Silberman, M.S. "Agency and Exploitation in Mechanical Turk" at Internet as Playground and Factory Conference at The New School. 2009. http://www.slideshare.net/lirani/agency-and-exploitation-in-amazon-mechanical-turk.

[13] Kittur, A., Chi, E.H., and Suh, B. Crowdsourcing user studies with Mechanical Turk. *Proc. CHI 2008*, ACM (2008), 453-456.

[14] Paolacci, G. and Massimo, W. Experimental Turk. http://experimentalturk.wordpress.com/.

[15] Sala, M.C., Partridge, K., Jacobson, L., and Begole, J. An exploration into activity-informed physical advertising using pest. *Proc. Pervasive Computing*, (2007), 73.

[16] Smith, V. New Forms of Work Organization. *Annual Review of Sociology 23*, 1 (1997), 315-339.

[17] Snow, R., O'Connor, B., Jurafsky, D., and Ng, A.Y. Cheap and Fast—But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks. *Proc. EMNLP 2008*, ACL (2008), 254-263.

[18] Sorokin, A. and Forsyth, D. Utility data annotation with Amazon Mechanical Turk. *Proc. CVPRW 2008*, IEEE (2008), 1-8.

[19] Zittrain, J. Ubiquitous human computing. *Phil. Trans. R. Soc. A.: 366*, 1881 (2008), 3813-3821.