Motivating Participation by Displaying the Value of Contribution

Al Mamunur Rashid[1]**, Kimberly Ling**[3]**, Regina D Tassone**[3]**,**
Paul Resnick[2]**, Robert Kraut**[3]**, John Riedl**[1]

[1]University of Minnesota ({arashid, riedl}@cs.umn.edu), [2]University of Michigan
(presnick@umich.edu), [3]Carnegie Mellon University ({kling, rdt, kraut}@andrew.cmu.edu)

ABSTRACT

One of the important challenges faced by designers of online communities is eliciting sufficent contributions from community members. Users in online communities may have difficulty either in finding opportunities to add value, or in understanding the value of their contributions to the community. Various social science theories suggest that showing users different perspectives on the value they add to the community will lead to differing amounts of contribution. The present study investigates a design augmentation for an existing community Web site that could benefit from additional contribution. The augmented interface includes individualized opportunities for contribution and an estimate of the value of each contribution to the community. The *value* is computed in one of four different ways: (1) value to self; (2) value to a small group the user has affinity with; (3) value to a small group the user does not have affinity with; and (4) value to the entire user community. The study compares the effectiveness of the different notions of value to 160 community members.

Author Keywords

Value of contributions, recommender systems, identity.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Since the advent of the Internet, an ever-increasing number of online communities have been formed. However, in most online communities the lion's share of contributions usually comes from a small fraction of the members [1,2]. Although it is not necessary for each member to contribute in order for online communities to flourish [6], in a number of cases, both the individuals and the community as a whole benefit from every contribution. We will discuss MovieLens (ML), an online movie recommendation community where individual ratings of movies help members receive better personalized recommendations from the system. Though ML now invites other types of contributions, including forum postings and actor and director information about movies, the present study focuses on *rating* as the mode of contribution. Although MovieLens receives about 30,000 ratings each week, MovieLens could benefit from additional ratings: more than 20% of the movies have so few ratings that MovieLens cannot make personalized recommendations for them.

This paper investigates ways to encourage additional ratings from users, by making it clear to them the *value* their ratings provide to themselves as well as to others. In this study, value can be measured either from the perspective of the user who provides the rating, who benefits from more accurate recommendations as MovieLens understands his taste better, or from the perspective of other members of the community, who benefit by receiving more accurate recommendations for the newly rated movie. This paper explores the effect of presenting different types of value to users. The results of this study are intended to aid community designers whose goal is to build successful online communities. The designers will benefit from understanding how to motivate contributions to their communities.

We utilize Karau and Williams' Collective Effort Model (CEM) [4] in our approach. The CEM suggests conditions under which people might increase contributions. These include believing that their effort is important to the group's performance, believing that their contributions to the group are identifiable, and liking the group they are working with. Other theories (e.g. [9]) suggest that the identity of those helped plays a big role in drawing people's support and contribution. By using these theories we formalize the following research questions with respect to the community of context:

1. Will MovieLens subscribers rate more movies if they receive feedback about the value that ratings have to themselves and others?

2. Does the *identity* of the beneficiary matter?
   - Will users be more motivated to help themselves or others?
   - Will users be more motivated to help the full MovieLens community or subgroups of it?

Figure 1. A part of a sample screenshot from a user in the *Similar Group* experimental condition.

- Will users be more motivated to help subgroups that are similar to them in some way?

Note that this work can be considered as a continuation of the research presented by Ling et al [5]. In their studies they addressed the under-contribution problem and used the same online community as the test-bed. Their approach included sending motivational email messages to MovieLens members conveying unique abilities and contributions of the members. Ling et al. were able to produce significant changes in participation behavior; however, the results were mixed and more research is needed in this area.

In Ling et al., telling people about the value of their contribution either to self or to others *depressed* their willingness to contribute. One possible explanation is that by explicitly giving people a rationale for contributing, a psychological reactance [3] was created. In the current research, we test our experimental hypotheses with a stronger manipulation by changing the current MovieLens interface. By building the estimate of value into the interface a) we hope to give people the value-information without it being part of a persuasive message and thereby reduce reactance; and b) we wanted to create an ongoing estimate of value, which could actually become part of the standard interface. As in the past studies, both the individual and the community as a whole will be highlighted to motivate contribution. In addition, subgroups of interest to the member will be made salient.

EXPERIMENTAL DESIGN
We provided the subjects with a list of movies they might be able to rate based on historical information about the types of movies each person rated in the past. Each movie was annotated with the value some beneficiary would receive if the participant rated it by including smilies next to each movie (figure 1), where three full smiley-faces suggested the maximum value and half of a smiley-face indicated the smallest value.

Experimental Groups. We designed four main experimental groups reflecting different types of beneficiaries of ratings. The fifth group, the control group was used as a

baseline. The groups are:

Self. In this group, the beneficiaries of the ratings are the subjects themselves. Therefore, the smilies indicated how much rating a given movie would benefit the subject.
Similar Group. In this case the beneficiaries are the Movie-Lens subscribers who like genres of movies that the subjects also like. In this regard, the smilies for this experimental condition indicated how much rating a given movie would benefit members of a particular genre-fan group (with which the participant could identify).
Dissimilar Group. This experimental condition is similar to the *Similar Group* except the subjects do not like the genres that the beneficiaries like.
All. The beneficiaries here are all MovieLens users.
Control. This experimental condition does not involve any beneficiary at all. The smilies here simply indicate the subjects' likelihood of having seen the movies in the past.

The Experiment. We designed a *between subjects* experiment; it was also a *one-shot* experiment. Users were invited to participate in the experiment. Upon accepting the invitation they were assigned to one of the experimental groups randomly. Subjects then saw a list of 150 movies annotated with the value to some beneficiary and rated zero or more movies. At the end they filled out a short survey. Note that we had 32 subjects for each of the five groups.

Hypotheses.

- **H1.** Value > No value. Users will be more likely to rate movies with more smilies than fewer ones, when the smilies indicated that the rating was valuable to a beneficiary. In the control condition, the number of smilies should not influence the likelihood of rating the movie. Moreover, users in the control condition will rate fewer movies compared to the groups that show values.

- **H2.** Self > Others (All, Similar Group, Dissimilar Group). Most economic utility theories assume that people are selfish and would rather help themselves than others. Therefore, smilies should have a greater influence on the likelihood of rating a movie if they indicate value to the subject than value to any of the other beneficiaries.

- **H3.** Subgroup (Similar Group, Dissimilar Group) > All. People feel greater concern towards others as the reference group that they are part of grows smaller. Therefore, smilies should have a greater influence on the likelihood of rating a movie if the smilies indicate value to a particular fan-group rather than all MovieLens users.

- **H4.** Similar Group > Dissimilar Group. Because people like others who are similar to themselves, smilies will have a greater influence on the likelihood of rating a movie if the smilies indicate value to a fan-group similar rather than dissimilar to themselves.

Post-Survey. At the end of the experiment, participants filled out a short survey. The goal of this survey was to evaluate, first, participants' assessments of how effective our manipulations were toward motivating contributions and,

second, their evaluation of how usable and desirable the user interface was for displaying value.

METHODS

Genre Fanship. For discovering which genre group a user would identify with, we ran a preliminary experiment in which we tested several algorithms for computing genre group for a user versus users' self reports of preferred genre groups. Approximately 150 MovieLens subscribers rated how much they liked each of 10 genres on a 7-point Likert scale. We developed a regression model that predicts users' genre preferences from the historical movie-rating data. Historical data included the number of movies in each genre subjects had previously rated, their average rating of movies in this genre, standard deviation of this user's ratings for this genre, and the *tfidf* (term-frequency × inverse document frequency) [7] of this genre.

The best linear combination of these variables explained 34% of the variance in subjects' preferences for the genres. The best algorithm (which we then used in the later experiments) did a pretty good job of selecting the top few genre groups for a user, but was sometimes wrong about the single top group. Therefore, in the study we decided to provide a list of the top (bottom) 5 genre groups the participants were likely (least likely) to identify with, and let them choose from that list.

Value of Ratings. We defined value of a rating as the improvement in the future accuracy of recommendations for some group of users. Note that, fundamentally, we displayed two broad types of values using the smiley-interface—value-to-self and value-to-others. Next we briefly mention the procedures we adopted for computing these values.

As of this writing, MovieLens uses an item-based algorithm [8] to produce movie recommendations. In this algorithm, first the similarities $w_{a,b}$ between each pair of movies a and b are computed. Then the prediction for the target user u_t and the target movie a_t is computed by taking a weighted average over the movies rated by u_t:

$$\hat{R}_{u_t,a_t} = \frac{\sum_{all\_similar\_items,d}(w_{a_t,d} * R_{u_t,d})}{\sum_{all\_similar\_items,d}(|w_{a_t,d}|)} \quad (1)$$

where $R_{u_t,d}$ is user u_t's rating on the movie d. From the equation it is clear that one of the 'controllable' factors leading to the accuracy of the prediction is the term $w_{a_t,d}$—similarity of a *neighbor* movie with the target movie. From this observation we postulate that rating movies that have high similarity with many other movies might improve recommendation accuracy. We verified this by offline data analysis, and therefore, we consider, for a particular movie: *value-to-self* $\propto \sum$ *similarity of the movie with other movies*.

For computing values-to-others we adopted an *all-but-N* approach described very briefly as follows. In this approach we hold out N ratings of a chosen movie and note aggregated gains (losses) in recommendation accuracy over all users. At the same time we record various features of the movie including number of ratings, popularity, average

rating value, and entropy of ratings. Once we repeated this process for 1,000 randomly chosen movies, we get a dataset to perform regression. Our objective is to learn a function so that given the same set of features on a new movie; the function is able to output the gain in accuracy if the movie receives an additional rating. We used an SVM non-linear regression since it gave the best performance (10-fold cross-validation $R^2 = 0.38$) among the regression methods we tried. We used the output from the regression model; in other words, the predicted gain (loss) in accuracy is an estimate for the value-to-others.

Movie List. We decided to present personalized 150 most ratable movies to each user. We did so because a) most users have seen only a fraction of the movies in MovieLens, and b) due to the one-shot nature of our experiment, we did not want inability to rate to mask the effects of *value* and *identity*. The most ratable movies for each user were compiled by utilizing an *Item-item* similarity algorithm, where a user's list of movies rated to date is used to identify other movies highly probable to have seen by the user. For each of these 150 movies, the smilies were computed according to the appropriate beneficiary as described above. The numerical benefit was converted into discrete smilies which was within the range of 0.5 to 3.0 with a 0.5 increment. Note that, the message delivered to the subjects in the control group was different, lacking any concept of *values* or *beneficiaries*.

RESULTS AND DISCUSSION

In this section we present results compiled from two sources: logged behavioral data and self-report survey data.

The surveys indicated that the the manipulation seemed moderately effective. More than 90% of the users reported having seen the movie was the most important determinant of whether they would rate a movie; regressions from the logged data show that ratability is indeed the best predictor of whether they rate a movie (p<.0001). According to self-reports, users indicate that the second most common attribute influencing whether to rate was the number of smilies (11.5%). Again, regression shows that the number of smilies predicts the likelihood of rating a movie (p<.0001).

Overall, users picked *somewhat/neutral* on a 5 point Likert scale to describe their understanding of smilies, value from smilies, and interest in having smilies in the future (figure

(a) Inferred probability of rating a movie from the logged data. (b) Preference for smilies in the regular MovieLens interface.
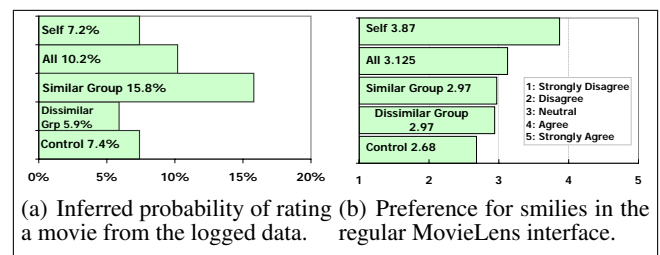
Figure 2. Notice the difference between (a) the #of movies rated in the smilies conditions and (b) the self-reported preference for smilies.

Table 1. Contrasts testing the four hypotheses.

Contrast	Least Sq Mean #1	Least Sq Mean #2	Prob $> \|t\|$	Hypothesis Confirmation
H1	0.105	0.068	$< .001$	confirmed
H2	0.069	0.117	$< .001$	disconfirmed
H3	0.124	0.102	$< .002$	confirmed
H4	0.161	0.086	$< .001$	confirmed

2(b)). The Self group was nearly a full point higher on each of these ratings. As shown in figure 2, self-report data indicate that users in the Self group had the most interest in adding smilies to the ML interface. In contrast, users in the Similar group were the ones most likely to rate movies.

Table 1 shows the result of the *contrast-tests* corresponding to our hypotheses. Note that H1 was supported—subjects were 3.7% more likely to rate a movie when smilies indicated value than when they indicated likelihood of having seen a movie. H2 was disconfirmed—subjects were 4.8% less likely to rate a movie when smilies indicated value to self than when they indicated value to another. H3 was supported—subjects were 2.1% more likely to rate a movie when smilies indicated value to a subgroup than when they indicated value to MovieLens subscribers as a whole. H4 was supported—subjects were 7.4% more likely to rate a movie when smilies indicated value to people who liked genres that the subject liked than when they indicated value to people who like genres the subject did not like.

Figure 3 shows the interactions plots between the number of smilies on a movie and experimental conditions. Overall, subjects were more likely to rate movies with more smilies. However, smilies had larger effects in the Similar beneficiary condition than in the other conditions and had a weaker effects in the Self beneficiary and Dissimilar beneficiary conditions than in the other conditions.

One thought about why participants that viewed the Self condition screens did so poorly is that users' experience may have indicated to them that we were simply wrong about their additional ratings adding much value to themselves. That is, they may have noticed in the past that once they had a couple of hundred ratings, additional ratings did not seem to improve their predictions. Therefore, they may not have believed our assertion that these additional ratings would benefit them.

One important possible confound was that even though subjects were always presented with the 150 movies they were most able to rate and we controlled for ratability in the analyses, smilies in the Dissimilar Group condition may have highlighted movies subjects did not have an opinion about, since they do not like that genre.

CONCLUSIONS AND FUTURE WORK
In this paper we approached the problem of under-contribution in online communities. Past research attempting to motivate contribution has produced mixed results when distinguishing
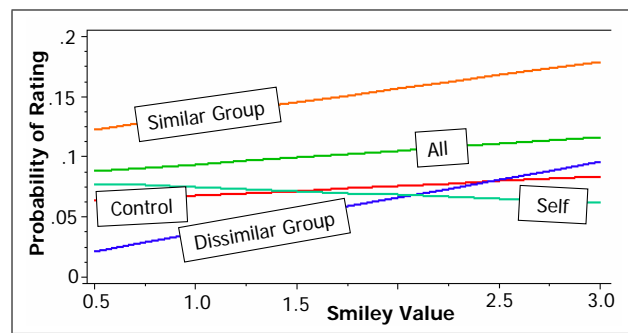

Figure 3. Interaction plots between the number of smilies on a movie and experimental conditions.

underlying mechanisms. In Ling et al., email messages explaining value of contribution caused members to contribute less compared to those whose messages did not mention value at all. Through presenting information about value in a more integrated and subtle manner, H1 confirmed that explaining value to the members increased contributions. As predicted by the CEM model, our interface motivated more contributions by highlighting how much the individual identifies with the group (H3) and how much they like the group that would receive the benefit (H4). These results illustrate ways that designers can use information about the beneficiaries of contributions to create subtle and integrated messages to increase motivation.

REFERENCES
1. E. Adar and B. A. Huberman. Free riding on gnutella. *First Monday*, July 2005.

2. B. Butler. The dynamics of cyberspace: Examining and modeling the dynamics of online social structure. *Ph.D. Thesis, Carnegie Mellon University, Pittsburgh*, 1999.

3. J J. Brehm. A theory of psychological reactance. *Academic Press*, 1966.

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

5. K. Ling, G. Beenen, P. Ludford, X. Wang, K. Chang, X. Li, D. Cosley, D. Frakowski, L. Terveen, A. M. Rashid, P. Resnick, and R. Kraut. Using social psychology to motivate contributions to online communities. *JCMC*, 10(4), 2005.

6. B. Nonnecke and J. Preece. Lurker demographics: Counting the silent. In *Proceedings of CHI*, 2000.

7. G. Salton and C. Buckley. Term weighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5):513–523, 1988.

8. B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the WWW10*.

9. D. A. Small and G. Loewenstein. Helping 'a' victim or helping 'the' victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, 26(1):5–16, 2003.