The Effect of Extrinsic Motivation on User Behavior in a Collaborative Information Finding System

Bracha Shapira
Department of Management Science and Information Systems (MSIS), Rutgers University, Levin Building, Room 245, 94 Rockafeller Rd., Piscataway, NJ 08854-8054. E-mail: shapira@business.rutgers.edu

Paul B. Kantor
School of Communication, Information and Library Science (SCILS), Rutgers University, 4 Huntington St., New Brunswick, NJ 08901-1071. E-mail: kantorp@cs.rutgers.edu

Benjamin Melamed
Department of Management Science and Information Systems (MSIS), Rutgers University, Levin Building, Room 256, 94 Rockafeller Rd., Piscataway, NJ 08854-8054. E-mail: melamed@rbs.rutgers.edu

In collaborative information finding systems, evaluations provided by users assist other users with similar needs. This article examines the problem of getting users to provide evaluations, thus overcoming the so-called "free-riding" behavior of users. Free riders are those who use the information provided by others without contributing evaluations of their own. This article reports on an experiment conducted using the "AntWorld," system, a collaborative information finding system for the Internet, to explore the effect of added motivation on users' behavior. The findings suggest that for the system to be effective, users must be motivated either by the environment, or by incentives within the system. The findings suggest that relatively inexpensive extrinsic motivators can produce modest but significant increases in cooperative behavior.

Introduction

Collaborative systems for information finding maintain knowledge bases that include users' evaluations of the relevance of data items (Konstan, Miller, Maltz, Herlocker, Gordon, & Riedel, 1997; Resnick & Varian, 1997). Collaborative systems recommend relevant information to users, based either on similarity between users (represented in profiles or demographic information) or on similarity between users' needs (called "quests" in the AntWorld system) (Kantor 1999a, 2000; Kantor, Melamed, Boros, & Menkov, 1999b). In operation, such systems use a similarity calculation between other quests and the current user's quest to estimate the relevance of data items for a current user or query. Some systems (Konstan et.al., 1997; Direct-Hit site) attempt to unobtrusively estimate the relevance of data items by observing and recording users' behavior (Morita & Shinoda, 1994), e.g., time spent, actions taken, or users' auxiliary activities and notations (such as their bookmarks). Automatic derivation of relevance has not yet been shown accurate enough (Oard, 1998); therefore, explicit contribution (such as evaluation) is needed to accurately acquire information on the relevance of the information. On the other hand, explicit evaluations (judgments) may have other problems, such as low quality and sparsity.

Evaluations, which are public goods and require user effort (even if minimal) to produce, are likely to be under-provided. Free-riding, in which users make use of others' evaluations, but do not contribute their own, in turn leads to sparse and unrepresentative evaluations. Unreliable system results (recommendations) may provoke further unreliable recommendations, in a destructive positive feedback loop. This makes it important to understand the determinants of valid judgment contribution. This article describes a study of users' collaborative behavior, using the AntWorld collaborative information finding system (Kantor et al., 1999a, 1999b). We find that to elicit evaluations and cooperative user behavior, there must exist some incentive (extrinsic motivation) to stimulate user behavior, above and beyond users' intrinsic (but second-order) motivation to improve the system's performance for their own benefit.

Received June 23, 2000; Revised January 31, 2001; accepted February 13, 2001

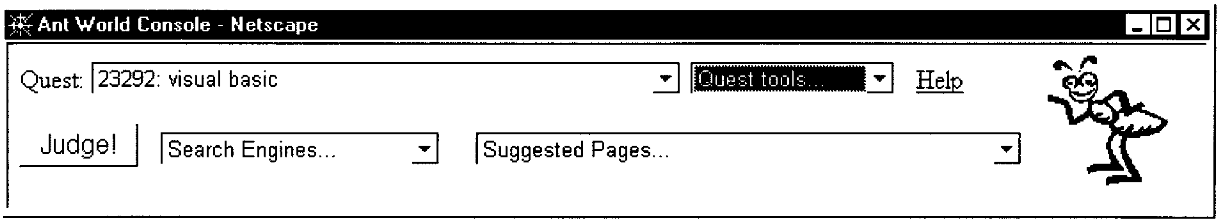

FIG. 1. Antworld console window.

There are several kinds of motivators. One option is "system internal incentive" that rewards the user for contributing evaluations with "scrip" that can be later used for services that the system offers (Avery, Resnick, & Zeckhauser, 1999). Conversely, users who do not provide evaluations might be charged for access to the database of evaluations. Another option is "external incentive" provided by the organization where the system is used. For intraorganization systems, users might be rewarded either as individuals or as a team when they contribute "enough" evaluations.

The present experiment simulates these alternatives. An experiment is described, which demonstrates the free-riding problem and supports the "required incentive for motivation" hypotheses. The experiment was conducted with students who used the AntWorld system to find specific information for their course assignments, in three trial sessions. The first session served as a pilot study to establish feasibility of the experiment, and for selection of certain experimental settings. In the second session the users were not extrinsically motivated, while the third session included an extrinsic incentive for providing evaluations. The results show that the extrinsically motivated group significantly exhibited a "contributing behavior" while the other group tended to be less active overall, and contributing an even lesser proportion.

The next section describes the AntWorld system, Section 3 discusses related work, Section 4 describes the experiment and its results, and Section 5 discusses the implications of the results for collaborative system design. Section 6 presents conclusions and suggests future research.

AntWorld

The AntWorld system developed at Rutgers University (AntWorld site; Kantor et al. 1999a, 1999b, 2000) implements collaborative information finding on the Internet. AntWorld works as an extension to a Web browser, which interfaces with the users through a small "console window," which may be positioned anywhere on the screen (Fig. 1). The system allows users to describe their information needs as "quests." A quest is initially represented by a "short description" similar to a search engine query, and an optional "long description," which is similar to a TREC query. The system asks users to evaluate the relevance of pages they reach during the search in the specific quest session(s). Users can evaluate a page on a five-point scale allowing evaluation of the page as an information source, or as a navigation pointer. The scale is presently converted to numbers as follows: *meets my need* (5), *adds information* (4), *helps navigation* (3), *not useful (2)*, and *no comment* (1). (Conversion of judgments to numbers is dependent on the specific configuration of the ''AntWorld'' client software.)

The evaluations provided by a user during a quest session serve as additional dynamic information about the quest. This information, together with the quest short and long descriptions, supports calculation of similarity between the current quest and other former quests stored in the database. The similarity calculation results in two pull-down lists on the "console window": (1) suggested pages: ranked URLs that were evaluated by other users as relevant to other similar quests. Higher rank for a URL on the suggested page list represents higher predicted relevance of that URL to the current quest "AntWorld" computed as a sum of prior judgments weighted by quest similarity. (2) Similar quests: the user can explore detailed activities (pages visited, and judgments) of quests that are selected by the system as most similar to the current quest.

Also, whenever the page currently viewed by the user contains a link to any of the documents on the suggestion list, a small colored "ant" appears on the user's screen next to this link to highlight the possibly relevant link. AntWorld can be used by private users with the public AntWorld database to locate relevant information as defined by Internet users. It can be also used to build secure organizational quest databases, to be used for intraorganization data and intelligence sharing.

Related Work

Most collaborative information finding systems use explicit rating acquisition to build their information space. The ratings are used for recommending potentially relevant information to users. However, requiring explicit evaluation introduces the problem of underprovision. Users tend not to contribute their evaluations, even if the effort is as low as a single keystroke. During trials with Grouplens, a collaborative Usenet News system (Konstan et al., 1997), users rated as few as 1 to 2% of articles. Consequently, systems that use explicit evaluations are vulnerable to the free-riding problem (Avery et al., 1997; Holt, 1997), that is, users take advantage of other users contributions without providing their own. The general problem is defined and studied in game-theory research (Fudenberg & Tirole, 1991), and has been empirically demonstrated in various studies of collaborative environments. For example, in collaborative learning, students tried to achieve grades assigned based on their teamwork without contributing any effort (Brownell & McArthur, 1996; Wlison & Whitelock, 1997).

For collaborative information-finding systems there are two approaches to the established problem of underprovision and the suspected problem of free riding. One approach is to covertly derive evaluations (Konstan et al., 1997; Oard, 1998) from aspects of the users' environment or behavior, such as time spent on a page or bookmarks, thus avoiding the need for user cooperation. This approach has advocates, but has not yet reported success for any collaborative information-finding systems. If successful, it could reduce the current reliance on explicit ratings. However, extraneous factors can affect the user's time on page, including interruptions. An enhanced browser, which could track the active motion of the mouse pointer, might reduce this problem, because the distracted user does not move the mouse. But then, neither does the engrossed reader.

In contrast to the covert approach, the motivational approach seeks incentives that motivate users to provide evaluations (Konstan et al., 1999). Avery et al. (1997) have suggested a pricing and subsidy mechanism that in effect creates a computerized market for evaluations, to induce the efficient provision of evaluations. Basically, in their model users "buy" and are "paid" for their evaluations, and the "money" serves as the incentive. Although the model has been formally defined, its feasibility for implementation has not yet been tested.

The need for incentives was also explored at the Collaborative Filtering Workshop in Berkeley 1996 (Arnhein, 1996). In this analysis, underprovision is related to the users' decision about whether to pay the cost of reading a document and evaluating it, or to pay the cost of waiting and acquiring the information later. The conclusion on the subject at that workshop was that, in the future, market pricing might eventually subsidize and coordinate evaluations. Such pricing schemes are likely to be complex, but could be adapted over time as systems learn more about users.

Because the current study examines the effect of the "motivational" approach on user behavior, we present here a brief background on motivation from the psychological point of view. In general, motivation is defined (Gange & Medsker, 1996, p. 168) as a "cognitive persistence, the drive, tendency, or desire to undertake or complete a task, expend effort and do a quality job." Intrinsic motivation is understood in contrast to extrinsic motivation. If intrinsically motivated, a person will perform the task even if there is no apparent reward except the activity itself (Deci & Ryan, 1985). Extrinsic motivation is present when an activity is rewarded by incentives not inherent in the task. Psychologists do not agree on the relationship between the two types of motivation. One opinion (Cameron & Pierce, 1997) is that intrinsic and extrinsic motivation combine in an additive fashion to produce overall motivation. For example, in work settings, some organizational psychologists have argued that optimal performance would occur when jobs were interesting and challenging and employees were externally rewarded (e.g., with money) for their work (Cameron & Pierce 1994, 1997; Porter & Lawler, 1968). This approach was widely accepted by industry (75 to 94% of U.S. companies are using some form of incentive) (Kohn, 1993, p. 13). Recently, other theorists challenged the additive assumption, suggesting instead that extrinsic rewards might interfere with intrinsic motivation (Deci & Ryan, 1985; Kohn, 1993; Spitzer, 1995). This approach contends that providing incentives ignores the benefits of the nonmotivational situation, instead of creating intrinsic motivation by, for example, making the task more interesting, or by making people understand its benefits. Incentives cause people to perform a task even if they do not like doing it, and thus is likened by Kohn (1993) to punishment.

However, it is overall agreed that extrinsic motivation is successful in getting people to perform tasks, and that it is easy to apply. The disagreement is about the extent to which incentives decrease creativity and productivity, or other concomitants of intrinsic motivation. Incentives are suspected to have negative effect on the quality of performance when sophisticated learning is involved, and to undermine interest in the task (Kohn, 1993). Yet, negative effects were not observed when incentives were used with low interest activities, or when given only for task completion (and not for task performance).

When the task considered is the provision of evaluations for a collaborative system, creativity and productivity do not really apply, and no "sophisticated" learning is involved. In addition, contributing evaluations cannot be considered an interesting task, at least in our experience, so that the afore-mentioned negative aspects of incentives are not relevant. Moreover, the only possible intrinsic motivation to provide evaluations could stem from understanding the benefits of the contribution. We found that this is not a sufficiently strong motivation, especially in the initial phases of the database, when evaluations contribute only to future performance of the system, and to the benefit of other users.

Intrinsic motivation (or nontangible incentive) might exist in collaborative systems only in environments where people could gain immediate benefits from the system. For example: Ackerman (1998) developed "Answer Garden," a collaborative system for "growing organizational memory." This system keeps records of questions that were addressed to experts by users, and builds a database of questions for later uses. Ackerman (1998) reported on high usage of the system during the experiments conducted. This is not surprising, considering that "Answer Garden" provides an immediate benefit to its users by saving them work. When using the system, the user might save the need to address a question to an expert, if the same question had been already addressed. For the expert, using the system to store the answers provided to users might save him future work, if other users might have the same question. Moreover, to use the system, users do not have to do any extra unneeded work (like providing evaluations), but they simply need to prefer the system over a telephone help desk.

This article demonstrates the free riding problem in collaborative information finding systems, and empirically shows that incentives may improve the provision of evaluations, at statistically significant levels.

Experiments

Conceptual Framework

In this study, we tentatively divide user behavior into "contributing behavior" and "benefiting behavior." "Contributing behavior" includes actions that contribute to the performance of, and to our evaluation of the system. These actions include providing evaluations, and filling on-line surveys (if applicable). "Benefiting behavior" includes actions that add to user knowledge, such as visiting pages and following suggestions (if applicable). Our experiments explore the effect of motivation on these two types of behavior.

Experimental Setting

The participants for the experiments were MBA, and undergraduate junior and senior students at the Rutgers University Business School, Department of Management Science and Information Systems (MSIS). The students were given course assignments requiring them to search for specific, nontrivial, information relevant to their course subjects (see Appendix for the detailed assignments). They were required to use the AntWorld system as their information-finding tool. The students were briefly introduced to the AntWorld idea, and to the shared benefits of providing judgments to improve system performance. During their searches with the AntWorld, the students were presented with occasional on-line surveys (see Appendix) providing feedback on user satisfaction with the system. To capture student feedback, the assignment included two additional questions regarding the AntWorld system: (1) Did you see ants during your searches with AntWorld? (2) If so, were the ants helpful? If so, in what way were they helpful?

The students' activities using the AntWorld did not affect their course grades, as to do so would violate university policy. Accordingly, one common user name was assigned to each group of students in each trial, so that we might track students' activities as members of a group, and not as individuals.

The experiment consisted of three trials. The first trial was conducted with a group of executive level, highly focused older MBA students. This first trial served as a pilot for the system and as a feasibility test for the evaluating and tracking software. Its results shaped the experimental design, regarding the following issues:

(1) *Suggestion recording*—on the pilot (first) trial, the system recorded the students' quests, judgments, and pages (URLs) visited. The results revealed that the students used many of the same pages. A reasonable assumption is that some of these common pages were reached by following links from the suggestion list that the system provided. To test this hypothesis, the system was modified so that it recorded specifically (starting on the second trial), all URLs reached by following the suggestion list. This recording supported the comparison of users' benefiting behavior with their contributing behavior.

(2) *On-line survey frequency*—in the pilot trial, the students were prompted with a first on-line survey after five pages had been visited and then, after each 10 additional pages visited. The purpose was to track whether satisfaction changes as users learn the system with increasing use. Not surprisingly, students complained that the on-line surveys were too frequent and annoying. Of the 10 surveys completed, only two were filled during the same quest (that is, by the same user). The highest

TABLE 1. Experiments performed.

Trial	Students type	N 1	N 2	Motivation	Activity record	On-line survey freq.
1	MBA Executive Level	7*	11	No	No suggestion list	** $x = 5$ pages $y = 10$ pages
2	Seniors	21 (early)	29 (late)	No	Suggestion list included	$x = 2$ judgments $y = 5$ judgments
3	Juniors	36 (early)	58 (late)	Yes (Pizza)	Suggestion list included	$x = 2$ judgments $y = 5$ judgments

* The class was divided into two groups (based on the last digit of the social security number [1–5, 6–0]). Each group had two questions to answer. Group 1 answered question 1 on week 1, and question 2 on week 2, and group 2 answered question 2 on week 1, and question 1 on week 2. This creates the "late" situation in which students can use data provided by other "early" students. This crossover design made each student "early" on one problem and "late" on another.

** The on-line surveys are shown after x pages visited, or after x judgments were provided. After the first appearance, the survey is displayed after a fixed interval, y.

TABLE 2. Experimental results.

Trials	Active quests	Pages visited	Judgments provided	Suggestion followed	On-line survey submitted	Assignment question: Saw ants?	Assignment question: Do ants help?
1	36	1163	120	No data	10 (8 users)	10	6
2	108	2205	89	390	1	17	9
3	195	5749	447	1073	34 (31 users)	45	27

number of judgments provided in a quest was 15, and the average number of judgments provided per quest was 3.3. For the subsequent trials, the on-line survey was triggered for the first time after two *judgments* were provided by the user, and then, after an interval of five *more judgments*, so as to reduce user annoyance. Users were thus prompted with one or two surveys at most.

(3) *Feasibility of system and experiment*—the pilot established that the system is usable, and understood by the users, as evidenced by the on-line surveys submitted. Seventy percent of the users who submitted surveys answered "How useful did you find AntWorld?" with "moderate" or better. However, many users characterized the user interface as being "not friendly." The interface was improved and adjusted in response to specific users' complaints and expert suggestions.

Trials 2 and 3 each included students from two different sections of the same course. The students were all from the same department (Management Science and Information Systems), and the assignments given had the same weight on their final grades. Trial 2 included senior students in their first semester of the senior year, and Trial 3 included junior students on their second semester of the junior year. We believe that the one-semester difference of experience in the university is not likely to have a significant effect on students' behavior. One section of each course formed an "early" group with the assignment due on an earlier date, and the other formed the "late" group, with a later due date. We hoped that the "late" group might find the system more useful, after the "early" group had provided relevant judged pages to the database. However, we cannot report on findings on this issue because we did not control the time of completing assignments, but only the time of submitting them. Members of the "late" group may have worked on the assignments before members of the "early" group. We, therefore, report only on the usage of visited and judged pages, without including the time variable.

The third trial also included an "extrinsic motivation" for providing judgments. It was announced that the group providing the highest number of judgments (without regard to quality or usefulness) would receive a reward (Pizza party). (We could not reward the students with grades, which would have been the best motivation in a student environment, due to university policies regarding research on human subjects).

Tables 1–3 summarize the experiments. Table 1 summarizes the experimental setting. Table 2 displays the overall data on the three trials.

More detailed information about the two later trials is given in Table 3, and in the figures.

Discussion

We are concerned primarily with differences in the behavior of groups 2 and 3. Table 3 shows that the mean number of judgments provided by users in the extrinsically motivated group is 2.3 per quest (versus 0.8 per quest in the unmotivated group). Using the two-sided t-test this difference is significant at better than 99% confidence (p = 0.00048) (Table 5). Motivated students provided a significantly higher number of judgments per quest than unmotivated ones. We also see that the motivated students visited more pages overall.

Table 4 shows that the combined contributing activity is significantly greater for the motivated group (Trial 3) at

TABLE 3. Descriptive statistics.

Trial	Contributing behavior: Judgments	Contributing behavior: Surveys	Benefiting behavior: Suggested pages	Benefiting behavior: Visited pages
1—pilot				
Mean	3.33	0.28	N/A	32.30
Max	15.00	1.00		115.00
Min	0.00	0.00		2.00
Std. dev	4.34	—		25.04
Median	1.00	—		23.50
Standard error	0.72	—		4.17
2—no motivation				
Mean	0.82	0.00	3.61	20.90
Max	12.00	1.00	35	178.00
Min	0.00	0.00	0	**1.00
Std. dev	2.09	—	5.36	27.20
Median	0.00	—	2	11.00
Standard error	0.20		0.52	2.62
Percentage %*	3.23	0.00	14.25	82.52
3—motivation				
Mean	2.29	0.17	5.50	29.50
Max	21.00	2.00	49.00	323.00
Min	0.00	0.00	0.00	1.00
Std. dev	3.82	—	7.36	38.10
Median	1.00		3.00	17.00
Standard error	0.27		0.52	2.72
Percentage %*	6.11	0.44	14.70	78.75

* Percentage is the specific activity out of all the activities.
** We considered only quests that have at least one visited page.

TABLE 4. Difference in the levels of contributing activity, trials 2 and 3.

	Contributing activity percentage	No. of quests
Trial 3	6.55%	195
Trial 2	3.23%	108

Difference significant $p = 0.047$.

TABLE 6. Significance of differences between trials 2 and 3.

	t-value	Df	p	F-ratio variance
Visited pages	2.18*	301*	0.0299*	1.92*
Judgments	3.69*	301*	0.0027*	3.32*
Suggested pages	2.37*	301*	0.0186*	1.88*

* Significant at the 95% confidence level.

over 6.5% of all activity recorded. Using a one-sided t-test, we find this to be significantly different from the value for Trial 2 (3.23%), at better than 95% confidence. This shift in the proportion of user activities is the most significant finding of our pilot study. That is, not only does extrinsic motivation increase the magnitude of "visiting activity" by almost 40% (29 per quest, versus 21 per quest, see Table 3), but it also differentially increases the portion of behavior that enhances the value-added of the collaborative system (i.e., contributing behavior), from 3.23% without extrinsic motivation to 6.55% with extrinsic motivation.

Another way of appraising the significance of these findings is to note that the t-value for the difference in judgments provided is almost double the t-values for the other activities (~4 compared to ~2), and the p-value is correspondingly more significant (Table 6). This supports the inference that the effect of extrinsic motivation on judgments is significant even when the incentive is as trivial as a pizza party (for well-fed students). The rate of providing judgments in the nonextrinsically motivated group did not even reach an average of one judgment per quest, and the median was less than 1, as seen in the histogram of judgments (Fig. 2a). That is, more than half of the students did not provide any judgments at all. This result clearly demonstrates the free-riding problem, as the other activities (visiting pages, and following suggestions) have nonzero values.

One way to visualize the effect of extrinsic motivation is to compare the ternary charts depicted in Figure 3. These charts show the allocation of activity in three distinct categories with each vertex representing one of the three categories. We have oriented the graph so that the desired behavior, "judging", appears at the top. In both the extrinsically motivated and the unmotivated groups, activity is concentrated at the lower left corner—visiting pages. However, the eye detects what the statistics have confirmed: in the extrinsically motivated group, the points scatter more towards the top—that is, towards providing judgments. In this plot each symbol represents a complete quest. Note also that the number of judgments should not exceed the number of pages visited, so that the median, drawn from the lower right vertex, represents the best possible outcome to be expected.

TABLE 5. Mean number of judgments trials 2 and 3.

	Mean	Std. dv	Judgments	p-value two-sided test
Trial 3	2.292	3.826	447	.000479*
Trial 2	.824	2.099	89	

* Significant at the 95% confidence level.

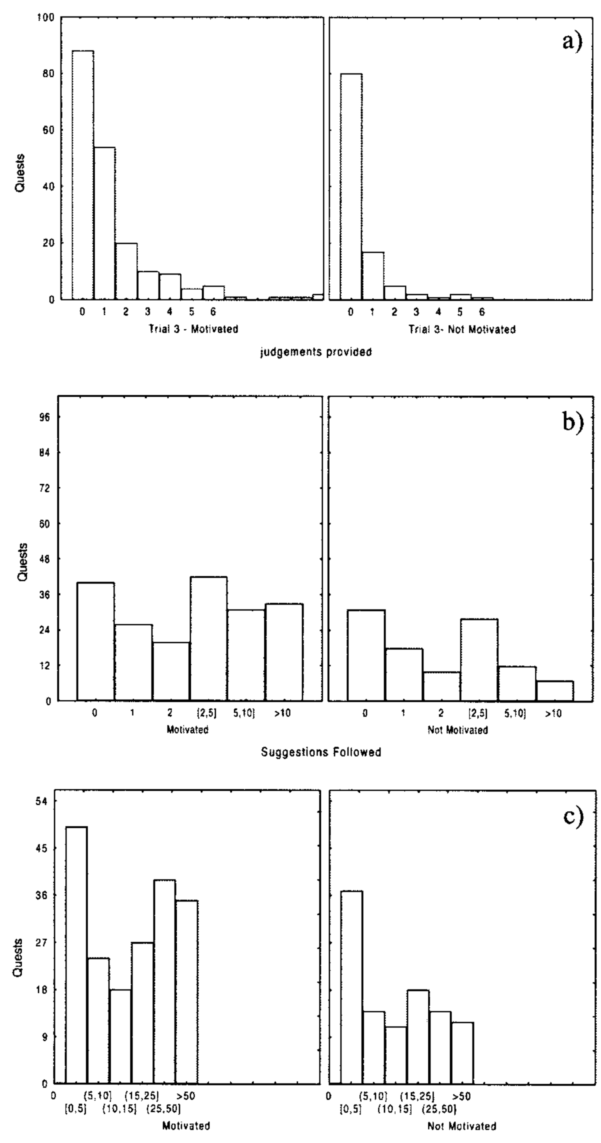

FIG. 2. (a) Judgments histograms. (b) Suggestions-Followed histograms. (c) Visited-Pages histograms.

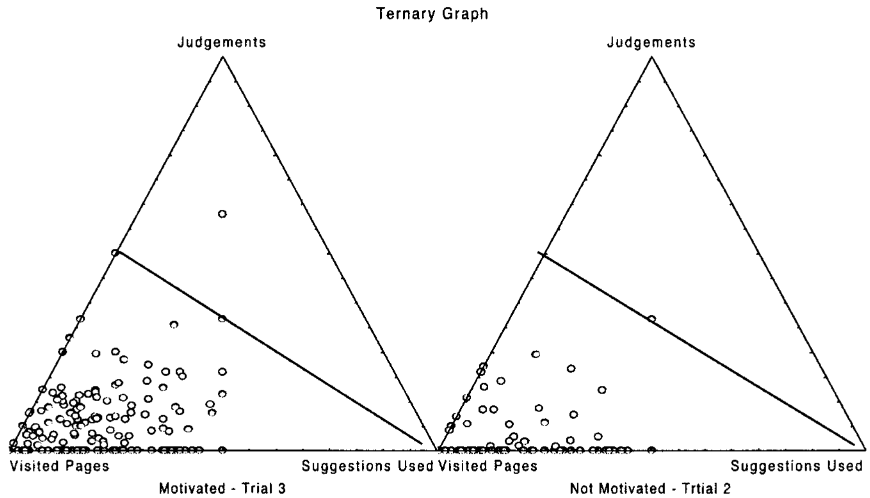

FIG. 3. Ternary graphs for activity of both trials. (One of the users from the motivated group provided three judgments to the only page he visited. This explains the outlier point above the median on the left-hand graph.)

As noted, for the extrinsically motivated group, judging activity exceeds two judgments per quest, and most of the users provide at least one judgment, as seen in the histogram (Fig. 2a). In addition, the benefiting activity, when not extrinsically motivated is also significantly lower, but the significance of that finding is not as great. That a user follows on average five suggested pages when not extrinsically motivated, compared to seven pages when extrinsically motivated, does not seem enough of a difference to justify taking (expensive) steps to reward users. But if most users provide absolutely no judgments when not extrinsically motivated, and provide even one judgment when extrinsically motivated, then some incentive is worth considering. Otherwise, the system will not accumulate value. In other words, collaborative systems are not harmed as much by lower levels of benefiting activities, as they are by lower levels of "contributing activities."

Regression analysis yields some relationships among pages visited, suggestions used, and judgments. The number of judgments in the motivated situation is estimated as a constant (0.58) plus approximately one judgment per 20 pages visited (0.047 per page) and approximately one judgment per 17 suggestions followed. For the nonmotivated group all three terms in this equation are lower: constant (0.17) judgments, plus one per 40 pages visited, plus one per 20 suggestions followed. Ideally, the system would accumulate one judgment for every suggestion followed, as this would ensure the currency and provide revalidation of those suggestions. Thus, a prime motivational problem is to increase the likelihood that users will judge those pages reached by following suggestions.

Conclusions and Future Research

This article reports on an experiment that tested users' behavior in a collaborative environment, and explored the effect of extrinsic motivation (incentive) on user behavior. The results demonstrate the *"free-riding" problem*, as users who were not given incentives contributed scarcely any evaluations, but tended to benefit from the system. We observed more activity when users were extrinsically motivated, and most significantly, more contributing activity, which is the crucial behavior that must be elicited in collaborative environments. This result suggests that a suitable motivation may improve users' contributing activities.

In considering the extrinsic motivation applied as the "treatments" in this experiment, we must recognize that there were in fact two extrinsic motivational factors. The stated incentive was the "pizza party," which is a pseudo-monetary extrinsic incentive, but the terms of the experiment made it appear that only one group would receive the party. Thus, two motivational factors are confounded in our experiment. One is the pizza itself, which we do not think has particularly high value for this group of students. The other is the intangible intragroup competitive motivation, because only one group was expected to "win." Thus, competition, which is known to have a positive effect on achievements (Hammond & Goldman 1961), is almost certainly part of the motivation. However, it is not always feasible to implement competitive motivation, especially in public domain systems. Without substantial further study, we cannot know whether the key factor here is an innate "competitiveness" among the students, or a belief that, despite assurances to the contrary, winning this competition would influence the teachers in assigning grades to the assignment or the course. To be more generally useful, the system should be exercised with noncompetitive extrinsic motivations. Further experiments on this issue are planned.

Another issue that must be further explored is whether two domains of implementation of the system, namely, local and homogenous as opposed to global and presumably heterogeneous, require different types of incentive. We believe that the type of incentive must be adapted to the particular group of users.

If the system is organization based, as is one version of AntWorld, and the organization is interested in the development of a shared knowledge base, the organization can determine the kind of incentive that would lead its members to provide evaluations. This is especially important in the initial phase of the system, until it collects enough data to become clearly beneficial. When the system reaches a beneficial state users may become intrinsically motivated and provide the evaluation without extrinsic motivation. In our experiment, the first trial session (pilot) with the "nonextrinsically motivated" MBA students yielded higher activity levels than were achieved with even the "extrinsically motivated" undergraduate students. We conjecture that this relates to the intrinsic motivation of more adult and responsible users. In an organization, this more mature behavior might well be the norm, as users might find it "natural" to contribute to their organization. We plan to test this issue by installing AntWorld in an organizational setting, to explore the effect of further "overmotivation" in such a setting. Within an organization the extrinsic motivation might be an "external incentive" that is not handled by the system. The role of the system in that case, as in our experiments, is simply to collect the data for allocating the extrinsic motivational rewards.

If the system is public oriented (e.g., on the Internet), an "internal incentive" that is managed entirely by the system is needed. One possibility is to integrate a "pricing mechanism" into the system; for example, the system might provide "points" for each evaluation. The points would enable the users to "purchase" other services, such as access to paid knowledge bases, or to digital libraries maintained by the system provider. Any such incentives are threatened by the possibility of pointless or spoofing behavior, which could still accumulate user credits. However, weighting credits by the apparent usefulness of the suggestions contributed (which would require changes to the AntWorld system) can counter this problem.

Overall, we have found that, at least in one particular example, a cooperative system for finding information in a networked environment benefits from the introduction of relatively inexpensive extrinsic motivational factors. Incentive seems to be necessary to achieve even a relatively low number of judgments (2.3 judgments per quest), and should be considered when designing such a system. As the Internet grows explosively, it is to be hoped that other such extrinsic motivators will be found and employed, to improve universal access to information.

Acknowledgments

This work is supported by the Defense Advanced Research Projects Agency (DARPA), contract number N66001-97-C-8537. We acknowledge the support of the AntWorld programming team, particularly Dr. Vladimir Menkov, Qin Shi, Insuk Oh, and David Neu, who provided specific tools to conduct our interviews and data analysis.

References

AntWorld site: http://aplab.rutgers.edu/ant.

Arnheim, L. (1996). Summary of proceedings of the collaborative filtering workshop, 1996. Available online: http://www.sims.berkeley.edu/resources/collab/collab-report.html.

Avery, C., & Zeckhauser, R. (1997). Recommender systems for evaluating computer messages. Communications of the ACM, 40(3), 88–89.

Avery, C., Resnick, P., & Zeckhauser, R. (1999). The market for evaluations. American Economic Review, 89(3), 564–584.

Ackerman, M.S. (1998). Augmenting organizational memory: A field study of answer garden. ACM Transactions on Information Systems, 16(3), 203–224.

Brownell, G., & McArthur, J. (1996). A preliminary report: Robotics and collaborative learning in sixth grade classroom. Proceedings of Society of Information Technology and Teacher Education, SITE 96, 270–300.

Cameron, J., & Pierce, W.D. (1994). Reinforcement, reward and intrinsic motivation: A meta-analysis. Review of Educational Research, 64(3), 133–151.

Cameron, J., & Pierce, W.D. (1997). Rewards, interest and performance: An evaluation of experimental findings. American Compensation Association (ACA) Journal, 6(4), 6–15.

Deci, E.L., & Ryan, M.R. (1985). Intrinsic motivation and self-determination in human-behavior. New York. Plenum Press.

Direct Hit site: http://directhit.com.

Fudenberg, D., & Tirole, J. (1991). Game theory. Cambridge, MA: MIT Press.

Gagne, R.M., & Medsker, K.L. (1996). The conditions of learning: Training applications. New York: Harcourct Brace College Publishers.

Holt, A. (1997). Prisoners, chickens, volunteers, free riders and suckers: Dilemmas on the internet. Engineering Science and Educational Journal, 6(2), 73–77.

Hammond, L.K., & Goldman M. (1961) Competition and non-competition and its relationship to individual and group productivity. Sociometry, 24(1), 46–60.

Kantor, P., Boros, E., Melamed, B., Menkov, V., Shapira, B., & Neu, D. (2000). Capturing human intelligence in the net. Communications of the ACM, 43(8), 112–115.

Kantor, P., Boros, E., Melamed, B., Neu, D., Menkov, V., Shi, Q.,& Kim, M. (1999a). AntWorld. In Proceedings of the the international conference on research and development in information retrieval (ACM SIGIR '99) (p. 323).

Kantor, P., Melamed, B., Boros, E., & Menkov, V. (1999b). A dynamic model of user's information needs. In Proceedings of the American Society for Information Science annual meeting (ASIS 1999) (pp. 536–545).

Kohn, A. (1993). Punished by rewards. Boston. Houghton Mifflin Co.

Konstan, J.A, .Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., & Riedl, J. (1997). Grouplens: Applying collaborative filtering to usenet news. Communications of the ACM, 40(3), 77–87.

Morita, M., & Shinoda, Y. (1994). Information filtering based on user behavior analysis and best match text retrieval. In Proceedings of the the international conference on research and development in information retrieval (ACM SIGIR '94) (SIGIR 94) (pp. 272–281).

Oard, D.W. (1998). Implicit feedback for recommender system. In Proceedings of the American Association for artificial intelligence workshop of collaborative systems (AAAI98) (pp. 80–83).

Porter, L.W., & Lawler, E.E. (1968). Managerial attitudes and performance. Homewood: Irwin-Dorsev.

Resnick, P., & Varian, H. (1997). Recommender systems. Communications of the ACM, 40(3), pp. 56–58.

Spitzer, D. (1995). Supermotivation. New York: Amacom.

Wilson, T., & Whitelock, D. (1997). Monitoring the on-line behavior of distance learning students. Journal of Computer Assisted Learning, 14(2), pp. 91–96.

Appendix

1. On-line survey displayed to users.
2. Students' Assignments

A. Pilot Trial

Use the AntWorld system to find relevant information for the following two questions:

1. Explain the following terms: Rapid Application Development (RAD), and Object Oriented Programming. Explain why RAD extensively uses the Object-Oriented approach, and why Visual Basic is one of the world's premiere RAD languages.
   Submit a list of the www main resources that your answer is based upon.
2. Explain the ActiveX technology. Find three examples of widely used ActiveX controls. Give an example of a new control that could have been developed.
   Submit the main WWW resources that you base your answer on.

B. Trial 2—Nonmotivated Group

Use the AntWorld system to answer the 2 following questions:

1. Examine any two popular case software that include upper and lower case tools. Compare them according to the following parameters:
   Supported methodologies,
   Supported Operating Systems
   Functionality (what they can "do")
   User friendliness
   Reliability
2. Give an exact definition for RAD. In what situation would you suggest RAD as the development methodology? Why is Visual Basic a popular RAD tool?
   Submit a list of the WWW pages that your answers (1,2) are based upon.

C. Trial 3—Motivated Group

Use the AntWorld system to answer the two following questions:

1. List the major players in the telecommunications industry in the USA. List the main URLs that you based your information upon.
2. What kinds of telecommunications capabilities would be required for a large multinational clothing company to provide any employee with instant access to any data in the entire company? Aside from technical capabilities explain why do you or do not believe this is practical. Use the AntWorld system to find an example of a multinational company that uses telecommunication; describe the kinds of telecommunications that it uses. List the WWW pages that you used for your answer.