Observed Behavior and Perceived Value of Authors in Usenet Newsgroups: Bridging the Gap

Andrew T. Fiore
Human-Computer Interaction Lab
Cornell University
209 Kennedy Hall
Ithaca, NY 14850 USA
+1 607 255 5530
atf2@cornell.edu

Scott LeeTiernan
Department of Psychology
University of Washington
Box 53525
Seattle, WA 98195 USA
+1 206 616 6086
slt@u.washington.edu

Marc A. Smith
Microsoft Research
One Microsoft Way
Redmond, WA 98052 USA
+1 425 936 6896
masmith@microsoft.com

ABSTRACT
In this paper we describe an evaluation of behavioral descriptors generated from an analysis of a large collection of Usenet newsgroup messages. The metrics describe aspects of newsgroup authors' behavior over time; such information can aid in filtering, sorting, and recommending content from public discussion spaces like newsgroups. To assess the value of a variety of these behavioral descriptors, we compared 22 participants' subjective evaluations of authors whose messages they read to behavioral metrics describing the same authors. We found that many metrics, particularly the longevity and frequency of participation, the number of newsgroups to which authors contribute messages, and the amount they contribute to each thread, correlate highly with readers' subjective evaluations of the authors.

Keywords
Social cyberspaces, social accounting, persistent conversations, discussions, behavioral indicators

INTRODUCTION
Public, online social spaces like Usenet newsgroups are frequently noisy and voluminous places. Larger groups regularly receive more than ten thousand messages a month; many of these messages are of limited value. Despite the many advantages groups of people gain when they interact through computer networks, enough poor-quality messages flood newsgroups and similar public discussions that finding the valuable content typically proves too difficult to be worthwhile.

Efforts to address this issue have primarily focused on methods for groups of people to share their opinions of the messages and authors contributing content to a social cyberspace [3]. These techniques, used by Web sites like online auctioneer eBay.com and Linux news provider Slashdot.org, face the dual challenges of attracting a critical mass of people willing to contribute evaluations and ensuring the quality of those evaluations – and even when they have managed to achieve this, they have not succeeded in using the information to reduce the noise and prevent the abuse that often prevails in online public spaces. The rampant misuse of these common spaces is one of the reasons many people who used to use Usenet no longer do [2].

Systems that require the active participation of their users to assess other users may not be necessary to find out which authors and messages are valuable, however. Here we present an alternative approach that brings to online spaces the social context and interactional history that help people navigate physical spaces but that are not generally available in online environments. Most Usenet news browsers provide limited, if any, information about the histories of the authors in a newsgroup, such as how long they have been active in the group, in what other newsgroups they participate, to what other threads of conversation they contribute, and which other participants they most often engage in discussion. Instead, the browsers present information about the messages themselves – their posting date, their subject, how many lines of text they contain – which forces the user to pay more attention to the structure of the medium than to the qualities of the participants, who would more naturally draw the user's focus.

We contend that a news browser would serve users far better if it allowed them to sort and search by such salient behavioral attributes as the number of days on which each author contributed a message to the newsgroup, or the fraction of each author's messages that were replies versus initiating turns [1, 4, 5]. Combined with information about the structure and development of threads, these metrics could be used to extract valuable content out of large, active social cyberspaces.

Such an interface, though, would be useful only if it presented behavioral descriptors of authors that correlated with the value of those authors as perceived by readers and other authors. In the following, we report our effort to identify those behavioral metrics that correspond to perceived value of authors in Usenet newsgroups. We

CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA.

find that a variety of metrics, particularly the longevity and frequency of participation, the number of newsgroups in which authors contribute messages, and the average size of the threads to which authors tend to contribute, correlate strongly with subjective assessments of the authors' qualities by our study participants.

Some previous studies have explored relationships among the quantitative characteristics of online social spaces [7], but to our knowledge none have attempted to close the gap between objective behavior and subjective social and informational perception.

Behavioral Metrics

We needed each author's activity in Usenet newsgroups to be quantified in a variety of ways, expressing various dimensions of posters' behavior and the reactions of others to this behavior.

The Netscan project, from which we chose to draw our data, has collected Usenet message headers since 1996 [6] and, by data mining a large collection of messages, generated extensive metrics about messages, threads, newsgroups, and authors from January 2000 through July 31, 2001. For each author, the system tracks such information as the date the author was first and most recently seen in specific newsgroups and the Usenet as a whole, the newsgroups to which they contribute, the total number of days on which the author has been active, the total number of messages the author has contributed, the number of those messages that were replies, and the number of replies from other people that the author's messages garnered. Netscan aggregates these measures at different levels, so researchers may examine them for a particular newsgroup or hierarchy of newsgroups at various temporal resolutions (e.g., day, week, month, year).

METHOD

To create a framework against which to evaluate the usefulness of these behavioral metrics, we gathered 22 people's evaluations of the authors of newsgroup messages that they chose to read during an observed reading session in the lab. The participants, 20 males and two females, identified themselves as frequent or expert users of Usenet newsgroups.

Table 1. Subject-selected newsgroups by number of unique authors for the period January 1, 2001 through July 31, 2001

Newsgroup	Authors	Repliers	Initiators	Returning authors	Posts	Replies	Thread starts*	Barren posts**	Cross-posts	Crosspost Targets***
alt.comp.periphs.mainboard.asus	14856	10211	7654	2634	74290	57236	13545	3509	5473	304
comp.sys.palmtops.pilot	8093	5391	3655	1788	44965	34991	6995	2979	11042	308
microsoft.public.windowsme.general	8086	4916	5401	881	56895	45649	9610	1636	4963	275
rec.woodworking	6347	4492	3324	1863	64544	54066	8330	2148	1467	131
comp.dcom.sys.cisco	5282	2912	2867	998	23114	15810	5034	2270	1323	216
rec.food.cooking	4905	3755	1969	1147	72710	65759	5203	1748	9241	471
microsoft.public.xml	3917	1939	1993	402	11384	6595	2887	1902	1532	308
rec.models.rockets	2083	1665	1112	758	45055	39487	4413	1155	367	46
rec.arts.sf.fandom	1602	1138	332	521	83501	79846	1647	2008	5333	190
alt.binaries.sounds.mp3.d	1596	1075	468	238	11313	8611	1359	1343	2152	231
rec.windsurfing	1497	1015	768	432	13013	10823	1631	559	76	50
comp.software.testing	1491	334	360	202	4082	1028	506	2548	162	108
microsoft.public.outlook.mac	1404	593	610	9	2892	1506	697	689	95	71
rec.equestrian	1327	959	562	437	24087	21373	2061	653	708	109
comp.sys.mac.programmer.help	1178	719	565	319	6382	4705	1170	507	1606	65
alt.sports.baseball.sea-mariners	1029	780	414	233	17341	14868	1806	667	1194	86
alt.aldus.pagemaker	483	232	165	67	1183	753	213	217	311	67
seattle.eats	481	375	172	146	2050	1713	252	85	99	36
sci.agriculture.beekeeping	458	293	230	146	2543	1902	440	201	101	28
microsoft.public.vb.vbce	436	228	193	55	1287	665	323	299	275	96
microsoft.public.certification.networking	370	184	120	28	688	379	136	173	200	150
alt.support.ex-cult	315	221	53	62	1307	1000	126	181	997	114
microsoft.public.platformsdk. mslayerforunicode	54	36	25	0	336	271	41	24	31	27

* Initial turns which received replies ** Initial turns which received no replies *** Distinct groups with which this one shared messages

Participants first identified their favorite newsgroups (listed in Table 1) and then answered a series of questions about their newsgroup browsing habits, including frequency of use, extent of participation in newsgroups as opposed to lurking, and types of use (e.g., socialization, technical support, commerce).

Next, we counter-balanced two newsgroup-reading scenarios, so that half of the participants did one first and half did the other. In the one, participants read messages from their self-identified favorite newsgroup; in the other, they read messages from a common group we selected for all of them to read, *microsoft.public.windowsme.general.* This newsgroup, which hosts wide-ranging discussions on issues related to Microsoft Windows Millennium Edition, received more than 8,000 messages contributed by more than 4,000 participants in the period between January 1, 2000 and July 31, 2001. Participants browsed both newsgroups with Microsoft Outlook Express 5.5, a news reader whose interface serves as a reasonable standard for modern, threaded news-browsing software. For the favorite newsgroup session, we set Outlook Express to display the 300 most recent messages in the group. For the *windowsme* session, we captured the group's 50 most recent messages on the morning of the first day of the study and presented every subject with exactly that set.

The 22 subjects' favorite newsgroups covered a variety of topics related to technology, sports, religion, and hobbies. The groups ranged from the massive to the minuscule in terms of population and message volume, but they shared some features. All exhibited a relatively high rate of reply (proportion of messages which garner replies), averaging 62 percent, far above the average of 13 percent over all newsgroups, indicating that they are all discussion-oriented groups, hosting predominantly conversational exchanges and not spam, automated postings, or binary files. The selected newsgroups also showed a higher than average rate of return for their participants – on average, 20 percent of the authors who posted in these newsgroups in 2001 also posted in 2000, as compared to 13 percent across all newsgroups.

As subjects browsed messages in their favorite groups, we made note of the author of each message they selected. When they had completed about 20 minutes of reading, we presented them with a survey concerning their opinions of each author whose messages they had selected.

We asked our subjects to respond to statements about each author on a seven-point scale where 1 equaled "strongly disagree" and 7 equaled "strongly agree." The statements (see Table 2) queried their satisfaction with the messages, their sense of each author's accuracy and value, and their likelihood to read other messages from the same author or to recommend them to others.

The procedure was the same for the *windowsme* scenario except that we gave subjects a shorter time (10 minutes) to read messages from the pre-selected set and then a longer time (45 minutes) to rate as many of the 31 authors who posted the 50 messages as possible.

Table 2. Subjective evaluation statements.

"I would read a message by this person again in the future."
"I would make a special effort to find this person's messages in the future."
"I would try to avoid this person's messages in the future."
"This person contributes information to the group that is useful to me."
"This person contributes information to the group that some people in the group find useful."
"This person's information is reliable."
"This person contributes to the cohesiveness of the community."
"This person behaves rudely or disruptively."
"This person is respected within the community."
"Most people in the community trust this person."
"I trust this person."
"This person seems like someone I might like as a friend."
"I am very familiar with this person's posts."

In both scenarios, subjects needed to return to the news browser to refresh their association between messages and the authors they were rating. This suggests (and many subjects confirmed in casual comments) that the author's name is not an element that is particularly memorable to many users of newsgroups, at least when using a standard interface like that of Outlook Express.

We collected evaluations for a total of 340 distinct authors, 309 of them from our subjects' 22 favorite newsgroups and 31 of them from the *microsoft.public.windowsme.general* newsgroup.

RESULTS

In the following, we present the results of our newsgroup usage survey and an analysis of the relationships within and between author evaluations and author behavioral metrics.[1]

Newsgroup Usage Survey Results

Subjects reported that they checked newsgroups a least a few times every week and that they spent on average

[1] The large number of simultaneous correlations used in our analysis demands a more stringent criterion for the statistical significance of any one correlation. Regardless, the strength of the effect as captured in the correlation metric is the most meaningful indicator of relationships within the data presented here.

between 15 and 30 minutes per session. Seeking technical support was the activity they said they performed most frequently, followed by discussing news and current events. They also sought social support and looked for music and images to download. Users reported that, in the course of their normal news browsing, they read other participants' contributions but rarely post messages of their own. They identified the subject line and the size of the thread as the most important properties to guide their selection of messages to read. To a much lesser extent, they also mentioned the name of the author and the date of the post as important features to readers.

Analysis of the survey responses showed that people who sought out technical support in newsgroups were more likely (r (20) = .637, p < .001) to buy or sell material through newsgroups. Seeking technical support also correlated highly (r (20) =.592, p = .004) with spending longer periods of time participating in newsgroups and negatively, though more weakly, with reading political newsgroups (r (20) = -.465, p = .029). Subjects who reported that they sought social support in newsgroups were more likely (r (20) =.589, p = .004) to seek out entertainment as well. Subjects who reported that they were interested in using newsgroups to find and discuss news events were also loosely related to those seeking political discussions (r (20) =.449, p = .036). Readers of news-related groups were more likely (r (20) =.553, p = .008) to pay attention to the date on which messages were posted. Also present were weaker patterns that related the length of time spent reading newsgroups with buying and selling goods in newsgroups (r (20) =.455, p =.033) and downloading images (r (20) =.466, p = .029).

Table 3. Behavioral descriptors of authors

- The total number of newsgroups to which the author posted at least a single message.
- The dates the author was first and last seen in any newsgroup across the Usenet as well as in the focal newsgroup.
- The number of days on which the author posted one or more messages to any newsgroup across the Usenet as well as in the focal newsgroup.
- The total number of messages the author posted to any newsgroup across the Usenet as well as the number posted to the focal newsgroup.
- The total number of messages the author posted to any newsgroup across the Usenet as well as the number posted to the focal newsgroup that were initial turns ("thread heads").
- The total number of messages the author posted to any newsgroup across the Usenet as well as the number posted to the focal newsgroup that were responses to other messages ("replies").
- The total number of other authors to whom the author replied in any newsgroup across the Usenet as well as the number replied to within the focal newsgroup ("reply targets").
- The total number of replies to messages posted by each author in any newsgroup across the Usenet as well as the number replied to within the focal newsgroup ("responses").
- The total number of other authors who responded to messages the author posted in any newsgroup across the Usenet as well as the number responded to within the focal newsgroup ("reply targets").
- The total number of threads to which the author posted at least a single message in any newsgroup across the Usenet as well as the number of threads in the focal newsgroup.
- The average number of messages each author posted in each thread they joined both in any newsgroup across the Usenet as well as within the focal newsgroup.
- The average number of generations ("depth") of turns and reply messages of each of the threads the author posted to in any newsgroup across the Usenet as well as in the focal newsgroup.
- The average number of sibling turns beneath each parent message ("breadth") in each of the threads the author posted to in any newsgroup across the Usenet as well as in the focal newsgroup.

Behavioral Metrics

We gathered behavioral metrics about authors whom our subjects evaluated for three time periods: the year 2000, this year to date from January 1, 2001 through end of July 2001, and the month of July 2001. Year-to-date 2001 provided the clearest relationships between our metrics and our subjects' evaluations, suggesting that a behavioral history longer than just a month but short enough to live in recent memory is most valuable[2]. Thus, all results presented here refer to authors' behavioral metrics for the time period from January 1, 2001 to July 31, 2001. (Our subjects viewed messages posted through the end of July 2001, when we conducted the lab portion of this study.)

The metrics we selected, listed in Table 3, describe various dimensions of each author's behavior in the newsgroup in which one of our subjects found them (the "focal newsgroup," the one that the subject identified as his or her favorite) and across the entire Usenet in any newsgroup to which they contributed.[3] We also extracted from our database additional information about each focal newsgroup (Table 1).

Reliability of Subjective Author Evaluations

Our subjects' evaluations of the 31 authors in the *windowsme* newsgroup provided a valuable opportunity to test the reliability of the subjective assessments across subjects, because they rated many of the same authors in the group.

[2] More than forty of the authors reviewed by our subjects had posted no messages at all in July 2001, making their behavioral descriptors for that month far less useful than the pattern of behavior they created over the first seven months of the year.

[3] The Netscan system receives messages from more than 50,000 active newsgroups. While it is difficult to evaluate the exhaustiveness of these data, they do represent a large portion of the mainstream Usenet feed.

Eleven of the 31 authors were rated by every participant. For these 11 authors, we assessed inter-rater reliability using Cronbach's alpha for each subjective author evaluation statement. Ten of these 12 measures[4] received inter-rater reliability scores of $\alpha > .70$; we consider these reliably rated. Two measures, "This person behaves rudely or disruptively" ($\alpha = .47$) and "This person seems like someone I might like as a friend" ($\alpha = .43$), were unreliably rated. The median alpha level for reliably rated subjective author evaluation measures was .85.

High inter-rater reliability indicates agreement among participants' perceptions of newsgroup authors in terms of our subjective assessment measures. Strong agreement is important to further analyses of participants' ratings of the 309 authors in their favorite newsgroups, in which only one participant rated each author.

Author Evaluation Interrelations

In addition to estimating the reliability of our raters, we also wanted to demonstrate internal consistency of our author evaluation measures. To achieve this, we examined the correlations between them, looking for expected relationships. One obvious example: authors whose messages would be read in the future should also be authors who would not be avoided in the future. Indeed, this is what we found. Also, unsurprisingly, level of interest in reading a message from an author in the future correlated highly ($r (306) = .76$, $p < .001$) with ratings of the usefulness of the author's post. The degree to which an author was considered rude correlated with avoidance of the author in the future ($r (307) = .58$, $p < .001$), and respectability and reliability of authors were highly correlated, $r (307) = .75$, $p < .001$.

Beyond the expected correlations, in looking just at these subjective evaluations of newsgroup authors, other interesting relationships emerge. For example, desire to read a message from an author in the future correlated highly with ratings of how much the author adds to group cohesion ($r (306) = .70$, $p < .001$) and how respected the author is in the newsgroup community, $r (306) = .63$, $p < .001$. These relationships indicate two community-based or social reasons for reading messages by particular authors.

The degree to which our subjects indicated personal trust in an author (as opposed to ratings of how much the newsgroup community trusts the author) correlated highly with how much they found the author's posts personally useful, $r (307) = .76$, $p < .001$. This suggests that authors that are trusted are considered useful, or that when an author is useful he or she tends to be more trusted. The degree of personal trust in an author also correlated highly with the degree to which an author is perceived as respected in the community ($r (307) = .74$, $p < .001$) and the likelihood that the author is someone participants would want as a friend, $r (296) = .73$, $p < .001$. The former relationship indicates that those authors that are trusted are perceived as more respected in the community, or that when an author is perceived as more respected in the community, he or she is consequently more trusted. The latter relationship is likely more unidirectional: trusted authors are more likely to be people participants would have as friends.

Behavioral Metric Interrelations

The quantitative behavioral descriptors for the newsgroup authors whom our subjects read exhibited some interesting interconnections. In fact, almost all of the behavioral metrics correlated significantly and positively with each other, probably because most of them measure some facet of posting activity, so they tend to trend upwards together. However, after we normalized these values against totals for individuals and whole newsgroups, we identified some correlations worth noting. For this analysis, we excluded outlying data points that were more than 3.5 standard deviations from that metric's mean, though the same effects were evident with outliers included in the set.

First, the number of messages that an author posted either to the focal newsgroup or to any newsgroup correlated positively with the average number of messages in the threads in which the author participated. (For the focal newsgroup, $r (296) = 0.308$, $p < 0.001$; for any newsgroup, $r (292) = 0.374$, $p < 0.001$.) The correlations between total number of posts and thread depth and breadth, secondary measures of thread size, were even stronger (between $r (290) = 0.415$ and $r (293) = 0.447$, $p < 0.001$). These relationships suggest that more active authors tend to participate in larger, more active conversations.

Second, logically but not obviously, the following metrics varied directly with the number of threads in which an author participates. (Correlations reported here derive from the author's behavior in all newsgroups; results were parallel for behavior in the focal newsgroup only.)

- Number of replies the author posts, $r (259) = 0.852$, $p < 0.001$
- Number of unique targets of his or her replies, $r (294) = 0.896$, $p < 0.001$
- Number of responses his or her messages garner, $r (279) = 0.801$, $p < 0.001$
- Number of distinct authors of those responses, $r (281) = 0.872$, $p < 0.001$

These correlations reveal that authors who joined in more threads, which we might term topics of conversation, tended to interact more with others in the group.

[4] Familiarity with the *windowsme* authors was not measured since no participant had ever read this newsgroup and was not likely to recognize any authors.

Finally, and perhaps most interestingly, the fraction of an author's total messages that were replies increased with the number of messages the author posted (in the focal newsgroups, $r(251) = 0.259$, $p < 0.001$; across all newsgroups, $r(259) = 0.297$, $p < 0.001$). The fraction of messages that were replies also rose in conjunction with the number of threads in which the author participated in the focal group ($r(248) = 0.288$, $p < 0.001$). These statistics underscore that more active authors also are more interactive; the more they post, the more they involve themselves with ongoing discussions and converse with other posters.

Relations Between Behavioral Metrics and Author Evaluations

When we considered together all of the evaluations created by our subjects, the correlations between the behavioral descriptors of authors and our subjects' ratings of those authors were tenuous. The only clear relationships involved the number of newsgroups in which an author posts and many of our subjective ratings: the more groups an author participated in, the less interested our subjects were in reading more messages by that author, and they were more like to rate him or her as rude. This is very likely because those authors who post advertisements and other undesirable messages tend to do so in many groups at once.

We suspected that our subjects' degree of familiarity with an author might influence the reliability of their evaluations of the author, so we computed correlations for our data set again with only those evaluations for which the subject reported being familiar with the author. The results tightened up, drawing out several patterns of interest.

First, our analysis revealed positive characteristics that paint a picture of an author whom our subjects were likely to want to read again. Some of the behavioral metrics that correlated positively with subjects' desire to read more by an author include:

- The number of days on which the author posted in the focal newsgroup ($r(64) = 0.376$, $p < 0.002$) and in all newsgroups ($r(64) = 0.412$, $p < 0.001$)
- The number of messages the author posted in the focal newsgroup ($r(64) = 0.288$, $p < 0.019$) and in all newsgroups ($r(64) = 0.262$, $p < 0.034$)
- The number of distinct threads in which the author participates ($r(64) = 0.321$, $p < 0.009$) in the focal newsgroup
- The depth and breadth of the threads in which the author participates
    - Depth of threads in the focal newsgroup: $r(64) = 0.427$, $p < 0.001$
    - Breadth of threads in the focal newsgroup: $r(64) = 0.388$, $p < 0.001$
    - Depth of threads in all newsgroups: $r(64) = 0.378$, $p < 0.002$
    - Breadth of threads in all newsgroups: $r(64) = 0.327$, $p < 0.007$
- The fraction of the author's messages which are replies, in the focal group ($r(64) = 0.346$, $p < 0.004$) and in all newsgroups ($r(64) = 0.342$, $p < 0.005$)
- The number of distinct people to whose messages the author replies in the focal newsgroup and in all newsgroups (for both, $r(64) = 0.335$, $p < 0.006$)
- The number of replies which the author's messages garner in the focal group ($r(64) = 0.281$, $p < 0.022$) and in all newsgroups ($r(64) = 0.333$, $p < 0.006$)
- The number of distinct people who reply to the author's messages in the focal group ($r(64) = 0.338$, $p < 0.006$) and in all newsgroups ($r(64) = 0.361$, $p < 0.003$)

Correlations with subjects' respect for authors and trust in them followed patterns very similar to those with subjects' desire to read more messages by them, though some of the relationships were less significant. Those that were still strong included correlations between statements of trust or respect and:

- The fraction of the author's messages which are replies in the focal newsgroup and across all newsgroups
- The number of replies, reply targets, responses, and response targets across all newsgroups
- The number of days on which the author posted in the focal newsgroup and across all newsgroups

We interpret this portrait as one of a poster who participates actively and regularly in a variety of in-depth conversations, in which he or she responds to other participants but does not overwhelm the discussion (or, as we explain next, participate in too many different newsgroups).

Second, by isolating authors rated as familiar, the negative relationship strengthened between the number of newsgroups in which an author posted and our subjects' interest in reading the author again ($r(64) = -0.575$, $p < 0.001$). A negative correlation also emerged between the subjects' desire to read more messages by the author and the average number of messages the author posted in the threads in which he or she participated ($r(64) = -0.646$, $p < 0.001$). The latter result indicates that those who

dominate the conversations that they join tend to be viewed unfavorably.

In evaluating our results and deciding whether to examine only those subjects' evaluations with high familiarity ratings, we also investigated what contributes to familiarity. One factor is the number of posts made to the newsgroup. Authors whom our subjects rated as familiar contributed a median of 153 messages each to all newsgroups and 20 messages each to the focal newsgroup. Authors rated as unfamiliar to our subjects contributed a median of 42 messages each to all newsgroups and 8 messages each to their respective focal newsgroups.

Another consideration: If we achieve strong results by examining only those cases where our subjects were familiar with the authors they were rating, we should consider which cases we eliminate in the process. Our criterion for familiarity was a response of 6 or higher on the 1-to-7, "strongly disagree" to "strongly agree" scale to the statement, "I am very familiar with this person's posts." Of our 22 subjects, 13 responded to this statement with a 6 or higher for at least one of the authors they evaluated. However, the distribution was not even among those 13. The two most prolific subjects generated 13 and 17 familiar evaluations, together constituting approximately 45 percent of the 66 assessments of a familiar author. This gives them inordinate weight in determining our conclusions, but we do not consider them to be so dominant as to interfere with the overall validity of our findings.

In an effort to further verify the strength of our findings with familiar-author evaluations, we ran the same correlations on evaluations of only unfamiliar authors, those for whom subjects responded with a 3 or lower on the same 1-to-7 scale to same statement about familiarity with the author's messages. We confirmed the need for subjects to be familiar with authors in order to evaluate them soundly. Evaluations of unfamiliar authors showed almost no significant correlations with the behavioral metrics for those authors. In contrast, evaluations of familiar authors correlated strongly and significantly with objective metrics, as we described above.

FUTURE DIRECTIONS

Ultimately, with this research, we hope to build a system to help readers quickly find messages that they would actually like to read and avoid messages they would prefer not to read. For casual newsgroup users, such a system would reduce the clutter of undesirable messages and point them to authors and messages of value. More serious newsgroup users, especially those who primarily seek technical solutions, could benefit greatly from a system that helps them efficiently find what they are looking for amid the scores or even hundreds of daily posts.

We believe this can be done in part with the behavioral metrics we have examined here. The results presented in this paper suggest that reliable ratings of newsgroup authors can be obtained from other group members provided they are familiar with the author in question. Further, many of these author evaluations correlate strongly with quantitative behavioral metrics. Establishing the link between these behavioral metrics and the subjective evaluations permits us the next step of using the behavioral metrics to predict the subjective evaluations.

Such predictions could be made in many ways. As an example, regression could be used to generate weightings for the behavioral metrics as predictors of authors a person may want to read. To get a sense of how successful such an approach would be, we generated some preliminary linear regression equations to predict the author evaluation, "I would read a message by this person in the future," from different combinations of behavioral metrics. One successful set of predictors involved just six behavioral metrics: the average depth of the threads the author is involved in, both in the focal newsgroup and across all newsgroups; the average number of posts the author contributes to the threads in which he or she participates, both in the focal newsgroup and across all newsgroup; and the number of threads in which the author was active, both in the focal newsgroup and across all newsgroups. Considering just the authors for whom we obtained reliable ratings (the familiar authors), these six behavioral metrics predicted 56 percent of the variance in responses to the likelihood of reading another message by each author ($R^2 = .56$), $F(6,68) = 14.355$, $p < .001$.

Since we already have access to behavioral metrics for an extensive population of Usenet users, one way to refine our understanding of the relationship between them and the subjective evaluations would be to recruit evaluations from a significantly larger subject population. To that end, we propose a field deployment in which hundreds or even thousands of newsgroup readers would be given a special news browser that occasionally prompts them to evaluate authors with whom they are familiar. We are planning such a deployment later this year.

We might also gain insight by adding analysis of the textual content of messages to our set of variables. Numerous techniques from natural language processing and information retrieval would provide us with additional descriptors of authors based on the material of their posts and their styles of writing to supplement the structural information we have used in this study.

CONCLUSIONS

Quantitative behavioral metrics, in particular those that capture aspects of an author's tenure in a newsgroup and level of interactivity with other authors, serve as reliable predictors of subjective evaluations of the author's social and informational value and, by extension, of readers' satisfaction with his or her messages.

These metrics, revealed in an appropriate manner in browsing and searching interfaces for newsgroups, should make it easier and quicker for users to identify those messages and authors whom they would be most interested in reading. They show great promise as a foundation for other forms of collaborative filtering and offer a potential remedy for the information overload emerging in many online discussion spaces.

ACKNOWLEDGMENTS

We thank our study subjects, the diverse population of the Usenet, the Microsoft Usability Recruiting group, JJ Cadiz, Jonathan Grudin, and Anoop Gupta. Also, we owe special thanks to Duncan Davenport, lead database developer of Netscan, for his efforts in preparing and extracting the behavioral metrics we required for this study.

REFERENCES

[1] Donath, J., K. Karahalios, and F. Viegas. "Visualizing Conversation," *Proceedings of the Hawaii International Conference on System Sciences 32* (1999).

[2] Kollock, P. and M. Smith, "Managing the Virtual Commons: Cooperation and Conflict in Computer Communities," in *Computer-Mediated Communication*, S. Herring (ed.), John Benjamins, Amsterdam (1996).

[3] Resnick, P., Iacovou, N., Sushak, M., Bergstrom, P., and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," *Proceedings of Computer-Supported Cooperative Work* (1994).

[4] Sack, Warren, "Discourse Diagrams: Interface Design for Very Large-Scale Conversations," *Proceedings of the Hawaii International Conference on System Sciences 33,* Persistent Conversations Track, (2000).

[5] Smith, M. and A. Fiore, "Visualization Components for Persistent Conversations," *Proceedings of ACM Computer-Human Interaction* (2001).

[6] Smith, Marc, "Invisible Crowds in Cyberspace: Measuring and Mapping the Social Structure of USENET" in *Communities in Cyberspace*, M. Smith and P. Kollock (eds.), Routledge Press, London (1999).

[7] Whittaker, S., Terveen, L., Hill, W., and L. Cherny, "The Dynamics of Mass Interaction," *Proceedings of Computer-Supported Cooperative Work* (1998).