

Understanding and Designing for Cultural Differences on Crowdsourcing Marketplaces

Gary Hsieh
College of Communication Arts and Sciences
Michigan State University
East Lansing, MI
garyh@msu.edu

ABSTRACT

Crowdsourcing marketplaces, such as Mechanical Turk, enable tasks to be distributed to workers from all over the world for completion. While this global quality of crowdsourcing provides numerous benefits to requesters of work (*e.g.*, more diverse labor force that is available at all hours of the day), it could, however, pose challenges in terms of labor management. Much prior work has demonstrated the effects of cultural backgrounds on individual's thoughts, value, and even behavior. Therefore, the diverse cultural backgrounds on crowdsourcing marketplaces may interact with incentive types, amounts, and task types to impact workers' task performance, engagement, and selection. My research program explores the effects of cultural differences on these online marketplaces in order to build more efficient global crowdsourcing marketplaces.

Author Keywords

Crowdsourcing, online marketplaces, cultural differences

ACM Classification Keywords

H5.3. Group and organizational interfaces: Web-based interaction.

INTRODUCTION

Crowdsourcing marketplaces allow people to employ workers to work on tasks ranging in complexity, from labeling images to writing product descriptions. In exchange, the workers are compensated with monetary payments offered by the requesters. These services are only increasing in popularity. On the requesters' side, more and more individuals and businesses have begun to embrace to these crowdsourcing marketplaces as valid labor pools. On the workers' side, more professions are projected to participate in crowdsourcing [8].

One of the key strengths of online crowdsources marketplaces is that they provide requesters with access to a

large pool of workers, across geographic boundaries. Mechanical Turk (MTurk), for example, a popular crowdsourcing marketplace, boasts more than 100,000 workers in over 100 countries [9]. It continues to increase in geographic diversity – analyses of Mechanical Turk demographics have shown that since 2008, the percentage of US workers has dropped from 70-80% to around 50% [7,10]. This quality enables requesters to have work done by workers at all hours of day, and it allows them to get work done by a more diverse group of workers than they could by soliciting workers locally. This type of diversity can be extremely beneficial to the requesters. For example, researchers using these marketplaces for studies can obtain a more generalizable sample, while businesses may target workers from specific regions for higher quality work on location-specific tasks, such as language translations.

While there are clear benefits in having more diverse pools of workers, the increase in diversity can also result in complications in labor management. Many previous studies show that there is a significant effect of cultural backgrounds on individual's thoughts, values, and behaviors [*e.g.*, 2,3]. Given that many workers outside from the US are actively participating in these online marketplaces, it is imperative to understand the influence of cultural differences on workers' behaviors in these services. In the same way that websites have developed different versions and interfaces to accommodate visitors from different areas of the world, online crowdsourcing marketplaces may also need to offer different interfaces and interactions to support the diverse pool of workers.

PRIOR STUDIES WITH USING MECHANICAL TURK

I have used Mechanical Turk for a number of research projects: to rate the types and characteristics of posts to question and answer services [4], to test the understandability of novel and animated icons [1], and to study the effects of incentives on task performance, social relationships and credibility [5,6]. It is through these experiments that I realized there is an important missing gap of knowledge on cultural differences on crowdsourcing marketplaces. From an academic researcher's perspective, understanding the effects of cultural differences can be extremely critical in controlling for experimental or survey results collected. However, filling this knowledge gap is important for more than just academic researchers. With

this knowledge, designers of crowdsourcing services can improve these services for workers and requesters.

ONGOING RESEARCH ON CULTURAL DIFFERENCES

A key assumption of this research program is that significant cultural differences exist. Therefore, one of the first steps is to test this assumption. Otherwise, it may be possible that while crowdsourcing marketplaces attract workers from all over the world, those workers who choose to participate may actually be very similar to each other.

Currently, I am in the process of surveying MTurk workers to collect their responses on established culture and work value scales. This survey include scales measuring the six dimensions of culture presented by Hofstede: small vs. large power distance, individualism vs. collectivism, masculinity vs. femininity, weak vs. strong uncertainty avoidance, long vs. short orientation and indulgence vs. restraint [2]. It also includes the Value Inventory developed by Shalom Schwartz. In addition, my survey also contains the standard barrage of demographic questions [11] along with workers' cultural heritage and location of current residence.

This data will, first and foremost, improve our understanding of overall MTurk demographics. This data will also help show worker differences across geographic boundaries. For example, do MTurk workers hold the same sets of cultural beliefs as was shown by Hofstede? Do the workers have about the same income and education level as others who live in the same city/town? Do the workers differ from other workers as predicted by prior research? The two general research questions are listed below, and answering these questions will be extremely valuable for researchers and practitioners alike in interpreting and using the responses they collect.

RQ1: Along what dimensions are MTurk workers from different countries of origin different?

RQ2: Are MTurk workers representative of their countries of residence? If not, how are they different?

Another key assumption of this work is that cultural differences among workers can lead to differences in workers' behaviors. To confirm this assumption, survey takers are also presented with a screenshot of a task posted on MTurk and are asked to rate their likelihood to accept the HIT to work on it. Different aspects of the task, such as the payment amount, are manipulated.

RQ3: Do cultural characteristics and work-related values affect the selection of tasks on MTurk?

This ongoing work should conclude well before the CHI Conference and I will be able to share the results with the workshop participants.

VISION AND CONTRIBUTIONS

By nature, successful crowdsourcing services will involve a large, and inevitably, diverse group of workers. However, in order for crowdsourcing services to fully utilize these workers, labor management is critical, and to do so requires a better understanding of workers' cultural differences and how these differences affect work.

However, understanding is merely the first step of my research program. The next step will be to design better interactions and interfaces to support or even leverage existing cultural differences among workers. For example, let us assume that China-based workers on crowdsourcing marketplaces are indeed more collectivistic than US-based workers, as suggested by Hofstede's prior surveys. Can we then recommend certain tasks or tailor certain site features to improve workers experiences? If so, that could greatly improve the quality of work provide by the workers and at the same time, make participating on MTurk more rewarding. While much of my work will be conducted on MTurk, knowledge generated and tools developed can improve general crowdsourcing services, and may even help support localization of traditional organizations.

I hope to use this workshop as an opportunity share my findings and experiences with other researchers and to seek their input on my studies and designs to manage cultural differences in crowdsourcing.

ACKNOWLEDGEMENTS

This is ongoing research with my students Jeong-woo Jang, Matthew Muscat, and my collaborator Vaughn Hester at CrowdFlower. This work is funded by CrowdFlower.

BIOGRAPHY

Gary Hsieh is a joint-appointed Assistant Professor at Michigan State University in the Departments of Communication; and Telecommunication, Information Studies and Media. He is also affiliated with the Health and Risk Communication Center in the College of Arts and Sciences. His broad interests are in Human-Computer Interaction, Computer-Mediated Communication and Behavioral Economics. His focus is on studying, designing and developing technologies to enable people to interact in ways that are efficient and welfare-improving. He has conducted research at a number of industry research labs, including Microsoft, IBM, Intel and Fuji-Xerox. He received his Ph.D. from the Human-Computer Interaction Institute at Carnegie Mellon University. He earned his B.S. in Electrical Engineering and Computer Science, with a minor in Business Administration, from University of California at Berkeley.

REFERENCES

1. Harrison, C., Hsieh, G., Willis, K., Forlizzi, J., Hudson, S.E. (2011). Kineticons: Using Iconographic Motion in Graphical User Interface Design. To Appear in *ACM*

- Conference on Human Factors in Computing Systems (CHI)*.
2. Hofstede, G. 1980. *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
 3. Hofstede, G. The cultural relativity of organizational practices and theories. *Journal of International Business Studies*, 1983, Fall, 75–90.
 4. Hsieh, G., Kraut, R.E., Hudson, S.E. (2010). Why Pay: Exploring How Financial Incentives are Used for Question & Answer. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 305-314.
 5. Hsieh, G., Kraut, R.E., Hudson, S.E. (2011). Donate for Credibility: How Contribution Incentives Can Improve Credibility. To appear *ACM Conference on Human Factors in Computing Systems (CHI)*.
 6. Hsieh, G., Kraut, R.E., Hudson, S.E., Counts, S., Weber, R. (in preparation). The Social Cost of Financial Incentives.
 7. Ipeirotis, P. (2010). Demographics of Mechanical Turk. CeDER Working Paper.
 8. Kimbarovsky, K. (December, 2010). “Ten Crowdsourcing Trends for 2011.” <http://blog.crowdspring.com/2010/12/2011-crowdsourcing-trends/>
 9. Potin, J. (2007). Artificial Intelligence, With Help From the Humans, *The New York Times*, 25 March 2007.
 10. Ross, J., Irani, L., Silberman, M.S., Zaldivar, A. and Tomlinson, B. (2010). Who are the crowdworkers?: shifting demographics in mechanical turk. In *Proceedings of the ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*. pp. 2863-2872.
 11. Schwartz, S.H. 1999. A Theory of Cultural Values and Some Implications for Work. *Applied Psychology: An International Review*, 48, 23-47.