Participation in an Online Mathematics Community: Differentiating Motivations to Add

Yla R. Tausczik, James W. Pennebaker
Department of Psychology
The University of Texas at Austin
{tausczik, pennebaker}@mail.utexas.edu

ABSTRACT

Why do people contribute content to communities of question-answering, such as Yahoo! Answers? We investigated this issue on MathOverflow, a site dedicated to research-level mathematics, in which users ask and answer questions. MathOverflow is the first in a growing number of specialized Q&A sites using the Stack Exchange platform for scientific collaboration. In this study we combine responses to a survey with collected data on posting behavior on the site. User behavior suggests that building reputation is an important incentive, even though users do not report this in the survey. Level of expertise affects users' reported motivation to help others, but does not affect the importance of reputation building. We discuss the implications for the design of communities to target and encourage more contributions.

Author Keywords

question-answering, online communities, scientific collaboration, motivation

ACM Classification Keywords

H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces—*Computer-supported cooperative work*

INTRODUCTION

Communities of user-generated content, such as Wikipedia, Flickr, and Yahoo! Answers, rely on their users to make contributions. Therefore an important and well-researched question is: What motivates users to contribute? We investigate motivation in a specialized question answering site, MathOverflow (mathoverflow.net) in which users ask and answer research-level mathematics questions.

Many previous studies have looked at cooperation in communities of user-generated content in which contributions are low-cost and in which ties between users are loose. In this paper we describe the motivations to contribute in a community in which participation has greater costs and benefits.

CSCW 2012, February 11–15, 2012, Seattle, Washington, USA.

RELATED WORK

Self-reported Motivations to Contribute

Many studies have investigated the reasons people give for why individuals contribute to communities of user-generated content (e.g. [18, 23, 24, 25, 19]). These studies have identified a wide range of motivations, ranging from fun to helping to reputation building. Dholakia and colleagues [9] with slight modifications by Lampe and colleagues [19] find evidence for six categories of benefits of participating: getting information, giving information, reputation building, relationship development, recreation, and self discovery.

Motivations vary based on the community type [21]. Building reputation and self-development may be especially important in career oriented communities. Contributors to open-source software were found to be more motivated by gaining reputation and self-development compared to contributors to Wikipedia, who were more motivated by altruistic reasons [25].

Two qualitative studies identify motivations that may be especially important to those users answering questions. Five major themes were identified in interviews of contributors to Knowledge-iN, a large South Korean Q&A site [22]. The themes were helping others, learning, promoting their business, recreation and accumulating points [22]. Contributors to Yahoo! Answers explained that they only answered questions when they thought their question would be well received by the question asker and would not get lost in the crowd [8].

In this study we describe motivations to participate on MathOverflow (MO) using the categorization developed by Dholakia and colleagues. We compare the relative importance of the motivations on MO to motivations analyzed in the studies described above.

Inferring Motivations to Contribute

One of the common problems with self-reports is that they are subject to response biases. In particular, individuals are likely to inflate reports of socially desirable items and deflate reports of undesirable ones [26]. This may be especially true when individuals are selecting between items such as helping others and enhancing ones reputation. An alternative to self-reports is to measure behaviors that can be used to infer attitudes or motivations (e.g. [11]).

Gaining recognition is one benefit that users are likely to under-report, but has been inferred to encourage participation by observing individuals' behavior. Measures of recognition, such as the number of comments in response to a submission, have been shown to be related to participation level; those that receive more recognition participate more [28]. Furthermore, individuals have been shown to change their participation in relation to the amount of recognition they receive. For example, during the time period when Chinese Wikipedia was blocked on mainland China, thereby reducing readership, unblocked contributors dramatically reduced their contributions [34]. Not only are individuals sensitive to the amount of recognition they receive, in addition recognition may consistently reinforce participation. Huberman and colleagues [14] found a consistent pattern over time in which YouTube contributors posted more videos if two weeks earlier the videos they had contributed had received more views. In fact, recognition may be so reinforcing that it can create a positive feedback loop accounting for large discrepancies that are seen between users who contribute at high rates and those that contribute at low rates [33].

Other studies have used this same technique of detecting the kinds of social feedback that encourages participation (eg [15, 4]). These studies have mostly focused on predicting whether first time contributors continue to participate or not.

Reputation Systems

Authors have argued that one of the reasons the Stack Exchange platform is successful is because it is designed to make use of a reputation system [20]. In the platform questions and answers are rated by community members. The ratings that authors receive for their work contributes to a reputation score that is listed on their user page and next to their name whenever posting.

One of the advantages of an explicit reputation system is that it may increase participation on the site. Content ratings can encourage users who post highly appreciated content to contribute more because they have received positive feedback from other users [17]. Qualitative descriptions of two Q&A platforms with explicit reputation systems support this prediction. Contributors to Knowledge-iN in interviews both suggested and disputed that the game like aspect of earning points from answering questions motivated them to contribute more answers [22]. An analysis of StackOverflow users suggested two kinds of users, those who are motivated to answer questions by achieving high reputations and earning privileges and those that are motivated by answering the questions in themselves [20].

Theoretical Context

In this paper we explain motivations to contribute by describing the specific benefits that drive users to participate. This perspective has the potential to explain why certain kinds of feedback reinforce participation. The perspective comes from a long history of theories of motivation. It would be considered a functionalist perspective in psychology, meaning that individuals engage in behavior for specific functions. It is best exemplified in the explanations for participation in social movements [16]. According to this theory, motivation is understood at the individual level, and social rewards such as gaining reputation, or individual rewards such as developing knowledge, must outweigh the costs of participating in order to motivate users. In addition, social and individual rewards must go beyond collective rewards, such as a useful encyclopedia, because these are available to everyone whether or not they participate. This theory has been shown to be useful for describing motivations to participate in online communities [13].

Benefits can go beyond what we might think of as traditional rewards, such as monetary compensation. Social exchange theory has described what might normally be identified as a social process, such as relationship formation, as an exchange of social benefits. For example, both building reputation and helping others can be thought of as social benefits that motivate users to share more knowledge in an online community [32].

It is more fruitful to think of the benefits that we discuss as intrinsically motivating users. This is because voluntary online participation is more likely to be driven by intrinsic motivation [18] and those motivations which are extrinsic are likely to have the same advantages as intrinsic motivations, because they are self-directed [29].

MATHOVERFLOW

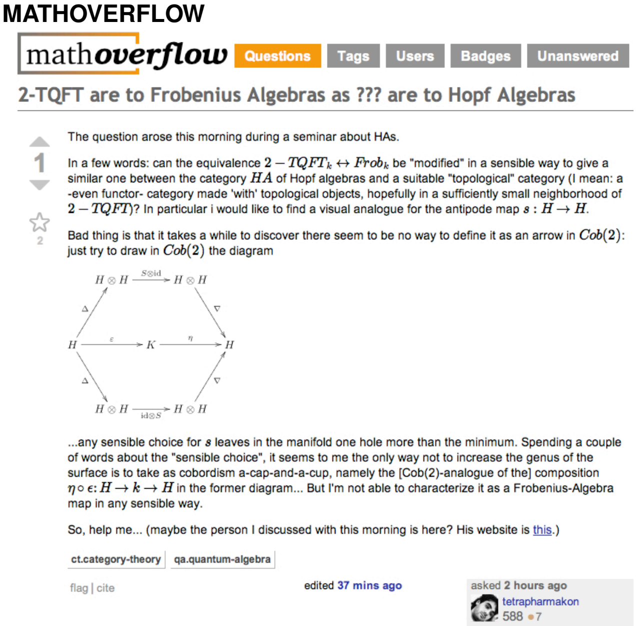

Figure 1. An Example Question

There is an expanding literature on large scale communities of question answering (e.g. [1, 22, 20]). MathOverflow is one of a new kind of community that lies at the intersection of participatory online communities and scientific collaboration. Two distinct features of MathOverflow, the higher costs and benefits of participation, make it an interesting community in which to investigate motivations.

There is a growing development of these specialized question answering communities. The Stack Exchange platform that was first developed for StackOverflow now hosts 51 other specialized Q&A sites. MathOveflow (MO), which was started in September of 2009, was the first to use the platform for scientific collaboration. News of MO spread through posts on popular math blogs, a notice in the American Mathematical Society publication, newspaper articles and by word of mouth and attracted large numbers of academic mathematicians.

Nearly a year later there were approximately 11,300 questions and 27,300 answers, contributed by 5,270 authors. In a survey conducted for this project we found that among more frequent users to the site 90.5% were in a mathematics degree program at some point. In this sample 96.6% were male and on average they were 32 years old (sd = 10.8). The site receives high traffic and regular users check the site frequently. The respondents reported, on average, visiting MO between every day and every two days. MO, like mathematics, draws users internationally although the content is in English and the site has the majority of individuals from English speaking countries.

One of the ways that the site is an interesting community to study is that many see it as an extension of the academic mathematics community. Users are encouraged to use their real names. Of those sampled 81.7% said they used their real name as their user name. It is difficult to quantify how many people on the site know each other offline, but individuals have the perception that behaviors on MathOverflow can affect their reputation in the larger mathematics community. One user stated that he used his real name because he was 'hoping a potential employer sees that I have interesting things to say'. Another user said even if he did not use his real name 'my research areas would narrow me down to one or two people anyway'. There are greater potential benefits of participating in MO. Developing professional ties and obtaining recognition are especially important for career development. Users may expect that by participating in MO they can develop professional relationships and gain recognition.

Expertise and career stage is an important individual difference in scientific collaboration and on MathOverflow. There was a wide range of the highest mathematics career stage that individuals had achieved or were working on: 14% Bachelors degree, 9% Masters degree, 34% PhD, 17% postdoctoral fellow or lecturer, and 26% professorship. Slightly more than 40.1% of users had published no math papers of those sampled, whereas at least one person reported publishing over 200 papers. The median number of papers published was 2 papers.

There are 9,536 registered users, and of those 5,268 have made at least one contributions. Contributions follow a power law distribution where the majority of the contributions are made by a small number of users. Fifty percent of the questions are contributed by 6.3% of the users and fifty percent of the answers are contributed by 2.4% of the users. Of those users who have made at least one contribution, the mean number of question per person is 2.14 (median = 1, sd = 5.39) and mean number of answers per person is 5.19 (median = 1, sd = 19.8). However the median number of days from when the person first logged into their account to the last time is long, 42 days (mean = 102, sd = 120).

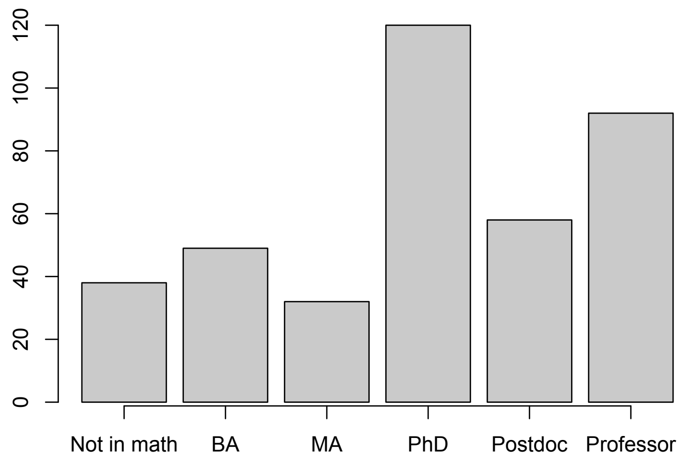

Figure 2. MathOverflow users' level of expertise. Shows the highest math degree that a user is working on or has obtained.

In order to keep the community limited to research level math questions there are extensive lists of the kinds of questions that are appropriate to MO and those that are not. Moderators and community members police questions and will downvote and close those questions below research level. For example an unidentified user posted a question:

> What are the applications of Postnikov approximation? As I understand it, it is dual to a cell decomposition and can, e.g. be used for computing some homotopy groups of spheres.

A commenter writes:

> Why not try reading any exposition of Postikov towers? The applications usually appear right after their definition. I've voted to close.

While some lower level questions like this one are made (often by newcomers) the site is mostly limited to very high level questions that have not been answered directly in other resources. Contributions on MO are not low-cost. Contributions often require original and new thinking which is effortful. Also, by contributing users potentially miss out on rewards from having published the same ideas in peer-reviewed journals. math.stackexchange.com was created for lower level math questions. Despite the high level, most questions on MO get answered, only 16.3% go unanswered and on average a question gets 2.42 answers.

There have been at least two other major online sites developed to aid research level mathematics through large scale online collaboration: polymath and Tricki. These other sites predate MO. While polymath1, the first and most successful of five projects had 39 contributors [7]; Tricki in September of 2010 had 100 users who have made at least one edit [10]; MathOverflow in October 2010 had 5,268 authors who had made at least one contribution. Already MO has a magnitude more contributors. We investigate the success of MO in gaining popularity.

THE CURRENT STUDY

In this study we undertake three major research objectives. We describe the motivations to contribute to MathOverflow; we compare the motivations of individuals with different levels of expertise; and we evaluate the importance of three kinds of social benefits in motivating contributions.

Objective 1: To describe the motivations for contributing to MathOverflow.

We investigate the reasons why users participate in MathOverflow. Other research has identified a variety of reasons why individuals contribute to communities of user-generated content, such as information seeking, entertainment, skill development, and altruism [19, 25]. We develop a full picture of the motivations to contribute by combining multiple sources. 1. Self reported motivations 2a. The kinds of participation that users undertake 2b. Which kinds of feedback encourage or discourage future participation.

Thus we combine two standard approaches to studying motivation in online communities. They each have their own advantages and disadvantages. The first method relies on gathering self-reported motivation from users (e.g. [23]). The second method relies on inferring user motivations from particular patterns of behavior, such as examining what feedback on past participation leads to future participation (e.g. [14]). The first method has the advantage of being able to describe user motivation in a deeper and more nuanced way through self-reports. The approach has the disadvantage of being susceptible to response biases. The second method has the advantage of being able to measure the motivations that drive actual behavior even if users are not consciously aware of them or do not report them. It has the disadvantage that it is more difficult to describe the true motivation behind the patterns in behavior. We combine both methods to make use of their complementary advantages.

Objective 2: To compare the motivations of individuals with different levels of expertise in mathematics.

One of the interesting aspects of MO being at the intersection of communities of user-generated content and scientific collaboration is the potential importance of expertise. The importance of level of expertise is something that is shared with software development and appears in the literature on contributions to open software. However expertise is more explicit and elevated within science. Level of expertise may influence the reasons for contributing to MO. Making helpful contributions will most likely be easier for individuals with more expertise [31]. Skill development may be less important to those with more expertise. Those with longer tenure in mathematics may have more organizational commitment that will translate into a stronger desire to benefit the community in MO [6]. We evaluate whether the motivations of senior professors differ or are similar to the motivations of undergraduate students. Findings about levels of expertise will likely generalize to other communities because expertise is an important part of knowledge sharing in general. The advantage of studying it in MO is that expertise is easier to measure.

Objective 3: To contrast the importance of three different kinds of social benefits.

Recognition has been identified as a potential motivator for contributions (e.g. [14]). We can be more specific in describing how recognition may matter. For example, recognition may be import because it is a proxy for social rewards. Social rewards may be especially important in motivating contributions. We investigate three possible kinds of rewards that may be motivating users to make contributions.

The first social motivation we investigate is ***reputation building***. Broadly our definition of reputation building also includes seeking social approval. Social exchange theory states that individuals choose to interact in order to receive specific social benefits [3]. If by contributing users can gain reputation in the community; building reputation can act as a social benefit that motivates contributions [32]. Related work has shown that social approval can also be an important motivator. When users are told that other community members like their past submissions they are more likely to make more contributions in the future [5].

The second motivation we investigate is ***relationship development***. The design of MO and other Q&A platforms creates an environment in which there is a direct recipient who benefits from the submission of an answer to a question. Sohn and Lekenby show that communities in which there is reciprocal exchange motivate more contributions [30]. One of the incentives of answering a question may be to help the question asker. By answering a question users build up social capital with the question asker in particular, this could be through impressing the asker or by inducing gratitude.

The third motivation we investigate is ***constructive feedback***. Constructive feedback is not a social benefit in of itself. Desire for constructive feedback is driven by a desire for skill development, however the process is social. Individuals may contribute to MO to obtain advice from a larger community on their ideas. The importance of whether the community critiques their answer may provide individuals with incentives to contribute. Skill development is an important motivation to contribute to other career focused communities [25].

We compare these motivations by examining user behaviors on MathOverflow. By using patterns of user behavior, Huberman and colleagues found that YouTube users' contributions were reinforced by getting recognition in the form of views [14]. We examine a broader set of community responses to assess which of these responses encourages participation. By looking at multiple kinds of community responses we can compare the importance of the three kinds of benefits that we have identified.

In this study we focus on answers. We record seven kinds of feedback authors can get on their answers and how many answers they later submit. There are two kinds of feedback that are important for reputation building. On MO community members can upvote or downvote answers to express their approval or disapproval of the answer. Ultimately the upvotes minus the downvotes make up the contribution's score. Together contribution scores make up the reputation score a user gets for the site. If higher scores on answers and fewer downvotes encourage more answers in the future we can infer that they are partially motivated by reputation building.

There are also two kinds of feedback that are important for relationship development. Question askers can select the answer that they like best, also question askers often directly comment on questions expressing interest or gratitude. If question answerers are encouraged by these two forms of response from question askers we can conclude that they are partially motivated by relationship development.

Finally community members leave comments on answers, in which we can measure whether individuals express either agreement or disagreement with the answer. If question answers are influenced by these comments we can infer that they are partially motivated by constructive feedback.

One possibility is that these three social benefits–reputation building, relationship development, and constructive feedback– may overlap conceptually. In terms of proximate motivations these social benefits are conceptually distinct. For example by receiving better scores users immediately get social approval and by receiving comments that agree or disagree with the answer users immediately get constructive advice on the content of their answer. These three social benefits may overlap in terms of deeper motivations. For example, if an individual is motivated by becoming higher status then they may be more inclined to being motivated by social approval or by constructive feedback. In this study we are only able to distinguish between proximate motivations. However, by distinguishing proximate motivations we have discovered something about the mechanisms through which motivations drive behavior.

METHOD

Information about users was collected by combining responses to a self-report survey with posting behavior on MO. Participants were recruited through two methods during September, 2010. All active participants (at least 30 reputation points) who included a personal website with a valid email address in their profile were contacted to take the web survey via email. The survey was also advertised on the main MO website (with a link to the survey) for four days. Posting behavior for users that completed the survey was compiled from publicly available dumps from MO servers. Posting behavior was collected for approximately one year from September 2009 (when the site began) to October 2010.

User Survey

A total of 401 participants completed the survey. A smaller sample of 217 users elected to give us their username on MO during the survey. This smaller sample was used in combined analyses of self-reports and user behavior. The users sampled contributed at the time about 27.5% of the posts on the site. This sample is most likely not representative of MO; it overrepresents the most frequent contributors to MO. However, we are most interested in what motivates the frequent contributors because their efforts result in the majority of content on the site.

Users were asked to report demographic information, including the highest mathematics degree they were working on or had obtained and to complete a motivations scale. **Level of expertise** was operationalized as the highest mathematics degree, which consisted of the levels: undergraduate, masters, PhD, post-doctoral fellow or lecturer, and professor. Masters students were classified as those in masters level programs; PhD students were classified as those in PhD programs even if they had not obtained a masters yet. To simplify figures, level of expertise was dichotomized based on whether a participant was above the PhD level. The motivations scale was adapted from a scale used on the online community of user-generated content, Everything2.com [19]. The questions were modified to be specific to the uses of MO and mathematics research in general. The self-discovery items were not included because they were not relevant to MO.

User Behavior

Survey data was matched with user behavior on the site. Two kinds of user behavior was recorded. First we recorded basic contribution information. For each user the **rate of questions** and the **rate of answers** was recorded; this was the total number of questions and answers divided by the time the user was active on the site. By looking at the rate of contributions we controlled for differences in when users discovered or left the site. We further broke down the contribution of answers into two kinds: answers to popular questions and answers to unpopular questions. Popular questions were defined as those that scored in the top 25% of all questions, unpopular questions were defined as those that scored in the bottom 25% of questions. The cutoffs were respectively greater or equal to 13 points and less than or equal to 3 points. For each user the **rate of answers to popular questions** and the **rate of answers to unpopular questions** was recorded.

Second we examined which kinds of community feedback encouraged or discouraged future contributions. We focused this analysis on the contribution of answers. We measured several feedback variables:

reputation building

score: The quality rating community members gave the answer (log transformed)

downvoted: Whether any user indicated that they disliked the answer

<u>**relationship development**</u>

accepted answer: Whether the question asker accepted the answer

question asker comment: Whether the question asker left a comment about the answer

<u>**constructive feedback**</u>

comments: The number of comments left by others about the answer (log transformed)

agreement: The amount of agreement expressed in the comments

disagreement: The amount of disagreement expressed in the comments

The comments were linguistically analyzed using Linguistic Inquiry and Word Count (LIWC), which measures categories of psychologically meaningful words [27]. The level of agreement was operationalized as the amount of positive emotion words (e.g. 'excellent') and assent (e.g. 'yes', 'agree'). The level of disagreement was operationalized as the amount of negative emotion words (e.g. 'difficult') and negations (e.g. 'no', 'not'). To create measures of agreement and disagreement z-scores of the two linguistic categories were taken and summed.

Using four distinct models we tested whether community feedback encouraged or discouraged the submission of answers. First we examined whether the feedback given to the first answer a user submitted predicted if the user would go on to post an answer again. Logistic regression was used with the posting of an answer again as the outcome variable and the feedback variables as predictors. The analysis was restricted to users who submitted an answer not a question first to avoid confounds related to contribution history. We report Nagelkerke pseudo R^2 for this model.

Next we examined which feedback encouraged submission of future answers for multiple periods of time. A method similar to the one established by Huberman and colleagues was used to determine which kinds of feedback encouraged or discouraged submission of answers [14]. Submissions were divided into two-week intervals. For each user, feedback variables were measured for periods in which there was at least one answer and were paired with the number of answers in the next two-week interval. Feedback variables were averaged across the two week period. This method allows observation of feedback for many time periods per user rather than focusing on a single contribution.

Multi-level models were used; two week time periods were nested within users. By using multi-level models we were able to control for individual differences in the rates of answers and quality of feedback. A two-step process was used to evaluate the significance of the feedback variables. All feedback variables except agreement and disagreement were included in the model. Then the model was restricted to those answers which received comments, the significant variables from the first model were included as well as agreement and disagreement. We report the percentage of variance explained when adding the predictor variables to the intercept only model; this is a common substitute for R^2.

This second method was used to examine feedback given to answers in general and to answers to popular and unpopular questions separately.

RESULTS

The results are grouped by the method used. The results relevant to the three objectives are intertwined and will be summarized in the discussion.

What reasons do users give for why they participate?

Confirmatory factor analysis (CFA) was performed with the same factor structure presented by Lampe and colleagues [19] excluding the omitted factor, self-discovery. CFA showed that this 5 factor model was a poor fit (RMSEA = 0.13). Exploratory factor analysis of the motivation scale was conducted. Observation of the scree plot revealed three factors. Oblique rotation was used; all items loaded above 0.30. The first factor, which we will refer to as recreation, included questions pertaining to entertainment and procrastination ($\alpha = 0.82$, mean = 2.67, sd = 1.08). The second factor, which we will refer to as self advancement, included questions pertaining to information seeking, reputation building, generating ideas, and relationship development ($\alpha = 0.74$, mean = 2.07, sd = 0.74). The third factor, which we will refer to as helping, included questions pertaining to providing information to others or contributing to a body of mathematics knowledge ($\alpha = 0.83$, mean = 2.29, sd = 1.11). All three factors have alpha reliability above the 0.70 acceptability cutoff, indicating that they are internally consistent. Confirmatory factor analysis with the new 3 factor model was a better fit (RMSEA = 0.10).

The strongest reasons for participating in MO were in order: recreation, helping, and self advancement. Breaking down the motivation questions into subscales provides more detail on individuals motivations (see table 1). Both reputation building and relationship development scored very low. Getting information scored the highest.

Table 1. Means and standard deviations for motivation subscales.

Factor	Subscale	mean (sd)
Recreation	entertainment	2.63 (1.21)
	procrastination	2.72 (1.22)
Self advancement	getting information	3.01 (1.34)
	generating ideas	2.04 (1.01)
	reputation building	1.51 (0.75)
	relationship development	2.22 (1.16)
Helping	helping	2.30 (1.11)

Those with higher levels of expertise differed significantly in their self-reported motivations from those with lower levels of expertise. Repeated measures ANOVA with level of experience (e.g. undergrad, masters ... professor) as a between subjects variable and motivation type as a within subjects variable showed a significant interaction between level of experience and motivation type (F(2,385) = 17.5, $p < 0.001$). Figure 3 shows that individuals with more expertise rated that they participated in MO to help others more than individuals with less expertise. There was no difference for the other kinds of motivations.

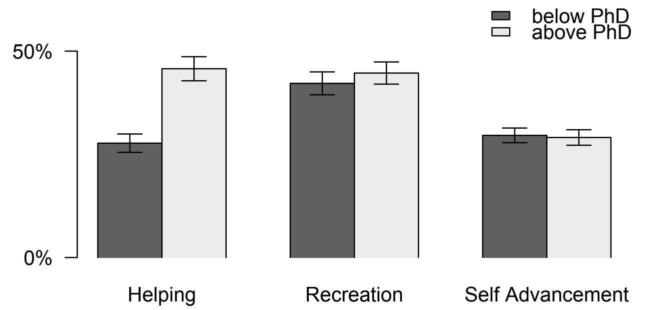

Figure 3. Self-reported reasons for participating in MathOverflow separated by level of expertise. Postdocs, lecturers and professors are grouped into 'above PhD'. Graduate students, terminating PhD students, and undergraduate students are grouped into 'below PhD'.

How do users participate?

The different ways that people participate in the site also give us insight into the reasons why they contribute. We found that users with higher levels of expertise answered more questions and asked fewer questions. Repeated measures ANOVA with level of expertise as a between subjects factor and participation type as a within subjects factor showed a significant interaction between level of expertise and participation type (F(1,194) = 22.7, $p < 0.001$). Figure 4 shows that those beyond the PhD level produced on average more answers (mean = 2.47, sd = 2.68) and fewer questions (mean = 0.40, sd = 0.48) than those at or below PhD level (answers: mean = 1.38, sd = 1.81; questions: mean = 0.62, sd = 0.85).

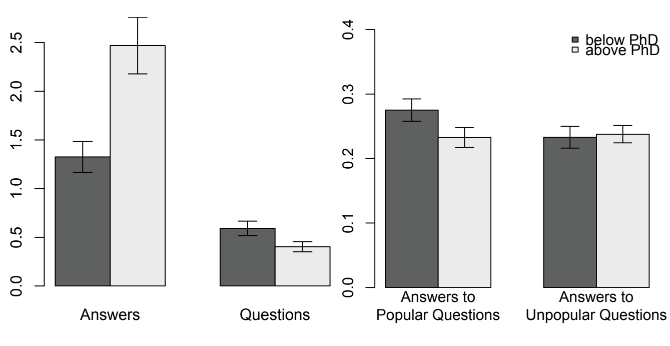

Figure 4. Patterns of participation separated by level of expertise. The left figure gives the rate of questions and answers per two weeks. The right figure gives the rate of answers to popular and unpopular questions controlling for overall rate of answering.

MO is limited to fact and advice oriented questions, so there are not distinct kinds of questions (unlike Yahoo! Answers which has both informational and conversational questions [12]). However there are differences in the popularity of questions. We found that level of experience predicted different rates of answering popular and unpopular questions. Figure 4 shows that users with less expertise had a higher rate of answering popular questions compared to users with more expertise controlling for answering rate overall ($\beta = -0.03$, df = 178, p = 0.006, $R^2 = 0.04$). There was no difference based on level of expertise in the rate of answering unpopular questions ($\beta = 0.01$, df = 178, p = 0.16, $R^2 = 0.01$)

What feedback encourages participation?

The kinds of feedback that encourage participation can reveal the incentives that drive users to contribute (see table 2).

Table 2. Four models identifying the kinds of feedback that encourage or discourage submitting answers. Model 1: does feedback to a users' first answer predict answering again? Model 2: does feedback to answers predict future contributions over multiple timeperiods? Model 3/4: does feedback to answers to popular/unpopular questions predict future contributions over multiple timeperiods? Pseudo R^2 and degrees of freedom are given for the models with and without language predictors (models with language predictors in parentheses).

	Model 1: First Answer odds	Model 2: Answers Overall β	Model 3: Answers Popular Questions β	Model 4: Answers Unpopular Questions β
intercept	0.00	-0.63	-0.05	-1.07
score	9.73***	0.34***	0.21***	0.56***
downvoted	0.53***	-0.05 †	-0.07 †	-0.11 †
accepted answer	1.36*	0.04 †	0.14**	0.41
question asker comment	1.11	0.00	0.09*	-0.01
comments	0.93	-0.03*	-0.06**	-0.01
agreement	0.96	-0.00	0.01	-0.00
disagreement	1.08 †	0.02**	0.02*	0.02*
Pseudo R^2	6.4% (6.8%)	4.5% (8.6%)	5.0% (5.8%)	6.0% (5.9%)
df	2126 (1029)	6076 (4236)	2451 (1492)	2484 (1540)

Signif. codes: '†' 0.10, '∗' 0.05, '∗∗' 0.01, '∗∗∗' 0.001

The results provide evidence that reputation building is an important incentive driving participation. In all four models, receiving a higher score on an answer was the strongest and most consistent predictor of future contributions. Users who received higher scores on the first answer they submitted were more likely to submit answers in the future. Users were more likely to submit more answers after periods in which they received higher scores. Receiving downvotes on one's first answer was strongly related to not answering another question. Periods in which a users' answers were downvoted was marginally related to making fewer future contributions for the other three models. The importance of answer scores and marginal effect of not being downvoted suggests reputation building is an important incentive.

Feedback related to relationship development was only partially related to encouraging future contributions. Of particular interest is the fact that this feedback mattered for answers to popular questions and not for answers to unpopular questions. Periods in which users' answers to popular questions were accepted by the question asker or in which the question asker commented on the answer encouraged future contributions. This was not true for answers to unpopular questions. Relationship development may be especially incentivizing when it is with high reputation individuals or when it occurs with a large audience.

Unexpected findings were found for community responses related to constructive feedback. Receiving comments discouraged participation in two of the models. When comments were given, three models showed that disagreement significantly encouraged participation and one showed it was marginally related to participation. Agreement in comments was not related to participation.

The direction of effects were consistent between what encouraged first answers and answers in general. Answer score was significant for both first answers and answers in general. Whether an answer was downvoted, accepted by the question asker or commenters expressed disagreement with it was marginally significant for one model and significant for the other. Many of the effects were consistent across all four models, which suggests there are consistent incentives to contribute. The largest differences were found between answers to popular and unpopular questions.

Finally, we tested whether level of expertise moderated the effect of answer score, the strongest predictor, on future contributions. Including an interaction effect between level of expertise and score was not significant (β = 0.04, p = 0.73); the main effect of score continued to be significant. There is no evidence that receiving higher answer scores is a larger incentive for those with lower levels of expertise.

DISCUSSION

In this study we described some of the motivations to participate in MO using a self-reported questionnaire and by inferring motivations from behavior on the site. Reputation building was found to encourage participation, although it was the lowest self-reported motivation. Other kinds of community responses also encouraged participation. Users with higher levels of expertise reported using MO to help others more frequently.

Motivations to use MathOverflow

Past work has categorized motivations to participate in online communities into six categories [9, 19]. On MO, three of these, getting information, reputation building, and relationship development were strongly correlated. Within MO these three factors are more strongly interrelated than in other communities. One explanation is that they are all necessary components of individual advancement as a mathematician. Moreover the self-discovery motivation was not relevant to this community. Thus, on MO these motivations factored into only three categories–recreation, helping, and self advancement. If however we examine the original categories, motivations on MO were similar to motivations on Everything2.com in that recreation was reported as most important and reputation building and relationship development were reported as the least important [19].

There was partial support from an analysis of user behavior on the site that reputation building, relationship development and constructive feedback motivated contributions. This will be explained in more detail in a later section.

Level of expertise

Level of expertise was related to only a few key differences in motivations. Users with higher math expertise reported that they use MO to help others more. These users also submitted answers at a faster rate. Users with higher levels of expertise may use MO to help others because they are better able to do so and/or because they feel a desire to share their expertise.

Despite the potential importance of informal recognition for users with lower expertise [2], we did not find that MO users with lower expertise valued reputation building more highly. Users with lower expertise did not report using MO for self advancement more frequently and moreover they did not contribute more in response to getting higher scores on answers. Although one should be cautious in interpreting null results, one explanation is that encouragement from gaining recognition may not plateau as users gain more expertise.

Reputation building, relationship development, and constructive feedback

User behavior was analyzed to complement findings from self-reports of motivations. We found evidence that reputation building did act as an incentive encouraging contributions. Partial support was found that other kinds of community responses also encouraged participation. All of these effects are weak; they explained 4% to 8% of the variance. However, even if these incentives are strong we would only expect to discover weak effects given the amount of error introduced by averaging feedback over time.

The strongest, most consistent evidence was found for the importance of reputation building in encouraging contributions. This is in contrast to self-reports which list it as the lowest rated motivation for participating in MO. The importance of reputation building is consistent with observations of other scientific collaboration. Reputation building on MO is a form of informal recognition and informal recognition can take the place of formal credit in large scale collaborations [2]. There is no formal credit (such as authorship on a paper) on MO, making informal recognition more important. In addition, MO provides the opportunity to gain interpersonal recognition that may not be as available through other means.

There was weak evidence that relationship development acted as an incentive to contribute. The evidence suggested that relationship development was more important when answering popular questions than unpopular questions. It may be of more value when the relationship is with high status question askers or when there is a large audience observing the connection. The fact that relationship development is more important for popular questions suggests it may be another form of reputation building.

There were mixed results for constructive feedback acting as an incentive. Looking first at all contributions, receiving constructive feedback (comments) acted as a disincentive. Second, when we restricted the analysis to those contributions which had comments, then comments that disagreed the most acted as an incentive. Disagreement shows that the commenters took the answer seriously. For example in the comments in response to an answer on MO the person who had asked the question first wrote: "Thanks. It will take some time to see if that's what I need, but it looks interesting." and then "No, this doesn't help. Thanks anyway."

It was this second comment that scored high for disagreement. This response was very polite, but it was critical of the answer which turned out not to be helpful. Even though the answer turned out to be a false lead, the response demonstrated that the questioner seriously considered the user's answer and in this case was glad they had contributed it. Users may be looking for more than simple approval from their audience. It may be important that their submissions were considered substantial contributions even if not ultimately helpful. Having one's answer accepted without comment is better still.

Limitations

The method of inferring motivations from user behavior has some limitations. Although we measure the effect of community response on subsequent contributions, we cannot be certain that it is the feedback that is causing the change in contributions. There are alternative explanations. For example, users may have had streaks in which they posted good answers that got positive responses and also answered many questions. Even if there is a causal relationship, there are multiple psychological reasons why community response might encourage participation. For example, receiving a high score on a question may be motivating because it increases a user's reputation score, or alternatively because it confirms that the user has expertise worth sharing.

Applications and future directions

Authors have argued that content ratings systems could encourage users who post highly appreciated content to contribute more [17, 20]. Our findings confirm that ratings can be used to encourage participation. Altruistic tasks, such as answering questions, have fewer direct benefits. Thus providing explicit social benefits may be important. 'Tricki', the wiki like project to create a repository of mathematics knowledge, was not as successful as MO. This may in part be because there were no social rewards for users who contributed knowledge. On MO users who write answers get credit both in terms of the rating they receive and comments from other users; on Tricki there was no direct credit or feedback from others.

Future work should investigate how to best design a content rating system to encourage participation. In the process of evaluating such a rating system, this work would further clarify the motivations to participate. There are a number of questions left to investigate. One, would answer ratings be more encouraging if they were more differentiated? For example, there could be separate ratings for quality, novelty, providing references, or being thought provoking. This would provide insight into whether users are rewarded by any kind of recognition or only specific kinds. Sites such as Slashdot already use ratings systems that have more clearly refined meanings. Two, would answer ratings be as encouraging if they did not contribute to a user's overall reputation score? Three, would quality ratings be more encouraging if individuals could tell who was rating their answer? If so, which raters (e.g. friends, popular users) would be the most encouraging? Sites such as Daily Kos have this design; content ratings also include a list of the users who cast the votes. The fact that a response from the question asker matters more for popular questions suggests that users may be encouraged by the approval of some MO users more than others.

CONCLUSION

MathOverflow is the first in what may become many participatory online communities of scientific collaboration. In this paper we described some of the motivations that encourage users to participate. We showed that users with more expertise report using MO to help more, but do not differ in using it for reputation building from those with less expertise. Multiple forms of community responses encourage participation, of these reputation building may be especially important.

REFERENCES

1. L. A. Adamic, J. Zhang, E. Bakshy, M. S. Ackerman, and A. Arbor. Knowledge sharing and Yahoo Answers: Everyone knows something. In *Proc. WWW 2008*, 665–674. ACM Press.
2. J. P. Birnholtz. 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*, 57:1758–1770, 2006.
3. P. M. Blau. *Exchange and Power in Social Life* Wiley, New York, 1964.
4. M. Burke, C. Marlow, and T. Lento. Feed me: Motivating newcomer contribution in social network sites. In *Proc. CHI 2009*, 945–954. ACM Press.
5. C. Cheshire. Selective incentives and generalized information exchange. *Social Psychology Quarterly*, 70:82–100, 2007.
6. D. Constant, L. Sproull, and S. Kiesler. The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science*, 9:119–135, 1996.
7. J. Cranshaw and A. Kittur. The Polymath Project : Lessons from a successful online collaboration in mathematics. In *Proc. CHI 2011*, 1865–1874. ACM Press.

8. D. Dearman and K. N. Truong. Why users of Yahoo! Answers do not answer questions. In *Proc. CHI 2010*, 329–332. ACM Press, 2010.

9. U. M. Dholakia, R. P. Bagozzi, and L. K. Pearo. A social influence model of consumer participation in network- and small-group-based virtual communities. *International Journal of Research in Marketing*, 21:241–263, 2004.

10. T. Gowers. http://gowers.wordpress.com/2010/09/24/is-the-tricki-dead/, 2010.

11. A. G. Greenwald, T. A. Poehlman, E. L. Uhlmann, and M. R. Banaji. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97:17–41, 2009.

12. F. M. Harper, D. Moy, and J. A. Konstan. Facts or friends? Distinguishing informational and conversational questions in social Q&A sites. In *Proc. CHI 2009*, 759–768. ACM Press.

13. G. Hertel, S. Niedner, and S. Herrmann. Motivation of software developers in Open Source projects: An internet-based survey of contributors to the Linux kernel. *Research Policy*, 32:1159–1177, 2003.

14. B. A. Huberman, D. M. Romero, and F. Wu. Crowdsourcing, attention and productivity. *Journal of Information Science*, 35:758–765, 2009.

15. E. Joyce and R. E. Kraut. Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication*, 11:723–747, 2006.

16. B. Klandermans and D. Oegema. Potentials, networks, motivations, and barriers: Steps toward participation in social movements. *American Sociological Review*, 52:519–531, 1987.

17. J. Koh, Y.-G. Kim, B. Butler, and G.-w. Bock. Encouraging participation in virtual communities. *Communications of the ACM*, 50:69–73, 2007.

18. K. R. Lakhani and R. G. Wolf. Why hackers do what they do: Understanding motivation and effort in Free/Open Source Software projects, In *Perspectives on Free and Open Source Software* J. Feller, B. Fitzgerald, S. Hissam, K. R. Lakhani, Editors, pp. 3–22. MIT Press, Cambridge, MA, 2005.

19. C. Lampe, R. Wash, A. Velasquez, and E. Ozkaya. Motivations to participate in online communities. In *Proc. CHI 2010*, 1927–1936. ACM Press.

20. L. Mamykina, B. Manoim, M. Mittal, G. Hripcsak, and B. Hartmann. Design lessons from the fastest Q&A site in the West. In *Proc. CHI 2011*, 2857–2866. ACM Press.

21. T. D. Moore and M. A. Serva. Understanding member motivation for contributing to different types of virtual communities: A proposed framework. In *Proc. SIGMIS-CPR 2007*, 153–158. ACM Press.

22. K. K. Nam, M. S. Ackerman, and L. A. Adamic. Questions in, knowledge iN? A study of Naver's question answering community. In *Proc. CHI 2009*, 779–788. ACM Press.

23. O. Nov. What motivates Wikipedians? *Communications of the ACM*, 50:60–64, 2007.

24. O. Nov, M. Naaman, and C. Ye. What drives content tagging: The case of photos on Flickr. In *Proc. CHI 2008*, 1097–1100. ACM Press.

25. S. Oreg and O. Nov. Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Computers in Human Behavior*, 24:2055–2073, 2008.

26. D. L. Paulhus. Measurement and control of response bias In *Measures of Personality and Social Psychological Attitudes*, J. Robinson, P. Shaver, and L. Wrightsman, Editors, pp. 17–59. Academic Press, Inc., San Diego, CA, 1991.

27. J. W. Pennebaker, R. J. Booth, and M. E. Francis. Linguistic Inquiry and Word Count: LIWC, 2007.

28. P. Russo and O. Nov. Photo tagging over time: A longitudinal study of the role of attention, network density, and motivations. In *Proc. ICWSM 2010*, 146–153. AAAI.

29. R. M. Ryan and E. L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25:54–67, 2000.

30. D. Sohn and J. D. Leckenby. A Structural Solution to Communication Dilemmas in a Virtual Community. *Journal of Communication*, 57:435–449, 2007.

31. M. M. Wasko and S. Faraj. It is what one does: Why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, 9:155–173, 2000.

32. M. M. Wasko and S. Faraj. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29:35–57, 2005.

33. F. Wu, D. M. Wilkinson, and B. A. Huberman. Feedback loops of attention in peer production. In *International Conference on Computational Science and Engineering*, 409–415, 2009.

34. X. Zhang and F. Zhu. Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *American Economic Review, Forthcoming*, 2010.