

Crowdsourcing and Creativity

Mira Dontcheva
Advanced Technology Labs
Adobe Systems
mirad@adobe.com

Elizabeth Gerber
Segal Institute of Design
Northwestern University
egerber@northwestern.edu

Sheena Lewis
Technology and Social
Behavior
Northwestern University
sheena@u.northwestern.edu

ABSTRACT

In studying creativity and techniques for enhancing the creative process, we utilized crowdsourcing through Amazon's Mechanical Turk. Through our studies we found that in its current form, the Mechanical Turk platform has some limitations in supporting creative tasks. In this position paper, we describe our studies and give recommendations for crowdsourcing platforms that foster creativity.

Author Keywords

crowdsourcing, creativity, mechanical turk

ACM Classification Keywords

H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces—*Web-based interaction*

General Terms

Design, Experimentation, Measurement

INTRODUCTION

Knowledge is not enough. Success is based on knowing how to think and act creatively with one's knowledge [20]. For this reason, researchers in many fields from psychology [10] to business [11] are increasingly interested in how to support creativity. As computers have grown to be a part of every aspect of our lives, computer science researchers have become interested in the creative process and tools and techniques that support the creative process [22]. For example, Lee et al. [17] present an interface for exploring alternatives during the design process, while well-known author Steven Berlin Johnson questions his own creativity, as he writes with the software DevonThink [15].

Crowdsourcing, or relying on a distributed network of individuals to reduce monetary and time costs, to complete a task [12] is an emerging paradigm that is changing the way creative work is done. Websites like 99designs [1] and CrowdSpring [2] connect designers from all over the world with potential clients and change the typical creative consulting model from one-to-one to one-to-many. Clients use

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

Copyright 2011 ACM 978-1-4503-0267-8/11/05...\$10.00.



Figure 1. source: <http://prestonkelly.com/blog/?p=622>

crowdsourcing platforms to publically solicit creative solutions and pick the best design from hundreds of alternatives. In a more playful setting, websites like Worth1000 [3] and LayerTennis [4] use crowdsourcing platforms to support creative work by encouraging creative professionals to compete with each other to show off their existing skills and learn new skills from one another.

Among academics Amazon's Mechanical Turk is the most popular crowdsourcing platform. Mechanical Turk is widely used for tasks that are hard for computers but easy for humans, such as labeling images and transcribing audio. There is an increasing interest in leveraging micro-task crowdsourcing platforms, such as Mechanical Turk, for creative tasks. For example, Bernstein et al. [8] use it to assist in writing, while Kittur [16] explores collaborative translation of poetry.

In this paper, we discuss our experience using Mechanical Turk as a platform for creative work and describe some of its limitations. Based on our own experience and a theoretical model designed to assess creativity in work environments [7], we offer recommendations for crowdsourcing platforms that better support creative tasks.

WORK ENVIRONMENT AND CREATIVITY

Psychological research has shown that people are most creative when they are intrinsically motivated by their interest, enjoyment, satisfaction, or challenge of the work, instead of extrinsically motivated by things like promotions or

money [6]. Research on work environments and how companies can foster innovation and creativity points to a number of environmental conditions that result in higher levels of creativity. Amabile et al. [7] show that environments that are supportive of creativity offer:

- task autonomy and *freedom*, which allow workers to have a sense of ownership over their work and ideas
- *intellectually challenging work*,
- *supervisory encouragement* including setting clear goals and frequent and open interactions between a supervisor and his/her team,
- *organizational encouragement* including encouraging workers to take risks, evaluating new ideas fairly without too much criticism, and offering rewards and recognition for creativity,
- and *work group supports* through team members with diverse backgrounds, openness to ideas, and a shared commitment to a project.

Mechanical Turk offers a very different working environment from a traditional company, which is the basis for these findings, because Turkers, or people who receive payment for completing tasks on Mechanical Turk, are independent contractors who work offsite, are compensated independently of merit, don't expect benefits, and don't have a long term vested interest in the success of the organization. Despite these differences, we expect that at least some of these environmental factors apply to crowdsourcing platforms when the desired outcome is creative work.

Since Mechanical Turk was not designed with creative tasks in mind, perhaps it is not surprising that as a work environment it does not explicitly embody work conditions that foster creativity and at this time does not include even one of the characteristics Amabile and her colleagues find encourage creativity.

Freedom. Turkers have some autonomy, as they can pick the tasks they complete, but they have little ownership over their work as the requesters are the ones who own the work after it has been completed.

Intellectual challenge. Although there are some intellectually challenging jobs on Mechanical Turk, such as offering feedback and generating new content, most Mechanical Turk jobs are trivial for humans, such as transcription and content labeling [13].

Supervisory encouragement. Generally, there is little interaction between the requesters and the workers and there is no notion of a supervisor that is an advocate for the worker. The HCI research community has raised questions around the ethics of this type of anonymous work relationship where the requester holds all the power [21]. Communities around forums like `turkernation.com` and tools like `Turkopticon` [5] have risen in response to this imbalance. Mechanical

Turk does provide a mechanism for offering extrinsic monetary rewards, but beyond this feature there is limited feedback back to the workers once they have completed a job. They are either paid or not, and they are given little or no feedback on the quality of their work.

Organizational and work group supports. Mechanical Turk does not explicitly support collaboration among Turkers. Turkers do communicate through external forums but do not typically work on tasks together.

Despite Mechanical Turk's apparent lack of support for creative tasks, its low cost and low time commitment benefits led us to experiment crowdsourcing creative tasks. Our goal was to determine whether affect-laden background images could influence creative performance. Previous work [14] shows that positive affect, or emotion, can have a positive impact on creative performance.

MECHANICAL TURK AND THE UNUSUAL USES TASK

Measuring a complex phenomenon such as the creative process is a difficult task. In the 1960's, the psychologists Guilford and Torrance developed a number of psychometric tests that evaluate convergent and divergent thinking. Their tests are widely used today and form a well established benchmark. In Guilford's Alternative Uses Test, participants are asked to write down as many unusual uses as possible for a common object, such as a brick or a paperclip.

We recruited 240 participants from Mechanical Turk that reside in the United States and had successfully completed at least 80% of prior Amazon Mechanical Turk tasks. The participants were paid \$0.05 for completing the task. Each participant was assigned to one of four conditions - positive, neutral, negative, and baseline. Using a database of images rated for affect, we selected a positive image (a laughing baby), a negative image (dead bodies after the 2010 earthquake in Haiti), and a neutral image (a hammer). The instructions for completing the task were as follows:

In this task, your goal is to think of as many unique and unusual uses for a common object. For example, using a paper clip as an earring is an unusual and unique use. However, using a paper clip to bind papers is not unique or unusual. Try to think of as many unique and unusual uses as possible. DO NOT use any external sources (e.g., websites, people) to complete this task.

To the right of the instructions was a 350 x 233 pixel image, which varied under the different conditions. For the baseline, we did not include an image. We did not provide information about why the image was present.

To proceed with the task, the participants had to click a next button. They were presented with a web form (Figure 2), which asked them to enter as many unique and unusual uses for an object (either a brick or a quarter) as possible. The participants were required to enter a minimum of 10 responses but had the option of entering up to 20 responses so that we could study the effort Turkers were willing to expend beyond the minimum requirements.

Enter as many unique and unusual uses
as possible for a **QUARTER** (the coin).
(Minimum: 10)

1.
2.
3.
4.
5.



Figure 2. Screenshot of Guilford’s unusual uses task

Results

We collected 2,463 responses from 226 turkers. All responses were rated for originality by expert raters from the design field on a scale of 1 to 5 (1- low originality, 5 - high originality). We found that the visual design of the task does have an impact on quality of ideas generated. The positively-laden background image (baby) led to responses that were rated significantly more original than the baseline (no image) and neutral-laden background image (hammer). More details on this can be found in our CHI 2011 paper on affective computational priming [18].

Despite this successful study, less than 13% of the participants gave more than the required 10 responses making it challenging to study the effect of visual design on the quantity of ideas generated. This means that almost all of the participants did the bare minimum to finish the task in order to get paid and therefore their motivation was most likely extrinsic rather than intrinsic. In a follow-up study, we plan to explore how additional reward might influence creative tasks and how our results might change if we remove the minimum 10 responses. Additionally, we found that participants were not following directions. Instead of generating ideas on their own, they used Web resources to search for answers to the unusual uses task. While search is considered a useful skill for creative work, the Turkers merely pasted solutions generated by others rather than using the search process to inform new creative solutions.

Based on these experiments and theoretical work on work environments, we form four recommendations for crowdsourcing platforms for creative tasks: build a community, align incentives with desired behavior, offer templates, and monitor use of other resources. These are not meant to be exhaustive but rather focus on aspects that may influence creative tasks.

RECOMMENDATIONS

Build a community

Amabile et al. [7] find that organizational encouragement and work group supports are some of the most important characteristics for supporting creativity at work. Both of these rely on a community of people. Crowdsourcing platforms primarily focus on the client-worker relationship rather than the worker-worker relationship. Crowdsourcing platforms can benefit from building communities of workers perhaps with reputation systems, that enable workers to learn from one another and share common experiences.

A community could also grow around collaborative tasks. Games such as FoldIt [9] demonstrate that people are willing to spend many hours and days competing to solve a puzzle. Kittur [16] describes a collaborative translation task and reports that workers were willing to come back and translate new content long after the task was finished. Games may offer a mechanism for intrinsic motivation that paid crowdsourcing platforms, such as Mechanical Turk, may never be able to reach.

Align incentives with desired behavior

Mechanical Turk incentivizes small tasks that can be done quickly over longer more involved tasks. Ipeirotis [13] shows that 90% of tasks cost only 10 cents. In our study, participants spent on average 6 minutes on each task. But creativity is not something that can happen quickly. Psychological research has shown that exploration of alternatives and time for that exploration directly correlate with creativity of task outcomes [19]. Interestingly, the time pressure on Mechanical Turk tasks is not explicitly controlled by the requesters but is rather an intrinsic motivation for the workers. They want to do as many tasks as possible quickly optimizing for a "good enough" response that will result in payment.

Crowdsourcing platforms that support larger jobs, such as entire projects (e.g., CrowdSpring [2]), offer high-quality results, but they also offer much higher compensation (starting at \$200) and are connected to a community of professionals. It remains to be seen whether micro-scale crowdsourcing markets can be effective for accomplishing creative tasks.

Offer templates and aesthetics guidelines

Choosing the right format and appropriate wording to make a Mechanical Turk task easy to complete is an art. Furthermore, our experiments show that the form and aesthetics of a task, i.e. which image it includes, can make a difference on users interactions with it and can influence the creativity of their responses. A crowdsourcing platform could offer templates that allow requesters to describe the types of responses they need. Perhaps a serious task may receive very different results than a whimsical task.

Monitor use of other resources

The fact that crowdsourcing platforms, like Mechanical Turk, are on the Web makes them accessible to many people, but it also means that it is hard to control the materials Turkers access during tasks. For creative tasks this can be particularly important as sources of inspiration are very important during creative problem solving and can lead to better or worse designs [17]. In our experiment, we did not want users to go to the Web to look for answers, but we could not control for this. In analyzing the responses, it was clear that many did search the Web instead of generating answers on their own. And since Guildord’s unusual uses task is well established there are many websites that list answers to unusual uses for common objects, such as a brick.

Perhaps a browser extension or desktop monitoring application can be a component of creativity focused crowdsourcing

platforms. Such a platform would allow requesters to monitor how workers use their existing desktop software to accomplish tasks and perhaps even suggest appropriate reference materials.

CONCLUSION

Crowdsourcing creative tasks is still in its early beginnings. Existing platforms are missing some important qualities to truly become supportive of creative tasks. We leverage psychological research and results from a creativity experiment to offer guidelines for future crowdsourcing platforms for creative tasks. We hope we can participate in this workshop, share our findings, and take part in discussing the future of crowdsourcing platforms.

REFERENCES

1. <http://99designs.com/>.
2. <http://www.crowdspring.com/>.
3. <http://www.worth1000.com/contests>.
4. <http://www.layertennis.com/>.
5. <http://turkopticon.differenceengines.com/>.
6. Amabile, T. *The social psychology of creativity*. Springer-Verlag, 1983.
7. Amabile, T., Conti, R., Coon, H., Lazenby, J., and Herron, M. Assessing the work environment for creativity. *The Academy of Management Journal* 39, 5 (1996), 1154–1184.
8. Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D., and Panovich, K. Soylent: a word processor with a crowd inside. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology*, UIST '10 (2010), 313–322.
9. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z., and players, F. Predicting protein structures with a multiplayer online game. *Nature* 466, 7307 (08 2010), 756–760.
10. Csikszentmihalyi, M. *Creativity: Flow and the Psychology of Discovery and Invention*. Harper Perennial, 1997.
11. Essentials, H. B. *Managing Creativity and Innovation*. Harvard Business Press, 2003.
12. Howe, J. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. Crown Business, 2008.
13. Ipeirotis, P. G. Analyzing the amazon mechanical turk marketplace. *XRDS* 17 (December 2010), 16–21.
14. Isen, A., Daubman, K., and Nowicki, G. Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology* 52, 6 (1987), 1122–1131.
15. Johnson, S. B. Tool For Thought, 2005. <http://www.nytimes.com/2005/01/30/books/review/30JOHNSON.html>.
16. Kittur, A. Crowdsourcing, collaboration and creativity. *XRDS* 17 (December 2010), 22–26.
17. Lee, B., Srivastava, S., Kumar, R., Brafman, R., and Klemmer, S. R. Designing with interactive example galleries. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, ACM (New York, NY, USA, 2010), 2257–2266.
18. Lewis, S., Dontcheva, M., and Gerber, E. Affective computational priming and creativity. In *CHI '11: Proceedings of the 29th international conference on Human factors in computing systems (to appear)* (2011).
19. Parnes, S., and Noller, R. Applied creativity: The creative studies project - part ii: Results of the two-year program. *Journal of Creative Behavior* 6 (1972), 164–186.
20. Resnick, M. Sowing the seeds for a more creative society. *Learning & Leading with Technology* 35, 4 (2007), 18–22.
21. Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: shifting demographics in mechanical turk. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, CHI EA '10, ACM (New York, NY, USA, 2010), 2863–2872.
22. Shneiderman, B. e. a. Creativity support tools: Report from a u.s. national science foundation sponsored workshop. *International Journal of Human-Computer Interaction* 20, 2 (2006), 61–77.

AUTHOR BIOS

Mira Dontcheva is a Senior Research Scientist at Adobe Systems. Her research focuses on search and sensemaking interfaces, end-user programming, and most recently creativity. Mira received her PhD in computer science from University of Washington and her BSE in Computer Engineering from the University of Michigan.

Elizabeth Gerber is an assistant professor of Mechanical Engineering at the Segal Design Institute at Northwestern University. Her research investigates the role of technology in creativity and innovation. She received her Ph.D. in management science and engineering from Stanford University.

Sheena Lewis is a Ph.D. student in the Technology and Social Behavior program at Northwestern University. Her research examines technology use to address violence in low-income Chicago communities.