Searching for the Goldilocks Zone: Trade-Offs in Managing Online Volunteer Groups

Loxley Sijia Wang*, Jilin Chen**, Yuqing Ren*, John Riedl**

*Carlson School of Management
University of Minnesota
321 19th Ave. S., Minneapolis, MN 55455
{wang1290, chingren}@umn.edu

**Dept. of Computer Science and Engineering
University of Minnesota
200 Union St. SE., Minneapolis, MN 55455
{jilin, riedl}@cs.umn.edu

ABSTRACT
Dedicated and productive members who actively contribute to community efforts are crucial to the success of online volunteer groups such as Wikipedia. What predicts member productivity? Do productive members stay longer? How does involvement in multiple projects affect member contribution to the community? In this paper, we analyze data from 648 WikiProjects to address these questions. Our results reveal two critical trade-offs in managing online volunteer groups. First, factors that increase member productivity, measured by the number of edits on Wikipedia articles, also increase likelihood of withdrawal from contributing, perhaps due to feelings of “mission accomplished” or “burnout”. Second, individual membership in multiple projects has mixed effects. It decreases the amount of work editors contribute to both the individual projects and Wikipedia as a whole. It increases withdrawal for each individual project yet reduces withdrawal from Wikipedia. We discuss how our findings expand existing theories to fit the online context and inform the design of new tools to improve online volunteer work.

Author Keywords
Online volunteer groups; Wikipedia; trade-off; productivity; withdrawal.

ACM Classification Keywords
H.5.3: Group and Organization Interfaces; K.4.3: Organizational Impacts

General Terms
Design; Human Factors; Management; Theory.

INTRODUCTION
The past decade has observed tremendous growth in the number of online volunteer groups that self-organize on the Internet to accomplish tasks that used to be performed in traditional organizations. One successful example is Wikipedia, the free online encyclopedia. In just ten years, millions of volunteer editors have collaboratively created more than three million articles in English – and nearly twenty million across all languages. Despite its success, Wikipedia and similar online collaboration efforts face challenges in recruiting and retaining active contributors. Active contributors are crucial to the success of online volunteer efforts. Without a steady rate of participation and a critical mass of contributors, online groups may suffer sustainability issues [10]. Contributors may leave for many different reasons. Some may leave due to outside influences in their lives, such as school, work and family. Others may leave due to conflict within the community [26]. Still others, who work on many articles or participate in multiple projects, may leave due to burnout.

Much of the existing research on Wikipedia and other online volunteer efforts has focused on understanding what motivates members to contribute and how they coordinate and resolve conflict (e.g., [14, 24, 26]). Our research, instead, focuses on the opposite problem: understanding why members stop contributing to a group and the implications of their withdrawal behaviors. Are newcomers or old-timers more likely to withdraw? Are active or inactive contributors more likely to withdraw? For what reasons? What about the impact of social connections between editors? In addition to retaining members, another crucial goal for online volunteer groups is to complete work. In the context of Wikipedia, continued edits to articles are important and have been used repeatedly to measure group activity and individual productivity [25, 39]. By comparing factors that affect individual productivity and withdrawal, we hope to explore a critical trade-off between the two: Can the goals of increasing productivity and reducing withdrawal be pursued simultaneously, or does one need to be sacrificed for the other?

The second key trade-off we investigate is between subgroups within a community and the larger community as a whole. Many online groups have a subgroup structure that allows members to participate in more intimate settings (e.g., sub-forums in web discussions and guilds in online games). In Wikipedia, these subgroups are known as WikiProjects, where members organize to work on Wikipedia articles related to specific topics. The relationship between the subgroups and the community is subtle. Some members of online groups may need to decide whether to focus their efforts on a specific subgroup or to split their efforts among multiple subgroups [17].
In this paper, we examine the two trade-offs in the context of Wikipedia and WikiProjects. We analyze data from 648 WikiProjects to explore the effects of member tenure, tenure dissimilarity, past performance, and involvement in multiple projects on members’ productivity and withdrawal behaviors at the project level. We then examine the effects of these factors on member productivity and withdrawal at the Wikipedia level. Our results provide evidence for both trade-offs and highlight the challenges in balancing multiple desirable outcomes of community health within a large online volunteer effort like Wikipedia.

The rest of our paper is organized as follows: (1) We review organization science and volunteerism literature on member productivity and withdrawal and propose how they may generalize to online groups. (2) We describe our research setting of Wikipedia and WikiProjects as well as how we assembled the data set. (3) We present our main results and discuss their implications for the management of online collaboration efforts.

THEORIES AND HYPOTHESES

We focus on studying productivity and withdrawal, two common ways to measure participation in online volunteer groups. In traditional organizations, productivity can be measured as the quantity of work output, such as the number of papers and reports produced in a research lab. In Wikipedia, number of edits is a common measure of productivity [23, 24, 25, 39]. In traditional organizations, withdrawal includes behaviors such as “psychological withdrawal, lateness, absenteeism, and turnover” [7]. In Wikipedia, the primary evidence of withdrawal is for a formerly active editor to cease contributing.

Research on productivity and withdrawal behaviors in traditional groups and organizations sheds light on studying similar behaviors in online groups [15]. At the same time, insights from offline groups may not be readily applicable to online groups for three reasons: 1) online groups have low entry and exit barriers. In traditional organizations, employees need to follow established procedures to enter or exit the organization or its subunits. In contrast, members of most online groups can enter and exit at will. 2) Online volunteer groups have little leverage to mandate active participation. Members are motivated by different incentives [36]. There is a wider variation in member participation [25], and factors that predict productivity in traditional organizations, such as tenure, may not predict productivity in online groups. 3) Compared to traditional organizations, online volunteer groups like Wikipedia often don’t have formal training and socialization programs, both of which could help new members learn the ropes and form relationships with other members to increase productivity and reduce withdrawal.

Furthermore, the voluntary nature of online groups like Wikipedia poses additional challenges in encouraging productivity and discouraging withdrawal. Similar to members of offline volunteer organizations, online volunteers also may suffer from burnout and subsequently withdraw from these groups [42]. Because withdrawal may happen suddenly without a formal process, project work may be left unattended until new volunteers take over [36].

Trade-Offs Between Productivity and Withdrawal

While Wikipedia strives to encourage productivity and reduce withdrawal, it is unclear whether the two goals can be accomplished equally well. Ideally, online volunteer groups should achieve both goals simultaneously. In reality, there may be tension between the two. For instance, research has shown that members working in larger groups tend to contribute less than members working alone or in smaller groups. There is a greater penchant for “free-riding” [1] in larger groups. Free-riding occurs when members retain their membership in a group and receive benefits or credit without providing the same amount of work as others. When free-riding behaviors are frequent, the group may have a seemingly large membership but does not actually benefit from these extra members.

The tension between productivity and withdrawal may also arise from overburdening the core set of active contributors. It is common for a core group of productive members to be the driving force behind a successful online collaboration effort [25]. Studies on Wikipedia have shown how important that core group is in helping to maintain the overall productivity of the free online encyclopedia, compared to the bulk of community members [23, 34]. Although these core members are highly committed and dedicated to the effort, shouldering a heavy workload may erode their energy and enthusiasm and increase their likelihood of withdrawal. An alternative explanation for why active members leave may be that they feel they have accomplished their mission by contributing all that they know, and that further contribution would require much more research and effort [40].

In summary, both productivity and withdrawal are important outcome measures to study, along with potential trade-offs between the two. A successful online collaboration effort requires both a “critical mass” of participants [28] and a reasonable level of contribution. Both the amount of activity within an online volunteer group and its size have a strong impact on the ultimate sustainability of the group [10].

Factors That Affect Productivity and Withdrawal

In this section, we identify a set of factors that have been linked to individual productivity and withdrawal in the organization science literature. We first summarize the insights from the existing literature and then speculate how they may generalize to online volunteer groups.

Tenure. Tenure has been conceptualized and measured as the amount of time that an individual has been part of a group or organization. Organization science literature posits a curvilinear, inverted-U relationship between tenure and employee productivity [38]. When a newcomer first joins an organization, productivity is expected to increase over time as the person acquires skills, accumulates experience, and becomes familiar with organizational routines and
policies. After a number of years of effort, participants are more prone to a burnout effect. The idea of job burnout is based on the potential buildup of stress and exhaustion that may cause workers to decrease productivity [13]. Research suggests that certain people may be especially prone to burnout over time [4], and burnout especially affects members with longer tenure who have felt frustration on the job [13]. Some of the nonlinear effects may be due to the aggregation of a plateau effect [22]. In essence, most workers improve at first but then reach a plateau, while others may start strong and begin to decline later.

Tenure has also been shown to be a strong predictor of withdrawal behavior across different professions [2]. Longer tenure allows for more experience, which would help increase productivity while reducing withdrawal [3]. In addition, newcomers have been consistently shown to be more likely to leave an organization than those with longer tenure [19]. We expect a similar effect of tenure in the context of Wikipedia collaboration, because newcomers who lack the experience of doing the work and interacting with other editors may feel frustrated or perceive a lack of fit with individual work groups [39]. In addition, research has shown that when newcomers’ edits are reverted by old-timers, the newcomers are more likely to leave Wikipedia permanently [20]. We thus expect tenure to increase productivity and decrease withdrawal.

**Tenure dissimilarity.** Another strong predictor of productivity and withdrawal is interpersonal similarity. Studies of traditional organizations have shown that members in the minority component of a group are discouraged from making substantial contributions due to assumptions that they may be weak performers and because they may have more trouble aligning their interests with the rest of the group [35]. These arguments hold true for demographic attributes such as age and tenure [41]. A newbie in a group that consists of mostly experienced members may feel uncomfortable or inadequately prepared to contribute. Similarly, an experienced member in a group with mostly newbies may have different goals and ideas that are hard to communicate to the rest of the group. Hence, we expect tenure dissimilarity to be negatively associated with productivity.

Tenure dissimilarity may also affect withdrawal [41]. The homophily literature suggests that people tend to interact with others who are similar to them on attributes like age, race, ethnicity, etc. Ties between people will dissolve if there is too much dissimilarity [31], while groups that are more homogeneous, in terms of age and tenure, have fewer members leaving than heterogeneous groups [33]. We expect that members whose tenure differs more from the rest of the group are more likely to withdraw.

**Past productivity.** In traditional organizations, past performance has been shown to be a reliable predictor of future performance. We therefore expect a strong association between an individual’s past and future productivity.

The effects of past productivity on withdrawal are more complicated. In traditional organizations, poor performers are generally more likely to leave than good performers, which implies that withdrawal may not be detrimental for the organization [29]. Although some past research suggests that high performers can find alternate opportunities and therefore be enticed to leave the organization, more recent studies usually show that poor performers are more prone to voluntary turnover. In addition, these members more often consider other factors, such as job satisfaction, in evaluating whether to remain in the organization [37].

We expect the effects of past productivity on withdrawal to be even more complicated in online volunteer groups. There are reasons to expect poor performers to voluntarily leave the project due to lack of contribution. However, there are also reasons to expect good performers to voluntarily leave, either because they have contributed what they know and accomplished their mission or because they are burnt out from maintaining a high level of contribution.

**Concurrent projects.** Being involved in multiple groups within a community affects member productivity and withdrawal. Similar to traditional organizations that compete for limited resources [5], online groups that are created on the same platform or have similar functions also compete for scarce resources like members’ time and effort. Online volunteer groups need dedicated members, who put in the most time and effort, in order to survive as a group [10]. However, from a resource-based view, the amount of time spent on one activity is time that cannot be spent on another [6]. Those who are involved in multiple subgroups need to decide how to allocate their time to each group, making them physically unable to spend as much time on any one group. We thus expect members who are involved in multiple projects in Wikipedia to do less work for each project. Being involved in multiple projects increases the demand for members’ time and effort and, thus, may increase the likelihood of withdrawing from the projects.

An alternative view on the effects of multiple subgroups or projects is job or social embeddedness theory [32]. Members who are affiliated with multiple projects are more socially embedded within Wikipedia, which should reduce their likelihood of withdrawing from Wikipedia as a whole, and may reduce their likelihood of withdrawing from the individual projects as well.

**Communication and social integration.** Social identity has been shown to be positively correlated with performance measures [35]. Members who strongly identify with a group are willing to exert greater effort and make more contributions to the group than those who do not identify with the group. This leads to greater social integration of members into the group. As members become more active within a group, they are more likely to be productive and display better performance [33]. Interpersonal relationships or psychological contracts are especially important to increase participation and reduce withdrawal intentions in not-for-profit volunteer organizations [18].

Communication...
with others is an important way for members of online volunteer groups to be socially integrated and feel like an essential part of the group [21].

Communication may occur internally within subgroups or externally across subgroups within a large community. Internal and external communication may have different effects on individual behavior. Social integration and frequent communication with members of one’s group is likely to focus members’ attention on the group goals and needs, and thus increase their contributions to the group and reduce their likelihood of withdrawal from the group [33].

In contrast, when members have many ties or frequent communication with those outside of a group, they are more likely to be pulled away from the focal group [30]. External connections have been shown to affect performance and withdrawal differently in traditional organizations. On one hand, an individual’s connections with external groups have been shown to improve performance due to access to novel and relevant information [16]. On the other hand, social networks research has shown that employees or community members who are on the outskirts of groups and have strong external connections are more likely to leave [30]. Increased communication and socialization outside of a project is therefore likely to increase both individual productivity and withdrawal.

Trade-Offs Between the Community and its Subgroups

Online groups often consist of subgroups, such as projects within Wikipedia. Membership in multiple subgroups, on one hand, creates tension around how an individual allocates time and effort to different subgroups; on the other hand, this strengthens connections between an individual and the larger entity by enriching the web of connections tying the member to the community. It is important to understand more precisely how this trade-off works and how individual commitment and contribution to subgroups might transfer to the community as a whole.

Two mechanisms may be at work in determining how an individual’s involvement in multiple subgroups affects the individual’s involvement with the larger community. The first mechanism is the competition argument. As mentioned earlier, members participating in multiple subgroups must dole out their time and effort to each subgroup [6]. Because members have limited time to spend on volunteer work overall, splitting their efforts is likely to cause decreased productivity and increased withdrawal for individual subgroups [17]. Reduced productivity and increased withdrawal for individual projects does not necessarily mean reduced productivity and increased withdrawal for the community. Reduced levels of contribution to multiple projects, when aggregated, may exceed the average level of contribution to a single project. Similarly, for members who are involved in multiple projects, leaving one project doesn’t conclude the person’s affiliation with Wikipedia.

The second mechanism is job or social embeddedness. Job embeddedness theory suggests that more investment in a job increases the quitting cost, which is negatively correlated with the likelihood of leaving the organization [32]. Someone who is highly embedded (or involved) in an organization would be less likely to withdraw. In the context of Wikipedia, involvement in multiple projects increases the extent to which individual editors are socially embedded within Wikipedia, which makes them less likely to leave the Wikipedia community as a whole.

Furthermore, affiliation with multiple projects can lead to effective knowledge transfer across the projects, and thus increase the amount of work done within each project [16]. In these instances, members increase their ability to draw from outside sources and gain new knowledge that is useful to their productivity within the group. Furthermore, in the online context, member participation is socially driven by perception of and interactions with other members [11, 18]. The social embeddedness perspective implies that involvement in multiple projects may increase member contribution to Wikipedia as a whole.

To explore these trade-offs, our research questions are:

R1: How do individual tenure, past productivity, project involvement, and communication affect productivity and withdrawal at the project level?

R2: How do individual tenure, past productivity, project involvement, and communication affect productivity and withdrawal at the Wikipedia level?

R3: Are there any trade-offs between productivity and withdrawal? Are there any trade-offs between these outcomes at the project level and the Wikipedia level?

WIKIPEDIA AND WIKIPROJECTS

In this section, we describe our research setting – Wikipedia and the WikiProjects subgroups. Wikipedia is a free online encyclopedia that anyone can edit. An edit is simply a revision, large or small, to any article, talk, or user page. If users perform edits, they may also choose whether or not to register for a user account. Those who have accounts may then interact with other registered users, or editors, to coordinate work and discuss articles. Each article page has an associated talk page enabling editors to collaborate while working on the article.

A WikiProject is defined as “a group of pages in the ‘Wikipedia’ article namespace which are devoted to the management of a specific topic or family of topics within Wikipedia; and, simultaneously, a group of editors who use those pages to collaborate on encyclopedic work” [31]. Since 2002, more than 20,000 Wikipedia editors have joined more than a thousand projects. WikiProjects provide a way to organize editors with the goal of improving a specific subset of articles in Wikipedia. Members may choose to join or leave a project by adding or removing their names on the project’s member list. The main page of a WikiProject typically includes a brief description of the project and its scope, a list of project members, guidelines, and tasks that require member contribution.

In this study, we focus on editors who have contributed to Wikipedia as a part of one or more WikiProjects. We chose to study WikiProjects because they have clearly defined goals and boundaries, which make it easy to assess individual productivity and withdrawal at the project level.

**DATA AND VARIABLES**

The dataset we use in this study is extracted from the January 2008 dump of the English Wikipedia, which includes the full text of all pages and their complete edit histories from the creation of Wikipedia to the end of 2007. To gather information about projects and their members, we traversed the main directory page of WikiProjects and included all projects that are topical (thus excluding projects such as WikiProject Citation Cleanup). We also excluded projects that never grew to have at least three members (the minimum size of a group), projects that do not have a member list to track membership, and projects whose scopes could not be estimated using categories. Our final data set has 648 WikiProjects and 14,464 individual editors who are or have been members of these projects.

We determined each WikiProject’s membership and scope following the approach in Chen et al. [12]. We used historical edits of a project’s member list to identify members of each WikiProject. We considered an editor to have joined a project when her username appeared on the member list and to have left when her username was removed or contribution stopped. To determine the scope of a WikiProject, we first found the Wikipedia category that matched the title (like category Computer science for WikiProject Computer science). We then traversed all subcategories of the matched category down to the 4th level and considered all articles in those categories to be within the scope of the WikiProject (see Chen et al. [12] for more detail).

We constructed a longitudinal dataset to temporally separate the independent and dependent variables. Each observation records the characteristics and activities of an individual editor as a member of a project for each quarter in that project’s lifespan. Within each project, each quarter is a 90-day period in a project’s lifespan, with the first quarter beginning immediately after its creation date. Within each quarter, every editor who was a member of the project in that quarter was measured once for that project. The level of the analysis is therefore project individual quarter, with quarters nested within individuals and individuals nested within projects. For any given WikiProject, the first quarter is the first 90-day period following its date of creation. If the project had 10 members in this quarter, the dataset would have 10 observations for the project during this quarter, with each observation measuring the activity of one editor within the project. In total we had 85,105 project individual quarters in the WikiProject dataset.

**Dependent Variables**

**Project-Level Productivity**: We measured an editor’s project-level productivity as the number of edits performed by the editor on articles within the scope of the WikiProject during the current quarter.

**Project-Level Withdrawal**: We measured an editor’s project-level withdrawal as a binary variable, i.e. either 1 or 0. The variable is 1 if and only if the editor was an active member of the project in the current quarter, but removed her username from the project member list or stopped contributing within the scope of the project by the end of the next quarter.² We considered a member to be active if the person had at least one edit during that quarter to any of the following: an article within the project scope, the talk page of such an article, any project organization page, or the user pages or user talk pages of another project member.

**Wikipedia-Level Productivity**: We measured an editor’s Wikipedia-level productivity as the number of edits performed by the editor on any and all Wikipedia articles during the current quarter.

**Wikipedia-Level Withdrawal**: We measured an editor’s Wikipedia-level withdrawal as a binary variable, i.e. either 1 or 0. The variable is 1 if and only if the editor had performed at least one edit in Wikipedia during the current quarter but made no edits in Wikipedia in the subsequent quarter.

**Independent Variables**

**Tenure**: We measured an editor’s tenure by how long the editor had been a member of Wikipedia, that is, the number of days elapsed from a member’s first edit in Wikipedia to the end of a quarter [12].

**Tenure Dissimilarity**: We measured an editor’s tenure dissimilarity from the rest of the project members using Euclidian Distance as follows [41]:

\[
\sqrt{\sum_{j=1}^{n} (S_i - S_j)^2 / n}
\]

where \( S_i \) is tenure for the editor in question, \( S_j \) is the tenure of the \( j \)-th member in the project, and \( n \) is the total number of members currently in the project.

**Past Productivity**: We measured an editor’s past productivity as the total number of edits performed by the editor on articles within the scope of the WikiProject before the current quarter.

**Concurrent Projects**: We measured an editor’s concurrent projects by the total number of projects of which the editor is currently listed as a member. A higher number means that the editor is involved in more projects at the same time.

**In-Project Communication**: We measured an editor’s in-project communication by the number of edits that other members of the project have made to the editor’s user page and user talk page. We only counted edits by others on the

² 11% of editors returned one or more times after withdrawing, with a majority returning after one inactive period. Excluding these editors from the analysis did not significantly affect our results. In the analysis below we consider only the last and final instance of withdrawal for each editor.
editor’s pages, not the editor’s own edits, for two reasons. First, it is easier to track one user page than two. Second, strong ties tend to be reciprocal, so the amount of inbound communication to an editor should be a good measure of how socially integrated the editor is within Wikipedia.

**Out-Project Communication:** We measured an editor’s out-project communication by the number of edits that non-members of the project have made to the editor’s user page and user talk page.

**Control Variables**

**Quarter Index:** The index of time within the project was measured in quarters (90-day periods), starting with quarter 0 from the moment the project is created until the last full quarter before the end of 2007.

**Project Scope:** Measured as the number of articles falling under the project scope. Project scope was determined using the same approach as Chen et al. [12].

**Project Size:** Measured as the number of project members during the current quarter.

**RESULTS**

Descriptive statistics and correlations of the variables in the dataset are displayed in Table 1. Most of the variables are highly skewed to the right so we performed base-2 logarithmic transformations for normality considerations. Because our data is nested by nature, we analyzed the data using Hierarchical Linear Models (HLM) [8], with member productivity and withdrawal as the dependent variables and project and individual characteristics as the independent variables. HLM is an advanced form of linear regression that allows us to examine the effects of independent variables on dependent variables while taking into account potential correlations across observations that are nested within a high-level entity (e.g., individuals nested within projects). Our dataset is cross-nested between projects and individuals, meaning an editor can belong to multiple projects. Thus, we ran the analysis using the lmer function in R.

Our main results are displayed in Table 2. The first two columns show how project and individual characteristics relate to individual productivity and withdrawal at the project level, and the second two columns show the relationships at the Wikipedia level. We standardized all independent variables for ease of comparing coefficients across variables. We estimated our HLM models using maximum likelihood estimation, random intercepts, and unstructured covariance structure. For each dependent

---

**Table 1. Descriptive Statistics and Correlations of Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Project-Level Productivity</th>
<th>Project-Level Withdrawal</th>
<th>Wikipedia-Level Productivity</th>
<th>Wikipedia-Level Withdrawal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.151**</td>
<td>-2.981**</td>
<td>5.343**</td>
<td>-5.495**</td>
</tr>
<tr>
<td>Quarter</td>
<td>-0.191**</td>
<td>0.473**</td>
<td>-0.567**</td>
<td>1.236**</td>
</tr>
<tr>
<td>Project Scope</td>
<td>0.464**</td>
<td>-0.111</td>
<td>0.127**</td>
<td>-0.446**</td>
</tr>
<tr>
<td>Project Size</td>
<td>-0.216**</td>
<td>0.104**</td>
<td>0.254**</td>
<td>-0.122</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.165</td>
<td>-0.129**</td>
<td>0.024**</td>
<td>-0.073**</td>
</tr>
<tr>
<td>Tenure Dissimilarity</td>
<td>0.085**</td>
<td>0.108**</td>
<td>0.029**</td>
<td>0.079**</td>
</tr>
<tr>
<td>Past Productivity</td>
<td>0.596**</td>
<td>0.362**</td>
<td>0.113**</td>
<td>0.133**</td>
</tr>
<tr>
<td>Concurrent Projects</td>
<td>-0.287**</td>
<td>-0.061**</td>
<td>-0.116**</td>
<td>0.034**</td>
</tr>
<tr>
<td>In-Project Communication</td>
<td>0.611**</td>
<td>-0.042**</td>
<td>0.290**</td>
<td>-0.047**</td>
</tr>
<tr>
<td>Out-Project Communication</td>
<td>1.006**</td>
<td>0.350**</td>
<td>2.390**</td>
<td>0.188**</td>
</tr>
<tr>
<td>BIC</td>
<td>317209</td>
<td>56144</td>
<td>347630</td>
<td>53816</td>
</tr>
<tr>
<td>Deviance</td>
<td>35867**</td>
<td>2561**</td>
<td>57519**</td>
<td>1185**</td>
</tr>
<tr>
<td>N</td>
<td>N = 85105</td>
<td>N = 65393</td>
<td>N = 85105</td>
<td>N = 65393</td>
</tr>
</tbody>
</table>

**Table 2. Predicting Member Productivity and Withdrawal Behaviors**

We use the following notation in tables to represent p-values: ** p < .01, * p < .05, + p < .1
variable, we ran two models: the base model with the intercept and quarter as predictors, and the complete model with two project level variables and six member level variables. Due to the large size of our data set, we examined Bayesian Information Criterion (BIC) across the models. BIC punishes models with a large sample size and a large number of parameters [9]. We included BIC and deviance between the complete model and the base model in Table 2. In all analyses, the deviance is greater than 10, meaning the complete model fits the data better than the base model.

The first two columns of Table 2, respectively, predict member productivity, measured as the number of edits performed by an individual member within a project in the current quarter, and member withdrawal, measured as either removing one’s username from a project or stopping contribution to the project in the next quarter. Our analysis revealed a negative relationship between tenure and both productivity and withdrawal ($p < 0.01$). Compared to newcomers, old-timers contributed fewer edits and were less likely to withdraw. Our analysis also revealed a positive relationship of both tenure dissimilarity and past productivity with productivity and withdrawal in the current quarter ($p < 0.01$). Members with tenure different from the group contributed more edits than those with similar tenure and were more likely to withdraw from the project. Members who had made more edits in the past continued contributing more but were also more likely to withdraw from the project.

Our analysis revealed a negative relationship of the number of concurrent projects with productivity and withdrawal ($p < 0.01$). Members who took on more projects contributed fewer edits than those with fewer projects. Contrary to our predictions, members involved in multiple projects were less likely to withdraw from any individual project. We also found a positive effect of communication both within and outside the focal project ($p < 0.01$). Members who engaged in communication, either with other project members or editors outside of the project, contributed more edits than those who engaged in less communication. Communication within and outside of the focal project had opposite relationships with withdrawal from the project, with more internal communication being associated with less withdrawal and more external communication with more withdrawal. These relationships are consistent with our predictions.

The last two columns of Table 2 show how project and member characteristics relate to productivity and withdrawal at the Wikipedia level, i.e. the total number of edits within Wikipedia during the quarter and whether an editor stopped contributing to Wikipedia after the next quarter. Most of the effects on productivity at the Wikipedia level are similar to those at the project level, with tenure being the only exception. While editors with longer tenure contributed fewer edits to individual projects, they contributed more edits to Wikipedia as a whole ($p < 0.1$). Contrary to our predictions, members who belong to multiple projects contributed less work, both to an individual project and to Wikipedia as a whole. The standard coefficient for project-level productivity is greater than the one for Wikipedia-level productivity, implying greater negative impact on the former.

The results on withdrawal at the Wikipedia level versus the project level are similar as well, with three exceptions. First, while project scope had no significant effect on project-level withdrawal, larger scope was associated with reduced withdrawal from Wikipedia as a whole. Second, members of projects with more total editors were more likely to withdraw from the project but not significantly more likely to withdraw from Wikipedia. Third, while belonging to multiple projects was associated with a lower likelihood of withdrawing from any single project, it was associated with a higher likelihood of withdrawing from Wikipedia as a whole ($p < 0.01$).

DISCUSSION

We set out to understand two critical trade-offs in online volunteer groups like Wikipedia: the trade-off between productivity and withdrawal and the trade-off between subgroups within a community and the community as a whole. Our quantitative analysis of WikiProjects provides evidence for both types of trade-offs. Table 3 summarizes some of the critical trade-offs revealed in our analysis.

<table>
<thead>
<tr>
<th>Tradeoffs between productivity and retention</th>
<th>Tradeoffs between projects and Wikipedia as a whole</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Characteristics</strong></td>
<td><strong>Mixed tradeoffs</strong></td>
</tr>
<tr>
<td>(+) project scope  → (+) productivity</td>
<td>(+) project scope → No effects on project retention</td>
</tr>
<tr>
<td>(-) project size  → (+) productivity</td>
<td>(+) project size → (-) project productivity</td>
</tr>
<tr>
<td><strong>Member Attributes</strong></td>
<td><strong>No tradeoffs</strong></td>
</tr>
<tr>
<td>(-) tenure  → (+) productivity</td>
<td>(+) tenure → (-) project productivity</td>
</tr>
<tr>
<td>(+) past productivity → (-) retention</td>
<td>(+) Wikipedia productivity</td>
</tr>
<tr>
<td><strong>Member Connections</strong></td>
<td><strong>Mixed tradeoffs</strong></td>
</tr>
<tr>
<td>(+) concurrent projects → (-) productivity</td>
<td>(+) concurrent projects → (-) project productivity</td>
</tr>
<tr>
<td>(-) out-project communication → (+) retention</td>
<td>(-) Wikipedia productivity</td>
</tr>
<tr>
<td>(+) in-project communication → (+) retention</td>
<td>(-) Wikipedia retention</td>
</tr>
</tbody>
</table>

Table 3. Summary of Main Findings
For the ease of illustrating the trade-off, we summarize our results as the impact on productivity and retention, where retention is the opposite of withdrawal. A trade-off exists when a factor increases one outcome while decreasing the other or increases the same outcome at one level while decreasing it at another level. The left-hand side of Table 3 refers to trade-offs between productivity and withdrawal, while the right-hand side refers to trade-offs between projects and Wikipedia as a whole. For instance, the second row in the left column shows some major trade-offs between productivity and withdrawal when tenure and past productivity are the independent predictors. The display of (−) tenure → (+) productivity and (−) retention means that shorter tenure is associated with higher productivity and lower retention.

In the rest of the section, we discuss these trade-offs, speculate on their underlying processes, and highlight implications for managing online collaboration.

**Trade-off between Productivity and Retention**

Our first main finding is a critical trade-off between how many edits an individual editor contributes to a project and her likelihood of staying with the project. A typical factor with such opposite effects is an editor’s past edits, which is associated with increased productivity but a reduced retention rate. Contrary to traditional organizations where poor performers are more likely to leave an organization [29], we find that good performers are more likely to leave or stop contributing in online volunteer groups. One possible reason is the “mission accomplished” effect, as illustrated in the following quote:

"Having done all I can on the Andorra rugby and womens sevens pages (aside from keep them up to date), I am going to see if I can help with the Shannara project."

Another possible reason is the burnout effect, as illustrated in the following quotes from conversations between active editors who have experienced “wikiburnout” or “wikistress”:

"I am suffering from wikiburnout and chronic Wikistress. [...] I probably won't be able to log in as frequently and contribute as much as I would like."

"On another point: I noticed your wikistress level is high, and your contributions may be dropping. [...] I'd like to add something else; avoid burnout. You are a very active contributor. It is easy for highly active contributors to get caught up in burnout."

Stress management seems to be a major issue among active contributors. For dealing with stress, the meta-wiki of Wikimedia (the mother organization of Wikipedia) lists 154 tips contributed by editors. The meta-wiki also suggests that stressed people leave Wikipedia for a short while so they can recover. Many active Wikipedia editors created “wikistress meters” on their user pages to indicate their stress levels to fellow editors. The use of these stress meters may help alleviate the burnout effect of productive editors by increasing awareness of stress levels among editors. However, the effectiveness of such a solution remains limited due to the effort and skills required to create and update the meters. In addition, some research has shown that focusing on the negative aspect of stress may lead to unintended effects such as depression [27].

A design opportunity lies in improving the ease of use and functionality of stress awareness tools like stress meters. For instance, software agents can be developed to automatically estimate stress levels from an editor’s recent activity and comparison with historical patterns. Stress information can be customized or hidden from the focal editor to avoid negative impact. Instead, it can be made available to other editors and project leaders who can use the information to proactively manage the stress of project members and avoid overburdening already-active members.

Tools like these can be promising for alleviating the stress of productive editors, thus helping to maintain high productivity while improving member retention.

**Trade-off between WikiProjects and Wikipedia**

Our second main finding is a critical trade-off between members’ continued contributions to individual projects versus Wikipedia as a whole. Membership in multiple projects reduces one’s likelihood of leaving an individual project but increases the likelihood of leaving Wikipedia. In contrast, membership in multiple projects reduces the amount of work editors contribute at both levels. The negative effects of multiple project membership on productivity have wide implications because about 45% of the editors in our data were involved in more than one project, with 10% being involved in five or more projects. Our results suggest the detrimental effect is not limited to individual projects. Instead, it may have spilled over to affect other types of work an editor does for Wikipedia. Anecdotal evidence we found on talk page conversations further highlights the challenge for active editors to take on and juggle too many projects and the risk for project leaders to over-drift from the same pool of active editors.

"Goodness[sic]: I should have abolished this article last month but got too many projects on my plate and forgot."

"I'm working on too many projects atm. I'm going to be moving slowly here."

"Ditto. PS. Considered joining us in this fine wikiproject? :)"

"Thanks. I can definitely occasionally lend a hand here and there, but I already am involved in too many projects for the limited time budget I am on."

The real trade-off lies between withdrawal behaviors at the project level and those at the Wikipedia level. Being involved in multiple projects showed decreased likelihood of leaving a specific individual project, possibly because multiple projects give members more opportunities to remain involved. At the same time, being involved in too
many projects may increase their stress level, which could eventually cause them to abandon the site altogether. The challenge is how to leverage the benefits of multiple project membership while minimizing its negative impact. Resolving this challenge requires both change to Wikipedia’s policies and guidelines and the development of software tools to improve awareness of editors’ activity and commitment across projects. Tools can be developed to share information across projects, such as how many projects an editor has joined, how many edits an editor has made for each of these projects in recent months, etc. Wikistress meters may also be an indication of instances in which editors have been overtaxed by too many projects. We expect such information to provide insights into an editor’s workload that may help coordination efforts to avoid competition among projects for member attention. Tools can also be developed to use a combination of signs such as involvement in a large number of projects and sudden or significant drops in recent editing behaviors to generate alerts for editors or project leaders.

**Does old wisdom on individual productivity and withdrawal apply to online volunteer groups?**

At a high level, many of the insights from social and organizational theories still apply to the online context. Both individual productivity and withdrawal are affected by factors such as group characteristics, individual attributes like tenure and past performance, and social connections within and across projects. We were able to replicate many relationships between variables in traditional organizations, such as the negative relation between tenure and withdrawal, the positive relation between internal and external communication and productivity, and the positive relation between external communication and withdrawal. However, there are subtle differences between our findings and what we expected from reviewing old theories.

One example is the negative relationship between tenure and productivity and the positive relationship between past productivity and withdrawal. Following the organization science literature, we expected old-timers to be more productive and members with low levels of productivity to be more likely to withdraw. Yet we found the opposite of these effects. The discrepancy can be attributed to the informal and voluntary nature of online groups compared to formalization and bureaucracy in traditional organizations. Members of online groups don’t have fixed roles (except those who become administrators) and self-select to take on certain tasks and responsibilities. The goal of many members may be to find or share information and contribute to a good cause, rather than sticking around and climbing corporate ladders. Therefore, members who have contributed much of their knowledge may either feel a sense of “mission accomplished” or become burnt out and leave or stop contributing.

Another discrepancy between our findings and the organizational literature is the effects of the number of concurrent projects on productivity and withdrawal. The variable itself is somewhat unique to the online context. Although employees of traditional organizations may work on multiple projects, they are limited in how many they can join, either by their billable hours or managerial oversight. The voluntary nature of online collaboration allows members to join as many projects as they wish, up to or even beyond what their time and effort allow. With greater control and less individual autonomy in choosing projects to join, we expect some differences in the effects of multiple project membership on individual behaviors. For instance, being involved in multiple projects does not necessarily increase one’s likelihood of withdrawal. Instead, individuals may develop social networks with different units within the organization, reducing their likelihood of withdrawing from these units though they may not have strong bonds with the organization itself.

Comparison with the organization science literature reveals a limitation in our measure of productivity. It does not consider the type of work or quality of work. This may help explain the interesting dilemma we found with old-timers – they stay longer with a project but do not contribute as much as newcomers. Because our measure of productivity only considers the quantity of contribution, it is possible that, with more experience, old-timers shift their focus to administrative work or more challenging tasks, which is not reflected in a simple edit count. Examining type and quality of work would be fruitful for future research.

Overall, our study confirmed the applicability of social science theories to at least one online group, while also highlighting the importance of reconsidering and modifying the assumptions and propositions to fit the online context.

**ACKNOWLEDGMENTS**

Thanks to the Wikimedia Foundation for providing the data. Thanks to Robert Kraut, members of GroupLens Research Lab at the Department of Computer Science and Engineering, and the OB Research Group at the Carlson School of Management for their feedback and support. This work is supported by National Science Foundation Grant IIS-0808692, IIS 0729344 and IIS 0534939.

**REFERENCES**