Toward a Quantitative Understanding of Teamwork and Collective Intelligence

Andrew Mao
Microsoft Research, NYC
mao@microsoft.com

Abstract

Teamwork and collective intelligence drive the production of goods and knowledge in our society and are an integral part of our increasingly networked lives. However, we are still far from understanding the causal mechanisms that underlie effective organization and communication in teams, and the study of human factors in many fields of computer science has tended toward abstractions of behavior over the understanding of fundamental social processes. I argue that the rapid growth of socio-technical systems on the Internet presents both an impetus and an opportunity for a more causal approach to understanding teamwork and collective intelligence. We can approach this goal through broader deployment of experiments, fostering closer ties between theory and empirical work, and by bridging the gap across different disciplines.

Introduction

The effectiveness of teamwork is fundamental to many forms of economic production and collective problem solving, and has become more common in everyday life through the way that the Internet connects people together. While traditional economic production took place through the industrial creation of physical goods in firms, today we also see collective knowledge gathered in socio-technical systems that would have been hard to imagine prior to their existence—Wikipedia, an encyclopedia curated and updated in real-time by volunteers; content aggregation and dissemination on social systems like Reddit, Twitter, and Facebook; and free, open-source software produced and maintained through decentralized processes. More than ever, much of social organization is not just characterized by top-down hierarchy, but also *collective intelligence*: the idea that groups of people can solve problems and produce useful content without any explicit central authority.

Yet, despite many decades of social science research and even the recent emergence of *computational social science* (Lazer et al. 2009), one important aspect that is largely missing is a deep understanding of the processes and mechanisms that allow teams and organizations to work together effectively (Watts 2013). A review of the study of social science reveals many competing theories. On the one hand, teamwork often enhances productivity through returns to specialization (Becker and Murphy 1992), learning from others (Mason and Watts 2012), and increased social engagement or social facilitation (Baek and Shore 2015). That is, except when teamwork reduces individual productivity—due to reduced effort from free riding (Latane, Williams, and Harkins 1979; Karau and Williams 1993), the overhead of coordination as groups become larger (Steiner 1972), and unintelligent behavior due to groupthink or herding behavior (Janis 1972; Anderson and Holt 1997) that inhibit the exchange of perspectives and information. The recent DARPA Request for Information (RFI) in 2015[1] advocates for renewed research of causal processes in social systems, noting how challenges of existing social science approaches have

> *... contributed to the significant gap between [social, behavioral, and economic science] theories and models and the actual patterns of emergent behavior of social systems as documented in the real world.*

The presence of many competing theories means that design and interventions for establishing effective teams are difficult to identify in practice. Observational studies of teams that are working together have a multitude of hypotheses to support any potential observation, leading to post hoc reasoning that any particular outcome should have been obvious to begin with (Watts 2014). On the other hand, many experimental studies of teams have so far been relegated to abstract, stylized tasks such as clapping, shouting, (Latane, Williams, and Harkins 1979) or rope-pulling (Ingham et al. 1974) that bear little resemblance to the real world. Moreover, many studies of teamwork and organization have been focused on the traditional economic organization of the firm, rather than novel types of decentralized, collective structure found online. As we often describe successful Internet systems by example rather than through any theoretical understanding, the design of many online socio-technical systems has largely been a process of trial and error.

Perhaps more concerning than the lack of a principled way to design social systems is the tendency to further abstract the human element in domains where people are a

[1] *New Capabilities for Experimental Falsifiability in Social, Behavioral and Economic Sciences*: See https://www.fbo.gov/spg/ODA/DARPA/CMO/DARPA-SN-15-71/listing.html.

fundamental component. A typical approach in computer science, this is exemplified by much of the work in microtask crowdsourcing, which assumes that humans can be modeled as noisy computers that make errors with some probability (e.g. Bachrach et al. 2012), and occasionally respond rationally to economic incentives but with many exceptions (Mason and Watts 2009; Yin, Chen, and Sun 2013; Ho et al. 2015). Furthermore, much work in human computation focuses on how *workflows* can be used to decompose complex problems into small tasks that can be done by workers without specialized skills in a short amount of time (Dai, Mausam, and Weld 2010; Kulkarni, Can, and Hartmann 2012).

An obvious question to ask in light of this work is whether *people* can simply be reduced via a layer of abstraction to noisy computers motivated by money. Perhaps the microtasking and workflow approach to collective problem solving effectively dismisses benefits from people being able to interact, coordinate, and collaborate on different tasks—fundamentally a building block of many collectively intelligent systems. Early experiments in collective intelligence have shown that groups of humans can collaborate and benefit from coordination mechanisms for working together and sharing knowledge, both in person and through electronic media (Woolley et al. 2010; Engel et al. 2014). There is further evidence that intelligent software systems can mediate and streamline communication within teams that are tackling a complex task (Zhang et al. 2012; Kim et al. 2013). By ignoring the social aspects of team collaboration, we may miss some of the intrinsic benefits of teamwork, as well as opportunities to further enhance and refine it.

Opportunities in online social systems

Currently, we are at an important crossroads in the study of teamwork and social organization. Not only are current theories in social science often insufficient for an understanding of how teams work together effectively, they are even more ill-suited for designing social systems on the Internet, where organization takes place in novel ways that are less well understood. However, the process of designing such systems and the data they capture also offer an opportunity to both engage in principled study of human behavior and contribute to a new wave of social science. Instead of using more layers of abstraction when people appear, we can instead tackle headfirst the challenge of understanding individual and group behavior at a deeper level.

In order to understand how to improve the effectiveness of teamwork, and of collective social systems at large, we need to be able to answer questions of the following form in a quantitative way:

> If we apply intervention **X** to a team working on task **Y**, how much more effective does the team become, and why?

Intervention **X** may take on many forms, such as a subtle priming of individuals, changes to the organization or communication structure of a team, or even extensive software facilitation of a certain aspect of teamwork. However, the fundamental property of an answer to this question is that it is *causal*: we can quantify the effect that **X** had on the performance of a team, and moreover through which mechanisms it took place. For example, we might imagine that arranging a large team to organize into smaller groups focusing on sub-tasks improves individual productivity by reducing the amount of coordination overhead required of each participant.

Answers to questions like these are fundamental to our ability to design socio-technical systems that function as intended, rather than being continually redesigned through a process of trial-and-error. In areas such as microtask crowdsourcing and human computation, people—rather than being a problem to abstract away—actually present an opportunity to study fundamental questions of social behavior. An understanding of the causal mechanisms that affect team performance will generalize and carry over to many different settings, and in turn allow for answers to broader questions such as how the decentralized organization emerges in collective intelligence settings, and why some teams and organizations are more effective than others.

Approaches to studying teamwork

The gold standard for establishing causality is the randomized controlled experiment, because of its ability to systematically eliminate unknown sources of variance across treatment groups—thus establishing treatment conditions as the only source of variance (Gerber and Green 2012). In the past, social science experiments on organization and communication in teams have often been limited in size and scope by the physical laboratories they are conducted in, leading to stylized tasks done by teams of limited size, and were eventually abandoned due to lack of progress—see (Shure and Meeker 1970) for one of many examples. Zelditch (1969) summarizes the era by asking the rhetorical question of whether large groups such as an "army" could be studied experimentally, concluding that it was both unnecessary and infeasible.

However, the growing prominence of online social systems and our primitive understanding of them provides not only an impetus, but an opportunity to study teamwork and collective intelligence experimentally. Software systems connect together many more people than could possibly be contained in physical labs, while also making it easier to collect data at scale. For example, Facebook has pioneered the technique of doing large-scale experiments within social networks (Eckles, Karrer, and Ugander 2014). In general, it is natural to study online organization by the Internet itself as a lab, as it both scales effectively and allows for generalization (Bainbridge 2007). As web applications routinely connect dozens or hundreds of users together in many different types of tasks, they will allow for new studies of behavior that were previously considered to be methodologically impossible. Our work, for example, uses the digital volunteer setting of crisis mapping as a novel approach to studying self-organization in teams of different size (Mao et al. 2016). We use this task to simultaneously test different theories of team effectiveness in a realistic way, and find that although social loafing and coordination costs result in reduced contribution from individuals in larger teams, the potential benefits of coordination can outweigh this loss in raw productivity.

Interdisciplinary research programs will be an integral part of studying teamwork. While there are many disciplines effectively studying human behavior, there is a distinct tendency for each field to hold to narrow tribal affiliations, focusing on the methods or discipline rather than the problem itself (Watts 2013). However, there is a significant benefit for researchers to overcome these barriers and understand the tools and perspectives employed by other fields. For example, many computer scientists are relatively new to the methods and practice of studying collective human behavior, and can learn a great deal from both aforementioned theories in the social psychology literature and about principles of proper experiment design, such as (Gerber and Green 2012).

Another important approach to studying teamwork is to more closely integrate the exchange of theory and empirical work. When purely theoretical work proceeds independently of empirical verification, we risk solving poorly framed problems or developing models that are far removed from reality. At the same time, conducting experiments without theory to guide hypotheses can lead to poorly designed experiments or worse, facilitate speculation of spurious post hoc hypotheses after observing the data. On the other hand, theory can be used to guide the design of experiments, the outcome of which can in turn inform better theories and models of behavior. This cycle can produce better experiment designs as well as more grounded models for teamwork, such as Kleinberg and Raghu (2015).

Conclusion

Through understanding the causal mechanisms that underlie team performance, we will gain a principled approach to design social systems that function as intended, and to enhance teamwork beyond existing forms of hierarchical organization or decentralized collective intelligence. I envision that a fruit of this vision will allow us to answer questions like the following:

1. **Optimizing beneficial versus detrimental effects.** How do we construct a team environment where social incentives stimulate productivity through engagement and learning rather than social loafing or decreased willingness to contribute?
2. **Computational team formation.** Given a large set of potential team members to choose from, how can we combinatorially assemble an "ideal team" to work on a given type of task? In other words, can we model and predict team performance from individual attributes and relational characteristics?
3. **Software facilitation of teamwork.** How can we design user interfaces and software systems that enhance information sharing and communication for teamwork while minimizing the disruption from coordination costs?

Although these are ambitious topics that are far beyond the reach of any one person or research group, I hope that the discussion from this workshop will steer us down a path toward answering these important questions with far-reaching consequences.

References

Anderson, L. R., and Holt, C. A. 1997. Information cascades in the laboratory. *The American economic review* 847–862.

Bachrach, Y.; Graepel, T.; Minka, T.; and Guiver, J. 2012. How to grade a test without knowing the answers—a bayesian graphical model for adaptive crowdsourcing and aptitude testing. *Proceedings of the 29th International Conference on Machine Learning (ICML)*.

Baek, J., and Shore, J. 2015. Group size and effort per person: a MOOC field experiment. *Working paper*.

Bainbridge, W. S. 2007. The scientific research potential of virtual worlds. *science* 317(5837):472–476.

Becker, G., and Murphy, K. 1992. The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics* 107(4):1137–60.

Dai, P.; Mausam; and Weld, D. S. 2010. Decision-theoretic control of crowd-sourced workflows. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*.

Eckles, D.; Karrer, B.; and Ugander, J. 2014. Design and analysis of experiments in networks: Reducing bias from interference. *arXiv preprint arXiv:1404.7530*.

Engel, D.; Woolley, A. W.; Jing, L. X.; Chabris, C. F.; and Malone, T. W. 2014. Reading the mind in the eyes or reading between the lines? theory of mind predicts collective intelligence equally well online and face-to-face. *PloS one* 9(12):e115212.

Gerber, A. S., and Green, D. P. 2012. *Field Experiments*. W. W. Norton.

Ho, C.-J.; Slivkins, A.; Suri, S.; and Vaughan, J. W. 2015. Incentivizing high quality crowdwork. In *Proceedings of the 24th International World Wide Web Conference (WWW)*.

Ingham, A. G.; Levinger, G.; Graves, J.; and Peckham, V. 1974. The ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology* 10(4):371–384.

Janis, I. L. 1972. Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes.

Karau, S. J., and Williams, K. D. 1993. Social loafing: A meta-analytic review and theoretical integration. *Journal of personality and social psychology* 65(4):681.

Kim, J.; Zhang, H.; Andr, P.; Chilton, L.; MacKay, W.; Beaudouin-Lafon, M.; Miller, R. C.; and Dow, S. P. 2013. Cobi: A community-informed conference scheduling tool. In *Proceedings of the 26th Symposium on User Interface Software and Technology (UIST)*.

Kleinberg, J., and Raghu, M. 2015. Team performance with test scores. In *Proceedings of the 16th ACM Conference on Electronic Commerce (EC)*, 511–528. ACM.

Kulkarni, A.; Can, M.; and Hartmann, B. 2012. Collaboratively crowdsourcing workflows with turkomatic. In *Pro-*

ceedings of the 15th ACM Conference on Computer Supported Cooperative Work (CSCW).

Latane, B.; Williams, K.; and Harkins, S. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of personality and social psychology* 37(6):822.

Lazer, D.; Pentland, A. S.; Adamic, L.; Aral, S.; Barabasi, A. L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. 2009. Life in the network: the coming age of computational social science. *Science (New York, NY)* 323(5915):721.

Mao, A.; Mason, W.; Suri, S.; and Watts, D. J. 2016. An experimental study of team size and performance on a complex task. *Working paper*.

Mason, W., and Watts, D. J. 2009. Financial incentives and the "performance of crowds". In *Proceedings of the 1st Human Computation Workshop (HCOMP)*, 77–85.

Mason, W., and Watts, D. J. 2012. Collaborative learning in networks. *Proceedings of the National Academy of Sciences* 109(3):764–769.

Shure, G. H., and Meeker, R. J. 1970. A computer-based experimental laboratory. *American Psychologist* 25(10):962.

Steiner, I. D. 1972. *Group Process and Productivity*. Academic Press Inc.

Watts, D. J. 2013. Computational social science: Exciting progress and future directions. *The Bridge on Frontiers of Engineering* 43(4):5–10.

Watts, D. J. 2014. Common sense and sociological explanations. *American Journal of Sociology* 120(2):pp. 313–351.

Woolley, A. W.; Chabris, C. F.; Pentland, A.; Hashmi, N.; and Malone, T. W. 2010. Evidence for a collective intelligence factor in the performance of human groups. *science* 330(6004):686–688.

Yin, M.; Chen, Y.; and Sun, Y.-A. 2013. The effects of performance-contingent financial incentives in online labor markets. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI)*.

Zelditch, Jr., M. 1969. Can you really study an army in the laboratory? *A sociological reader on complex organizations* 528–39.

Zhang, H.; Law, E.; Miller, R.; Gajos, K.; Parkes, D.; and Horvitz, E. 2012. Human computation tasks with global constraints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, 217–226. New York, NY, USA: ACM.