Feed Me: Motivating Newcomer Contribution in Social Network Sites

Moira Burke
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
moira@cmu.edu

Cameron Marlow and Thomas Lento
Facebook
156 University Ave.
Palo Alto, CA 94301
{cameron, lento}@facebook.com

ABSTRACT

Social networking sites (SNS) are only as good as the content their users share. Therefore, designers of SNS seek to improve the overall user experience by encouraging members to contribute more content. However, user motivations for contribution in SNS are not well understood. This is particularly true for newcomers, who may not recognize the value of contribution. Using server log data from approximately 140,000 newcomers in Facebook, we predict long-term sharing based on the experiences the newcomers have in their first two weeks. We test four mechanisms: social learning, singling out, feedback, and distribution.

In particular, we find support for social learning: newcomers who see their friends contributing go on to share more content themselves. For newcomers who are initially inclined to contribute, receiving feedback and having a wide audience are also predictors of increased sharing. On the other hand, singling out appears to affect only those newcomers who are not initially inclined to share. The paper concludes with design implications for motivating newcomer sharing in online communities.

Author Keywords

Social network sites, SNS, online communities, motivating contribution, production incentives, sharing, feedback, distribution, social learning, singling out

ACM Classification Keywords

H.5.3 [Information Interfaces]: Group and Organization Interfaces - Collaborative computing, Web-based interaction, Computer-supported cooperative work

INTRODUCTION

Social media services are dependent on user contributions to provide value to their products, and as a result designers of such systems build features targeted at increasing the amount of content a given user contributes. One such feature is a *content feed*, which publishes stories about a user or set of users and makes the stories available to others. Such feeds may cause users to increase their rate of content contribution, either by increasing user awareness of product features and the socially acceptable means of using them, encouraging users to contribute content to attract the attention of their peers, or a combination of these effects.

This paper examines the relationship between initial user behavior and content production in a social network environment. Using a set of approximately 140,000 Facebook users who joined in March 2008, we examine the newcomers' initial content contribution and their friend networks to assess the effects of friends' behavior, feedback, and audience size.

MOTIVATING CONTRIBUTION IN SOCIAL MEDIA

Previous studies of participation in online communities have focused on two types of social systems: *information commons*, where many individuals contribute to the construction of a small number of shared artifacts, and online *discussion groups*, where individuals exchange messages on a given topic. In both of these environments, content contribution can be described through the interactions between a few abstract elements: the *people* who are involved, the *content*, or artifacts they produce and share, *feedback*, or engagement between people around content, and *distribution*, the way in which people discover and consume content.

The success of social media environments rests on the correct balance of these elements. In information commons, such as open-source software projects, Wikipedia, and MovieLens, a critical mass of production must exist around a set of artifacts [5,13]; while anyone is allowed to participate around a given piece of information, the structure of the content largely dictates which artifacts get attention [13]. In discussion groups, the success of the community is dependent on motivating participation from enough people [5]. The topic of the forum and the group's interaction norms predict how this engagement will occur [21].

For a social network, the success of the system is tied to the amount of contribution any one member's social contacts have produced, an outcome that is dependent on the eventual participation of a large portion of the user base. Engagement

CHI 2009, April 4–9, 2009, Boston, MA, USA.

is limited by the privacy constraints of the system. The content any one member can access is limited by the privacy settings of those who produced it, meaning that members often only see content produced by the people they know [26]. Because the structure of engagement and types of social relationships vary so widely, we expect that many of the incentives that drive contributions in other social media may not be applicable to the SNS environment.

Theories about participation can be grouped into three high level categories: what a user sees other users doing (*social learning)*, effects that other users have on the newcomer (*feedback*), and the general structure of content and exposure achieved through participation (*distribution*). Before introducing our study, we will briefly review work in these areas.

Social learning

From the perspective of a user, a social network site is primarily comprised of a set of friends and the content they produce. Therefore, one likely influence on newcomers' behavior is the behavior of their friends. Social learning theory [2] suggests that people learn by observation in social situations, and that they will begin to act like people they observe even without external incentives. Though the initial studies dealt with children observing violence, the effects have been generalized to many other domains, including pro-social behavior in adults [2,3].

Social learning theory indentifies several necessary steps in the learning process: *attention*, that people need to be able to observe the behavior without distraction; *retention*, or the need to remember the behavior; *reproduction* or the ability to perform the action; and *motivation*, including past, promised, or vicarious reinforcement, which influence us to reproduce what we have learned [2]. Social network sites provide all of the necessary steps for social learning to occur, particularly when friends' actions are aggregated in a content feed. The feed allows newcomers to view friends' actions, recall them later, and may make links to the tools for content contribution more salient.

In previous research applying social learning theory to an online newsgroup, Slashdot, the number of pages newcomers viewed before posting their second comment was significantly correlated with the quality score of the comment, though the quality of their first comment was not affected by page views [37]. There are a couple of possible explanations for the discrepancy: First, newcomers with poor initial comments self-selected away from the site, and so others' actions had little effect on them, while those that stayed had more opportunity to observe other members. Alternatively, using page views to represent social learning may not capture subtleties like the quality of the content viewed or the social relationships between the actors. Social networking sites offer the opportunity to fine-tune the social learning metric, by taking into account friends' actions and exactly which actions the newcomers were exposed to. The fact that all modeled behavior is also produced by users' friends is an additional benefit offered by studying SNS, as friends have a baseline level of interest to the user.

Bandura emphasizes learning by observing the outcomes of others' actions (e.g., Newcomer A observes Friend B's reaction to Friend C's photos) [2]. As some of this reinforcement may be outside the SNS, or not entirely visible to the newcomer, a first step is to simply examine the first-degree impact of friends' content production on newcomers, to see whether newcomers model their friends. We discuss second-degree reinforcement in the Future Work section.

H1. Social learning: Newcomers whose friends share more content will go on to contribute more content themselves.

Social learning is not limited to isolated behaviors that newcomers observe their friends doing from afar; it also applies to behaviors where newcomers are singled out by their friends. For example, a friend might "tag" a newcomer in a photo, engage the newcomer in a chat session, or refer to the newcomer in a public status update. In all of these cases, the friend directly engages the newcomer with some content, and the new user may be both more likely to notice the content and come to understand the value of participation. This may lead to long-term engagement on the part of the newcomer. These singling-out actions may also highlight some social connection between newcomers and their friends, potentially providing an added effect of in-group membership [32]. Overall, we expect that newcomers who are singled out in content will go on to contribute more content.

H2. Singling out: Newcomers who are singled out in content will contribute more content.

Feedback

For a newcomer, feedback from fellow members could lead to future participatory behavior. Theories of reciprocity [12,24], reinforcement [19], and the need to belong [4] all suggest that feedback from other users should predict long-term participation on the part of the newcomer. Feedback differs from social learning, particularly singling out, in that feedback requires the newcomer produce some initial content, while newcomers can be singled out without taking any actions themselves.

In studies of online newsgroups, receiving a response to one's first posted message was significant in motivating ongoing contributions from newcomers. However, length, tone, content, and personal affirmation were not found to be significant predictors of long-term engagement [34]. Newcomers to the online news community Slashdot whose first comments received positive numeric ratings returned significantly faster to the site to post a second comment, and when their first comment received a reply they also tended to return more quickly [37]. Controlled experiments also show that social approval in the form of messaging increases a subject's number of contributions (e.g., [10]).

With these results in mind, we expect that feedback in the social environment of a SNS will increase the participation rate of newcomers.

H3. Feedback: Newcomers receiving more feedback on their initial content will go on to contribute more content.

Distribution

In the case of other online environments, it has been shown that reputation is a common motivation for participation. For open-source software, competitive motivations in the form of reputation and status attainment have been cited as a primary incentive for continued participation [30]. Similarly, bloggers cite the intent to affect their professional reputation as being among their top motivations for blogging [39]. In both of these cases, the distribution of attention received by the author is important independent of the particular feedback she receives.

For this reason, we should also consider the benefits derived directly from having a wider audience. This distribution might also affect other forms of feedback that are exogenous to a social media system (e.g. conversations in the hallway, email, etc.).

H4. Distribution: Newcomers whose initial content is distributed widely will go on to contribute more content.

NEWCOMER CONTRIBUTION IN FACEBOOK

Within the HCI and CSCW communities, social networking sites like Facebook and MySpace have received much attention as platforms for studying social psychological phenomena because friendships are articulated and interactions are logged. Recent topics include identity management and signaling [15,36], social capital [17], trust and privacy [16], and social use differences between demographics [7,28,33,35]. With over 150 million active users worldwide [18], including hundreds of thousands of new registrants daily, Facebook is strongly suited for studying newcomer engagement and participation in social media systems.

Within Facebook, content contribution takes many forms. For example, members post brief status messages, upload photos, or write on semi-public, free-text areas of their friends' profiles (a feature Facebook calls the Wall). Users' activities are listed on their own profiles, and users see their friends' recent activities and content on a dynamic list known as the News Feed. Figure 1 shows a sample News Feed.

The present paper focuses on how early experiences impact newcomers' long-term photo uploading behavior. The most common form of content contribution within Facebook is photo uploading; the photo application draws more than twice as much traffic as the next three largest photo sharing websites [18]. Not only is photo uploading common, there are also many mechanisms on the site for social learning, singling out, feedback, and distribution. Furthermore, photos appear in many parts of the site, including newcomers' own photo albums, their friends' albums, their News Feeds, and

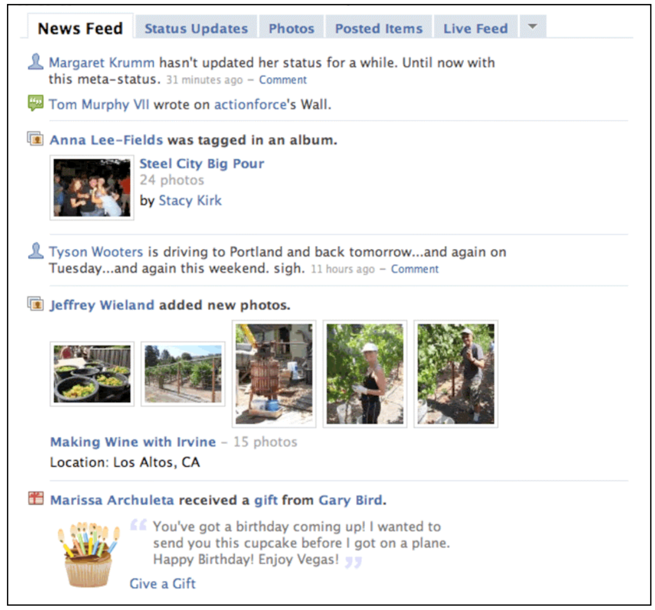

Figure 1. Sample Facebook News Feed showing the viewer's friends' actions.

their Profiles, and thus there are many opportunities for both future controlled experimentation and improving the photo sharing experience.

METHOD

Data

To test the impact of learning, singling out, feedback, and distribution on the extent to which newcomers share content with friends, we selected a cohort of all 254,603 users who registered for Facebook on a randomly chosen weekday in March 2008. The cohort includes members from 207 countries with 24% registering from the United States. We collected all model variables during the newcomers' first two weeks and then predicted their content sharing in the subsequent three months. To mitigate the effects of fake accounts and members who never returned, any users who did not log in at least once during their third month were removed from the data, leaving a set of 140,292 newcomers.

All variables were aggregated from server logs using the Hadoop distributed computing system [27]. The data were analyzed in aggregate so that no individual user's actions, friend network, or personally identifiable information were used in analysis.

Additionally, we performed semi-structured face-to-face pilot interviews with seven users who had been members of Facebook for less than eight months, and who had varying levels of photo activity. Participants responded to a classified ad and came to a lab in the Bay Area. They logged into their Facebook accounts and demonstrated how they typically use the site. We probed mentions of their own content production—such as status updates, wall posts, or photos—or lack thereof, but we did not ask directly about their motivations for contributing content. Interviewees generally talked about their friends' activity, what they considered socially acceptable, and privacy concerns. The present paper focuses on the quantitative model described above; quotes from the interviews appear in the Discussion section.

Measures

Dependent variable

The outcome predicted by the model is the number of photos uploaded by the newcomers between their third and fifteenth weeks on the site.

Controls

We control for basic demographics and initial interest in the site in our model. As users get older, their usage of the site is expected to change; for this reason, age (in years) is included as a control. Similarly, users who chose to set their gender may be more willing to disclose private information than those who do not fill out this basic profile field. Furthermore, men and women may behave differently, so we use dummy variables to represent the three levels of gender: male, female, and no gender specified. "No gender" is the omitted level in the models. Users who look at many pages may be more active or interested in SNS, and thus we control for the number of pages the newcomers view. Similarly, users with many friends may be more extraverted, or simply more engaged in the site, so we control for the size of their friend network. Finally, some newcomers may inherently enjoy viewing or sharing photos more than others, so we control for the percentage of pages they view that are photo-related, the number of photos they upload (if any), and the number of comments or photo tags that they write in their first two weeks.

Independent variables

Learning is represented by the number of photo-uploading stories the newcomers saw in their News Feeds during their first two weeks. The results were qualitatively similar when using the number of photos uploaded by the newcomers' friends, and the two measures are highly correlated (r=0.715). However, the latter variable does not ensure that the newcomers knew about their friends' photos, while the former at least guarantees that the newcomer was exposed to a story with photo thumbnails. We should also note that the number of photo stories about the newcomer's friends is also highly correlated with the number of friends the newcomer has (r=0.725), and including both photo stories consumed and number of friends in a single model leads to inflated standard errors caused by multicollinearity. Therefore, the number of friends is dropped from the affected models.

Singling out is represented by a binary variable indicating whether the newcomer was tagged in a photo during his or her first two weeks. Unlike tagging in other social systems like Flickr or de.licio.us, where members use descriptive terms like "sunset" or "cool," tagging in Facebook is the linking of a face in a photo with a registered user. Friends tag photos by clicking on a face and selecting a name from a list of their friends (see Figure 2). Tagged photos are then

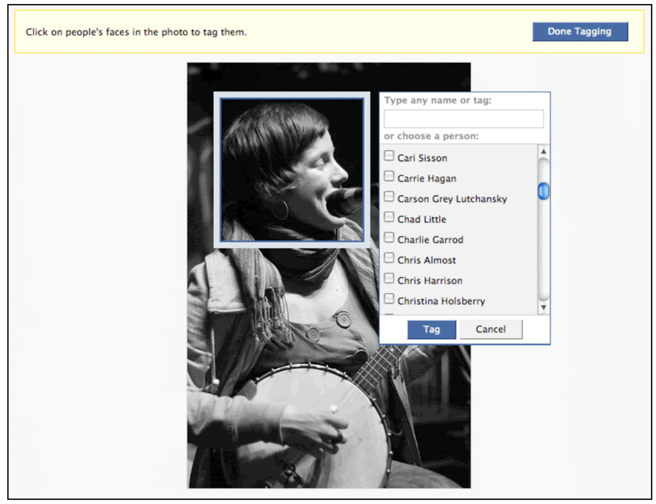

Figure 2. Tagging interface in Facebook. Once a user is "tagged" in the photo, the photo is linked to her profile.

linked from the tagged person's profile. Tagging may occur in any photo, not necessarily one posted by the newcomer, and thus is a way for friends to both demonstrate a feature of the site, and to draw newcomers into photo viewing and sharing. Some newcomers tagged themselves (5.8%), and these self-tags were excluded.

Feedback is measured by a binary variable indicating whether or not the newcomer received any comments on his or her initial photos during the first two weeks. Self-comments, where newcomers commented on their own photos, were excluded. As with singling out, approximately 7% of the newcomers made self-comments, although many of these were in response to comments made by others. This is a conservative measure, as feedback regarding a posted photo may occur in many other channels, including wall posts, private messages, email, and face-to-face conversation.

Distribution is measured as the number of News Feed stories shown to friends about the newcomer's photos. Note that newcomers would not know how many stories were generated about their own photos, as the appearance of any particular story depends on the relationship between the newcomer and the friend, and how many actions the friend's other friends performed recently, all of which are competing for space in the News Feed. However, newcomers may generally infer the size of their audience based on their number of friends, the number of stories in their own News Feeds, or from past comments by friends who have seen the newcomers' content. Based on Hypothesis 4, we expect that newcomers would be motivated to upload more photos if they believe they have a large audience, even if they do not know its precise size. While the purest measure of distribution might simply be the number of friends the newcomer has, this value depends on many exogenous factors for users who recently joined the site, such as the popularity of the website or growth in particular countries or demographics at the time. It could also reflect underlying

	Early uploaders (N=50,929)		Non-early uploaders (N=89,363)	
	Mean	SD	Mean	SD
Age (in years)	25.4	9.83	28.1	11.9
Male (0/1)	0.36	0.48	0.31	0.46
Female (0/1)	0.48	0.50	0.32	0.47
Pages viewed✧	789.6	1080.2	223.6	479.1
Photo pages (%pageviews)✧	0.11	0.10	0.06	0.09
Photo comments written✧	1.5	6.4	0.19	1.50
Tags written✧	2.7	14.8	0.01	0.43
Initial photo uploads✧	20.8	46.6	0.00	0.00
Friends' photo stories appearing in own News Feed✧	36.4	53.9	8.0	23.4
Got tagged in a photo (0/1)	0.23	0.42	0.08	0.27
Got a photo comment (0/1)	0.38	0.48	NA	NA
Own photo stories appearing in friends' News Feeds✧	43.4	170.5	NA	NA

(0/1)=binary variable, ✧=logged[1], NA=not applicable

Table 1. Descriptive statistics for "early uploaders" (newcomers who uploaded >1 photo in their first 2 weeks), and "non-early uploaders" (those with 0 or 1 photos).

Note that gender had three levels: Male, Female, or None Specified. Variables are continuous unless indicated.

personality characteristics like extraversion. Therefore, distribution is measured in terms of News Feed stories viewed by friends rather than pure friend count.

RESULTS

To determine the factors associated with newcomer content sharing, we performed a least squares regression on the outcome variable, the number of photos uploaded by the newcomer between his or her third and fifteenth week (Mean=17.2, Standard Deviation=53.5). All continuous variables, with the exception of age, follow power law distributions, and thus we use the logarithm of these measures in the models[1] to control for skew. The age variable was left in raw units. Results were qualitatively similar when using standardized variables, and thus are left in non-standardized form for interpretability. Raw means and standard deviations are reported in Table 1, and all variables were centered at their means. The choice to operationalize each variable as continuous or binary was made based on its frequency in the data. Two of the independent variables—photo comments and tags from friends—were rare, with medians of 0 and means of approximately 2. Therefore, it was more appropriate to test whether the newcomer was tagged or commented on at all, rather than count the number of tags or comments received. The other two independent variables (photo stories in the two newsfeeds) had non-zero values for nearly all of the newcomers, and so were more appropriately operationalized as continuous variables.

We present the results in two separate models. Hypotheses 3 and 4 can only be tested on newcomers who upload photos—those who have content to distribute or receive feedback on—so we divide the newcomers into two subsets: "early uploaders" who uploaded more than one photo during their first two weeks (N=50,929), and "non-early uploaders" who uploaded zero or one photo (often a profile picture) in their first two weeks (N=89,363). Model 1 (see Table 2) tests all four hypotheses on the early uploaders. As the early uploaders may be qualitatively different than the non-early uploaders (e.g. younger or with more active friends), there is the potential for the coefficients to reflect a selection bias. Therefore, we use a two-stage Heckman correction across both sets of newcomers [29]. In the first stage (not shown), we predict whether the newcomer will be an early uploader. In the second stage (shown), only early uploaders are included in the model, and we predict how many photos the early uploaders will go on to share after their first two weeks. All controls were significant predictors of early adopting in the first stage of the Heckman model. The second-stage coefficients then control for possible selection bias. All coefficients are reported in terms of the percent change in the dependent variable when the independent variable is increased by one unit (e.g. one year of age, or one doubling of the number of photos uploaded by friends), or when a binary variable is changed from 0 to 1 (e.g. the newcomer is tagged in a photo).

The intercept in Model 1 represents a newcomer of mean age (25 years old), who did not specify a gender, with the mean number of page views, etc. This person's expected number of photo uploads in the following thirteen weeks is 1.2. Women who specified their gender would be expected to upload an additional 131.2%, or a total of 2.8 photos. Men would be expected to upload 39.3% more than those newcomers who didn't set their gender, or 1.7 photos.

Hypotheses 1 and 2 are tested across all newcomers in Model 2 (see Table 3). We test for interactions with the binary variable "early uploader" to see if the impact of learning or engagement differs depending on the newcomer's initial uploading behavior. Note that the effects for early uploaders in Model 2 are consistent with the results of Model 1, but the effect for learning is slightly higher. The main effect for "early uploader" in Model 2 indicates that early uploaders are expected to upload 30.6% more photos in the following thirteen weeks than the non-early uploaders.

To check for inflated standard errors due to multicollinearity between controls and independent variables, we calculated variance inflation factors (VIFs). All VIFs are well below 4, indicating low collinearity between factors [31].

[1] All logarithmic normalization is base 2, after adding a start-value of 1.

Intercept		1.2 photos
Controls	**Coef**	**% change from intercept**
Age (in years)	-0.01	-1.0% ***
Male (0/1)	0.48	+39.3% ***
Female (0/1)	1.21	+131.2% ***
Pages viewed✧	0.24	+18.4% ***
Photo pages (% of pageviews)✧	2.80	+597.4% ***
Photo comments written✧	0.15	+11.2% ***
Tags written✧	0.10	+6.9% ***
Initial photo uploads✧	0.30	+22.8% ***
Independent variables		
(H1) Friends' photo stories appearing in own News Feed✧	0.09	+6.1% ***
(H2) Got tagged in a photo (0/1)	0.03	+2.1% ns
(H3) Got a photo comment (0/1)	0.09	+6.2% ***
(H4) Own photo stories appearing in friends' News Feeds✧	0.04	+2.6% ***
*** p < .001 ✧=logged	N=50,929	R^2=0.20

Table 2. Model 1: OLS regression predicting number of photos the newcomer would upload during weeks 3-15 based on factors during weeks 1 and 2. "No gender specified" is the omitted level for gender. Non-early uploaders were censored using a Heckman correction, described in the text.

Learning from Friends

Friends' behavior during newcomers' first two weeks modestly impacts the newcomers' eventual sharing. Newcomers who initially uploaded more than one photo themselves had a mean of 27 friends who uploaded an average total of 220.2 photos during the two-week window. These newcomers saw an average of 36.4 stories about their friends' photo uploads in their own News Feed. Every doubling of these photo upload stories was associated with a 6.1-10.7% increase in sharing (see Models 1 and 2). Newcomers who did not initially upload more than one photo had far fewer friends (Mean=9.8), with far fewer photos across them (Mean=69.8), and saw far fewer photo stories (Mean=8.0) but results are consistent: doubling the photo-upload stories in these newcomers' News Feeds is associated with a 2.2% increase in photo sharing.

Singling Out

Tagging results are somewhat surprising. After taking learning into account, being tagged in a photo is not significantly associated with an increase in subsequent sharing for the early uploaders. However, for the non-early uploaders, tagging was associated with a significant increase in sharing. Tagging a newcomer who is not uploading photos themselves is associated with a 7.2% increase in subsequent photo sharing. Results were qualitatively similar using number of tags.

Intercept		1.9 photos
Controls	**Coef**	**% change from intercept**
Age (in years)	-0.01	-0.7% ***
Male (0/1)	0.84	+79.6% ***
Female (0/1)	1.43	+169.8% ***
Pages viewed✧	-0.02	-1.6% ***
Photo pages (% of pageviews)✧	2.35	+408.3% ***
Photo comments written✧	0.24	+17.7% ***
Tags written✧	0.17	+12.6% ***
Early uploader (0/1)	0.39	+30.6% ***
Independent variables		
(H1) Friends' photo stories appearing in own News Feed✧ X non-early uploader	0.03	+2.2% ***
(H1) Friends' photo stories appearing in own News Feed✧ X early-uploader	0.15	+10.7% ***
(H2) Got tagged in a photo (0/1) X non-early uploader	0.10	+7.2% ***
(H2) Got tagged in a photo (0/1) X early-uploader	-0.05	-3.6% ns
*** p < .001 ✧=logged	N=89,363	R^2=0.20

Table 3. Model 2: OLS regression predicting number of photos the newcomer would upload during weeks 3-15 based on factors during weeks 1 and 2. "No gender specified" is the omitted level for gender. All newcomers were included in this model, as well as an interaction with "early uploader" status.

Feedback

For newcomers who uploaded more than one photo during their first two weeks, receiving feedback in the form of photo comments predicted the number of photos they would go on to upload in their second and third months. Only 38% of these newcomers received a photo comment, and receiving even a single comment was associated with a 6.2% increase in subsequent photo sharing ($p < .001$).

Distribution

Results for the distribution hypothesis are modest. Having more friends view a newcomer's content is associated with a small increase in sharing. The mean number of stories about the newcomer's photos appearing in friends' News Feeds is 43.4, and every time that number is doubled, the expected number of newcomer photo uploads increases by 2.6%.

DISCUSSION AND FUTURE WORK

We find support for Hypotheses 1, 3, and 4, and partial support for Hypothesis 2 (see Table 4).

Both models support the social learning hypothesis. This suggests that users are closely monitoring and adapting to what their friends are doing. This effect is even more pronounced for users who were initially attuned to photo participation, where the impact of seeing more stories about friends' photos was three times larger. Social comparison theory [20] suggests that if many of your friends contribute content, you would want to "keep up with the Joneses" and contribute more as well.

These results are reflected in the comments of the qualitative interviewees:

> "I don't usually become a fan of something myself [first]; I usually find someone else becomes a fan of something, like bacon or Obama, and I'll become a fan of it. . . . I rely almost solely on the News Feed [to learn about what friends are doing]" (P7, a man in his early 30s)

> Describing how she got started posting maternity pictures: "My friend had a baby in February, and she was posting all of her baby pictures here. And so she sent me direct links to the photos." (P3, a woman in her early 30s)

> "I like to see what groups my friends join, and then if there are any interesting [ones], I might join, too." (P4, woman in her late 30s)

The present model does not test traditional reinforcement, a key component in Bandura's social learning theory [2]. For example, we know that newcomers are motivated to upload photos when they see their friends uploading, but we don't know if seeing their friends being rewarded with positive photo comments is also motivational. The next study will examine the impact of newcomers viewing friend-to-friend feedback.

These effects should also be explored more deeply with respect to the social environment a newcomer enters. For instance, users with active friends might be connected to a few heavily active users, or a large number of moderately active users; our current model does not take this distinction into account.

We also found mixed support for our hypothesis regarding singling out. While one might expect tagging to have a profound effect on newcomer behavior, especially due to the uniqueness of this feature, it was not a significant predictor across our entire sample. However, for those users without the initial inclination to upload photos, this feature did impact their eventual use. One possible explanation for the differential effect is that the early uploaders already understood system functionality related to photo uploading—including tagging—and thus did not need the extra demonstration of tagging by their friends to become involved. Or, they may simply have not thought of tagging as anything special, while newcomers who were not initially engaged with photo uploading were more affected by being singled out by their friends.

Hypothesis	Early uploaders	Non-early uploaders
H1. Learning	✓	✓
H2. Singling out	NS	✓
H3. Feedback	✓	NA
H4. Distribution	✓	NA

Table 4. Summary of findings. NA=Not applicable to this subset of users. NS = Not significant

Additionally, newcomers may have incorrect or confused expectations about the functionality of tagging. In our qualitative interviews, participants expressed some confusion over social norms of acceptability, the purpose of tagging, and how the feature works:

> "'Tag,' to me, connotes graffiti. And that's a negative to me. Kind of a violation of privacy . . . that's kind of stalking." (P2, a woman in her late 40s)

> "Generally no [tagging] in family photos. Why would you bother tagging, because we all know who they are?" (P6, a woman in her late 40s)

> "I'm just assuming that like a game of tag, I'm It, and now I've tagged someone else. I'm done, and that means [my friend's] It, and I'm no longer It." (P1, a man in his early 40s).

> "I thought that only other people could tag you in their photos" (P5, a woman in her mid 20s).

In the qualitative interviews, participants indicated that they liked receiving feedback; as one participant said:

> "I like it when people comment on my pictures . . . I try to comment on my friends' blogs because it's nice to come see that somebody's been looking at it" (P3).

The present model does not take feedback valence into account: Feedback should only motivate newcomers if they feel positively about it. Systematic sentiment analysis to distinguish positive and negative comments is outside the scope of this study, but would further clarify how the feedback mechanism works. However, unlike other communities such as Slashdot, where users are generally strangers, feedback on Facebook photos is typically positive or neutral. An informal analysis of comments indicates that friends tend to be positive ("Awesome"), neutral ("I was at that concert, too"), or if negative, done in an empathetic way ("Finals, ick!").

In addition to the impact of the feedback channels we measured, there were many that were exogenous to our models: wall posts, chat on the site, private messages, and many we can't measure, including private email, instant messenger, phone, and face-to-face conversations. Interviewees frequently mentioned receiving or giving feedback over other channels:

> "[I've commented] just in passing. I haven't sent an email to anyone in Facebook saying I've seen their photos, mainly because I'm more inclined to talk to them [in person or on the phone]." (P6)

> Describing what happened after sending his cousin a photo of his garden in Facebook: "My cousin just emailed yesterday to say [her] garden is blossoming . . . and [her family is] doing really well." (P1)

> Commenting on her friend's photo of her daughter: "I wouldn't have posted [my comment] on a wall, though. I would do it privately in Eudora [email]." (P2)

> "It's usually over IM that we have a conversation about [my photos or status]" (P7)

Our model could be improved by looking at other channels within the site, such as messages and wall posts. We may be able to encode some of this additional feedback by considering only those messages occurring shortly after the newcomer posts photos and containing keywords such as "photo" or "picture." Similarly, we could observe the social properties of the feedback-givers, such as the popularity, sex, strength of friendship, or other dyadic features. Previous studies of feedback in online communities use feedback from strangers (e.g. in Slashdot [37] or discussion groups [34]). We show that feedback from friends also has a significant effect. Future analysis could look at the effect of feedback from strangers within social networks, for instance for those newcomers whose permissions allow strangers to view and comment on their photos.

Baumeister and Leary's theory of the need to belong suggests that frequent, pleasant feedback from a few people is a fundamental human motivation [4]. Future studies should determine if receiving feedback frequently from a few close friends has greater impact than receiving a little feedback from many friends, or even from strangers.

The effect of distribution was modest. This was reflected in the comments from interviewees, who did not mention taking into consideration the size of their audience, or a desire to express themselves. This might be because distribution is an indirect phenomenon, where a newcomer does not know the number of individuals observing her photos, but instead feels the effects of distribution via other channels, such as feedback from several distant friends. Further studies are needed to determine if knowing the actual distribution size (e.g. the number of times one's content has been viewed) impacts newcomers' long-term production.

Though the number of friends a newcomer had was removed from the models due to high collinearity, the social movement literature suggests that it's not how many friends a person has, it's how that person's friends behave that predicts an individual's participation. Threshold models (e.g. [9,11,25,42]) are based on the fundamental assumption that individuals are influenced by the actions of their social connections, not the simple fact of having connections. Empirical work (e.g. [14,40,41]) supports this assumption.

These analyses are the result of an initial longitudinal study of existing behavior. While the temporal arrangement of these correlations suggests a causal link, further analyses are needed to verify this relationship. For instance, a sensitivity analysis observing multiple cohorts of new users across several time periods should be conducted.

Interpreting the size of the effects found in this analysis can be difficult, given the size of the sample. The practical effect of the significant predictors (e.g. measures of social learning) may be small at the individual level, but at a population level, an additional one or two photo uploads per newcomer translates to hundreds of thousands of photos in the system. These photos may appear in any single user's network, creating virtuous cycles of content production. Measuring the systemic effects of one cohort of newcomers is beyond the scope of this paper, but additional analysis could explore the extent to which such cycles of influence and contribution operate. In the meantime these results should be interpreted with care – the effects found here have a small local impact, but may be more important at the system level.

Furthermore, we analyze just one form of content contribution: photo sharing. Photos are the most heavily used feature of Facebook and photos are shared in many other online communities, and thus they are a good starting point. Other forms of content require different amounts of effort; for example, writing a one-line status update takes much less time than photographing an event and uploading the pictures. Subsequent studies should test similar hypotheses on different kinds of content contribution, particularly if newcomers see their friends performing relatively "cheap" contributions. Social learning may have the same effect, or newcomers could perceive their friends as being less committed to the community, and thus less worthy of learning. Photos, especially those with other people in them, may also serve different signaling functions than other self-presentational elements in the profile, e.g. adding a favorite movie, which are easier to fake [36]. Therefore, results may differ for other forms of content.

The mechanisms investigated in the current study—social learning, singling out, feedback, and distribution—are present in various forms in other sites, including RSS feeds of blogs, photo comments on Flickr, and "barnstar" awards in Wikipedia, and thus we would expect the results to generalize well to other social communities. However, further studies are needed in communities with less of a social component—such as open-source software projects or product recommendation forums—to determine if the mechanisms are effective across relative strangers, as well.

Finally, the present study focuses on newcomers, and does not determine whether the mechanisms affect more established members in the same way or to the same degree.

Previous research of online discussion groups finds that some initial experiences affect the commitment of newcomers more than that of established members [1], and so we might expect experienced users of social network sites to be less affected. However, further studies are needed to determine this.

CONCLUSION

The experience that users of social networking sites have is primarily a function of the content their friends contribute. If a user's friends post photos, compose blog entries, or exchange public messages on each other's walls, she can consume continually refreshing content. This provides an incentive for that user to continue logging in to the site, and might encourage her to contribute more content of her own. While information commons and online discussions can afford to have a handful of highly engaged users contributing the bulk of the content in a specific area, social networking sites require a more widely distributed set of regular contributors. Therefore, it is vital for developers of social networking sites to encourage users to contribute content, as each individual's experience is dependent on the contributions of that person's particular set of connections. It is particularly important, if rather difficult, to encourage continuing contributions from newcomers. Newcomers in social media systems may be unwilling or unable to make contributions, either because they do not understand the norms and values of the community, they do not fully understand how to use the technology, or both.

In this analysis of newcomers' motivations for contributing content on Facebook, we find the mechanisms connected with continued participation vary depending upon a given newcomer's initial engagement with the site. For those who do not initially upload photo albums, social learning and singling out are important mechanisms. A relatively inactive newcomer who sees stories about her friends' photo uploads during her first two weeks on the site is more likely to increase her photo sharing over the next three months. Similarly, if a relatively inactive newcomer is tagged in a photo she will be more likely to increase her photo contributions. Newcomers who are more engaged initially are also affected by social learning, but singling out through photo tags does not appear to have an effect. In addition, these more active newcomers are affected by feedback and distribution. An initially engaged newcomer who receives comments on her early photos is more likely to increase her rate of photo contribution in the future. The same relationship was observed between the size of an initially active newcomer's audience and her propensity to upload photos in the following three months.

These results suggest possible courses of action for designers of social networking sites. Design elements which facilitate learning from friends, singling out, feedback, and content distribution can help increase the level of engagement for new users, leading to further content contributions and an overall better user experience.

The most consistent result we found was for learning from friends. An increase in visible friend photo activity was always predictive of increased newcomer contribution. This suggests that showing new users information about the content contributions of their friends makes them more comfortable with contributing themselves. As newcomers see the contributions their friends make, they may become more aware of a particular feature on the site, and may come to understand how that feature is used, both in terms of what is technically possible and what is socially acceptable.

Designers of social networking sites should also find ways to support newcomers with varying behavioral patterns. For newcomers who are active, highlighting opportunities for others to leave them feedback and allowing the newcomers to increase the size of their audience may be particularly effective. For newcomers who are relatively inactive, designers might want to encourage their friends to pay more attention to them, whether through singling out in a public fashion or sending more directed private communication.

ACKNOWLEDGMENTS

We would like to thank Itamar Rosenn, Adam Kramer, Bob Kraut and Ding Zhou.

REFERENCES

1. Arguello, J., Butler, B., Joyce, E., Kraut R., Ling, K, Rosé, C., and Wang, X. Talk to me: foundations for successful individual-group interactions in online communities. In *Proc. CHI 2006*, ACM Press (2006), 959-968.
2. Bandura, A. *Social Learning Theory*. New York: General Learning Press, 1977.
3. Bandura, A., Ross, D., and Ross, S. Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, 63 (1961), 575-582.
4. Baumeister, R., and Leary, M. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin* 117, 3 (1995), 497–529.
5. Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., and Kraut, R. Using social psychology to motivate contributions to online communities. In *Proc. CSCW 2004,* ACM Press (2004), 212-221.
6. Bierhoff, H. W. *Prosocial Behavior*. Psychology Press, 2002.
7. boyd, danah. "Friendster and Publicly Articulated Social Networks." In *Proc. CHI* 2004. ACM Press (2004), 1279-1282.
8. Butler, B., Sproull, L, Kiesler, S., and Kraut, R.(2002). *Community effort in online groups: Who does the work and why?* Lawrence Erlbaum Associates, 2002.
9. Centola, D. and Macy, M. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology* 113 (2007).
10. Cheshire, C. Selective incentives and generalized information exchange. *Social Psychology Quarterly,* 70, 1 (2007), 89–100.
11. Chwe, M. S.-Y. Structure and Strategy in Collective Action. *American Journal of Sociology* 105, 1 (1999), 128-56.
12. Cialdini, R. B. *Influence*. New York: William Morrow and Company, 1984.

13. Cosley, D., Frankowski, D., Terveen, L., and Riedl, J. SuggestBot: Using intelligent task routing to help people find work in Wikipedia. In *Proc*. *IUI 2007*, ACM Press (2007), 32-41.

14. Dixon, M. and Roscigno, D. Status, Networks and Social Movement Participation: The Case of Striking Workers. *American Journal of Sociology* 108, (2003), 1292-1327.

15. Donath, J. Signals in social supernets. *Journal of Computer-Mediated Communication* 13, 1 (2007).

16. Dwyer, C., Hiltz, S. R., and Passerini, K. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. In *Proc. AMCIS 2007*. http://csis.pace.edu/~dwyer/research/DwyerAMCIS2007.pdf

17. Ellison, N., Steinfield, C., and Lampe, C. The benefits of Facebook "friends": Exploring the relationship between college students' use of online social networks and social capital. *Journal of Computer-Mediated Communication*, 12, 3 (2007).

18. Facebook Statistics. http://www.facebook.com/press/info.php?statistics

19. Ferster, C., and Skinner, B. *Schedules of Reinforcement*. Appleton-Century-Crofts, 1957.

20. Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2) 117-140.

21. Fisher, D., Smith, M., Welser, H. You are who you talk to: Detecting roles in usenet newsgroups. In *Proc. HICSS 2006,* IEEE, 2006.

22. Gilbert, E., Karahalios, K., and Sandvig, C. 2008. The network in the garden: an empirical analysis of social media in rural life. In *Proc. CHI 2008*, ACM Press, (2008), 1603-1612.

23. Gould, R. Multiple Networks and Mobilization in the Paris Commune, 1871. *American Sociological Review* 56, 6 (1991), 716-729.

24. Gouldner, A. The norm of reciprocity: A preliminary statement. *American Sociological Review* 25, 2 (1960), 161–178.

25. Granovetter, M. Threshold Models of Collective Behavior. *American Journal of Sociology* 83, 6 (1978), 1420-1443.

26. Gross, R., Acquisti, A. Information revelation and privacy in online social networks. In *Workshop on Privacy in the Electronic Society*, 2005.

27. Hadoop Project. http://lucene.apache.org/hadoop/

28. Hargittai, E. (2007). Whose space? Differences among users and non-users of social network sites. Journal of Computer-Mediated Communication, 13, 1, (2007).

29. Heckman, J. "Sample selection bias as a specification error," *Econometrica*, 47, (1979), 153-161.

30. Hertel, G., Niedner, S., and Herrmann, S. Motivation of software developers in open source projects: An internet-based survey of contributors to the linux kernel. *Research Policy* 32 (2003), 1159–1177.

31. Hocking, R. Collinearity in multiple regression. Chapter 5 of Methods and Applications of Linear Models: Regression and the Analysis of Variance. Hoboken, NJ: Wiley-Interscience, 2003.

32. Hogg, M. Social categorization, depersonalization, and group behavior. In M. Hogg. & T. S. Tinsdale. *Blackwell Handbook of Social Psychology: Group Processes*. Blackwell, Malden, MA, 2001, 57-85.

33. Joinson, A. N. Looking at, looking up or keeping up with people?: motives and use of facebook. In *Proc. CHI 2008.* ACM Press (2008), 1027-1036.

34. Joyce, E. and Kraut, R. Predicting continued participation in newsgroups. *Journal of Computer Mediated Communication* 11 (2006).

35. Lampe, C., Ellison, N., & Steinfield, C. A Face(book) in the crowd: Social searching vs. social browsing. In *Proc. CSCW 2006,* ACM Press (2006), 167-170.

36. Lampe, C., Ellison, N., & Steinfeld, C. (2007). A familiar Face(book): Profile elements as signals in an online social network. In *Proc CHI 2007,* ACM Press (2007), 435-444.

37. Lampe, C., and Johnston, E. Follow the (slash) dot: effects of feedback on new members in an online community. In *Proc. GROUP 2005,* ACM Press (2005), 11-20.

38. Macy, M. Chains of Cooperation: Threshold Effects in Collective Action. *American Sociological Review* 56 (1991), 730-747.

39. Marlow, C. Linking without thinking: Weblogs readership and online social capital formation. In *Proceedings of the International Communication Association* 2006 (Dresden, Germany, 2006).

40. McAdam, D. Recruitment to High-Risk Activisim: The Case of Freedom Summer. *American Journal of Sociology* 92, 1 (1986), 64-90.

41. McAdam, D. and Paulsen, R. Specifying the Relationship between Social Ties and Activism. *American Journal of Sociology* 99, 3 (1993), 640-667.

42. Watts, D. A Simple Model of Global Cascades on Random Networks. In *Proc. National Academy of Sciences* 99 (2002), 5766-5771.