Mining Social Theory to Build Member-Maintained Communities

Dan Cosley
University of Minnesota

Abstract

Online communities need regular maintenance activities such as moderation and data input, tasks that typically fall to community owners. Allowing all members to perform maintenance might make communities more robust and valuable. A key challenge in creating member-maintained communities is that we have little knowledge of how to design interfaces to motivate contributors. Social science theories of motivation may be helpful for making and evaluating design choices. We have used Karau and Williams' collective effort model to develop and evaluate designs that motivate members by making salient the unique contributions they could offer to a group. We also show that using oversight to check contributions to a community database of movies increases members' motivation to contribute and that peers can provide motivation as well as experts in this domain. Our experiments suggest a number of design elements and guidelines for increasing members' motivation to maintain their communities.

The Problem

Max is the movie czar, the de facto dictator of which movies get into the MovieLens recommender system. He's the guardian of quality, the finder of movie facts, the defender of decency–the final authority on film. This suits Max well; he's pleased with the quality of the movie database. Not all MovieLens users agree, however. Max is busy and adds movies slowly, and members sometimes disagree with his criteria for including movies. Useful information (DVD releases, MPAA ratings, etc.) is also not in the database.

Max's community is not unique—many groups have a Max. Most members' contributions to an online community pertain to what Preece calls its *purpose* (Preece 2000). Members post to discussion groups, rate movies, receive recommendations, and read each others' blogs. These contributions are visibly important and constitute the day to day business of the community. However, communities need inputs besides conversation, such as moderation, governance, and the maintenance of databases (i.e., members, movies, FAQs, and histories). These duties usually fall to the owners of the community (Butler *et al.* 2005), the people who envisioned its existence and bought the machines.

Copyright © 2005, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

Max doesn't have to be the only person who adds movies. MovieLens could be designed to allow all members to contribute information. By taking advantage of all members' knowledge and effort, *member-maintained communities* can reduce their dependence on key members while increasing their overall value to everyone and reflecting everyone's desires. However, letting members contribute is not a panacea. Most systems have no interface for contributing. Members need to be motivated to do the job and trained to do it well—and perhaps watched also, because some will make mistakes while others may sabotage the group.

Social scientists, particularly in the fields of social psychology and economics, have tried to understand what motivates people to contribute to groups. Theories that address motivation to contribute can be a valuable tool for designing interfaces for member-maintained communities. We have identified a number of potentially useful theories, including social capital (Putnam 2000), communities of practice (Lave & Wenger 1991), the problem of public goods (Hardin 1982), and economic models of why people contribute (e.g., (Rabin 1993)). The primary tool we have used so far is Karau and Williams' collective effort model of motivation in group contexts (Karau & Williams 1993).

In this paper we report on our use of the collective effort model in designing interfaces that motivate people to contribute to an online community. We provide a high-level look at how designers can make use of the model, as well as brief reports on two experiments using the model to solicit contributions of ratings and of discussion posts. We then report in more detail on how we used the model to reason about how editorial oversight would affect the quality and quantity of movie information members contributed to MovieLens. As predicted, oversight improved database quality, reduced antisocial behavior, and increased the quantity of contributions. We close with a sketch of our plans to use the model, along with economists' theories of why people contribute to public goods, to intelligently match tasks with people who are likely to do them.

The Collective Effort Model

Karau and Williams' *collective effort model* (Figure 1) integrates theories based on a number of earlier studies showing

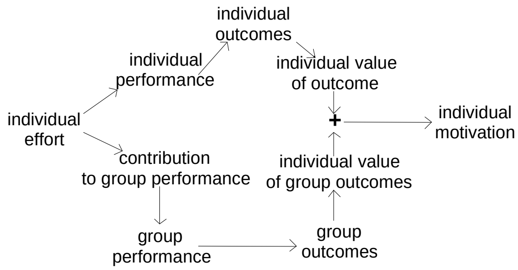

Figure 1: A slightly simplified version of Karau and Williams' collective effort model. Motivation depends on how well effort translates to performance, how effectively performance generates outcomes, and how much people value the outcomes. People derive motivation both from individual goals and group goals.

that people tend to contribute less in group contexts than they would on their own (e.g., (Kerr 1983; Harkins 1987)), a phenomenon called *social loafing* (Latane, Williams, & Harkins 1979). The model builds on Vroom's expectancy-value model (Vroom 1964), which suggests that one's motivation to make a given effort depends on how well one's effort will translate into good performance, whether good performance will lead to good outcomes, and how much one values those outcomes. The collective effort model posits that people reason about this both for their own outcomes and for the group; further, in groups, people also consider whether their contribution matters to the group. All other things being equal, as elements in the model increase, so does motivation. For example, a better interface that allows people to enter twice as many movies for a given amount of effort should increase their motivation to contribute.

In principle, every arrow and node in Figure 1 provides a way for designers to evaluate or critique their designs. Some critiques, such as making more efficient and usable interfaces to increase the performance for a given effort, are fairly obvious[1]. Others, however, are more subtle. People will be more likely to contribute if they value good outcomes for the group. This suggests that people who are more attached to a group are more likely to contribute to it—thus designs that induce people to form groups or teams such as the ESP Game (von Ahn & Dabbish 2004) may increase motivation. Further, a number of studies in the sub-area of persuasive computing have shown that people react socially to computers much as they do to people (Nass & Moon 2000). It might be possible to build a credible computer player for the ESP Game using standard information retrieval techniques; persuasive computing suggests people would likely be just as motivated to contribute picture labels when playing the computer as they are when playing with another human.

As another example, consider the top reviewers feature at amazon.com, shown in Figure 2. Amazon encourages members to write reviews of the items it sells, essentially

[1]However, designers regularly forget the basics. For example, the interface for updating a restaurant's information in ChefMoz, http://chefmoz.org/cgi-bin/update.pl, is complex.

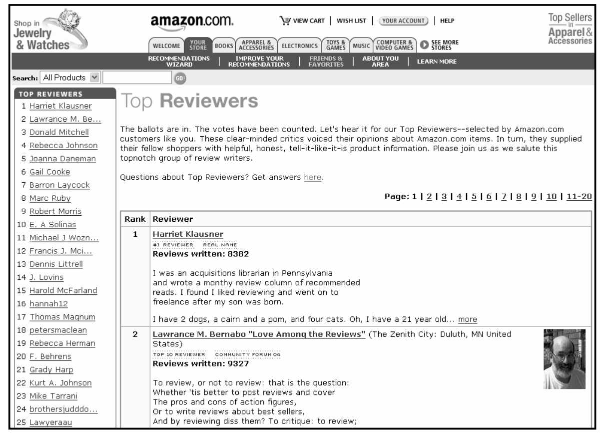

Figure 2: Amazon's top reviewers list. The interface makes contributions salient by showing a list of most helpful reviewers, as well as providing individual recognition through icons next to each top reviewer's name.

getting recommendation content for free from its members. Amazon readers can then vote on whether a particular review is helpful. These votes are used to identify amazon's most helpful reviewers, who receive recognition on a list of top reviewers as well as a "top N reviewer" icon next to the reviews they write. Naively applying the collective effort model suggests that recognizing top contributors will increase people's motivation to contribute. First, by showing people that they are among the most helpful members on amazon, it makes it clear that their contributions matter to the group. Second, the icons provide recognition and visible status, an individual outcome many people value.

However, newcomers are at a distinct disadvantage. They might look at the top reviewers list and decide that no matter how hard they work, they could never catch up to people who had been there for a while—no matter how good their performance, they will never achieve the outcome of individual recognition. Seeing that top reviewers have contributed hundreds of reviews, newcomers might also decide that their contribution of a few reviews will not matter to the group.

In other words, the model predicts that systems that recognize all-time top contributors may motivate long-time members but reduce the motivation of newcomers. Recognizing both all-time and recent valuable contributors might motivate a larger part of amazon's audience to write reviews.

Experiments Using the Model

We have conducted several experiments where we have used the collective effort model to design interface elements and predict how people will react to them. Below, we report briefly on two experiments showing that people contribute more when they believe their contributions are unique. We then report in more detail recent work showing how the checking of contributions affects people's motivation to contribute to a community database. All three experiments use the MovieLens recommender system as an experimental platform. MovieLens (see Figure 3) is a relatively loose but

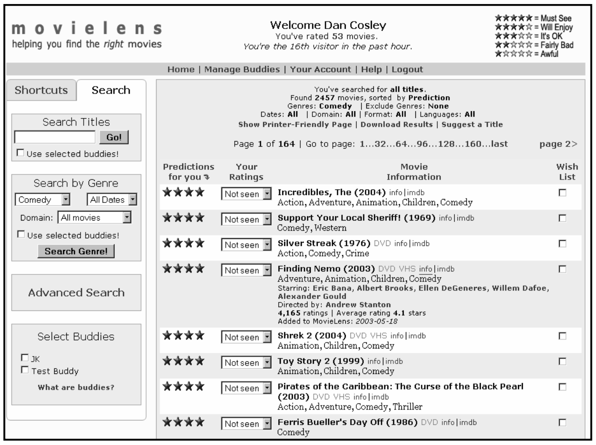

Figure 3: Getting a recommendation for a comedy using the MovieLens movie recommender.

active community, with several thousand members who log in regularly to rate movies and receive recommendations.

Uniqueness Motivates Contributions

The collective effort model predicts that if people believe their effort makes a significant contribution to the group's performance, they will be more motivated to contribute. One way to help people see their effort is important is to identify unique skills a contributor brings and make them aware of their special value to the group.

In one experiment, we invited MovieLens users to participate in online discussions of movies (Ludford *et al.* 2004). Subjects were assigned to one of eight small (about 30 member) discussion groups. The experiment lasted for five weeks; subjects were asked to contribute at least once per week and encouraged to discuss anything related to movies. The experimenters also posted a feature topic every week and notified subjects about the feature topic by email.

Feature topics were used to test the effect of providing information about uniqueness. In half of the groups, the system found three movies that each member had rated that were relevant to the feature topic (e.g., romances for the topic "What makes a love story click?") and that few other members of their group had rated[2]. The weekly email mentioned the movies chosen for each subject, along with how many other discussion group members had rated the movie, saying that this information might be relevant to the featured topic. Subjects who received this information on their potential unique contributions made about twice as many discussion posts on average as subjects who received unpersonalized mail that simply introduced the feature topic.

A second experiment used uniqueness information in an email campaign to motivate people to contribute more ratings to the community (Beenen *et al.* 2004). All subjects were told that they had been specially chosen for the experiment because of their past contributions to MovieLens. Subjects in the unique condition were told that they were chosen because people with unusual tastes are especially valuable to MovieLens. Subjects in the non-unique condition were told they were chosen because people with typical tastes are especially valuable to MovieLens. The system also listed three movies the subject had rated; the unique group saw movies few others had rated, while the non-unique group saw movies many others had rated. Subjects whose mail emphasized their unusual tastes were more likely to log in during the experiment and gave 18% more ratings than subjects whose mail emphasized their typical tastes. These results further support the prediction that designs which emphasize uniqueness are more likely to motivate contributors.

[2]The experimenters chose the set of relevant movies for each topic by hand. Algorithms for automatically suggesting relevant movies based on the conversation would be interesting future work.

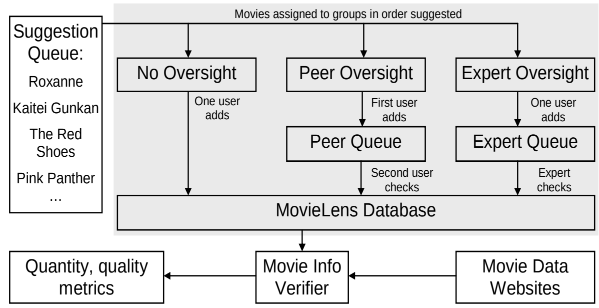

Figure 4: An overview of the oversight experiment.

Editorial Oversight Affects Motivation

The previous two experiments focused on the general problem of soliciting contributions to a group. In this section, we provide a more detailed description of an experiment where we asked MovieLens members to help maintain the database by contributing information for new movies (Cosley *et al.* 2005). One key problem when building community repositories of information is that they are more valuable if the information is of high quality. This means we would like like mechanisms that motivate people to contribute both more and better quality information.

We chose to look at editorial *oversight*—having someone check the quality of contributions—as a potential mechanism. Many communities use oversight in an effort to achieve higher quality information. We wanted to find out what effect oversight would have on people's motivation to contribute information, as measured by quantity and quality. We also wanted to find out if peers would be as effective at providing oversight as experts.

Figure 4 shows an overview of the experiment. Subjects were assigned to one of three oversight mechanisms: no oversight, oversight from peers, or oversight from a movie expert. We told half of the subjects in each group about the level of oversight. Subjects were asked to add and verify movie information. At the end of the experiment, we compared subjects' contributions to movie information available on other websites to evaluate quality and counted total, useful, and hijacked contributions to evaluate quantity.

We used the collective effort model to make a number of predictions about the effect of oversight. For example, people were asked to enter information for movies that other members had suggested in the past. Nothing prevented them from *hijacking* suggestions, adding movies they personally cared about instead of those assigned by the system. The model predicts that since oversight will reduce the link between performance (doing the hijacking) and outcome (having the cared-about movie in MovieLens), oversight will reduce the motivation to hijack movies. The model also suggests that transparency is important. Telling people about oversight gives people more information to help them determine the outcomes their effort will provide, which should increase the effect of oversight on motivation. Further, we predicted that oversight would reassure contributors that their efforts would matter, instead of being drowned in a sea of bad information, and thus that oversight would motivate people to contribute more.

In all, we generated five hypotheses about the quality and quantity of information members would enter:

- More oversight leads to fewer hijackings.
- More oversight leads to higher quality initial contributions.
- More oversight leads to more contributions overall.
- More oversight leads to higher final quality information in the database.
- Knowing about oversight will increase its effect.

Group	Subjects	Contrib.	Total Work	Live Movies	Hijacks
NoneVis	31	10	61	20	41
NoneNV	33	14	88	67	21
PeerVis	38	19	130	63	11
PeerNV	35	13	64	38	2
ExpertVis	32	17	140	109	11
ExpertNV	35	14	92	83	7

Table 1: A summary of participation in the experiment.

The experiment lasted six weeks. A total of 204 users signed up; 87 made at least one contribution. Table 1 shows how many subjects were assigned to each group, how many contributed, and the total number of contributions made and movies added by each group.

Figure 5 shows the average quality of contributions made by each group at two stages:

- *Initial quality*: The quality of a given movie's information when it was first added by a member. Looking only at the initial contribution for a given movie let us isolate the effect of oversight on individuals' motivation to enter accurate information from the effect oversight would have on final database quality because of the double-checking of information.
- *Final quality*: The quality of a given movie's information when it was inserted into the database. This allowed us to look at how oversight's checking of information affected the bottom line quality of the database.

Differences on initial quality, though large, were not statistically significant because variation in quality between

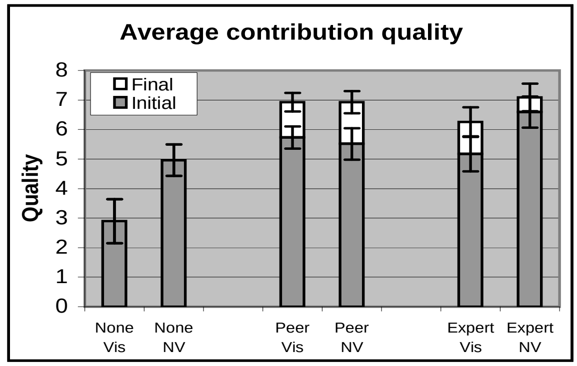

Figure 5: Initial and final contribution quality for a given movie. For groups with no oversight, initial and final quality were the same. No oversight is worse on final quality, especially in the group that knew they had no oversight.

users was huge. However, final quality in the no oversight group was significantly worse than the quality of the peer and expert oversight groups. Groups that knew about oversight did the most work and the group that knew there was no oversight did the least. Groups with no oversight were also much more likely to hijack movies whether we told them about the lack of oversight or not. These differences were both large and statistically significant.

A survey backed up these results. We told MovieLens members about each of our mechanisms and asked them to rate how well each mechanism would work and how likely they would be to add movie information under that mechanism. People preferred expert oversight to peer oversight and peer oversight to no oversight. Willingness to participate and estimated quality correlated strongly.

However, although people preferred expert oversight to peer oversight in the survey, there was no difference in their actual behavior. Peer and expert groups had essentially the same initial and final quality and no differences in quantity of contributions. The expert group did add more movies, because every contribution from a member was matched by a contribution from our movie expert.

Based on our findings, we developed several guidelines for using oversight in maintaining community databases:

- Oversight improves outcomes and increases contributions. Use oversight to improve quality and reduce antisocial behavior.
- We found no difference between peer and expert oversight in quality or quantity. Use peer oversight to share the load.
- Telling people about oversight may increase their motivation to contribute. If you use oversight, say so.
- Major differences in quality can be attributed to individuals. Identify skilled contributors, and improve the capabilities of individual users, e.g., through training.
- A number of users said they didn't see the experiment invitation. Make opportunities to contribute obvious.

There are a number of caveats. We were excited that peers and experts performed about as well at providing editorial oversight in MovieLens because this means letting members provide oversight is a reasonable strategy for building member-maintained communities. However, a frequent criticism of the Wikipedia collaborative encyclopedia[3] is that non-expert contributors harm its quality. It may be that as the perceived need for expertise increases, people will only accept oversight from experts. On the other hand, Wikipedia has generated over 400,000 articles by letting peers provide oversight. Contrast this with the extensive editorial review process for the Nupedia collaborative encyclopedia—and the 100 articles it produced.

Oversight done badly may reduce motivation. Consider this response to someone who takes a minute to update a restaurant's phone number in the ChefMoz community restaurant directory:

> Your comments will be reviewed by an editor, who will decide whether to include them in [ChefMoz]. This may take as little as a week, or up to a few months...

It would not be hard to use the collective effort model to analyze the ways this message might affect one's motivation to contribute in the future.

Finally, although people liked having oversight in the domain of adding information to MovieLens, many people are philosophically opposed to other forms of oversight, such as moderated discussion groups. More work is needed to determine the contexts in which oversight increases motivation.

From Guidelines to Algorithms

Our work thus far has focused on using the collective effort model at a high level and producing guidelines based on theory-driven designs tested in field experiments. Condensing theory down to design guidelines is a useful step in making theory more accessible to designers. We would like to make theory even more useful by designing algorithms for reasoning about motivation that can be directly deployed when building communities. Below, we look at how economics research might help us model and affect people's motivation to contribute to communities and briefly point to other social science theories that designers of member-maintained communities might find relevant.

Economists often model the problem of contribution to groups as a *public goods* problem. Public goods are resources that are *non-excludable* and *non-rivalrous*; that is, once produced, everyone can consume them and one person's consumption does not affect another's consumption. National defense and open source software are two real-world examples of public goods. Public posts, moderation, FAQs, and contributions to a group-specific database are examples of public goods in online communities.

A key characteristic of public goods is that they are susceptible to *free riders*, people who consume the good without helping to produce it. For instance, a small percentage of public radio listeners contribute money; non-contributors free ride on the few. Selfishly rational people might reason that if others care enough to provide the good, they need not expend the effort (Kerr 1983). A number of lab experiments have demonstrated this tendency to free ride in public goods situations (see (Ledyard 1995)).

However, not everyone free rides. Many experiments have shown that people contribute to public goods under some conditions (Dawes & Thaler 1988). In other words, people sometimes do not maximize their own individual utility. Economists have responded by modelling factors in addition to the value and cost of a contribution. For example,

- *Reciprocity*: people make or withhold contributions to others based on how others treat them (Rabin 1993).
- *Inequality aversion*: people adjust their effort based on their perception of what others contribute on average (Fehr & Schmidt 1999).
- *Increasing social welfare*: people sometimes make decisions that increase the welfare of all, especially those who are worse off (Charness & Rabin 2002).

This line of research holds promise for building member-maintained communities. Most likely, different people are motivated to different degrees by each of these factors. An algorithm might be able to model how much a particular person is motivated by reciprocity, inequality aversion, social welfare, rational self-interest, and so on by observing their behavior in the community. It could then use the model to present interfaces that are most likely to motivate that person. For example, the COMTELLA peer-to-peer file sharing system tries to model whether people are altruistic, rational, or selfish based on their file-sharing behavior. The interface shows graphs of contribution and consumption as well as motivational images and text, choosing which graphs, images, and text to show based on its inferences about each user's motivation (Bretzke & Vassileva 2003).

It may also be profitable to consider aspects of tasks that theories predict might affect motivation (such as task difficulty or the task's value to the individual and to the group). If we can compute proxies for difficulty and value, we can build models that use information about both tasks and people to algorithmically match people with tasks they are most likely to do. One reason Wikipedia works is that it highlights recently changed pages so members can review others' contributions. However, Wikipedia gets thousands of contributions per day, making it hard for people to find contributions they might care about. By intelligently routing changes to people who are most likely to care about checking them, we can increase motivation to contribute and the quality of our database while reducing contributors' workload.

Other social science theories may offer insights regarding contribution that the collective effort model and public goods theory do not provide. *Social capital* is the idea that relationships between people are a valuable resource that provide access to otherwise privileged information or interactions (Putnam 2000). Contributions to both specific members of a community (Lesser 2000) and to the collective interest of the community (Coleman 1988) can help build a reservoir of social capital that encourages further contributions and allows members to make requests of others in the community. Social capital is an attractive idea for designing online communities because, although it can be hard to measure offline, online communities can track and measure

[3] http://www.wikipedia.org/

members' activities. They can then use the data collected to help members of the community see each others' contributions. Technologies such as reputation systems (Resnick *et al.* 2000) and tools for revealing social information (e.g., (Smith 1999)) are examples of systems that can use measures of social capital to provide recognition and status.

The theory of *communities of practice* (Lave & Wenger 1991) addresses the evolution of groups that share common knowledge. A community of practice facilitates the growth of members in the course of "assuming an identity as a practitioner, of becoming a full participant, an oldtimer" (Lave 1993). One way to analyze the kinds of contributions people make to communities is through the roles they assume (Preece 2000). The theory suggests designs that structure maintenance tasks to provide multiple roles that require increasing levels of expertise and commitment may prove useful. In a community with discussions, for example, members may move from lurkers to question askers, then to answering questions, and from there to building FAQs, moderating discussion, and mentoring new members.

Conclusion

Although social science theory is hard to apply directly to system design, our experience suggests that careful consideration of these theories can improve the chances of success when asking people to contribute to a community. We have conducted several experiments that use the collective effort model to generate and evaluate mechanisms for increasing people's motivation to contribute to a group. By providing successful examples of the use of theory and identifying guidelines based on theory-driven field experiments, it should be easier for designers to incorporate theory into their design processes. Future work that explores other promising theories and distills theory into directly usable computational tools promises to further increase our ability to build successful member-maintained communities.

References

Beenen, G.; Ling, K.; Wang, X.; Chang, K.; Frankowski, D.; Resnick, P.; and Kraut, R. E. 2004. Using social psychology to motivate contributions to online communities. In *Proceedings of CSCW 2004*.

Bretzke, H., and Vassileva, J. 2003. Motivating cooperation on peer to peer networks. In *User Modeling*, 218–227.

Butler, B.; Sproull, L.; Kiesler, S.; and Kraut, R. 2005. *Community Building in Online Communities: Who Does the Work and Why?* Leadership at a Distance. Mahwah, NJ: Lawrence Erlbaum Publishers, Inc.

Charness, G., and Rabin, M. 2002. Understanding social preferences with simple tests. *The Quarterly Journal of Economics* 117(3):817–869.

Coleman, J. 1988. Social capital in the creation of human capital. *American Journal of Sociology* 94:95–120.

Cosley, D.; Frankowski, D.; Kiesler, S.; Terveen, L.; and Riedl, J. 2005. How oversight improves member-maintained communities. In *Proceedings of CHI 2005*.

Dawes, R. M., and Thaler, R. H. 1988. Anomalies: Cooperation. *Journ. of Economic Perspectives* 2(3):187–197.

Fehr, E., and Schmidt, K. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114:817–868.

Hardin, R. 1982. *Collective Action*. Baltimore: Johns Hopkins University Press.

Harkins, S. G. 1987. Social loafing and social facilitation. *Journal of Experimental Social Psychology* 23:1–18.

Karau, S. J., and Williams, K. D. 1993. Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology* 65(4):681–706.

Kerr, N. L. 1983. Motivation losses in small groups: a social dilemma analysis. *Journal of Personality and Social Psychology* 45:819–828.

Latane, B.; Williams, K.; and Harkins, S. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology* 37:822–832.

Lave, J., and Wenger, E. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge Univ. Press.

Lave, J. 1993. *Situating Learning in Communities of Practice*. Perspectives on Socially Shared Cognition. American Psychological Association.

Ledyard, J. O. 1995. *Public Goods: A Survey of Experimental Research*. The Handbook of Experimental Research. Princeton, NJ: Princeton Univ. Press. 111–194.

Lesser, E. 2000. *Leveraging social capital in organizations*. Knowledge and Social Capital. Butterworth-Heinemann.

Ludford, P. J.; Cosley, D.; Frankowski, D.; and Terveen, L. 2004. Think different: increasing online community participation using uniqueness and group dissimilarity. In *Proceedings of CHI2004*, 631–638.

Nass, C., and Moon, Y. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 60(1):81–103.

Preece, J. 2000. *Online Communities: Designing Usability, Supporting Sociability*. John Wiley & Sons, Ltd.

Putnam, R. 2000. *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster.

Rabin, M. 1993. Incorporating fairness into game theory and economics. *Am. Econ. Review* 83(5):1281–1302.

Resnick, P.; Kuwabara, K.; Zeckhauser, R.; and Friedman, E. 2000. Reputation systems. *Communications of the ACM* 43(12):45–48.

Smith, M. 1999. *Invisible Crowds in Cyberspace: Measuring and Mapping the Social Structure of USENET*. Communities in Cyberspace. Routledge Press, London.

von Ahn, L., and Dabbish, L. 2004. Labeling images with a computer game. In *Proceedings of CHI2004*, 319–326.

Vroom, V. H. 1964. *Work and Motivation*. New York: Wiley.