The Crowdsourcing Design Space

Yasuaki Sakamoto, Yuko Tanaka, Lixiu Yu, and Jeffrey V. Nickerson

Center for Decision Technologies, Stevens Institute of Technology
{ysakamot,yuko.tanaka,lyu3,jnickerson}@stevens.edu

Abstract. Crowdsourcing is a new kind of organizational structure, one that is conducive to large amounts of short parallel work: thousands of individuals may work for several minutes on tasks, their outputs aggregated into a useful product or service. The dimensions of this new organizational form are described. Areas for future research are identified, focusing on open-ended tasks and the coordination structures that might foster collective creativity.

Keywords: Crowdsourcing, distributed cognition, organizational design, peer production, collective creativity, human computation.

1 Introduction

Crowdsourcing, the assembling of strangers to accomplish a task [1-4], has the potential to transform the nature of work [5]. Many companies are sensing this, and adapting it to solve problems and provide services [6-8]. Crowdsourcing is becoming popular because the Internet-based infrastructure to support the management of crowds has grown and strengthened. This infrastructure provides opportunities not just for businesspeople, but also for researchers who want to study distributed cognition [9]. Specifically, crowds can participate in experiments: crowds can be simultaneously assembled in both control and experimental conditions, so the effects of social networks, organizational structure, and information flow can be studied (e.g., [10, 11]). Because of its potential for both business change and fundamental research, the crowdsourcing phenomenon deserves systematic study through analysis and experimentation.

There are several ways crowdsourcing differs from other forms of organization. First, the assembly of a crowd can happen quickly, and need only persist for short amounts of time – minutes, hours, or days. In contrast, social institutions tend to persist for many years, and take time to build. For example, the hiring of one thousand individuals into a company is a large undertaking, but a crowd of 1000 can be pulled together and asked to perform a simple task, all in the space of less than one hour. Moreover, a crowd can be assembled from around the world, constituted from different nationalities, cultures, and professions.

What motivates crowds to participate? Some sites provide money, others provide reputational incentives, and still others provide neither of these. But crowds will perform tasks they find interesting, whether or not they receive monetary incentives. In particular, games have proved popular: game players have helped label images [12] and discover ways to fold proteins [13]. These successes suggest that crowds might be

assembled to solve a wide variety of problems, if such problems can be broken down into fun-to-play activities.

In fact, for many tasks, a crowd might be more efficient and effective than an expert. For example, Galton showed that a crowd could guess the weight of an ox better than a farmer [14]. Other work showed that groups of people performed better than individuals on tasks such as traversing a maze if their decisions were aggregated [15, 16]. Such techniques are effective probably because aggregation cancels out individual errors and reinforces correct solutions [17].

As a way of synthesizing crowdsourcing research, we will describe the dimensions that define a design space for crowdsourcing systems. We then discuss four dimensions in particular: the nature of the task, the communication methods, the levels of hierarchy, and the workflow. Along the way we point out possibilities for structuring the crowd in new ways, and opportunities for understanding more about collective cognition.

2 Dimensions of Crowdsourcing

A design space identifies the range of possible solutions for a design problem [18]. The space is delineated by set of dimensions that structure and constrain the set of decisions made in creating a design.

Malone et al. [19] analyzed a related space – collective intelligence – in terms of four broad questions: *Who*, *What*, *How*, and *Why*: applying this to crowdsourcing leads to a breakdown of the space into the users of the system, the tasks to be performed, the method for crowdsourcing, and the motivation of the crowd. The authors also identify more specific dimensions they call *genes* that can be combined into *genomes* in order to constitute a new application.

In contrast, Quinn and Bederson [2] use a different set of dimensions to describe another related space they call human computation: *Motivation*, *Human Skill*, *Aggregation*, *Quality Control*, *Process Order*, and *Task-request Cardinality*. These can be reconciled with the four broad questions of Malone et al.: *Aggregation* and *Process Order* address a big part of *How*; *Task-request Cardinality* and *Human Skill* relate to *What*. *Motivation* focuses on how incentives are used to encourage participation. Table 1 shows the four broad dimensions, and several specific dimensions from Quinn and Bederson.

We add several dimensions to the mix. In the *Who* category, we include demographics and level of expertise. Participants vary along demographic dimensions, such as age and nationality, as well as along the level of expertise they have in the task domain. These additional dimensions may become important when allocating tasks and assessing crowd performance. In the *What* category, we add the task domain (e.g., image labeling) the nature of the task (e.g., generation) and the output (e.g., list of terms). These dimensions highlight the context for crowdsourcing. The *How* category includes an evaluation method. Evaluation processes can increase the quality of responses. *How* also includes incentives as distinct from workers' motivations, an aspect of *Why*. That is, one can offer cash but still find that workers perform the task for different reasons: there is a rich literature on the relation between motivation and incentives that can be brought to bear on these two dimensions [20, 21]. We further

include visibility of outputs as a dimension of *How* – for example, one might restrict crowd members from seeing each other’s work. We will elaborate other attributes of *How* in the following sections: protocols for communication between workers, hierarchical levels within the crowd, and the sequence of work. We distinguish two kinds of motivation in *Why*: the reasons for crowdsourcing a task, and the reasons for people to work on a task. For example, companies might use crowdsourcing to optimize profit [8], while researchers might do so to increase knowledge in a domain [13] or to understand an aspect of distributed cognition [22, 23].

Table 1. The dimensions of crowdsourcing (italics flag sections in the paper)

Broad dimension	Specific dimension	Example attribute
Who	Demographics	age, country
	Level of Expertise	novice, expert
What	Domain of the task	protein folding, image labeling
	<i>Nature of the task</i>	recognition, generation
	Output	sequence of protein folds
How	Incentives	contest prize for the best task
	Aggregation method	collection, combination
	Evaluation method	vote, expert opinion
	Visibility of outputs	opaque, transparent
	<i>Communication</i>	mediated through the tasks
	<i>Levels of hierarchy</i>	single, multiple
	<i>Workflow</i>	evaluation following collection
Why	Requester’s motivation	profit, knowledge
	Worker’s motivation	money, fun

Next, we focus on four areas in this design space that offer great promise as topics for future investigation. The authors have used crowdsourcing to perform generative, creative tasks [22, 23]. This topic often interests both industry and academe, so we address it first.

3 Nature of the Task: Creativity and the Crowd

Views of creativity reflect the times: the romantic era’s mystique of the lone inventor inspired late in the night has been eclipsed by a view of the creative act as a social one, made by teams of closely knit people [24]. Art schools and design studios are social places: so are engineering institutes and research laboratories. The view, then, has moved from individual to group, and the groups are seen as cohesive and close knit. Nevertheless, people remain uneasy with the idea of collaborating across great distances. Indeed, some studies of geographically-distributed collaboration have identified a tradeoff between convenience and cost [25]. The Internet can reduce the cost of coordinating work among geographically-distributed, but otherwise traditional, teams. However, the Internet also makes possible a new model of collaboration: combining the work of isolated strangers.

Inspired by this new model of collaboration, we decided to assess the sequential application of crowds to both generate and aggregate design ideas. If this approach to idea generation works well, then many difficult design challenges might be attacked in parallel. If the approach has limits, then in finding these limits we will provide a more focused path for research on peer production (cf. [5]).

In one experiment, we combined the work of many designers working in parallel to create a composite design [26]. We found that taking elements from many different designs according to a simple heuristic returned a strong design, as judged by experts. This suggests that Galton's insight into the crowd's ability to estimate [14] also applies to the crowd's ability to design, perhaps for the same reason: individual errors disappear, and strong connections reinforced. Conceptual combination [27] can be controlled at a finer scale, with pair-wise combination techniques such as those used in genetic algorithms [28-31]. We have performed experiments using these techniques [22, 23]. To do so, we needed to understand and implement mechanisms for connecting crowd output. We turn to these next.

4 Communication between Crowd Members

A member of an online crowd may have no connection to another member except through participation on the same project. On the other hand, in physical crowds, those in the immediate vicinity often influence one another. For example, members of a crowd at an outside concert will be aware of who is around, and will move in order to fill in gaps or create more breathing room. There may be no verbal interaction, but there may be awareness. Moreover, there is mediation: members of a crowd streaming toward an exit may affect members more distant through a wave of local adjustments.

Such awareness can exist electronically – for example, in virtual worlds. But in other cases, the crowd can't directly see others situated in space, whether cyber or physical. There have been many approaches to understanding how ideas flow and people interact in such environments. Coordination theory [32] focused on the resources through which people work toward goals. Electronic brainstorming looked at the way computers mediate communication [33, 34]. A general theme in such research is that team collaboration is not always useful. So collaboration between crowd members may or may not be a good idea.

In a strictly separated crowd, one might only let people interact through intermediate productive output. That is, one member of a crowd may be asked to edit the output of another crowd member. There is some coordination, but the interaction is very limited.

In more loosely separated environments, one might see more of what others have done, and one can modify – remix – that output. For example, in the Scratch environment, youths find examples of each other's work, modify it, and upload it [35]. In the Yahoo! Answers environment, users doesn't modify the contributions, but instead comment on them. Users make critical comments on answers if they think that the answers are wrong, and other users may provide new answers after reading the comments. In crowdsourcing environments, even very self-conscious individuals who have difficulty providing critical comments in the presence of others may feel comfortable doing so [36].

Both of these are forms of collaboration – commenting and remixing – can be combined. In Scratch, youths have engaged in a kind of recursive crowdsourcing – by using open-ended discussion forums, they initiate their own contests ask their own peers to enter, and provide their own prizes [37]. The above examples provide crowd members a way to react to or modify a single idea. What about providing participants with two ideas?

In a human *genetic algorithm* two ideas are combined [22, 30]. There is an integration of the crowds' ideas, but the integration works in one direction, with members never collaborating interactively with each other. The ideas can be textual, or involve images, and can be combined across many generations [23].

Games also provide a way for pairs of people to interact and integrate ideas. In *games with a purpose* the participants perform a collaborative task [12]. For example, in an image labeling game, they try to predict how the other person would label an image. Their responses provide a list of labels for images that are difficult for computers to come up with. These tasks involve a structured kind of collaboration, and are often restricted to pairs of people.

Larger groups can also communicate through games. Specifically, lab experiments have been conducted, in which participants are configured into social network structures and asked to perform tasks with known outcomes. For example, graph-coloring problems performed by a group connected one way will produce different results from a similar group connected another way [38]. Similar experiments have been run online, in which members of a crowd are connected in different ways, and given the opportunity to exchange information to aid in a search task [11].

Most of the previously described work doesn't allow people to interact with each other using natural language. Such interaction can itself be experimentally controlled: for example, one person can look at input and describe it to another [39]. There is opportunity to perform such work at large scales by providing crowd members with the ability to have structured dyadic conversations with each other around a goal, such as understanding a new concept or recognizing a threat. Moreover, experiments might also look at the ways crowd members evolve their own ways of communicating. For example, studies have shown that even when verbal and iconic communication are removed, pairs of people often invent their own task specific signaling systems [40].

What other coordination techniques can be used in open-ended tasks that lend themselves to creative outputs? In a project called *SwarmArt*, each member of the crowd can add a line to a drawing, and then adjust the strength of other participants' lines [41]. Each member of the crowd can see all previous members' inputs, and each member is both a generator and an evaluator of the input to date.

In sum, there are many ways crowds can be connected, as shown in Table 2.

Open-ended creative tasks might or might not benefit from the collaboration methods just discussed. The more one sees and hears of one's neighbors, the more one might fixate on the previous solutions. On the other hand, the more one sees, the more one might be inspired to produce a novel ideas triggered by someone else's idea. How much to reveal to members of the crowd, and when, is a question open for more research. Given the use of a strong collaboration method, it is possible to create simple, flat organizations, or complex, nested organizations, the next section's topic.

Table 2. Crowd connection schemes

Connection through:	Example action	Reference
virtual proximity	Approach someone in a virtual world	[42]
an overall task	Guess the weight of an animal	[14]
seeing two examples	Combine	[22, 23]
modifying another's work	Remix a computer program	[35], [6]
being a member of a team	Collaborate with team members	[43], [8], [13]
focus on one example	Critique	[35] [36]
playing a two-person game	Guess image labels	[12]
playing a many-person game	Solve NP-Complete problems	[38], [11]
dyadic conversation	Recognize threats	[39]
discussion boards	Code software	[35], [44]

5 Levels of Hierarchy

In organizations, tasks are broken up and assigned to different people. These people can sometimes in turn break up the tasks and assign them. Multiple levels of hierarchy allow an organization to scale work, and adapt flexibly to new demands. On the other hand, such hierarchies introduce bottlenecks.

Most paid crowdsourcing uses one level of hierarchy: the task originator asks people to contribute, and the originator performs the consolidation. Since workers are usually being paid *after* performing a task, they as workers may have little appetite for creating paid tasks for others to perform, which would introduce credit risk.

Still, it is possible to imagine recursive crowdsourcing, where members of the crowd ask others to accomplish a task. This would be useful in addressing open-ended creative problems: groups could create their own strategies for performing the task, and then assemble a crowd to follow their particular strategy.

Indeed, some sites are exploring similar structures. InnoCentive allows someone to form a team around an idea, and specify up-front design the compensation division between the leader and the team members [8]. By forming a new team, a leader participates in recursive crowdsourcing. The user-initiated contests of Scratch have a similar feel: there is an overall running contest, the reward of which is to see one's project displayed on the front page. Unexpectedly, participants have created their own contests [37]. Indeed, some groups of Scratch users describe themselves as "companies" [45]. Furthermore, sites related to citizen science and volunteerism allow for subgroups to be formed, creating at least two levels of hierarchy [46]. Finally, the winning team in the red balloon contest encouraged people to recruit others to find a set of balloons, and compensated them based on the structure of the recruiting tree [47].

In order to study the use of hierarchy in crowds, one needs ways of creating recursive crowdsourcing experiments. If money is involved, then perhaps some combination of pre and post payment might be used, to both reduce credit risk and to provide incentive for the completion of subtasks. For non-monetary crowdsourcing,

ways of making announcements and coordinating work are needed, available to not just the initial requesters, but also to the workers who may want to also crowdsource their own tasks.

This area of research, while complex, would allow for the solution of complex open-ended problems that might require not just many workers, but also many different approaches.

6 Workflow

Crowdsourcing uses workflow: tasks are split up, modified, and then completed. Most companies use workflow: for example, expense vouchers are routed to managers and accounting departments for approval and payment. There has been much research on how to describe workflows formally [48, 49]. In an alternative approach, Malone et al. suggest a metaphor: the subtasks of crowdsourcing are genes, and these can be combined to form genomes [19].

This metaphor suggests that a crowdsourcing task design involves not just picking one attribute for each different dimension, but might involve picking more than one attribute that might be realized in different stages of crowdsourcing activity. For example, Threadless uses the crowd to vote on T-shirts, but then uses the site operators, presumably experts, to make the final determination of what to manufacture [7].

One area for research might involve taking the metaphorical idea of a genome, and combine it with formal definitions of workflow. Then, the design space includes different ways of assigning and sequencing recognition, generation, and evaluation tasks.

The crowd itself might be useful in designing and evaluating ways to structure work on a particular topic. For example, the crowd may first participate in two alternative processes, and then suggest ways of combining them to create a more effective process.

7 Concluding Thoughts

Crowdsourcing can become many things. Right now, the space of possible ways of coordinating large numbers of people is relatively unexplored. There is an opportunity to structure the crowd in new ways. There is also an opportunity to understand more about collective cognition by designing crowd-based experiments. Several areas seem particularly promising for research. First, the crowd can perform open-ended, creative tasks. Through this, we may solve large-scale social problems, and gain insight into social creativity. Second, crowd communication can run the spectrum from total isolation to open collaboration, and through experiments we may begin to understand how much and what kind of collaboration is fruitful for specific tasks. Third, crowd members can potentially crowdsource, creating flexible ad hoc hierarchies: unleashing this potential may be important for handling creative tasks. Fourth, crowds can be organized in new and interesting ways by selecting dimensional attributes from the design space and designing workflows that allocate and sequence tasks accordingly. Exploring this space may yield new organizations; eventually, the crowd will design itself.

Acknowledgements. This research was supported by the National Science Foundation under awards IIS-0855995 and IIS-0968561.

References

1. Howe, J.: Crowdsourcing: Why the power of the crowd is driving the future of business. Three Rivers Pr. (2009)
2. Quinn, A.J., Bederson, B.B.: Human Computation: A Survey and Taxonomy of a Growing Field. In: CHI. ACM Press, New York (2011)
3. Kittur, A.: Crowdsourcing, collaboration and creativity. Crossroads, XRDS (2010)
4. Little, G., Chilton, L.B., Goldman, M., Miller, R.C.: Exploring iterative and parallel human computation processes. In: Proceedings of the ACM SIGKDD Workshop on Human Computation, pp. 68–76. ACM, New York (2010)
5. Benkler, Y.: The wealth of networks: How social production transforms markets and freedom. Yale Univ Press (2006)
6. Gulley, N.: Patterns of innovation: a web-based MATLAB programming contest. In: CHI 2001 extended abstracts on Human factors in computing systems, p. 338. ACM, New York (2001)
7. Brabham, D.: Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. Information, Communication & Society 13, 1122–1145 (2010)
8. Jain, R.: Investigation of Governance Mechanisms for Crowdsourcing Initiatives. In: AMCIS 2010 Proceedings, p. 557 (2010)
9. Hollan, J., Hutchins, E., Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research. ACM Transactions on Computer-Human Interaction (TOCHI) 7(2), 174–196 (2000)
10. Salganik, M.J., Dodds, P.S., Watts, D.J.: Experimental study of inequality and unpredictability in an artificial cultural market. Science 311(5762), 854 (2006)
11. Mason, W.A., Jones, A., Goldstone, R.L.: Propagation of innovations in networked groups. Journal of Experimental Psychology-General 137(3), 422–433 (2008)
12. Von Ahn, L., Dabbish, L.: Designing games with a purpose. Communications of the ACM 51(8), 58–67 (2008)
13. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z.: Predicting protein structures with a multiplayer online game. Nature 466(7307), 756–760
14. Galton, F.: Vox Populi. Nature, 450–451 (1907)
15. Gurnee, H.: Maze Learning in the Collective Situation. The Journal of Psychology 3, 437–443 (1937)
16. Knight, H.C.: A Comparison of the Reliability of Group and Individual Judgments. Master's Thesis, Columbia University (1921)
17. Estes, W.K., Maddox, W.: Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. Psychonomic Bulletin & Review 12(3), 403 (2005)
18. Brooks, F.: The Design of Design: Essays from a Computer Scientist. Addison Wesley, New York (2010)
19. Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing crowds: Mapping the genome of collective intelligence. MIT Sloan School Working Paper 4732-09 (2010)
20. Boudreau, K., Lacetera, N., Lakhani, K.: Parallel search, incentives and problem type: Revisiting the competition and innovation link. Harvard Business School, Working Paper 09-041, 2008 (2008)

21. Deci, E.L., Koestner, R., Ryan, R.M.: A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125, 627–668 (1999)
22. Nickerson, J.V., Sakamoto, Y.: Crowdsourcing Creativity: Combining Ideas in Networks. In: *Workshops on Information in Networks* (2010)
23. Yu, L., Nickerson, J.V.: Cooks or Cobblers? Crowd Creativity through Combination. In: CHI. ACM Press, New York (2011)
24. Amabile, T.: *Creativity in context*. Westview Press (1996)
25. Hinds, P., Kiesler, S.: *Distributed work*. The MIT Press, Cambridge (2002)
26. Nickerson, J.V., Corter, J.E., Tversky, B., Zahner, D., Rho, Y.J.: The Spatial Nature of Thought: Understanding Systems Design Through Diagrams. In: *ICIS 2008 Proceedings*, p. 216 (2008)
27. Thagard, P.: *Conceptual revolutions*. Princeton University Press, Princeton (1992)
28. Goldberg, D.E.: *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading (1989)
29. Holland, J.H.: Building blocks, cohort genetic algorithms, and hyperplane-defined functions. *Evolutionary computation* 8(4), 373–391 (2000)
30. Kosorukoff, A.: Human based genetic algorithm. In: *IEEE International Conference on Systems, Man, and Cybernetics*, vol. 5, pp. 3464–3469. IEEE, Los Alamitos (2002)
31. Gero, J.S.: Computational models of innovative and creative design processes. *Technological Forecasting and Social Change* 64(2-3), 183–196 (2000)
32. Malone, T.W., Crowston, K.: The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)* 26(1), 87–119 (1994)
33. Dennis, A., Williams, M.: Electronic Brainstorming. In: *Group creativity: Innovation through collaboration*, pp. 160–178 (2003)
34. Nunamaker, J.F., Dennis, A.R., Valacich, J.S., Vogel, D., George, J.F.: Electronic meeting systems. *Communications of the ACM* 34(7), 40–61 (1991)
35. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B.: Scratch: programming for all. *Communications of the ACM* 52(11), 60–67 (2009)
36. Tanaka, Y., Mochizuki, T., Manalo, E., Kusumi, T.: Cultural differences between Asian students regarding judgments about using critical thinking. In: *14th International Conference on Thinking* (2009)
37. Nickerson, J.V., Monroy-Hernandez, A.: Appropriation and Creativity: User Initiated Contests in Scratch. In: *Hawaii International Conference on System Sciences* (2011)
38. Kearns, M., Suri, S., Montfort, N.: An experimental study of the coloring problem on human subject networks. *Science* 313(5788), 824 (2006)
39. Voiklis, J.: *A Thing Is What We Say It Is: Referential Communication and Indirect Category Learning*. Ph.D., Columbia University (2008)
40. Galantucci, B.: An experimental study of the emergence of human communication systems. *Cognitive Science: A Multidisciplinary Journal* 29(5), 737–767 (2005)
41. Boyd, J.E., Hushlak, G., Jacob, C.J.: SwarmArt: interactive art from swarm intelligence. In: *Proceedings of the 12th annual ACM international conference on Multimedia*, pp. 628–635. ACM, New York (2004)
42. Boellstorff, T.: *Coming of age in Second Life: An anthropologist explores the virtually human*. Princeton University Press, Princeton (2008)
43. Bell, R.M., Koren, Y.: Lessons from the Netflix prize challenge. *ACM SIGKDD Explorations Newsletter* 9(2), 75–79 (2007)

44. West, J.: How open is open enough? Melding proprietary and open source platform strategies. *Research Policy* 32(7), 1259–1285 (2003)
45. <http://blog.scratch.mit.edu/2010/01/scratch-companies-by-guest-blogger.html> (2010)
46. Raddick, J., Lintott, C.J., Schawinski, K., Thomas, D., Nichol, R.C., Andreescu, D., Bamford, S., Land, K.R., Murray, P., Slosar, A.: Galaxy Zoo: an experiment in public science participation. *Bulletin of the American Astronomical Society* 38, 892 (2007)
47. Pickard, G., Rahwan, I., Pan, W., Cebrian, M., Crane, R., Madan, A., Pentland, A.: Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy. *J Arxiv preprint arXiv:1008.3172* (2010)
48. van Der Aalst, W.M.P., Ter Hofstede, A.H.M., Kiepuszewski, B., Barros, A.P.: Workflow patterns. *Distributed and parallel databases* 14(1), 5–51 (2003)
49. Milner, R.: *Communicating and mobile systems: the pi-calculus*. Cambridge University Press, Cambridge (1999)