A Contingency View of Transferring and Adapting Best Practices within Online Communities

Haiyi Zhu
GroupLens
University of Minnesota
200 Union Street SE
Minneapolis, MN, USA
haiyi@cs.umn.edu

Robert E. Kraut, Aniket Kittur
Human Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA, USA
robert.kraut@cmu.edu; nkittur@cs.cmu.edu

ABSTRACT
Online communities, much like companies in the business world, often need to transfer “best practices” internally from one unit to another to improve their performance. Organizational scholars disagree about how much a recipient unit should modify a best practice when incorporating it. Some evidence indicates that modifying a practice that has been successful in one environment will introduce problems, undercut its effectiveness and harm the performance of the recipient unit. Other evidence, though, suggests that recipients need to adapt the practice to fit their local environment. The current research introduces a contingency perspective on practice transfer, holding that the value of modifications depends on when they are introduced and who introduces them. Empirical research on the transfer of a quality-improvement practice between projects within Wikipedia shows that modifications are more helpful if they are introduced after the receiving project has had experience with the imported practice. Furthermore, modifications are more effective if they are introduced by members who have experience in a variety of other projects.

Author Keywords
Best practice adaptation, Practice modification, Contingency view, Propensity Score Matching (PSM).

ACM Classification Keywords
H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces, Web-based interaction; K.4.3 [Computers and Society]: Organizational Impacts.

INTRODUCTION
Online communities, like companies in the business world, often need to transfer best practices internally from one unit to another to improve their performance. For example, communities in the Stack Exchange network of question and answer websites use a common reputation system modeled on Stack Overflow’s original one. Similarly, many non-English language Wikipedia versions have borrowed policies and procedures originally developed in the English Wikipedia. Barnstars, the badges Wikipedia editors give to each other to reward meritorious work and motive each there, originated in the MeatballWiki and were imported into Wikipedia in 2003 [47]. Since then Wikipedia has developed over 100 distinct Barnstars and thousands of Wikiprojects have created their own specialized Barnstars. Similar tales could be told of Wikipedia’s various quality improvement programs, such as Collaborations of the Week (CotW), a practice designed to increase the quality of under-developed content areas that has diffused across hundreds of Wikiprojects [46,51].

While the effectiveness of particular practices has been studied in isolation [10,27,33,46,51], we are aware of no research that examines how the process of acquiring and changing these practices influences their effectiveness. Understanding the factors that determine how practices are internally transferred and effectively adapted could provide insights into community success that go beyond individual practices. This is also one of the central topics in the field of organization research in the last two decades [1,30,41].

As organization scholar Szulanski noted, “Identification and transfer of best practices is emerging as one of the most important and widespread management issues” [43].

One important question regarding best practice transfer within organizations is the extent to which recipients need to modify an original practice to make it effective in a local context [49]. Organization scholars have a long-standing debating on this topic. According to the re-creation perspective, strict replication leads to incompatibility between the new practice and the recipient’s environment, rendering the imported practice less effective [1,11,36,37,25]. The recipient units need to continuously modify the original practice and create their own practice that better fits with their culture, structure and approach. For example, according to this approach, McDonalds, which sells billions of beef-based burgers in the US, needed to change its menu by introducing localized products like...
McVeggie™ to appeal in India, where half of the population is vegetarian [24].

In contrast, the replication perspective argues that modifying a successful practice for a new environment increases the risk that the modifications will harm performance (e.g., [1,16,34,40,48,49]). Some empirical evidence shows that in a large franchise organization changing a successful practice (by selling non-standard products) harms franchisees’ survival. A one-standard-deviation increase in revenue derived from non-standard products more than doubles a franchise unit’s hazard of failure [49, p. 678].

In this paper, we propose that in online communities neither replicating an original practice without modification nor freely implementing modification is a successful approach to transfer best practices. Instead, we propose a contingency perspective and hypothesize that modifications are most successful if they are introduced after the receiving unit has had experience with the imported practice. This allows for a form of iterative organizational design, in which a receiving site can tweak an imported practice based on experience. We also hypothesize that modifications will be more effective if they are introduced by people who are core members of the receiving unit and who participate in a variety of other communities. These are the people who likely to be knowledgeable about what their unit needs and about alternative practice tweaks used by others.

To test these hypotheses, we analyzed historical data about Collaborations of the Week (CotW) in Wikipedia. A Collaboration of the Week is quality-improvement practice in Wikiprojects, which organizes editors collaboratively to improve a designated article in a limited time period. Collaborations of the Week spread from project to project and are often modified before they are imported and then as they are used. We collected the history of CotW in 146 Wikiprojects and measured how different types of modifications influenced their success, in terms of the length of time the CotW continued to be used in a project, the amount of work they elicited from project members and the number of unique editors who contributed to them. The results generally supported the hypotheses.

**THEORY AND HYPOTHESES**

**The Best Practice Adaptation Dilemma**

Practice refers to an organization’s routine use of knowledge for conducting a particular function [43]. According to organization scholars, the ability to transfer best practices internally within a firm provides a competitive advantage [2] and is one reason they can be more effective than other institutional arrangements such as markets [5,28]. The benefits of transferring good practices between parts of a single organization have been documented in many different organization settings (see [2] for a review). For example, Darr et al. showed how pizza franchises benefited from learning from other franchise stores how to place pepperoni [12]. Similarly, Baum and Ingram [7] found that hotels within a single chain benefited from the experience of other hotels in their chain that were in the same environment.

An important question is the extent to which units within a larger organization benefit by modifying practices received from another parts of the organization to fit their local environments. On one hand, modifying a successful working practice increases the risk that the modifications will harm performance. However, on the other hand, strict replication might lead to incompatibility between the imported practice and recipient’s environment, reducing the benefit derived from the imported practice. In this section we review existing evidence on both the replication perspectives and re-creation perspectives of best practice transfer. Based on the prior research, we suggest a contingency perspective to understand best practice modifications and develop testable hypotheses about the conditions under which source practices should be modified and re-created in order to be more successful.

**Not to Modify: The Replication Approach**

Winter and Szulanski [48] claimed that knowledge transfer is maximally effective when only necessary value-creating facets of the knowledge are replicated, and no time or effort is devoted to the creation of addition features, which could harm performance. There is evidence showing that attempting to modify a successful working practice could be harmful, even when they initially seemed sensible, promising, or desirable. Work in population ecology has found negative survival effects of modifying core features of organizations in a variety of contexts, including voluntary social service organizations [40]; Finnish newspapers [1]; U.S. medical diagnostic imaging firms [34]; U.S. bicycle manufacturers [16]; and French, German, and British auto manufacturers [15]. Recent work on franchise provides empirical evidence supporting the replication perspective. There results showed that deviation from a franchisor template (i.e., a source practice) has negative consequence on the survival of franchise units within a large franchise organization [49]. According to the replication perspective, modification of a working practice introduces risks, and the risk increases when the practice is complex. Modification of complex practice can lead to unanticipated deleterious interaction effects that are causally ambiguous and difficult to interpret [49,31].

**Modify: The Re-creation Approach**

However, the problem of the replication approach is that practice might encounter incompatibility problems when moving from a source environment to the recipient one. According to Argote and Ingram [2], practice is often embedded in structural elements of an organization, such as its people and their skills, technical tools, or other routines and systems used by the organization, as well as in the networks formed between and among these elements. Failure of practice transfer thus often results from...
incompatibility with the new context. And the risk of failure caused by incompatibility increases when the practice is more complex [2,18].

In contrast to the replication approach that emphasizes accurate replication, the re-creation approach focuses on modifying and adapting the source practice in the recipient site to reduce incompatibility. The re-creation perspective on practice transfer is influenced by literature in organization innovation, technological adaptation and organization routine [11,17,25,43]. Kim and Nelson examined learning and innovation in newly industrializing economies and proposed that knowledge transfer is a dynamic learning process where organizations continually interact with customers and suppliers to innovate or creatively imitate. Wanda Orlikowski [36] explored the introduction of groupware into an organization to understand the changes in work practices and social interaction it facilitated. She found that people’s mental models and an organization’s structure and culture significantly influenced how technology is actually used. She further proposed that change is endemic to the practice of organizing and is enacted through the situated practices of organizational actors as they improvise, innovate, and adjust their work routines over time [37]. Feldman and Pentland [17] challenged the traditional understanding of organization routines as creating inertia in organizations. They argued that organization routines are a source of change that create on-going opportunities for variation, selection and retention of new practices. Synthesizing these perspectives, practice is seen as being continuously modified in the transfer process. Practice transfer is a dynamic learning process, involving the continuous modification, re-configuration and re-creation.

**Contingency view of best practice modification**

Prior research suggests that modifying best practice can ameliorate the incompatibility between a source practice and the local environment, but increases the risk of introducing deleterious features to a successful working practice. Both the risk of incompatibility and unanticipated deleterious modification increases when the practice is more complex.

We suggest that not all modifications are equally effective. Either strictly replicating an original practice without modification or freely implementing modifications is unlikely to optimize the utilization of the imported practice. Instead, we need to understand the conditions under which modifications are more or less effective. In the following sections, we develop testable hypotheses about when and who should make modifications in order to achieve optimal utilization of the imported practice. Specifically, we propose hypotheses about the effectiveness of modifications at an early stage (i.e., pre-implementation) versus later (i.e., post-implementation), and the influence of characteristics of the people involved in the modification on their success.

**When to modify: Effectiveness of Pre- versus Post-implementation Modification**

Tyre and Orlikowski's [45] examined the temporal pattern of modifications to a new technology in organizations. The authors found modifications disproportionately occurred when the technology was first introduced (and even before its official use). Thus, they suggested that there exists a relatively brief window of opportunity to explore and modify new technology. However, the authors only examined the temporal pattern of the modifications, not their effectiveness at different stages.

We propose that modifications at early stages are often based on people’s presumptions (i.e., predictions about which components of the new practice might go wrong) and therefore may be wrong because they are not based on evidence. In contrast, modifications after implementation are based on experiences with using the practice and can respond to actual compatibility problems between the imported practice and the receiving site. This allows for a form of iterative organizational design, in which a receiving site can tweak an imported practice based on experience. Therefore, we hypothesize that post-implementation modifications are less likely to introduce deleterious changes compared to pre-implementation modifications, and thus will be more effective than pre-implementation modifications.

The idea that experience-based, post-implementation modifications are effective is consistent with the organization learning and knowledge creation literature (see [4] for a recent review). According to organization learning theories, new knowledge is iteratively created as experience interacts with context. We propose to use an iterative organization design model to depict the post-implementation modification of source practice as an ongoing use-mismatch-create cycle. In this cycle, the recipient site adopts and implements the new practice, uses it, detects mismatch, fixes the mismatch, and creates a new iteration. Each iteration results in more effective utilization of the practice. The re-creation process does not end when the new practice achieves satisfactory results at the recipient site. Even after successfully implementing the new practice for a period of time, any change in the local context at the recipient site (e.g., environmental change, member turnover, introduction of new tools or policies) might result in a new mismatch and thus prompt a new iteration.

The process of post-implementation, organizational iterative design is analogous to the iterative user-interface design [35,39]. Nielson proposed that software improves more rapidly when users use the interface and developers learn from their feedback, rather than designing and iterating without evidence [35]. He provided data to show that redesigning user interfaces on the basis of user testing substantially improved usability [35].

This hypothesis might reconcile the difference between the replication and re-creation perspective discussed above.
Szulanski and Jensen [42] and Winter et al. [49] provided empirical evidence showing that deviation from the corporate templates negatively affect the survival chances of franchise units within a large organization. However, those studies only focused on the presumptive modification (i.e., ones based on managers’ non-evidence-based assumptions about what should work) [42] or conflated presumptive modifications and post-implementation modifications [49]. We suggest that modification made before implementation (presumptive modification) will generally not lead to successful use of the practice, while the post-implementation modifications should significantly improve its successful utilization.

**H1. Modifications made after implementing the practice are more effective than modifications made before implementation.**

**Who to modify: Effectiveness of Modifications Created by Different People**

The next hypothesis considers the individuals who are eligible to propose and implement new iterations in the recipient site. Specifically, we ask: which characteristics of people in the modification process affect successful modification?

First, we hypothesize that central members in the local site are more likely to create better modifications because these central people know more about the local environment. Central people are more likely to identify a mismatch between the new practice and local needs, and craft a good solution to fix the mismatch.

Second, we propose that members’ social network might also affect whether they will create successful post-implementation modifications. Prior research has examined how social network ties affect practice transfer. It is natural that external ties will benefit the search of available knowledge/practice and initial implementation of the new practice at the recipient site [20,21]. However, we propose that external ties will also benefit successful post-implementation modifications at the recipient site.

To support this view, we draw on the concept of “learning in a world of learners” from Levitt and March [32] and adopt an ecological view to understand the role of external ties in successful post-implementation modification. The key element of creating an effective modification is to resolve the mismatch between the local environment and the new practice in the new iteration. Note that each recipient site attempts to fix the mismatch of the source practice. It is possible that other recipient sites, especially those that are similar to the local site, have encountered and solved similar mismatch problems. Members with external ties with other sites that have also adopted the new practice can better search for solutions from other sites. Furthermore, according to work on analogical reasoning [44], even though mismatch problems are not identical in other recipient sites, exposure to the mismatch-fixing cycle in other recipient sites might inspire good solutions at the local site.

Although people who have external ties with other recipient sites are more likely to generate good solutions for mismatches at the local site, acceptance of their solutions cannot be taken for granted. Gruenfeld et al. [19] investigated the consequences of temporary membership changes for itinerant members (i.e., those who leave their group of origin temporarily to visit a foreign work group) and indigenous members of those origin and foreign groups. They found that, although itinerant members produced more unique ideas than indigenous members, their ideas were significantly less likely to be utilized by the group. Kane et al. [23] later found that groups were more likely to adopt the ideas from a rotator when they shared a superordinate social identity with that member than when they did not. Therefore, our final hypothesis is that people with external ties who are also central in the local units can generate good solutions that result in a higher acceptance rate. Those persons, therefore, are most likely to create more effective modifications.

**H2a. People who are central at the recipient units are more likely to create effective post-implementation modifications.**

**H2b. People who have external ties with other recipient units are more likely to create effective post-implementation modifications.**

**H2c. People who have external ties with other recipient units and are central in the focal unit are most likely to create effective post-implementation modifications.**

**STUDY PLATFORM**

We conduct our studies in the context of Wikiprojects (subgroups organized around different topics in Wikipedia). Particularly, we investigate a widely adopted project-based practice called Collaboration of the Week (CotW).

**Collaborations of the Week (CotW)**

CotW is a mechanism that designates one or two articles to be improved within a defined time period. Previously, CotW was a Wikipedia-wide activity that was not restricted to any specific project. Since 2004, hundreds of Wikiprojects have adopted this practice and created their own CotW, which often have dedicated project pages. Figure 1 shows the CotW project page in Wikiproject Video Games (WVG).

CotWs have two phases: selection and collaboration. In the selection phase, project members nominate candidates and then elect members to collaborate. During the collaboration phase, the project tags the chosen article(s) with a special template in its talk page. In addition, the project typically announces the targets of the collaboration on its project pages.

CotW is an important practice to direct volunteer editors’ attention to articles that are important to the group but which may not attract individual members’ interests. As
discussed in Zhu et al.’s study [51], editors may want to work on popular articles, and thus neglect less popular articles. CotW can effectively direct contributions to these less popular, but important, articles. Research also showed that, in addition to increasing contributions on important but less popular articles, CotWs have other benefits. For instance, the effects of CotWs carry over to non-CotW-target articles. Contributions on non-CotW-target articles also increased during the CotW period. Furthermore, editors exposed in CotW were more likely to perform similarly to their role models in the project and increased their contributions on assessment and anti-vandalism.

Despite the benefits of CotWs, their utilization in Wikiprojects varies widely. Among 146 Wikiprojects that adopt CotWs, 74 Wikiprojects have hosted more than a single collaboration, and 55 Wikiprojects successfully hosted more than five collaborations. The significant discrepancy in CotW utilization proves the need to further understand the process of transferring and adapting best practices in online communities.

CASE STUDY: COTW IN WVG
We conducted an in-depth case study on the Wikiproject Video Games (WVG)’s Collaboration of the Week, named “Gaming Collaboration of the Week” (GCOTW). The case study can help us better understand the hypotheses in the context of Wikipedia and CotW.

Method
We analyzed the complete revision history of GCOTW project page (3431 revisions) and discussions on WVG’s talk page that mentioned GCOTW. We also cross-linked key participants’ activities in GCOTW and other parts of Wikipedia during the given time period. Wikipedia records almost every single activity and provides data and API for researchers to conveniently retrieve and analyze the activities. We rely on the complete records to reconstruct WVG’s experience of using CotW.

Findings
On 3 Oct 2004, editor pie4all88 started a discussion thread on WVG’s talk page, and expressed an interest in developing a WVG-specific CotW similar to those of Wikipedia’s many other projects. After receiving supportive messages from two other members within 24 hours, pie4all88 created a CotW page on 4 Oct 2004 called “Gaming Collaboration of the Week” (GCOTW).

1. Illustrate the goal of CotW. For instance, this page says: “Each week a Gaming Collaboration of the week will be picked using this page”…“The aim of this project is to improve the quality of Wikipedia's computer and video game articles through widespread cooperative editing.” “The project is also used to fill gaps in Wikipedia, to give users a focus, and to give us all something to be proud of.”

2. Template designed to announce targets of the collaboration each week. The template shows “the current focus of collaboration of the week is XX. The last article was XX – see how it improved.”

3. Policies and guidelines about running the collaborations. The policy on this iteration includes five parts: how to vote, how to deal with vote ties, how to nominate a candidate, what to consider before nominations, and how to prune nominations that do not receive enough votes. For instance, the policy for voting says “Please vote for as many of the following candidates as you like. Please add only support votes. Opposing votes will not affect the result, as the winner is simply the one with the most support votes (see Approval voting). Remember: Any registered user is encouraged to vote.”

4. This is the area for editors to participate in the nomination and voting. They post the title (with a link) of the article they nominate and reasons why they want to nominate this article. Other users will support the nominations or leave comments about the nominations.

Figure 1. The page for the collaboration of the week in Wikiproject Video Game on Oct. 5th 2004.
Modifications of GCOTW

Table 1 shows five iterations of GCOTW as examples to illustrate what we mean by “modifications” in the context of CotW. The first example discusses the guidelines for nomination. The original guideline inherited from the source CotW simply reminded people to justify their chosen candidates. Editor pie4all88e had a concern that members of WVG might be enthusiastic about a particular niche topic yet not consider its importance for the whole gaming community. Therefore, in the new iteration, a new guideline was added by pie4all88 to remind nominators to consider the impact of their desired articles to the wider gaming community.

The second modification example considers the pruning policy, which defines the threshold to prune unsuccessful nominations (i.e., those that fail to receive adequate support). After implementing the original pruning policy for a while, users stated that the threshold of receiving votes in a week was too high. In the talk page, people proposed to lower the number of needed votes per week because “this CotW does not get as much traffic as the original CotW gets.” That change is reflected in the new iteration.

The third example relates to the voting policy. The original policy encouraged members to “vote for as many of the following candidates as you can.” That policy, however, allowed people to vote but not contribute. As such, articles selected as GCOTW targets received little contribution during the collaboration period. One member expressed this problem in the discussion and suggested that the weekly improvement drive (itself a variant of the source CotW) create a template to remind voters to contribute. As a result, two changes were made in the new iteration. First, the description was changed to “A vote … shows your commitment to support and aid in collaborating on that specific article if it is chosen.” This change highlighted the meaning of votes as a commitment to contribute as opposed to a simple social gestures. Second, a new template was created to remind voters when the articles they voted for were chosen.

The fourth example also concerns voting policy. The original policy stated that any registered user is encouraged to vote. To increase the likelihood that their preferences would be selected, some members created “sockpuppets” to cast false votes. In the new iteration, sockpuppets were forbidden from voting.

The final example relates to the selection mechanisms in GCOTW. After implementing GCOTW for over four years, member enthusiasm eroded. Low participation frustrated members who were still actively organizing the nomination and voting. To address the problem, the nominate-vote-select schema was changed to a bot-selecting schema. Each week, a bot would randomly select an article from the low-quality-high-importance category and post it as GCOTW.

In the discussion, people claimed that the goal of the change was to remove the stress caused by nomination and voting and focus on the contribution. Also, the random nature of the selection was more enjoyable. After implementing the new bot-selecting schema, GCOTW ran successfully for another 2.5 years.

Pre- and Post-implementation Modifications

The first example modification was made before the WVG officially implemented the GCOTW (i.e., the date of announcing the first GCOTW). The remaining four example modifications were made after the GCOTW was officially implemented. Prior to the official implementation, the modifications were created based on people’s predictions about which component might go wrong. For instance, in the first example, editor pie4all88e predicted that members of WVG might be enthusiastic about a niche topic without considering its importance for the whole gaming community. No discussion found related to the problem of proposing a niche topic. In other words, it was uncertain whether nominating niche topic articles would be problematic. In contrast, the remaining four examples were all based on lessons learned from previous iterations, such as the high pruning threshold, the lack of contributions despite the number of votes, false votes, and decreased enthusiasm. We found discussion histories related to each of these four examples. The post-implementation modifications are more targeted to actual problems compared to pre-implementation modifications.

People in the modification process

The third example about the voters not contributing shows how people with external ties can generate good solutions to resolve problems of using new practice at the local site by borrowing solutions. The editor (Jacoplane) mentioned that another project created a template that “gets put on every user’s talk page that vote”. The editor suggested borrowing this solution: “I think we should do something similar to remind people that they voted to remind people that they voted.” We checked Jacoplane’s editing history and found that this editor participated in nine other Wikiprojects that hosted CotWs that year. Despite the multiple project participation, the editor was based in WVG (87.7% of his/her project page contributions are devoted to WVG at that year). In WVG, the editor was a top 3 contributor among the group’s 347 members. The central role of this editor in WVG might make it easier for him/her to identify the problem. Second, the external relationship with other projects was an advantage for him/her to find a solution. Finally, the central role of this editor made it easier for his/her suggestions to be accepted.

The case study provides real examples to help better understand the hypotheses about modification of best practice in the context of CotW in Wikiprojects. In the following section, we conduct quantitative analysis to test the hypotheses.
<table>
<thead>
<tr>
<th>Old Iteration</th>
<th>Discussion</th>
<th>New Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guidelines for nominations - Giving reasons as to why an article should become the COTW may assist others in casting their vote.</td>
<td>No discussion found specifically related to this change.</td>
<td>Guidelines for nominations - Giving reasons why an article should become the GCOTW may convince others to support your nomination. - Can the wider gaming community easily contribute to the article? Or is it something only a small number of people will know about?</td>
</tr>
<tr>
<td>Pruning policy: Nominations will be moved to /Removed if they have not received 5 votes after 7 days on the list, 10 votes after 14 days, 15 votes after 21 days, and so on.</td>
<td>5 votes per week? &quot;I propose we lower the needed votes per week to 4 or even 3, as this CotW does not get as much traffic as the original CotW gets.&quot;</td>
<td>Pruning policy: Nominations will be moved to /Removed if they have not received 5 votes after 3 days on the list, 9 votes after 14 days, 12 votes after 21 days, and so on.</td>
</tr>
<tr>
<td>Voting policy: Please vote for as many of the following candidates as you like. Please add only support votes. Opposing votes will not affect the result, as the winner is simply the one with the most support votes</td>
<td>People voting but not contributing &quot;I've noticed that there seems to be a lot more people voting in the GCOTW lately, but the number of contributors hasn't really seemed to increase much. Is the idea that anyone can vote, or only people who intend to contribute? With the Weekly improvement drive, the Template:AIDvotes gets put on every user's talk page that voted. I think we should do something similar to remind people that they voted.&quot;</td>
<td>Voting policy: A vote or a show of support for an article shows your commitment to support and aid in collaborating on that specific article if it is chosen. Although you are not required to fulfill that commitment, we ask that you only support articles that you are able to contribute to so that this collaboration's goals of expanding and improving articles can adequately be achieved. Feel free to vote for as many of the following candidates as you like. Add template to remind voters:</td>
</tr>
<tr>
<td>Voting policy: Remember: Any registered user is encouraged to vote.</td>
<td>Fake votes &quot;It seems that someone is adding other people's signature to the nomination XXX&quot;</td>
<td>Voting policy: Any registered user is encouraged to vote so long as you abide by the policies of Wikipedia, especially Wikipedia:Sockpuppets.</td>
</tr>
<tr>
<td>The selection of collaboration article is based on nomination and voting.</td>
<td>GCOTW is big letdown this week &quot;This week's Wikipedia:Gaming Collaboration of the week was Prima Games. It's been rather a poor show.&quot; No longer working? &quot;So, is Gaming Collaboration of the week now nonfunctional? As is, no one working on it.&quot; Reactivating Collaboration of the Week - with ROBOTS!! (Propose the plan of having robots randomly select one article from the category of low quality but high importance as collaboration) &quot;Removing the stress of nomination and voting will reduce frustration, and make participation the focus, not bureaucracy (this isn't an RfA). The random nature will make it more fun, as part of it is wondering which article will be chosen. &quot;</td>
<td>Introduction: The WikiProject Video games collaboration is a collective effort to improve related articles covered by the project's scope. An article is chosen every Monday, by a bot that randomly selects one video game-related article that is rated Stub or Start or C class, and Top or High priority for WP:VG. The bot then updates Template:Collab-gaming with the pick, and the collaboration begins. If there is consensus that a selected article is not felt to be suitable for collaboration, then the bot will be requested to &quot;re-roll&quot; and select a different article. Articles that have previously been chosen for collaboration will not be chosen again. Previous collaborations can be found at /History.</td>
</tr>
</tbody>
</table>

Table 1. Example modifications in Wikiproject Video Games.
QUANTITATIVE ANALYSIS

Method
We ran a quantitative analysis on 146 Wikiprojects that adopted CotW. The first step is to identify the modifications of CotW in these projects.

Automatically identify modifications in CotW
We want to automatically identify modifications from the CotW pages’ historical revisions. Modifications are defined as the changes to the practice, which is modifying the way of operating and organizing CotW. Not all the historical revisions of CotW pages were “modifications”. The goal of this section is to automatically identify the modifications.

We found that a large proportion of the historical revisions on the CotW pages are actually candidate nominations or votes to select collaboration articles, rather than modifications to the CotW rules. To rule out these nomination and voting activities, we excluded the revisions that only modified the sections of nomination and voting. Results show that 88.6% of the revisions on the CotW pages are the nomination and revision activities.

To further detect the modifications in the remaining 11.4% revisions we used a machine-learning approach in which we hand-coded 335 non-nomination-voting revisions from two Wikiprojects’ CotWs as a training set. We then created a feature set containing nine different features (see Table 2 for details). We trained statistical models (rule-based model generated based on our domain knowledge, decision-tree, and SVM) on the training set and evaluated them using a separate set of hand-coded data (113 non-nomination-voting revisions from another two Wikiprojects). Details of the feature set and model shown in Table 2.

We compared the performance of rule-based model, decision-tree and SVM. Results are shown in Table 3. The rule-based model and decision tree outperformed SVM on both the training set and test set. On the training set, the decision-tree performed slightly better than the rule-based model. However, in the test set, the rule-based model performed slightly better than the decision-tree model. Because the rule-based model performed the best in the test set and is easy to interpret we used it in the following analysis.

Analysis overview
This analysis seeks to identify the effects of different types of modification on the successful utilization of CotWs. We measured the success of CotW according to three criteria: (1) the survival of CotW (i.e., the likelihood that projects continuously use CotW), (2) the number of contributions on CotW target articles during the collaboration period, and (3) the unique contributors to CotW target articles during the collaboration period. The analysis was conducted on the project-collaboration-period level. We predicted outcomes (i.e., survival, contribution, and participants) in the current collaboration period according to whether the project made a new iteration in the last collaboration period.

Since we used observational data to run the analysis, the creation of a new iteration is not a true experimental treatment. New iteration creation (i.e., modification on CotW), as with most events in the real world, is endogenous in the sense that it is caused by other factors inside the system. In our data, Wikiproject activity correlates to project members’ participation in CotW, as well as their likelihood to modify its procedures of CotW. Not controlling for confounding factors that influence both the treatment (CotW modifications) and the outcome (CotW utilization) can lead to biased estimates of the treatment effects. To ameliorate the endogeneity problem, we used propensity score matching (PSM). We will discuss the details of PSM method later.

Data preparation
The data were longitudinal, following the same project across multiple collaboration periods. The data comprised 1588 project-collaboration-period observations.

Dependent variables
• Practice Death. We defined a CotW as near abandonment (i.e., dying) if the project did not have at least two collaborations after a focal collaboration period (a sensitivity analysis with differing threshold values showed no difference in the pattern of results). This variable is assigned to 1 if the project’s CotW was dying (i.e., had no more than two collaborations in the future); it is assigned 0 if the project’s CotW was still active (had more than two collaborations in the future).

• Contributions. We measured the number of revisions to the target articles during the collaboration period, controlling for the number of revisions on these articles during the non-collaboration period. Particularly, we divided the number of revisions on the target articles during the collaboration period by the number of revisions on the target articles during the pre-

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Number of total inserted characters, Length of the longest inserted word sequence, Number of total deleted characters, Length of the longest deleted word sequence, Add templates, Add sections, Maintenance, Being reverted in the next revision, Revert previous revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>• Rule-based model generated based on our domain knowledge: 1) Must have the length of the longest inserted word sequence no less than five or add new sections or add new templates but excluding the Wikipedia’s maintenance templates; 2) Must not being reverted in the next revision or reverting previous revision.</td>
</tr>
<tr>
<td></td>
<td>• Decision-tree</td>
</tr>
<tr>
<td></td>
<td>• SVM</td>
</tr>
</tbody>
</table>

Table 2. Feature set and model to classify modifications

<table>
<thead>
<tr>
<th></th>
<th>Rule-based</th>
<th>Decision Tree</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>93%</td>
<td>93%</td>
<td>94%</td>
</tr>
<tr>
<td>Test</td>
<td>93%</td>
<td>92%</td>
<td>91%</td>
</tr>
</tbody>
</table>

Table 3. Performance of three models on training & test set.
collaboration period. The pre-collaboration and collaboration periods lasted the same length (e.g., normally a week to a month).

- **Participants.** We measured the number of unique contributors who edited the target articles during the collaboration period, controlling for the number of unique contributors during the non-collaboration period. Particularly, we divide the number of contributors during the collaboration period by the number of revisions during the pre-collaboration period. Both periods lasted the same length.

**Independent variables**

- **Post-implementation modification.** We measured the number of modifications the project’s CotW had in post-implementation periods.

- **Pre-implementation modification.** We measured the number of the modifications the project’s CotW had in the pre-implementation period (i.e., the preparation period).

We further divided the modifications according to which editors would implement the modifications.

- **Modification made by core members in the recipient project versus Modification made by non-core members in the recipient project.** We defined core members as those whose overall contributions to the project are among the top 10%. We then divided the modifications into two groups: those made by core members versus those made by non-core members.

- **Modification made by members with more external ties versus Modification made by members with fewer external ties.** We measured external ties as multiple memberships in other projects that also adopt CotWs. If a member participates in three projects in addition to the focal project, he/she has three external ties. We defined members with more external ties as those participated in more than the medium number (the medium number in the data is 3). Similarly, we define member with fewer external ties as those with ties less than medium number.

In addition to the above variables designed to measure the main effects of core-ness in the focal project and external ties, we also measured the interaction effects. We defined four more interaction measurements: (1) modifications made by core members in recipient project and have more external ties, (2) modifications made by core members in recipient project but do not have many external ties, (3) modifications made by non-core members in recipient project but have more external ties, and (4) modifications made non-core members in recipient project and do not have many external ties.

In addition, we measured the popularity of the source.

- **Popularity of the source.** In our data, we observed that Wikiprojects have different sources. Many of the earliest projects learned and copied rules and policies from the Wikipedia-level CotW (which has since been terminated). Some projects started by copying other Wikiprojects’ CotW. The very first revision of the CotW page is likely to be the source CotW. We calculated the popularity of the source by comparing the structural similarity of the given project’s first CotW page revision with all the other CotWs in other projects at that time period. Higher similarity indicated that more projects were using the same structure, and that focal project was starting with a more popular “branch”.

**Propensity score matching**

The basic idea of PSM is to pair the treated project and the control project. For a given project that had modifications, we selected a comparison project that was most similar on confounding variables but did not have modifications. We used Propensity score matching (PSM) to pair the projects (more precisely, project-collaboration-periods).

PSM involved three steps. In the first step, we estimated the propensity score (i.e., the probability of having modifications) from a set of conditioning variables. We chose four variables indicating the activity level of the project listed below as conditioning variables. In the second step, we matched each project that had modifications in a particular week with another project that did not have modifications, but which had the most similar propensity score based on four activity indicators. Propensity scores allow researchers to control for many variables simultaneously by matching on a single scalar variable. To conclude the second step, we tested whether the treatment group and control group were well matched in terms of the conditioning variables. In the third step, we ran fixed effects regression analyses to estimate the effect of modifications on the treated groups and matched controls.

**Step 1: Estimate propensity score**

We first used logistic regression to estimate the probability of having modifications based on the project activity level. The estimated probability is the propensity score. The four predictors are listed below.

- **Active members.** We measured the number of active members during the period of time.

- **Number of CotW hosted before.** We measured how many CotW were hosted. The logarithmic transformed number of CotW was added in the regression to represent the baseline hazard function in the survival analysis.

- **Project page activities.** Project pages are places where Wikiproject organize activities. CotW is one of activities organized through project pages. We measured the amount of contributions on the project pages during the given period, indicating whole project activity during the given period of time.

- **Number of project pages.** We measured the number of pages the project had during the given period, which indicates the size of the project.
Step 2: Matching based on propensity score.
In this step, we matched projects that modified their CotWs with projects that did not, based on the estimated propensity score. To do this, we ordered the treated and control projects according to their propensity scores. For each treated project, we then selected a control project with the closest propensity score within a maximum distance.

Figure 2 reports the histogram of the propensity score (i.e., the likelihood of making modifications) for treated groups and control groups before and after matching. Here the treated group contains projects that did not make modifications at the given time period and the control group contains projects that did make modifications at the given time period. Figure 2 shows that the treated group and control group are balanced on the likelihood of making modifications after matching.

<table>
<thead>
<tr>
<th>N of proj pages</th>
<th>Full mean</th>
<th>Control mean</th>
<th>Bias %</th>
<th>Reduced bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>83.3</td>
<td>89.0</td>
<td>-5.2</td>
<td>-59.9</td>
</tr>
</tbody>
</table>

Table 4. Comparison between treatment projects that made modifications (Treat) and control projects that did not make modifications (Control) before and after propensity score matching (Full vs. Match).

bias in % = 100(\bar{x}_T - \bar{x}_C)/\sqrt{(s_T^2 + s_C^2)/2}, where \bar{x}_T and \bar{x}_C are the sample means in the treated and control groups and s_T^2 and s_C^2 are the corresponding sample variance.

Figure 2. The distributions of propensity score for treated group (i.e., projects made modifications, indicated by blue solid lines) and control group (i.e., projects that did not make modifications, indicated by red dot lines) before matching (left) and after matching (right). This figure shows that after matching, the treatment group and control group has more similar distribution of propensity score.

Table 4 reports the details of the matching process. Note that variables that correlate highly with the treatment (also having higher risk to introduce bias) will be balanced better than variables with lower correlation with the treatment. This explains why PSM tends to favor page activities, active members and previous CotWs over the number of project pages during balancing.

There is an interesting observation that the bias (i.e., unbalance) between the treatment group and control group is not that serious even before matching. In Zhu et al.’s [51] study where they used PSM to match an editor who received messages with editors who did not receive messages, the bias was 79%-110% before matching. In this analysis, the bias is only 5%-8% before matching.

The statistical results are consistent with our observations. We observe that project activity and project size do not correlate with the number of modifications made on the CotWs (and the success of CotWs). For example, Wikiproject Military History is considered the largest and most active Wikiproject, with eight times as many active members and five times more project pages than Wikiproject Oregon. But Wikiproject Military history only made four modifications in total while Wikiproject Oregon made 77 modifications. Wikiproject Oregon hosted 89 CotWs while Military history only hosted 24 CotWs, although the latter project generally much more active than the former project.

Step 3: Run the analysis on the match sample
Using the matched sample, we then examined the effects of modifications on the outcomes (survival, contributions and participants). We used fixed effects linear regression to predict outcomes, with each treated control pair as a group.

Results
The temporal patterns of the modifications are shown in Figure 3. The results are consistent with Tyre and Orlikowski’s (1994) findings (Figure 4) that a substantial proportion (about 30%) of modifications happened in the pre-implementation stage. Far fewer modifications happened in each post-implementation CotW period.

Table 5 shows the main findings of the analysis on the effectiveness of the modifications. Models 1-4 test how modifications affect the survival of CotW in Wikiprojects. Each coefficient in Models 1-4 represents the hazard ratio. A hazard ratio is the ratio of the risk of a CotW being abandoned in a given time period associated with a one-unit change in the explanatory variables. A hazard ratio smaller than 1 indicates decreased rate of abandonedness (i.e.,

Table 4 reports the details of the matching process. Note that variables that correlate highly with the treatment (also having higher risk to introduce bias) will be balanced better than variables with lower correlation with the treatment. This explains why PSM tends to favor page activities, active members and previous CotWs over the number of project pages during balancing.

There is an interesting observation that the bias (i.e., unbalance) between the treatment group and control group is not that serious even before matching. In Zhu et al.’s [51] study where they used PSM to match an editor who received messages with editors who did not receive messages, the bias was 79%-110% before matching. In this analysis, the bias is only 5%-8% before matching.

The statistical results are consistent with our observations. We observe that project activity and project size do not correlate with the number of modifications made on the CotWs (and the success of CotWs). For example, Wikiproject Military History is considered the largest and most active Wikiproject, with eight times as many active members and five times more project pages than Wikiproject Oregon. But Wikiproject Military history only made four modifications in total while Wikiproject Oregon made 77 modifications. Wikiproject Oregon hosted 89 CotWs while Military history only hosted 24 CotWs, although the latter project generally much more active than the former project.

Step 3: Run the analysis on the match sample
Using the matched sample, we then examined the effects of modifications on the outcomes (survival, contributions and participants). We used fixed effects linear regression to predict outcomes, with each treated control pair as a group.

Results
The temporal patterns of the modifications are shown in Figure 3. The results are consistent with Tyre and Orlikowski’s (1994) findings (Figure 4) that a substantial proportion (about 30%) of modifications happened in the pre-implementation stage. Far fewer modifications happened in each post-implementation CotW period.

Table 5 shows the main findings of the analysis on the effectiveness of the modifications. Models 1-4 test how modifications affect the survival of CotW in Wikiprojects. Each coefficient in Models 1-4 represents the hazard ratio. A hazard ratio is the ratio of the risk of a CotW being abandoned in a given time period associated with a one-unit change in the explanatory variables. A hazard ratio smaller than 1 indicates decreased rate of abandonedness (i.e.,

Note that here we do not use the traditional interaction model (e.g., with modification, modification X pre-post, and modification X pre-post X the types of people as explanatory variables in the regression) but divide the number of modifications into different groups. Our analysis is essentially the same as the traditional interaction method but is easier to interpret.
increased survival rate), while a hazard ratio larger than 1 indicates increased rate of abandonment (i.e., decreased survival rate). Models 5-8 test how modifications affect the amount of contributions received by CotW target articles. Models 9-12 test how modifications affect the number of unique contributors in CotW. Models 5-12 report the regular coefficients.

Model 1 shows that a one-unit increase in pre-implementation modification decreases the hazard ratio by 3%, while a one-unit increase in post-implementation modification decreases the hazard ratio by 62%. The difference between the pre- and post-implementation modification is significant ($\chi^2=14, P < .01$). The results confirm Hypotheses 1, showing that post-implementation modifications have a much stronger positive effect on the practice survival. Models 2-4 show that modification effectiveness is influenced by editor type (e.g., core vs. non-core member and strong external ties versus weak external ties). Model 2 shows that the modification created by core members were more effective in decreasing hazard rate (68%) than non-core members (24%) and the difference is marginally significant ($\chi^2=3.0, P = .09$). Model 2 confirmed Hypothesis 2a partially. Model 3 shows that the modifications introduced by contributors with more external ties were more effective (decreasing the hazard rate by 83%) than modifications introduced by people with fewer external ties (decreasing the hazard rate by 13%). This difference is also statistically significant ($\chi^2=14, P < .01$). The results of Model 3 confirmed Hypothesis 2b. Regarding the interaction effects of being a core member with external ties, Model 4 provides mixed results. The modifications introduced by core members with more external ties (V7) significantly decrease the hazard rate by 82%. The modifications introduced by the other three types of contributors (core members with fewer external ties—V8, non-core members with more external ties—V9 and non-core members with fewer external ties—V10) did not significantly decrease the hazard rate. Also, core members with more external ties tend to create more effective modifications than those with fewer external ties ($\chi^2=8.5, P < .01$), which indicates that external relationships help core members create effective modifications. However, among the people with external ties, the difference between being core members and non-core members is not significant ($\chi^2=6.2, P = .43$). The results support Hypothesis 2c partially.

Models 5-12 present similar patterns as Models 1-4. The results collectively support Hypotheses 1, and 2b, and provide partial support for Hypothesis 2a and 2c.

**DISCUSSION**

**Modification timing of imported practice**

Research by Tyre and Orlikowski [45] as well as our own, although conducted in different organization settings, reveal similar patterns of new practice modifications (see Figure 9, top and bottom). Specifically, we find that a substantial proportion of modifications were made relatively soon after receiving the new practice and far fewer modifications were made afterwards. The underlying psychological process might be as follows: when the recipient site receives a new

<table>
<thead>
<tr>
<th>Survival Hazard Ratio</th>
<th>Contributions Coefficients</th>
<th>Participants Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source Popularity</strong></td>
<td><strong>Model 1</strong></td>
<td><strong>Model 2</strong></td>
</tr>
<tr>
<td>Pre-implementation Modifications (V1)</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Post-implementation Modifications (V2)</td>
<td>0.97**</td>
<td>0.97**</td>
</tr>
<tr>
<td>Modifications by core (V3)</td>
<td>0.38**</td>
<td>0.17**</td>
</tr>
<tr>
<td>Modifications by non-core (V4)</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Modifications by external (V5)</td>
<td>0.17**</td>
<td></td>
</tr>
<tr>
<td>Modifications by non-external (V6)</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Modifications by core &amp; external (V7)</td>
<td>0.18**</td>
<td></td>
</tr>
<tr>
<td>Modifications by core &amp; non-external (V8)</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Modifications by non-core &amp; external (V9)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Modifications by non-core &amp; non-external (V10)</td>
<td>1.13</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Effectiveness of the modifications.
practice, people are excited to adopt it yet believe that they can improve its potential contribution value by modifying it. However, after implementing the practice for a while, people tend to become reluctant to make changes. When the imported practice does not achieve expected performance, they might simply abandon rather attempt to further modify the practice.

However, empirical analysis reveals that modifications introduced before implementation are less effective than those introduced after implementation. Results show that the benefits of pre-implementation modifications are one order of magnitude lower than post-implementation modifications. A one-unit increase in pre-implementation modification decreased the hazard of failure by only 3%, while a one-unit increase in post-implementation modification decreased the hazard of failure by 62%. Similarly, a one-unit increase in pre-implementation modifications increased member contributions on targeted articles by only 0.7%, while a one-unit increase in post-implementation modifications increased the contributions by 17%.

The results suggest an alternative way to treat an imported practice. It might be better for a recipient unit to change the imported practice only slightly—if at all—before trying it because pre-implementation modifications (although initially deemed sensible and promising) minimally improve practice utilization. In contrast, more resources should be devoted to modifying the practice after the receiving units have experienced it.

**Effects of modifications introduced by core members**

Hypotheses related to core members (2a and 2c) are weakly supported by the data. For instance, modifications created by core members decreased hazard rate of CotW by 68% and those modifications created by non-core members decreased hazard rate by 24%, but the difference is only marginally significant (p=.09).

One possible reason why the effects are not as strong as anticipated is that the operationalization of core members—top 10% contributors—might be arbitrary. According to this operationalization, some peripheral members might be labeled as core members or vice versa, which might explain the relatively low significance.

Second, the current core-ness measurement, which essentially measures people’s contribution levels, might not be a good proxy. There are two possible underlying mechanisms of the effects of modifications introduced by core members. The “expertise-based” mechanism suggests that core members are more experienced and better understand the local project. Thus they can better identify or proactively search for effective modifications. The “influence-based” mechanism suggests that core members are more influential in the project and thus their modification suggestions are more likely to be accepted by other project members. Contribution levels might be a first order of approximation of the expertise or influence people have in the projects. However, this study will benefit from a closer examination on the roles of core members play in the practice adaptation process and more nuanced and precise measurements of member core-ness. Future work should attempt to address these aspects.

**Generalization to offline organizations**

This chapter proposes a contingency theory aimed at answering one management question that applies to any online community or offline organization that attempts to transfer best practices from one unit to another. The empirical study presented in this chapter provides evidence that the theory holds in the context of online communities. However, it remains unknown to what extent the findings may be generalized to an offline context.

One conjecture is that the findings might translate to offline organizations that share the some of the same features as online communities, especially those “organic organizations”. Roughly fifty years ago, Burns and Stalker [9] proposed the concept of “organic management system” as an alternative to bureaucratic management systems (what
they called a “mechanistic system”). They suggested that organic systems and mechanistic systems represent two poles of organizing forms: a mechanistic system is highly formal, rigid and centralized, while the organic system is informal, dynamic and flat. Organic management systems feature “the contributive nature of special knowledge and experience to the common task” and “lateral rather than a vertical direction of communication through the organization” ([9], Page 121). Organizations fall on different positions on the organic-mechanistic spectrum. For example, universities, offline volunteer organizations, design studios and research labs are more organic and thus more similar to online communities in terms of organization structures than, for example, the military and government, which are more mechanistic. Recently, there has been an increasing trend to adapt organizations to be “more organic” [13, 14, 29, 38].

Given the similarity between organic organizations and online communities, we conjecture that the findings of this chapter might be easier to transfer to organic offline organizations as compared to mechanistic organizations. However, this conjecture must be regarded with caution until it is confirmed by empirical work. Our intent of connecting online communities and organic organizations is to stimulate readers to bridge the CSCW and organization science areas, and consider new perspectives in studying important organizational phenomenon in both new and traditional organization forms.

**Internal versus external practice transfer**

This chapter focuses on examining the best practice transfer within the same community or organizations. Practice transfer across different communities or organizations is a different story. External practice transfer is often hindered by confidentiality and legal obstacles [43], which makes it difficult or even impossible for the recipient site to accurately replicate the original practice. Sometimes, it is legitimacy rather than effectiveness that becomes the priority concern in the recipient site [32]. For example, firms adopt the ISO 9000 Quality Certificates primarily to legitimate themselves and ensure public and customer support, which leaves little room for modification.

**CONCLUSION**

In this paper, we propose a contingency perspective to understand the process of incorporating and adapting best practice within online communities. We conducted quantitative analysis on the transfer of a quality-improvement practice between 146 Wikiprojects within Wikipedia. The results show that modifications were more helpful if they were introduced *after* the receiving project already had experience with the imported practice. Modifications were more effective if they were introduced by people who had experience in a variety of other projects.

**REFERENCES**


