Collaborative Problem Solving: A Study of MathOverflow

Yla R. Tausczik
Human-Computer Interaction Institute
Carnegie Mellon University
ylataus@cs.cmu.edu

Aniket Kittur
Human-Computer Interaction Institute
Carnegie Mellon University
nkittur@cs.cmu.edu

Robert E. Kraut
Human-Computer Interaction Institute
Carnegie Mellon University
robert.kraut@cmu.edu

ABSTRACT
The Internet has the potential to accelerate scientific problem solving by engaging a global pool of contributors. Existing approaches focus on broadcasting problems to many independent solvers. We investigate whether other approaches may be advantageous by examining a community for mathematical problem solving – MathOverflow -- in which contributors communicate and collaborate to solve new mathematical "micro-problems" online. We contribute a simple taxonomy of collaborative acts derived from a process-level examination of collaborations and a quantitative analysis relating collaborative acts to solution quality. Our results indicate a diversity of ways in which mathematicians are reaching a solution, including by iteratively advancing a solution. A better understanding of such collaborative strategies can inform the design of tools to support distributed collaboration on complex problems.

Author Keywords
Problem solving; scientific collaboration; crowdsourcing; Q&A sites

[redacted]

INTRODUCTION
The Internet has enabled scientific problem solving on a global scale, where individuals can contribute their expertise to solve challenging problems in domains ranging from finding a red balloon [38] to software engineering [25] to R&D [16]. Increasing our collective ability to tackle such problems could significantly impact progress in science, technology, and innovation. Many approaches, such as Innocentive, the Climate Collaboratorium, or TopCoder, focus on broadcasting a scientific problem to many contributors under the assumption that at least one may have a valuable solution. In such systems the "solvers" are generally individuals or small teams who work independently [16,25].

However, despite the large number of "open call" approaches to scientific problem solving online, examples of deeply interactive and collaborative problem solving on the internet remain few and far between. Notable exceptions in which contributors communicate with and build on each others' work include the Polymath Projects and MathOverflow, in which anyone with an internet connection can be involved in the collaborative solution to an unsolved mathematics problem. Such collaborative problem solving approaches raise challenges when involving many contributors, including reaching a shared understanding of the problem, which itself may be ill-defined, decomposing an interdependent problem into subproblems, and coordinating the efforts of many contributors, each with varied expertise and commitment levels. Previous work has begun to characterize such collaborative problem solving communities, describing the leadership structure and showing that even peripheral contributors make meaningful contributions [6]. In this paper we contribute a detailed process-level understanding of the collaborative activities that happen in one such community, MathOverflow, and quantify the effects of different collaborative activities on solution quality. Our results have implications for systems aimed at supporting large-scale, collaborative scientific problem solving.

RELATED WORK
Collaboration in science is becoming even more important, and our ability to tackle scientific questions may be aided by online systems to improve and open up scientific problem solving. We know from the literature on small group problem solving and peer-production that carefully designing and structuring interactions is important to elicit good work. Yet, it is not clear how to structure interactions for the specific domain of scientific problems. Below we describe this related work and why it motivates our process-level examination of MathOverflow.

Scientific Collaboration
Studies of global patterns of scientific collaboration have shown that science is increasingly being driven by collaboration [7, 43], with more multi-authored grants and papers being published more [14] with more successful outcomes than single-authored papers: they are cited more often and are published in more prestigious journals [31, 41]. However, these studies have focused on global patterns of

CSCW '14, February 15–19, 2014, Baltimore, Maryland, USA.

ACM 978-1-4503-2540-0/14/02$15.00.
http://dx.doi.org/10.1145/2531602.2531690

collaboration in science, such as who collaborates with whom, and have not investigated the mechanics of the collaboration process -- e.g., how scientists jointly pose questions, develop ideas, and refine solutions.

Another branch of social studies of sciences have examined the cognitive processes that take place among scientists working together. For example, Dunbar and colleagues identified the use of analogy as an important cognitive process in developing hypotheses in a new experimental domain [11]. However, these studies typically focus on the practices in a single lab or a few labs and are limited to explicit discussions between scientists rather than following the entire trajectory of an idea (e.g., [26]).

Finally a third branch of studies has questioned the degree of true collaboration in science. They find that scientific collaboration often involves large multi-institutional collaborations, in which there is little communication and coordination [9]. Institutional barriers and reward structures impede the ability for true large-scale collaboration even when co-authoring papers, or sharing data or equipment [9]. Entrenched status hierarchies and competition among individuals for recognition create non-collaborative environments in which being helpful and making small contributions without acknowledgement as an author is costly [2]. Even physical distance can present a barrier to traditional forms of collaboration [29]. Current scientific collaboration may only make use of limited forms of collaboration that in particular don't scale.

Small-Group Problem Solving

The costs and inefficiencies of communicating in small groups highlight the importance of structuring collaboration and assistive technology. Early work on small group decision-making found discussions to be fraught with unexpected inefficiencies and biases [19]. Despite expected gains from multiple-perspectives and unique expertise, experiments in which information was distributed among group members showed that groups spent the majority of time discussing common knowledge rather than unique knowledge [36]. Structuring the group's interaction, for example by making individual's area of knowledge known, has been shown to reduce some of the inefficiency [37]. The medium through which a group communicates can also reduce inefficiencies; using computer-mediated communication has been shown to keep students' problem-solving discussions more task-focused [18].

Collaborative Sensemaking, Peer-production, and Crowdsourcing

Peer production and crowdsourcing have emerged as powerful mechanisms for sharing information (e.g. Usenet [42], StackOverflow [27]), building rich artifacts (e.g. open-source software [10], Wikipedia [20]), and accomplishing complex tasks (e.g. CrowdForge [22]) by facilitating collaboration among many individuals [5]. These systems highlight the importance of structuring collaboration to reduce the costs of coordination [20], focus discussion [23], and represent knowledge in ways that facilitate making inferences [34]. Below we focus on two strategies that are most relevant to collaborative problem-solving on MathOverflow: broadcasting difficult problems to a large audience of potential experts and dividing work into sub-problems to involve a large number of people while keeping coordination costs low.

Platforms such as Innocentive, TopCoder provide good models of broadcasting difficult problems to a large crowd [16, 25, 24]. These systems work by running competitions among individuals or small teams to solve complex, difficult problems. They have been very successful at solving difficult problems quickly by reaching individuals who have the insight or unrealized expertise to solve the problem. However, these systems primarily focus on a broadcast model with little communication or coordination between individuals or small teams of individuals.

Another strategy focuses on decomposing a problem into subtasks that can be assigned to different workers. Crowdsourcing systems such as CrowdForge, Soylent, and Turkomatic, make use of a crowd of novices to do complex tasks by breaking them into sub-tasks and using workflows to handle dependencies between tasks [22, 3, 24]. These implementations have successfully enabled novices to do complex tasks that require interdependence, like writing an article, but are challenging for more complex unsolved problems in which the problem may be underspecified, there is no clear way to go about solving the problem, and partial answers may be incorrect, incomplete, and not on a path to a final solution.

Polymath and MathOverflow

Mathematicians have started to make use of user-generated content platforms to build large-scale collaborations to solve research-level mathematics problems. The Polymath Projects began as an attempt by Timothy Gowers, a senior mathematician at Cambridge University, to create massively collaborative mathematics [6]. During the Polymath 1 Project a new proof to an important theorem, was solved in 155 involved blog posts made by a total of 39 unique users [6]. On MathOverflow (mathoverflow.net), an online community of mathematicians, members solve small novel mathematical problems or "micro-problems". By micro-problems we mean relatively small problems which may not themselves qualify as publishable results, but are nonetheless novel to the mathematicians involved, enriching the participants' own knowledge and work practices and occasionally contributing to published research.

While past research has described these projects and attempted to evaluate whether everyone really makes meaningful contributions [6,39,40] no one has yet examined contributions to these collaborations at a process-level. In this paper we develop a deeper, process-level understanding of collaboration by examining individual contributions to

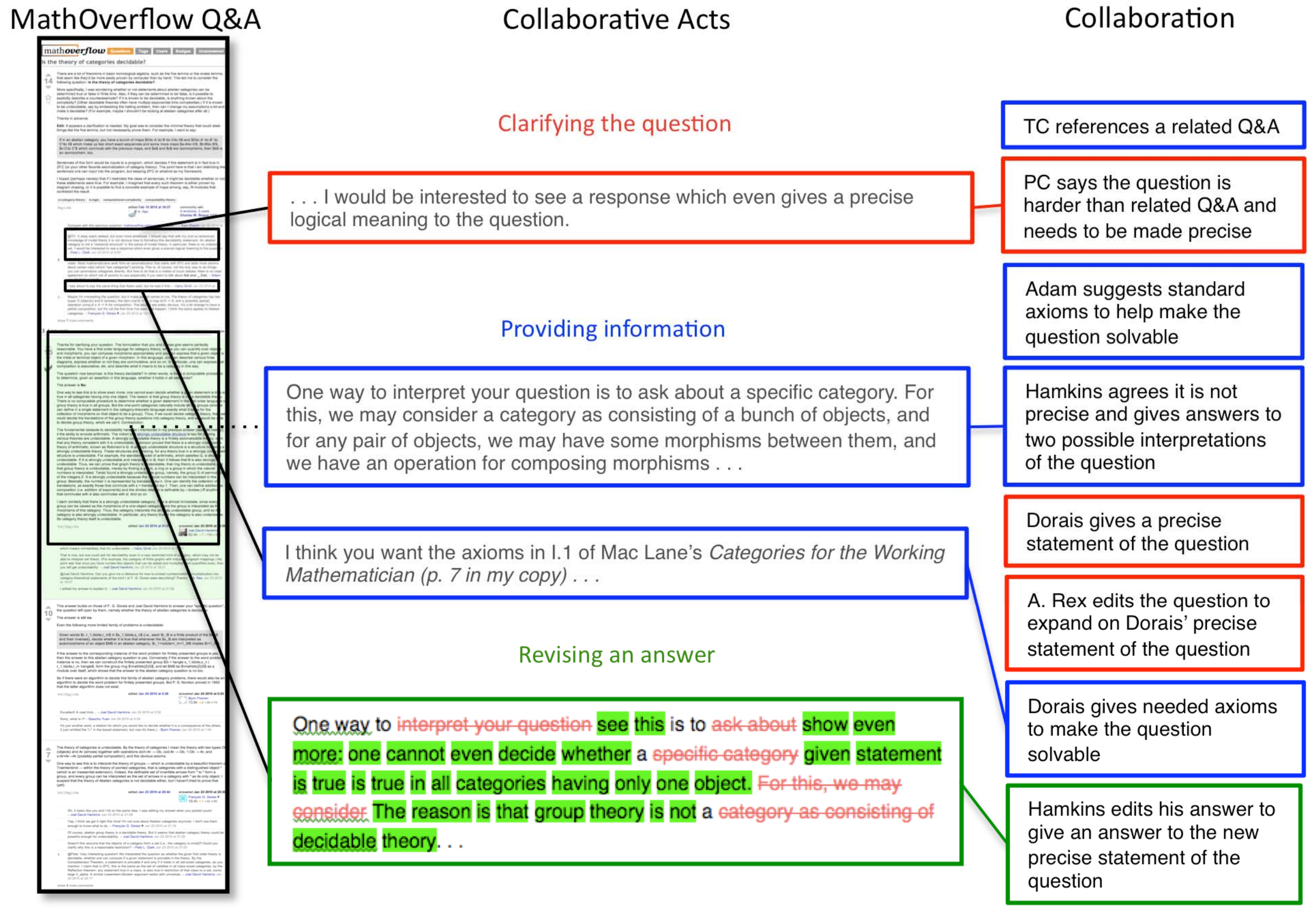

Figure 1: MO question "Is the theory of categories decidable?" The left column shows the entire MO Q&A (comments are threaded within the question and answers). The middle column shows a few examples of collaborative acts (excerpts). The right column shows a description of early contributions in chronological order and shows how a collaboration unfolds.

solving a problem on MathOverflow. MathOverflow provides rich archival data to survey many examples of distributed collaborative problem solving in practice.

STUDY OVERVIEW

As discussed above, while large-scale collaborative scientific problem solving has significant potential, little research has to date examined the low-level mechanisms by which collaboration takes place and their impact on the resulting quality of the solution. In the studies below we begin to build up our understanding of online collaborative problem solving at a process level, and investigate the impact of those processes on solution quality. In Study 1 we use a grounded theory approach on a sample of MathOverflow questions to build up a taxonomy of "collaborative acts" which describe the collaboration processes involved. We leverage structured interviews with active participants in the collaborations to provide insights on how the collaborative acts fit together and lead to the development of a final solution. In Study 2 we use this taxonomy and insight from interviews to quantify the impact of collaboration processes on resulting solution quality. Finally, we discuss implications of our results for the design of large-scale collaborative problem solving systems.

RESEARCH SITE

MathOverflow (MO) is a mathematical question and answer site started in September of 2009. Contributors ask and answer research level mathematics questions through an implementation of the StackExchange platform (stackexchange.com). MO is primarily an academic community [40] and is viewed as a professional outlet for mathematics research.

MO provides a number of advantages that make it an ideal platform for examining collaborative problem solving. (1) The problems require higher order cognition to solve. Many prior studies of collaborative problem solving are artificial in the sense they examine the process by which groups go about solving well-defined problems with pre-defined solutions, such as insight problems or constraint satisfaction problems [18]. However, many real-world problems that are done collaboratively such as scientific discovery, design, or innovation are difficult and ill-defined. (2) Because all the communication in MO is text-based and archived, it preserves the record of the collaboration practices used by ad hoc groups, practices which are often invisible in studying other types of collaborations. MO shares this quality with other recent online collaborations such as the Polymath Projects [6]. However, MO has further benefit lacking in these other collaborations (3) MO of hosting scores of groups and providing many data points.

As of September, 2011 there were 22,531 questions posed on MO. The questions on MO are very different from other Q&A sites. The intention of the site is for researchers to ask and answer small, novel problems that arise in doing mathematics research. Thus, questions are difficult enough that they require original work, but not so difficult that they are unsolvable. Successful questions on MO are small steps in the process of building a larger mathematical contribution. For example, a MO Q&A might represent one section of a mathematical publication. Despite being small problems, most of the problems are difficult and only experts in the field can solve them. At the time of the writing of this paper 66 published journal article or preprints acknowledged contributions made on MO.

STUDY 1: IDENTIFYING COLLABORATIVE ACTS

How does collaborative problem solving occur on MathOverflow? And what strategies are most successful? There are a number of possibilities, ranging on a spectrum of interdependence to independent. At one end of the spectrum there is a somewhat romantic view of collaborative problem solving in which there is a special connection between collaborators and ideas emerge from deep conversations. This view would suggest a highly interactive and interdependent process of collaboration on MathOverflow.

On the other end of the spectrum, MathOverflow might succeed because it is a good platform for broadcasting a problem to a large enough crowd of people where at least one person with appropriate expertise can solve the problem by him or herself. Indeed, its structure as a Q&A site (based on the StackOverflow Q&A system) might suggest exactly this: users submit independent answers to questions, with more popular answers receiving upvotes.

To understand global patterns of collaboration on a problem, in Study 1, we first examine each contribution to a problem individually. We categorize contributions into what we call types of collaborative acts, or a single contribution contributing to a larger collaboration. For example, a collaborative act might be providing a critique of an answer, in which the larger collaboration is made up of many answers, critiques, and revisions. We find that features supporting meta-commenting (i.e., commenting on others' answers, or even on the question itself) provide an avenue for more interactive collaboration. Indeed, as we will discuss below, such features support a rich and varied methods of collaborating. By leveraging interviews with the most active members of the site we provide context to understand the collaborative acts we identified as well as uncovering ways that the acts may lead to better answers.

Background

Collaboration on MO begins with a question. For example, in Figure 1, A. Rex asked the question "Is the theory of categories decidable?" which he explained in a few paragraphs (MO question 12732). His question was a broad big-picture question. In order to be addressed it required an exploration and elaboration of the relevant mathematical concepts.

The original question was ambiguous and too broad in scope. Early respondents, in trying to provide information to solve the problem, pointed to problems with phrasing. For example, PC, a professor, mentions in a comment: *"it is not obvious how to formalize this decidability statement"* and says *"I would be interested to see a response which even gives a precise logical meaning to the question."*

Joel David Hamkins, another professor, reiterated PC's concerns in the first proposed solution. *"I believe that there are several reasonable interpretations, totally different in nature"*. He then outlines two interpretations of the question and provides solutions to both interpretations.

Next, the problem began to take shape. Following attempts at clarifying the question, François G. Dorais a professor, provided what was being requested. He wrote *"Maybe I'm misreading the question, but it made perfect sense to me."* He then gave a precise logical statement of the question. The original question asker A. Rex in a major edit to the question expanded on this new understanding of the problem. At this point, although the problem was finally taking shape, it was still not clear that this was the right way to think about the problem.

Hamkins and Dorais then grappled with whether this was the best understanding of the problem as they worked to solve it. Hamkins wrote *"@Dorais: Yes, I think this must be the best way to think about it"* and then connected the new understanding of the problem to part of his original solution. Dorais then countered *"@[Hamkins]: I don't think it's that easy"*. The back-and-forth continued and eventually Hamkins and Dorais simultaneously posted solutions that were conceptually equivalent.

Although, Hamkins and Dorais posted solutions that answered the general case of the problem, they were unable to answer the more specific case asked by A. Rex. The specific case required different mathematical machinery in an area in which neither was an expert. A day later this specific case was filled in by Bjorn Poonen, a professor with more expertise in that area. At the beginning of his answer he wrote *"This answer builds on those of F. G. Dorais and Joel David Hamkins to answer your 'specific question', the question left open by them"*. Finally, the problem was fully solved.

By the end, 9 users made contributions, with substantial solutions from 3 users, Hamkins, Dorais and Poonen. We will return to this example later. We analyzed a random sample of 150 MO collaborations similar to this one to conduct our study.

Methods

Open Coding

Because there is little systematic and comprehensive research on the mechanisms by which collaborative problem solving takes place, a bottom-up approach was taken to fully explore the possible ways that individuals collaborate. To begin we focused on the most basic element of collaboration, a single contribution, which we call ***a collaborative act***. Collaborative acts were identified using an open coding by multiple coders. A deeper understanding of the collaborative acts was developed by conducting semi-structured interviews.

A grounded theory approach was used to identify collaborative acts inductively [13]. 150 collaborations on MO, which encompassed 737 contributions, were studied in detail by four independent judges.

Question-answer posts were randomly selected from the publicly accessible database dump generated September 2011. These posts were screened to ensure they met the inclusion criteria. Those questions that were critiqued by community members as too easy (e.g. closed, redirected to a lower level site), related to the discipline of mathematics but not a mathematics problem (e.g. career advice, questions about teaching math), or requests for known information (e.g. reference request, or a request for what is known in an area) were excluded. This left research level questions that were about a specific problem. The first 150 that met the criteria were selected for the final coding.

Four coders (including the first author) coded the contributions. Each coder had undergraduate level mathematics expertise. Two of the raters had or were working on a B.A. in mathematics and had taken upper division undergraduate coursework in mathematics. Two of the raters were M.A. level engineering students and had taken lower division undergraduate coursework in mathematics. Each contribution was coded by at least 3 raters. Every comment, answer, or edit was considered as a potential contribution to a collaboration. In total there were 737 contributions. Coders viewed each contribution individually and in the order they had been published on MO using a specially designed website.

During open-coding many sub-categories were identified, such as providing a complete answer (e.g. proof), supporting the previous contribution, and changing the focus of the question. In later phases of coding these subcategories were grouped into larger categories and a questionnaire was developed to help code contributions. Classification into concrete sub-categories was used as a technique to guide the creation of more abstract categories to describe collaborative acts, and to reinforce category definitions to ensure inter-rater reliability was high. Using this questionnaire the coders rated whether the contribution:

1. provided information (“Does this contribution provide new information that contributes or could contribute to a solution, whether or not correct*?”)*

2. clarified the question (“Does this contribution modify the original question or ask that it is clarified?”)

3. critiqued an answer(“Does this contribution evaluate, ask for clarification?”)

4. revised an answer (“Does this contribution improve on, or answer a related question about all or part a previous contribution?”)

5. extended an answer (“Does this contribution contribute a new answer that incorporates previous ideas?”)

For each question about the contribution the coders rated the question on a 5 point Likert scale from 0 to 4 labeled as ‘Not at all’ to ‘Extensively', as well as indicating the category subtype if applicable (e.g. “It provides a complete answer (e.g. proof)”. Collaborative acts were rated on a 5-point scale to assess the magnitude of the contribution for use later in quantitative analyses. There was significant variation in the magnitude of contributions; for example, one contribution might provide a useful definition while another might provide a complete proof, these were rated as 1 and 4 respectively on a 5-point scale. In the questionnaire, critiquing and revising an answer were grouped together; in the final analyses they were separated using subcategory ratings as it became apparent that the two represented distinct types of contributions.

Semi-structured Interviews

The first author also conducted semi-structured interviews with active MO contributors to better understand the collaborative acts, the role they played in the collaborations, and how they contributed to the development of a final solutions. Thirty-five of the most active users from MO who had listed an email address or a website with an email address on their MO profile page were contacted by email. All of these community members had made at least 350 contributions. Of those contacted 22 agreed to be interviewed, and 16 were eventually interviewed. The most active contributors were contacted because they have the most familiarity with the site and have been exposed to a variety of different collaborations.

The interviewees were all male; they were current Ph.D. students or graduates (n = 6 interviewees), postdoctoral fellows (n = 1), or professors (n = 9) in mathematics. Interviewees were asked a series of questions, including to briefly explain why they contributed to MO, to discuss the kinds of collaboration that they had observed on MO, and about specific contributions from their personal collaborations on MO. The interviews were conducted by phone or skype call with at minimum audio (n = 9); instant messenger (n = 2); or email (n = 5). The interviews by phone and skype were recorded and transcribed (M = 42 min); the interviews by instant messenger (M = 1,803 words) and email (M = 980 words) were saved.

The interviewer presented interviewees with examples of contributions from Q&A collaborations that the interviewees had participated in. Interviewees were sometimes the author of the selected contribution and sometimes peripherally involved. Interviewees were asked to explain in detail the process by which the question was solved and how the specific contribution fit into the final solution. Transcripts of the interviews were divided into sections based on the type of collaborative act represented by the specific contribution and quotes were pulled from the transcripts to illustrate different ways interviewees described the collaborative act and how it fit into the larger search for a solution.

Results

Collaborations on MO questions were small to moderate in size (1 to 14 distinct contributors per question, M = 4.23, Mdn = 4). However, many more users viewed a question than contributed (42 to 2,993 views, M = 384, Mdn = 298.5). Almost all questions received some contributions (0 to 39 comments, answers, and edits, M = 9.51, Mdn = 7). This work usually resulted in at least one answer (0 to 8 answers, M = 1.68 answers, Mdn = 1). Only 5 out of 150 questions in the corpus received no answers.

Five categories of contributions emerged; evidence from coding confirmed that the identified collaborative acts form a good taxonomy. Independent raters were able to reliably identify the collaboration acts. Intraclass correlations were calculated as a measure of inter-rater agreement among the 3 raters; agreement was moderate to high (see Table 1). This suggests that the collaboration acts represent real and detectable ways in which individuals are working together.

The collaborative acts were common enough to describe the important ways in which individuals contribute. Each type of collaborative act identified was present in at least 20% of question-answers examined and at most 91% of question-answers (see Table 1). While the collaborative acts are conceptually distinct, representing unique categories developed during open-coding, in practice they often co-occurred in solving the same problem. All of the collaborative acts were highly correlated with providing information ($r(148) = 0.44 - 0.69$, $p < 0.001$). Controlling for the relationship between information and collaboration, critiquing an answer was correlated with revising an answer and clarifying the question (*partial* $r(147) = 0.39$, $p < 0.001$; *partial* $r(147) = 0.20$, $p = 0.01$); and extending an answer was correlated with revising an answer and clarifying the question (*partial* $r(147) = 0.20$, $p = 0.02$; *partial* $r(147) = -0.30$, $p = 0.001$).

Collaborative Acts	Inter-rater agreement (ICC)	Frequency
Providing information	0.887	91.3%
Clarifying the question	0.817	38.0%
Critiquing an answer	0.829	45.3%
Revising an answer	0.733	42.7%
Extending an answer	0.651	22.0%

Table 1: Collaborative acts, building blocks of collaborations, and the percent of the 150 Q&A collaborations in which they are present.

Semi-structured interviews with frequent contributors provided confirmation that the collaborative acts identified by non-participants were sensible to active participants as well as a deeper understanding of when and how the collaborative acts contributed to a final solution. Below we describe the identified collaborative acts with descriptions from interviews and examples from the dataset to provide context.

Providing Information: was the most common type of contribution; 91.3% of questions received a contribution that provided information. Providing information can mean providing a complete solution to the problem. Interviewee 14 wrote *"Usually, I participate in MO as an individual, reading and then answering a question completely by myself."* However, often provided information was not a solution in itself but background information that was useful for understanding the problem and needed for the solution. For example, in response to the question "Is the theory of categories decidable?" discussed above, many users provided information about the definitions needed to understand the problem. Adam described a common way of defining the axioms that were needed to clarify and solve the problem. His comment received 2 votes indicating others thought is was important. Dorais later cited a reference giving specific axioms that could be used, he wrote *"I think you want the axioms in I.1 of Mac Lane's Categories for the Working Mathematician (p. 7 in my copy)."*

Clarifying the Question: There were a few different reasons the users clarified the question. Understanding the problem was often a necessary step before the problem could be solved. Many interviewers stressed the importance of having a precise statement of the problem. Some questions on MO have mistakes, do not provide enough information, or are vague. Interviewee 2, a graduate student, mentioned that *"a fairly common thing is that people will ask a question but will word it in such a way that it is completely not clear what they want. And then at the best the comments and requests for clarification really are kind of 'I'd like to answer your question I don't know what your question is. Here is one possible question. Is this it?'"*

At other times, a question may have been clear but it not the question that the question asker meant to ask or should have asked. Interviewee 3, a professor, mentioned that *"often times when people ask questions they are not even really sure of what question they want to ask"*. The asker may have needed help clarifying the question because he or she did not have enough expertise to frame the question in the right way. The interviewee described a case where *"the literal answer to the question [the question-asker] asked [was] not nearly as interesting as a nearby question, so [the interviewee and others] sort of pushed him toward that question and [the interviewee and others] tried to explain why that other question was really the one he had lurking in his mind even though it [was] not what he actually asked."* There were often multiple ways to phrase or describe a problem. Coming up with the right way of thinking about it early on helped make it solvable. Clarifying the question draws on community members expertise to fill in gaps left by the question asker. The asker may have needed additional information in order to ask the "right" question. Even after a question was clarified, there may have been a period of uncertainty and refinement, until a new understanding of the problem led to a solution.

Critiquing an Answer: Solutions were evaluated by the larger community on MO. Corrections, critiques and comments about an answer filtered out bad solutions. Critiques also helped guide the original answerer to improve his or her solution. Answers were critiqued when others believed the solution or partial solution was incorrect or could have been improved. In many cases this led to a much better solution when the problems were addressed. Interviewee 6, a professor, described a situation in which he had edited his answer. He said *"I was getting into this question partially to learn the material ... I looked up some references and learned some more. Based on the comments I hadn't really learned the specific material that well, so my first answer was not that great so I revised it to make it better. Draft 1 was a B minus paper I got some comments and the second draft was an A minus paper so that was better. ... my first answer was sloppy and it needed to be rewritten."* In this case, he acknowledged that his first solution was not very good. However, without being critiqued, he would not have known that his solution needed more work.

In other cases, the corrections may not have corrected the substantive nature of the solution. For example, Interviewee 2, a graduate student, submitted an answer that included some mistaken assumptions about a theorem, someone pointed out the mistake and he fixed his answer by being less specific. About the changes Interviewee 2 wrote *"rather than trying to fix the details of my answer I basically decided to give what amounts to the same answer at the conceptual level"*. In this case, the errors in his answer were not central to the argument he was making. Even though minor critiques like this one did not challenge the main substance of the solution, they were important in preventing incorrect statements. In the classification scheme critiques to an answer were distinguished from revisions of an answer and extensions of an answer, by limiting it to contributions that pointed out errors or suggested ways to improve an answer without making these improvements or modifications themselves (e.g. suggesting a way to close a hole in a proof without actually carrying out the steps to fix the proof).

Revising an answer: Critiques of an answer often led to substantive revisions of an answer. As we mentioned above, Interviewee 6 described how his answer progressed from a B minus paper to an A minus paper because of substantial revisions, which were inspired by critiques of his original answer. Multiple revisions were common on MO. Interview 3 described an answer to one of his questions. He said *"I remember that [MO user's answer] maybe changed the most. I think [MO user] had maybe some incorrect answers at the beginning . . . I think that the first version was problematic. Or something like that. Later versions were useful."* Through revisions a more valuable solution emerged. In the classification scheme revising an answer was distinguished from extending an answer by limiting it to answers that corrected an existing answer rather than took an answer in a new direction (e.g. revising an answer to close a hole in a proof as opposed to taking a related but different approach to reach a new proof).

Extending an Answer: Sometimes a final solution emerged from the ideas of multiple users, each with his own insights. Interviewee 13, a professor, describes this type of collaboration as an idealized prototype in which *"one person might suggest a refinement of a definition, or suggest a vague idea which is subsequently given legs by another researcher,"* and later described an instance of this collaboration. He and another user had submitted similar answers, then a third person extended their answer *"[third person]'s answer came later, as I recall, and extends the same idea to give a more striking instance of the phenomenon. I think that his answer builds on the idea we used, but required his insight to see it through."* In this example, the question could not have been solved without the insights of multiple users. This is similar to Poonen's role in the collaboration surrounding the question "Is the theory of categories decidable?" Using his expertise, he was able to solve the original question posed by A. Rex, after Hamkins and Dorais had worked out a slightly different (and less specialized) case. The final solution was a combination of work by all three.

Extending an answer also occurred when users added to an answer by providing insight from another perspective. This insight was not needed to solve the problem, but it made the solution better by elaborating on the idea. Interviewee 12, a professor, described why he sometimes adds to a solution, *"one thing that happens a reasonable amount is me reading an answer, saying `Hmm, that's not wrong, but not how I would have put it either.' and writing an answer which tries to put what I think is a better spin on things."* Compared to revising an answer extending an answer added substantive content the Q&A collaboration that was considered to be a significant portion or a complete answer in itself.

Summary

Collaboration was diverse and fell on the spectrum between independent and interdependent generation of answers. Providing information was the most common collaborative act. However users also made a variety of other kinds of contributions that built on existing work including clarifying the question, critiquing answers, revising answers and extending answers.

Contributions often built on existing work by revising or extending answers; we will refer to these contributions as *secondary, additive work* because they append complete or partial answers to existing work. The design of the Q&A platform used for problem solving on MO suggested that collaboration might take place primarily through independent contributions that provide information; we will refer to these contributions as primary additions. Despite the limitations of the Q&A platform we found multiple ways in which users made use of the limited features to revise or extend existing work. Interviewees suggested that these secondary additions were often as important as primary additions in developing a solution. For example, in solving the question "Is the theory of categories decidable?" the answer that Poonen gave, which extended the answers of Hamkins and Dorais, was as important as the answers given by Hamkins and Dorais in solving the problem. Remember Hamkins and Dorais were able to answer the question posed by A. Rex in the more general case, but only Poonen was able to answer the specific case asked by A. Rex.

Contributions also built on existing work by evaluating and critiquing that work; we will refer to these contributions as *indirect, evaluative work*. A substantial amount of work was in the form of critiquing the question or answers, identified as clarifying the question and critiquing an answer respectively. These contributions did not directly provide answers or parts of answers, but instead indirectly contributed to answers by asking questions, providing clarification, pointing out errors, and making suggestions. Interviewees suggested that these contributions were important in encouraging better answers. For example, in solving the question "Is the theory of categories decidable?" there was a lengthy discussion among several contributors about how the question should be reformulated to be logically precise. Without the clarification of the question, Hamkins, Dorais, and Poonen's final answers would not have been given. Despite the importance of this indirect, evaluative work, it also created some confusion. In the case discussed above, there was a period of a few hours where the correct version of the question was unclear, leading to uncertainty as to whether the original version of Hamkins' answer was sufficient.

In Study 2 we directly test whether these contributions add value to collaborations.

STUDY 2: THE IMPACT OF COLLABORATIVE ACTS ON SOLUTION QUALITY

Collaboration is often assumed to be valuable and there is evidence showing that collaborative work in aggregate when compared to independent work is often superior. Pairs outperform individuals in lab tasks simulating scientific discoveries [28]. Co-authored scientific papers are cited more often and appear in more prestigious journals than single-authored papers [14,31,41]. Teams developing inventions create more influential patents and fewer very poor patents [35]. However, less is known about the value that different types of contributions make to a collaborative outcome. In this study we evaluate what value, if any, each type of collaborative act identified in Study 1 added to the solution quality.

In Study 1 we found that solutions often grow organically and become better through additions of answers or parts of answers. *Primary additions*, which are defined as providing information independent of existing work on a problem, are expected on average to increase a Q&A's solution quality by providing the first solution or a better solution. *Secondary additions*, such as revisions and extensions of answers, are expected, based on Study 1, to be as important as primary additions and to increase a Q&A's solution quality by improving, completing, or surpassing existing solutions. Thus we predict that all additions, primary and secondary, will increase a Q&A's solution quality.

In Study 1 we found that *indirect, evaluative contributions*, such as clarifying the question and critiquing an answer, were common and important in encouraging better answers. These contributions are expected to increase a Q&A's solution quality by making the question easier to answer or by pointing out ways to fix an answer. They are expected to encourage and enable the creation of better solutions.

While both additive and indirect, evaluative contributions might improve solution quality, they might do so in different ways and under different conditions. For example, while additive contributions might immediately affect solution quality by providing a better solution, indirect contributions rely on subsequent contributions to have an impact, and so they might not cause an immediate improvement in the solution. In fact, it is possible that indirect contributions could even *decrease* perceived solution quality in the short term, because clarifying a question or critiquing an answer might reduce the relevance or highlight flaws in existing solutions. In Study 1 we found that these evaluative contributions can sometimes create a period of confusion by calling into question and devaluing accepted information.

Below we operationalize quality in MathOverflow and examine the association of collaborative acts on quality over time.

Method

By tracking changes in solution quality over time we were able to evaluate the impact of each type of collaborative act on the change in solution quality.

Outcome Variable

Users vote on the quality of answers by voting an answer score up or down. **Solution quality at time t** was operationalized as the maximum score an answer to a question received during hour t. For example, if a question had no answers at hour t it was given a score of 0, if a question had two answers scoring 2 and 5 during hour t, it was given a score of 5. This operationalization involves a few assumptions. First, users' ratings of answer quality approximate objective answer quality. Solution quality is difficult to measure objectively, particularly for research level mathematics because there are few qualified judges. The answer score represents a good approximation because MO users are some of the few judges who can judge the quality of an answer [39]. Second, the solution to a problem is best judged by the quality of its best answer. There are a few alternative metrics including sum of the quality of all answers and the average quality of all answers to a question. The score of the best answer seemed the most logical because it does not penalize for extraneous wrong content, while also not rewarding extraneous beneficial content. In this case it was the most stringent test that the additional collaborative activity was actually adding value to the final best answer. Third, studying changes in the score of an answer in units of hours is meaningful. While one hour is a short interval votes accrue quickly and predictably on MO. An answer in its first hour typically receives 15% of the votes it will receive overall and more importantly the score in the first hour is significantly correlated with its final score ($r(250) = 0.44$, $p < 0.001$). Due to the constraints of the data on MO, scores are not available in units shorter than one hour. Thus an hour is the shortest unit of analysis to measure changes in solution quality. For simplicity the analysis was limited to the first two days after a question was posted, because this was the period in which the majority of activity occurs (81.5%).

Predictor Variables

Providing information at time t was measured as the sum of the score given to each contribution categorized as providing information during hour t (contributions were rated by coders on a scale 0 to 4, scores were scaled to be from 0 to 1 and the average was used, see Study 1). For example if there were two contributions in the hour classified as providing information and given scores of 0.5 and 0.75, then the total score for the hour for providing information was 1.25.

Clarifying the question at time t was measured as the sum of the score given to each contribution categorized as clarifying the question during hour t. Each contribution was rated and the scores were scaled to be from 0 to 1 to indicate how much it clarified the question (see Study 1).

Critiquing an answer at time t was measured as the sum of the score given to each contribution categorized as critiquing an answer during hour t. Each contribution was rated and the scores were scaled to be from 0 to 1 to indicate how much it critiqued an answer.

Revising an answer at time t was measured as the sum of the score given to each contribution categorized as revising an answer during hour t. Each contribution was rated and the scores were scaled to be from 0 to 1 to indicate how much information it provided.

Extending an answer at time t was measured as the sum of the score given to each contribution categorized as extending an answer during hour t. Each contribution was rated and the scores were scaled to be from 0 to 1 to indicate how much information it provided.

Collaborative Acts	Mean # (SD)	Avg. Score
Providing information	2.57 (2.18)	0.44
Clarifying the question	1.32 (1.89)	0.38
Critiquing an answer	1.26 (1.97)	0.46
Revising an answer	0.73 (1.35)	0.28
Extending an answer	0.25 (0.50)	0.45

Table 2: Mean (standard deviation) number of collaborative acts in a Q&A and average score per collaborative act from 0-1.

Covariates

Several control variables were included in the models as covariates.

Solution quality at time t-1. The outcome of interest was the change in solution quality in a given hour. Solution quality in the previous hour was included to control for solution quality up until the hour of interest [8].

Question. A dummy variable for each question was included to control for unmeasured differences at the question level that did not vary with time, such as poor quality question or a question attracting more contributors. In other words the question was included as a fixed effect to control for differences between questions. Models including the question as a random effect instead resulted in results that are substantively the same as those we report below; we report the results with question as a fixed effect because it is a stricter control.

Time. The hour since the question was posted was included to control for systematic changes in a collaboration that depended on how long it had been active on the site. For example, questions receive more attention shortly after they are posted rather than later.

Linear regression models were constructed with predictors, covariates, and the outcome variable. We had hypothesized that clarifying the question and critiquing an answer would have delayed impact on solution quality. Therefore, we tested two regression models, one that did not consider downstream impact of contributions and one that did. Model 1 considered only immediate increases in solution quality that occurred in the same hour as the contribution. To examine possible delayed impact clarifying the question and critiquing an answer Model 2 examined increases in solution quality up to three hours later. Partial R^2 were calculated for the addition of the collaborative acts into the models and Akaike Information Criterion (AIC) were calculated to compare Model 1 and 2.

Results

Primary and secondary additions had a direct and immediate impact on solution quality. Both Model 1 and Model 2 showed that providing information, revising an answer, and extending an answer all significantly predicted an immediate increase in solution quality (See Table 3). Over the course of an hour, one additive contribution typically increased the score of the best solution by 0.35-0.51 points. Sometimes this meant a new best solution replacing the old best solution. On average the best solution's quality score increased by 0.06 points per hour (*SD* = 0.30 points), so the value of an addition is relatively large.

Indirect, evaluative contributions had a delayed impact on solution quality. A comparison of Model 1, which only included immediate effects of contributions on solution quality and Model 2, which included up to 3 hour delayed effects of contributions for clarifying the question and critiquing an answer showed that Model 2 was superior. Model 2 explained slightly more variance and had a better AIC value (see Table 3). A model comparison test revealed that Model 2 was significantly better. Model 2 showed that clarifying the question had a negative effect on solution quality in the first two hours—a clarification of the question lowered the best solution quality by 0.08 points in these two hours—while critiquing the answer had no effect (see Table 3). Model 2 also showed that clarifying the question and critiquing the answer had a positive effect on the best solution quality after three hours. The inclusion of a clarification of the question or critique of an answer after three hours increased the best solution's quality by 0.16 and 0.08 points respectively.

Overall the collaborative acts only explained a small amount of variance in change in solution quality, 3.3%. This may be largely due to the large amount of random chance and unexplained factors in the ratings of answers on MO and is a limitation of this method. The amount of variance explained is comparable to other studies of answer quality on MO [39].

Summary

Past literature has shown that, in aggregate, collaboration improves solutions compared to independent work. Study 2 tested at a process level the specific types of collaborative acts that lead to increased solution quality. We had predicted and found that adding a new answer or parts of an answer, whether independent of existing work in the case of providing information or building on existing work, in the case of revising an answer or extending an answer increased solution quality. Further, we found that revising and extending an answer had an equal effect on solution quality as providing information. This finding demonstrates that building on existing work is as valuable as adding an original contribution.

All the collaborative acts increased solution quality, but through different ways. While additions had a direct effect on solution quality, evaluative contributions, such as clarifying the question and critiquing an answer, had a delayed impact on solution quality. One explanation is that clarifying the question and critiquing the answer improve quality by inspiring other improvements, such as revisions or new answers within a few hours. But in the short-term clarifying a question had little effect and clarifying the question had a negative effect on solution quality, because they bring into question the adequacy of current solutions. Although critiques and clarifications eventually had a positive impact on solution quality the effect was smaller than the effect of direct contributions.

	Model 1: Change in Solution Quality	**Model 2: Change in Solution Quality**
Provided information	0.35***	0.35***
Clarified the question	-0.12**	-0.08*
an hour ago		-0.08*
two hours ago		0.16***
Critiqued an answer	0.04	-0.04
an hour ago		0.04
two hours ago		0.08*
Revised an answer	0.51***	0.51***
Extended an answer	0.43***	0.43***
Partial R^2	3.1%	3.3%
AIC	674.7	659.2
Model Comparison	$F(4) = 5.88, p < 0.001$	

Table 3: Regression models showing the impact of collaborative acts on solution quality, with and without delayed effects. * p < 0.05, ** p < 0.01, *** p < 0.001

DISCUSSION

We examined collaborative scientific problem solving in the context of MathOverflow, a site in which anyone connected to the Internet can pose, answer, critique, or improve a problem. By characterizing a sample of posts to the site we proposed a simple taxonomy of collaborative acts in which users engage, including both direct contributions (e.g., providing answers) as well as indirect contributions (e.g., clarifying the question, critiquing answers). A quantitative analysis relating collaborative acts to solution quality showed that quality improved for both direct and indirect contributions, but that indirect contributions had a delayed effect on quality. Below we discuss the implications of these results for existing Q&A systems which are being repurposed for scientific problem solving, as well as the design of novel systems that could better support such activities.

Repurposing Q&A Platforms for Problem Solving

Increasingly Q&A platforms are being repurposed for other uses, including as a repository of knowledge, in the case of Stack Overflow [1], and as a problem-solving platform, in the case of MathOverflow. Q&A systems are designed to support one form of collaboration very well. They support the selection of the best answer among many, by broadcasting a question to a large group, allowing multiple independent answers, and rating of answers through voting by a large audience. Our analyses of the ways mathematicians have been collaborating suggest some concrete changes to Q&A platforms that could enhance other forms of collaboration.

First, Q&A platforms could include more suitable tools to facilitate building on existing work. A major finding of this paper is that MO users often advance a solution by building on others' work. Better tools might make it easier to building on others' work making it even more common. For example, better versioning tools that could connect related ideas or answers together and allow joint ownership of answers might encourage building on existing work. At the moment answers can be revised, but there is no tool to link an answer that develops as an extension of another answer to its precursor. Also, in practice other users are hesitant to directly edit someone's answer because of the perceived ownership over an answer. Tools to gather ideas into an answer and share ownership might make people less hesitant about adding to other people's work. Another possible change would be to create a section for partial work to elicit more contributions that could be built upon. At the moment users sometimes include definitions, related work, partial ideas in the comments sections by convention; if there were an explicit section more content might be added in a more systematic way. Better tools to support building on existing work might encourage even more additions and increase the prominence of these later contributions.

Second, Q&A platforms could include better incentives to encourage indirect contributions. Our results provide quantitative evidence that good answers come about not just from more people providing answers, but also by fixing errors, filling in gaps, being inspired by existing work, and adding additional perspectives. These secondary additions are as valuable in reaching a final solution as primary additions. Thus, figuring out ways to support and inspire iterating on solutions is important. However, Q&A platforms generally award reputation points to answers and not to indirect contributions. One way of supporting this indirect work may be to provide rewards for downstream effects that are not immediately recognizable as useful. For example, a flat number of reputation points could be awarded to evaluations of either the question or answers or an evaluation could receive the same number of reputation points as a contribution it inspires but from a separate "indirect" reputation pool.

Third, Q&A platforms could include mechanisms to keep people involved in a Q&A for longer. The importance of building on existing work and indirect, evaluative contributions suggest that continued interaction and revisiting a Q&A may lead to improvements in solution quality. Q&A platforms could use alerts and content prioritization to bring authors of the question, authors of answers, and potential experts back to a Q&A multiple times as work progresses. It also suggests that there may be value recommending specific tasks to specific users, such as bring in a user to review a particular answer that is in their area of expertise.

Large-scale Collaborative Problem Solving Systems

Given better tools to collaborate we might see a proliferation of larger scale collaborations. Even with the limitations of the Q&A platform we see people eking out ways to collaborate in complex ways. Q&A platforms support basic ways to enable large scale collaboration, such as the ability to find experts within a crowd and to select the best solutions among many. More support is needed to enable users to build on each others' work such as ways to iterate on solutions, ways to inspire new solutions from partial ideas, and ways to elicit and represent multiple perspectives. Trying to shoehorn such functionality into existing Q&A platforms (such as MathOverflow) or blogs/discussion boards (such as The Polymath Projects) may not be the most effective approach. New systems for collaborative problem solving could be built from the ground up to address not only the issues identified above but also fundamental crowd coordination issues such as expertise identification, support for subteams, decomposition of tasks while maintaining interdependencies, task routing, and quality control [21]. Matching the motivations for participating in such systems with the already-existing reward structure for professional advancement in science will be a key challenge to overcome as well.

Limitations

In this paper we developed identified a set of collaborative acts that describe the types of contributions to collaborations at a process-level. However, the observations were built from the examination a single site, MathOverflow, with problems all from the single domain of mathematics. To ensure that the set of collaborative acts and their assessed value generalizes and is not site or domain specific future research will need to examine the set of collaborative acts in other domains of problem solving in which collaboration occurs through other media. Nor are the collaborative acts that we describe comprehensive, and we anticipate that future research will expand on this work to better test which are most effective and how they fit together.

CONCLUSION

In this paper we identified a set of collaborative acts that describe at a process-level how collaboration transpires in solving complex, difficult problems. A quantitative analysis relating collaborative acts to solution quality suggested that some types have an immediate effect while others have a delayed and indirect effect on improving solution quality. Together these findings inform the design of tools to aid large-scale collaborative problem solving.

[redacted]

REFERENCES

1. Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. Discovering value from community activity on focused question answering sties: A case study of Stack Overflow. In *Proc KDD'12,* ACM Press (2012), 850-858.
2. Birnholtz, J. P. What does it mean to be an author? The intersection of credit, contribution, and collaboration in science. *Journal of the American Society for Information Science and Technology 57,* (2006), 1758-1770.
3. Bernstein, M. S., Little, G., Miller, R. C., Hartman, B., Ackerman, M. S., Karger, D. R., Crowell, D., & Panovich, K. Soylent: A word processor with a crowd inside. In *Proc UIST 2011*, ACM Press (2011), 313-322.
4. Blumenstock, J. E. Size matters: Word count as a measure of quality on Wikipedia. In *Proc. WWW 2008*, ACM Press (2008), 1095-1096.
5. Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B., and Mytkowicz, T. Augmented social cognition. In *AAAI Spring Symposium: Social information processing,* (2008), 11-17.
6. Cranshaw, J. and Kittur, A. The Polymath Project: Lessons from a successful online collaboration in mathematics. In *Proc. CHI 2011,* ACM Press (2011), 1865-1874.
7. Cronin, B. *The Hand of Science: Academic Writing and its Rewards.* Scarecrow Press, Lanham, MD, 2005.
8. Cohen, J., Cohen, P., Wes, S. G., & Aiken, L. S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Science*. Hillsdale, NJ: Lawrence Erlbaum Associates (2003).
9. Cummings, J. N., & Kiesler, S. Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science 35* (2005), 703-722.
10. Dabbish, L., Stuart, C., Tsay, J., and Herbsleb, J. Social coding in GitHub: Transparency and collaboration in an open software repository. In *Proc. CSCW 2012,* ACM Press (2012), 1277-1286.
11. Dunbar, K. How scientists think. On-line creativity and conceptual change in science. Eds. Ward, T. B., Smith, S. M., Vaid, J. In *Creative thought: An investigation of conceptual structures and processes.,* (pp. 461-493). American Psychological Association, Washington DC, 1997.
12. Fiske, D. W., and Fogg, L. But the reviewers are making different criticisms of my paper! Diversity and uniqueness in reviewer comments. *American Psychologist 45* (1990), 591-598.
13. Glaser, B. G., and Strauss, A. L. *The Discovery of Grounded Theory: Strategies for Qualitative Research,* Aldine, New York, 1967.
14. Gordon, M. D. A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication. *Scientometrics 2* (1980), 193-201.
15. Hesse, M. B. *The Structure of Scientific Inference.* Macmillan, London, (1974).
16. Jeppesen, L. B., and Lakhani, K. R. Marginality and problem solving effectiveness in broadcast search. *Organization Science 21* (2010), 1016-1033.
17. Jonassen, D. H. Instructional design models for well-structured and ill-structured problem solving learning outcomes. *Educational Technology Research and Development 45,* (1997), 65-94.
18. Jonassen, D. H., and Kwon, H. Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development 49,* (2001), 35-51.
19. Kerr, N. L., and Tindale, R. S. Group performance and decision making. *Annual Review of Psychology 55,* (2004), 623-655.

20. Kittur, A., and Kraut, R. E. Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In *Proc. CSCW 2008,* ACM Press (2008), 37-46.
21. Kittur, A., Nickerson, J. V., Bernstein, M. S., Gerber, E. M., Shaw, A., Zimmerman, J., Lease, M., and Horton, J. J. The future of crowd work. In *Proc. CSCW 2013,* ACM Press (2013), 1301-1318.
22. Kittur, A., Smus, B., Khamkar, S., & Kraut, R. E. CrowdForge: Crowdsourcing complex work. In *Proc. UIST 2011,* ACM Press (2011), 43-52.
23. Klein, M. How to harvest collective wisdom on complex problems: An introduction to the MIT deliberatorium. (2011) *Working paper.*
24. Kulkarni, A. P., Can, M., & Hartmann, B. Collaboratively crowdsourcing workflows with turkomatic. In *Proc. CSCW 2012,* ACM Press (2012), 1003-1012.
25. Lakhani, K. R., Garvin, D. A., & Lonstein, E. Top Coder (A): Developing software through crowdsourcing. *Harvard Business School Case* (2012), 610-632.
26. Latour, B., and Woolgar, S. *Laboratory Life: The Construction of Scientific Facts.* Princeton University Press, Princeton NJ, 1979.
27. Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G., Hartmann, B. Design lessons from the fastest Q&A site in the west. In *Proc. CHI 2011,* ACM Press (2011), 2957-2967
28. Okada, T., and Simon, H. A. Collaborative discovery in a scientific domain. *Cognitive Science 21* (1997), 109-146.
29. Olson, G. M, & Olson, J. S. Distance Matters. *Human-Computer Interaction 15,* (2000), 139-178.
30. Pennington, N., Lee, A. Y., and Rehder B. Cognitive activities and levels of abstraction in procedural and object-oriented design. *Human-Computer Interaction 10* (1995), 171-226
31. Persson, O., Ganzel, W., and Danell, R. Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluation studies. *Scientometrics 60* (2004), 421-432.
32. Schoenfeld, A. H., and Herrmann, D. J. Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology 8* (1982), 484-494.
33. Shin, N., Jonassen, D. H., and McGee, S. Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science and Teaching 40* (2003), 6-33.
34. Shrager, J., Billman, D., Convertino, G., Massar, J. P., & Pirolli, P. Soccer science and the Bayes community: Exploring cognitive implications of modern scientific communication. *Topics in Cognitive Science 2,* (2010), 53-72.
35. Singh, J., and Fleming, L. Lone inventors as sources of breakthroughs: Myth or reality? *Management Science 56* (2009), 41-56.
36. Stasser, G., and Titus, W. Pooling of unshared information in group decision making: biased information sampling during discussion. *Journal of Personality and Social Psychology 48,* (1985), 1467-1478.
37. Stasser, G., and Titus, W. Hidden profiles: A brief history. *Psychology Inquiry 14,* (2003), 304-313.
38. Tang, J. C., Cebrian, M., Giacobe, N. A., Kim, H., Kim, T., & Wickert, D. Reflecting on the DARPA Red Balloon Challenge. *Communications of the ACM 54,* (2011), 77-85.
39. Tausczik, Y. R., and Pennebaker, J. W. Predicting the perceived quality of online mathematics contributions from users' reputations. In *Proc. CHI 2011,* ACM Press (2011), 1885-1888.
40. Tausczik, Y. R., and Pennebaker, J. W. Participation in an online mathematics community: Differentiating motivations to add. In *Proc. CSCW 2012,* ACM Press (2012), 207-216.
41. Thagard, P. Collaborative knowledge. *Nous 31* (1997), 242-261.
42. Turner, T. C., Smith, M. A., Fisher, D., and Wlser, H. T. Picturing Usenet: Mapping computer-mediated collective action. *Journal of Computer-Mediatied Communication 10* (2005).
43. Wuchty, S., Jones, B. F., and Uzzi, B. The increasing dominance of teams in production of knowledge. *Science 316* (2007), 1036-1039.