Knowledge Refinement Effectiveness

Completed Research Paper

Tingting (Rachel) Chung
Carlow University
Center for Leadership & Management
Pittsburgh PA 15213
tchung@carlow.edu

Dennis Galletta
University of Pittsburgh
Katz Graduate School of Business
Pittsburgh PA 15213
galletta@katz.pitt.edu

Abstract

Electronic knowledge repositories represent one of the fundamental tools for knowledge management (KM) initiatives. This research examines organizational and dyadic factors that determine the effectiveness of knowledge refinement, the process of evaluating, analyzing and optimizing the knowledge object to be stored in a repository. A survey study with a matched-triad design was conducted with 318 authors, validators and users of a global knowledge repository system. Results revealed that perceived procedural justice, expertise gap, communication frequency, and shared understanding significantly influenced the quality of refined knowledge, while perceived procedural justice also positively influenced the quality of initial knowledge contributions. These findings are discussed with respect to future research and managerial practices.

Keywords: knowledge management, knowledge refinement, knowledge quality, procedural justice, shared understanding, communication frequency, expertise gap
Introduction

In today’s knowledge economy, companies are constantly seeking ways to gain sustainable competitive advantage by leveraging their knowledge assets (Davenport et al. 1998). One of the information technology tools widely deployed to support knowledge management (KM) strategies is the electronic knowledge repository. Defined as “databases allowing the storage and retrieval of explicit research and technical and management knowledge” (King et al. 2002, p. 93), electronic knowledge repositories represent one of the fundamental tools for KM efforts (Alavi 2000; King et al. 2002), particularly for sustaining organizational memory (Olivera 2000).

The success of a repository system is contingent on the extent to which their content is actively utilized by organizational members, either for replication (Dixon 2000; Markus 2001) or for innovation (Majchrzak et al. 2004). Surprisingly, the literature suggests that valuable content of repository systems is not always used as frequently as desired (Goodman et al. 1998; Gray et al. 2005). For instance, in an empirical study of various knowledge sources, repositories receive the lowest level of usage (Gray et al. 2005). KM efforts to motivate knowledge sharing and contributions to repositories may be effective, but the KM process that they support will not be successful if organization members are reluctant to access and apply the contributed knowledge to their own work (Kulkarni et al. 2006).

Prior research speculates that low levels of repository usage or ineffective use can be attributed to questionable content quality (Goodman et al. 1998; Gray et al. 2005; Haas et al. 2005). “[I]n the absence of the ability to monitor and control quality for each contribution, [knowledge sharing] systems can spark a flood of low-quality information contributions” (Fulk et al. 2004, p.571).

There is some suggestion in the literature that having effective knowledge refinement mechanisms in place is crucial for optimizing content quality (Cho et al. 2008; Zack 1999). An increasing number of organizations commission “knowledge intermediaries” to refine and develop contributed knowledge for maximal reusability (Markus 2001). The effectiveness of such refinement mechanisms, and their impact on contribution and user behaviors, have rarely been examined empirically (Lampe et al. 2004). The present study focuses on the role of refinement mechanisms in the supply-demand relationship between knowledge contribution and usage in the context of repository systems. This study is designed to explore antecedents to effective knowledge refinement both theoretically and empirically.

The balance of this paper provides a literature review, a description of antecedents to knowledge refinement effectiveness, the methodology employed in our empirical study, a discussion of the results, implications for researchers and practitioners, limitations, and conclusions.

Literature Review

Knowledge refinement refers to the process of evaluating, analyzing and optimizing the knowledge object to be stored in a repository (Alavi 2000; Cho et al. in press; Markus et al. 2002; Qian et al. 2005; Zack 1999), supporting the knowledge creation process of externalization, the articulation of tacit knowledge into explicit forms (Nonaka 1994). In contrast to environments where knowledge contributions are made in response to direct requests for assistance (e.g., discussion forums), contributions to knowledge repositories require the initiation of proactive effort to raise awareness about an opportunity to contribute, to search for a solution that matches a potential problem, and to formulate and deliver the solution as a contribution (Olivera et al. 2008). The knowledge refinement process serves as an intermediary that supports these mediating mechanisms for knowledge contribution in the absence of specific requests for help.

Repositories are increasingly implemented with a refinement process that selects, verifies and improves contributions to be included. For example, submissions to the best practices repository at Shell are subjected to careful validation by a dedicated team (Hicks et al. 2002-2003). These refinement processes serve as filters that allow only a subset of knowledge to be collected for permanent storage (Walsh et al. 1991). By evaluating and selecting content given some quality criteria, the refinement process enhances the perception of the knowledge repository system’s accuracy (Hicks et al. 2002-2003), credibility, and
legitimacy (Olivera 2000), providing “justification” to the knowledge created and gathered via the system (Nonaka 1994, p. 26). Knowledge refinement processes have been specified as a fundamental component of a life-cycle model of knowledge management (King et al. 2008).

Refinement could address the increasing concern about information overload. Information technology has significantly lowered the barriers for accessing information and knowledge. However, the overabundant availability of information and knowledge could create a barrier for effective use of such resources as human attention is limited in capacity. Research that examines the relative importance of information quantity versus quality has begun to demonstrate that users derive value from high quality content (Gu et al. 2007). Large quantities of content of unknown quality, in contrast, can raise the costs for accessing and using such resources and subsequently decrease usage levels (Butler 2001).

Many repository systems implement well-defined refinement processes. A common approach is to commission a centralized review committee of domain experts to select, refine, and approve knowledge that enters repository systems (Goodman et al. 1998; Markus 2001; Zack 1999). At the other extreme is a decentralized system that assigns randomly chosen peers for refinement (Cho et al. 2008) or an open-access system allowing anyone interested to participate in the refinement effort on a completely voluntary basis, as seen in the Wikipedia project (Voss 2005). Corporations such as Dell, Starbucks, and Harrah’s have implemented a similar “beauty-contest” approach to knowledge refinement using Salesforce.com’s online voting service “Ideas” (Greenfield 2008). Other companies implement a “stock-market” approach to refinement, allowing employees to buy and trade stock shares of ideas contributed by employees (Greenfield 2008; Grosslight 2008). Both approaches engage employees through all ranks in selecting optimal ideas. They differ only in the extent to which the employee is personally invested in the quality of the decision – employees can obtain personal gains or losses depending on the stocks they choose to invest in with the “stock-market” approach.

Somewhere in the middle is Xerox’s refinement mechanism for its Eureka system: More than 600 globally-dispersed validators, or refiners, are responsible for refining knowledge submissions from more than 20,000 users (Bobrow et al. 2002; Boucher 2006). Refinement processes also differ in the extent to which they support collaboration among refiners and authors, and whether or not the refinement can be performed directly on the knowledge artifact (Chung et al. 2007).

Universal to these various refinement approaches is the joint participation of the refiner and the author in the knowledge creation process. The author externalizes tacit ideas into an explicit format, while the refiner helps the author optimize the quality of the explicated knowledge. Such a refinement process often involves the collaboration of the refiner and the author with varying degrees of interaction between the two.

The refinement process is conceptualized here using the input-process-output (I-P-O) framework of organizational teams and groups (Hackman 1987; McGrath 1984; Steiner 1972), with the author’s contribution as the input, the collaboration between the refiner and the author as the process, and the refined knowledge object as the output of the collaborative process. This view suggests that the goal of effective refinement is to produce knowledge objects of the quality that meets both the producer’s and the user’s criteria (see Figure 1). Knowledge refinement effectiveness, based on this view, is defined as the degree to which the refinement process produces quality knowledge.

---

This conceptualization of the knowledge refinement process is built upon two assumptions. First, the refined knowledge quality (KQ1) should correlate with the initial knowledge quality (KQ0) because the improvement brought by the refinement process should be rather incremental. The refinement process may enhance the quality of the initial submission; however, it is unlikely to turn a poorly written submission around into an outstanding publication. Moreover, when a knowledge refinement process is implemented as part of an organization’s knowledge management efforts, the refinement procedure presumably accomplishes its goal of improving the quality of the contribution over time. Specifically, the quality of the knowledge content should be higher after the refinement process than before it. In other words, the output knowledge quality that eventually gets published should be greater than the initial knowledge quality. These basic premises are expressed as the following formal hypotheses:

Hypothesis 1a (H1a): Refined knowledge quality (KQ1) is positively related to initial knowledge quality (KQ0).

Hypothesis 1b (H1b): Refined knowledge quality (KQ1) is greater than initial knowledge quality (KQ0).

Organizational and Dyadic Antecedents of Knowledge Refinement Effectiveness

To explain effectiveness of the knowledge refinement process, factors are explored at both organizational and dyadic levels. At the organizational level, we focus on the role of procedural justice, a dimension of organizational justice. Organizational justice, defined as “members’ sense of the moral propriety of how they are treated” (Croppanzano et al. 2007, p. 34), is an appropriate theoretical frame because justice explains employee motivation above and beyond short-term economic gains (Croppanzano et al. 2007), which often characterize voluntary participation in KM efforts. At the dyadic level, we examine refiner-author factors that facilitate the effectiveness of refinement efforts. These factors include the extent of the author and the refiner’s shared understanding, the frequency of their communications, and their expertise level difference.

Perceived procedural Justice

In many organizations, the creation and use of electronic knowledge repositories rely on employees’ voluntary participation and are not necessarily linked directly to compensation or other reward systems. Even when participation is linked to specific reward mechanisms, reward does not necessarily lead to desirable organizational outcomes (Haas et al. 2005). In other words, pure economic incentives are insufficient to create effective KM processes. When behaviors are not directly linked to economic gains, employees care about whether they would be treated fairly, particularly in the long run (Croppanzano et al. 2007). Employees often participate in KM efforts such as refinement for long-term or indirect benefits, as opposed to immediate or direct gains. The extent to which organizational procedures such as the refinement process operate fairly, therefore, should influence the effectiveness of the process.

When a knowledge repository is implemented with a refinement process, knowledge contribution is the subject of refinement decisions. Authors make contributions knowing that their efforts will be evaluated with the risk of potential rejection for publication. Thus, the decision to contribute may be determined not just by the author’s motives, but also by his or her perception of the refinement process. One of the main questions an author may ask is whether the process is fair.

Fair processes are important managerial tools that influence organizational outcomes (Kim et al. 1997). The concept of fairness, or justice, has been examined extensively in the organizational justice literature (Greenberg 1987; Greenberg 1990; Leventhal 1980; Lind et al. 1988; Thibaut et al. 1975). Procedural justice, in particular, refers to the perceived level of fairness of procedures used in decision making (Lind et al. 1988). In contrast to distributive justice, which concerns the fairness of the outcome of a decision, procedural justice focuses specifically on the fairness of the decision process independent of the outcome of the process.

The organizational justice literature shows that justice perceptions have significant impact on how people behave in organizations. Fairness perceptions promote organizational citizenship behavior, such as interpersonal helping and individual initiative (Moorman et al. 1998; Organ et al. 1993). Contribution to a knowledge repository has been characterized as an altruistic helping behavior (Wasko & Faraj, 2005), and therefore can be considered as a form of organizational citizenship behavior (Constant et al. 1996). It is
often an extra-role behavior because repository contribution is commonly voluntary rather than mandatory in performance evaluation. If fairness perceptions promote organizational citizenship behavior, fairness in KM processes should presumably promote knowledge contribution behavior. In other words, if the process through which repository contribution is judged as fair, individuals are more likely to initiate the extra-role behavior of knowledge sharing through a repository.

When individuals feel that they are treated fairly by a group and its authority, they feel more respect from the group members, and they take more pride in being a member of the group. These feelings of pride and respect, in turn, motivate individuals to go beyond the call of duty, and to take action that serves the group’s interests in addition to their personal interests (Tyler et al. 1996), such as the extra-role behavior of knowledge contribution. Indeed, procedural justice perceptions have been consistently linked to organizational citizenship behavior that involves discretionary and extra-role actions to promote organizational effectiveness (Podsakoff et al. 2000; Tepper et al. 2003).

Procedural justice is critical for building successful repositories. When organizational members feel they are treated fairly, they develop more positive affect (Tyler 1994), such as trust in management (Chen et al. 2004; Kim et al. 1993) and commitment to a group or organization (Colquitt 2001; Kim et al. 1993; Masterson et al. 2000). These positive attitudes towards the organization as a result of fair process perceptions are crucial in developing voluntary cooperation in knowledge workers (Kim et al. 1993; Kim et al. 1997; Kim et al. 1998). Getting organizational members to contribute to a knowledge repository depends vitally on such voluntary cooperation as “creating and sharing knowledge are intangible activities that can neither be supervised nor forced out of people” (Kim et al. 1997, p.71). Therefore, when procedural justice is exercised, organizational members feel that their knowledge and expertise receive intellectual and emotional recognition. Such recognition not only helps individuals protect their self-interests, it also supports their commitment to the group or organization. As people receive intellectual and emotional recognition, they become more willing to share knowledge and less likely to hoard their unique expertise (Kim et al. 1993). An empirical study of knowledge sharing provides support for the positive impact of an organizational climate characterized by fairness on organizational members’ intention for knowledge sharing (Bock et al. 2005). The following hypothesis is based on these ideas:

**Hypothesis 2 (H2):** Perceived procedural justice leads to higher levels of initial knowledge quality (KQ<sub>0</sub>.

Because procedural justice represents the degree to which the process, rather than the outcome, is perceived to be fair (Lind et al. 1988), its magnitude should predict the effectiveness of the refinement process. A fair process for reviewing and improving knowledge contribution may not necessarily produce outcomes that are favorable to a particular author, refiner, or user. However, the refinement process, if it is fair, should produce outcomes of maximal utility to the community to which it serves. In other words, procedural justice of the knowledge refinement process should positively enhance the quality of refined knowledge, or KQ<sub>1</sub>:

**Hypothesis 3 (H3):** Perceived procedural justice leads to higher levels of refined knowledge quality (KQ<sub>1</sub>.

**Shared Understanding**

Shared understanding is defined here as the extent to which the refiner and the author dyad share a common knowledge base and comprehension with respect to the knowledge target of the refinement process (Gerwin et al. 1997; Ko et al. 2005; Nelson et al. 1996). Shared understanding is crucial for collaborative work. However, when teams from different backgrounds work together, the differences in their functional knowledge base and conceptual perspectives often create tension that prevents them from collaborating effectively (Gerwin et al. 1997; Nelson et al. 1996). This lack of shared understanding forms barriers for collaboration partners to appreciate the unique value of team partners’ contribution and to develop synergy among team members (Nelson et al. 1996). Empirical studies of collaborative work have consistently demonstrated the importance of shared understanding in team performance (Gerwin et al. 1997; Ko et al. 2005; Nelson et al. 1996).

To the extent that the refinement process represents a collaborative effort between a refiner and an author, the level of their shared understanding should influence the effectiveness of their collaborative effort to optimize the quality of the author’s contribution. Collaboration in new product development, for example, depends critically on the level of shared understanding among otherwise autonomous teams
(Gerwin et al. 1997; Wheelwright et al. 1992). The development of a knowledge object for repository storage is similar to new product development in two ways. First, like the creation of a new physical product, the creation of a new idea involves a certain level of ambiguity in terms of what the end product should be. Participants in the creation process must reduce the level of ambiguity through constant negotiation and communication. A shared understanding provides a common framework that helps the participants make decisions and resolve discrepancies more effectively (Wheelwright et al. 1992). Secondly, the criteria for success involve a certain level of subjectivity for both product and idea creation processes. What constitutes a successful new product or idea can be debatable. When collaborative partners in product development lack a shared understanding, disagreement is more likely to occur, which may require intervention by management that reduces the autonomy of the collaborative partners and consequently the development team’s performance (Gerwin et al. 1997).

The literature presents some preliminary evidence that shared understanding improves knowledge refinement effectiveness. In a laboratory experiment, Cho et al. (2008) demonstrated that when the refiner and the author are more similar in their levels of expertise, the quality of the refined knowledge, in terms of idea novelty, logical rigor and structural flow, is higher than when the refiner has significantly much more expertise than the author does. Therefore,

_Hypothesis 4 (H4): Shared understanding leads to higher levels of refined knowledge quality (KQ1)._  

**Communication Frequency**

Collocated work environments benefit tremendously from the amount of trust and mutual influence developed over time through frequent interactions (Nelson et al. 1996). However, these relationship-based factors may be less relevant in the context of electronic repositories. Authors and refiners who support an electronic repository might have little or no prior experience working together. Knowledge refinement represents a form of computer-mediated collaborative work between organizational members who otherwise may not work together on a regular basis.

Specifically, the refinement process often involves frequent interactions and multiple iterations between the refiner and the author (e.g., Cho et al. 2008). Sharing an idea through a knowledge repository is a process of externalization (Nonaka 1994), the product of which is often less than optimal because the owner of the idea may not be able to effectively convey the full range of information that someone else lacking the same expertise needs in order to understand and apply the same idea. The refinement process, therefore, often involves communication between the refiner and the author in an attempt to fully explicate relevant information needed to codify the idea comprehensively. Frequent communication allows the refiner to clarify the author’s intended perspective and to establish a more accurate understanding of the submitted idea. In other words, communication frequency should improve the effectiveness of knowledge refinement:

_Hypothesis 5 (H5): Communication frequency leads to higher levels of refined knowledge quality (KQ1)._  

**Expertise Gap**

In addition, the refiner needs to understand the author’s view before she or he can help the author communicate the idea to the target audience. This task becomes particularly challenging when the refiner and the author differ significantly in their levels of expertise (Cho et al. 2008). A significantly large expertise gap makes communicating and articulating knowledge extremely difficult, as it reduces the expert refiner’s ability to understand the non-expert contributor’s perspective (Hinds et al. 2003), which in turn compromises the establishment of common ground for collaborative work. In contrast, two parties with more similar levels of expertise can establish a common understanding more easily, which allows the refiner to more effectively facilitate the author in the codification process. This effect has been demonstrated empirically in an experimental setting (Cho et al. 2008). We anticipate the same effect in the refinement process in a field setting. In the knowledge refinement literature, the refiner is usually more expert than the contributor (Markus 2001; Zack 1999). However the same impact of the expertise gap on the development of shared understanding should be expected whether the refiner or the author is more expert than the other. Therefore, we anticipate the negative impact of the expertise gap on refinement effectiveness in both directions.
Hypothesis 6 (H6): The larger the expertise gap between the author and the validator (in either direction), the lower the level of refined knowledge quality (KQ).

These hypotheses are summarized visually in the research model presented in Figure 2.

![Figure 2. Summary of the Research Model](image)

**Research Methodology and Data Collection**

An empirical study was designed and conducted with participants of Eureka, a knowledge repository system of the Xerox Corporation\(^2\) that has supported knowledge sharing initiatives at the firm since 1994.

Eureka is a knowledge repository system that supports knowledge sharing among Xerox service technicians (Bobrow et al. 2002; Boucher 2006; Hickins 1999). More than 20,000 users worldwide access Eureka for service “tips” that are authored by Xerox service technicians who are also Eureka users, and refined by more than 600 expert technicians who serve as “validators,” both improving and approving or rejecting tip submissions (Boucher 2006). Validators are members of the Eureka user community. Like the others, they are service technicians, but they have completed training and certification provided by Eureka management. They are, therefore, more like peers to fellow Eureka authors and users, and less like external expert auditors or inspectors.

When a tip is submitted, validators from the same country as the submitter are notified, and they can volunteer to work with the submitter in developing and refining the tip. After 30 days, validators from any country can evaluate the tip as long as they are experts in the same product family. The validator may decide to reject a tip submission on several grounds: The tip is already included in Eureka, the tip is part of another tip already present in Eureka, the tip duplicates information available in other repositories

---

\(^2\) The firm’s identity is disclosed with permission.
Knowledge Management and Business Intelligence

(e.g., Xerox's Electronic DOCumentation (EDOC)) or the tip is invalid. If the tip is indeed original, innovative, and worthy of inclusion, the validator goes on to work with the author on refining it. The refinement process could include one or more revisions of the submission. The refinement process could last as short as a day and as long as 687 days (mean = 60.8 days; median = 21 days) for the 3,485 tips refined and validated in 2006.

Survey Study Design

Validated measurement items were adapted from the existing literature for the survey study. The survey instruments were designed to be administered online with three versions: the user version, the author version, and the validator version. To support a matched-triad design of the survey study, participants were recruited in two waves, similar to the recruitment procedure implemented by Constant et al. (1996). The first wave consisted of the user survey, during which all Eureka users were invited to complete the user version of the online questionnaire with respect to a Eureka tip of their own choice. In the questionnaire, the user was asked to nominate a tip for inclusion in the study, and specify the tip by specifying the six-digit ID number that uniquely identified a tip in the Eureka database. Tips identified by the users were included in the subsequent surveys targeting authors and validators.

The user survey was first featured as an announcement on the internal service site of the Eureka system and remained on the site for over two months. A follow-up invitation was sent to 800 randomly selected users. Users clicking on the hyperlink embedded in the announcement were directed to the online questionnaire. By the end of the study, a total of 483 users clicked through the link to the questionnaire, of which 179 participants answered the survey questionnaire. Because an indefinite number of potential respondents only might have seen the initial announcement, an exact response rate cannot be determined. Considering only the 800 personal invitations, however, we have at most an effective response rate of 37.1%.

After the users completed the survey, authors and validators of the nominated tips were invited to complete the author and validator surveys (both available upon request), respectively. For each tip, the goal was to receive responses from both the author and the refiner. A total of 175 user survey responses were collected that contained a tip ID number. Excluding duplicate tip ID numbers (i.e., tips that were nominated by more than one user), the Eureka coordinator generated a list of unique author and validator names and contact information on 145 tips (78.95%) with valid ID numbers (excluding ID numbers that were invalid according to the current Eureka database.) These authors and validators were contacted via email and invited to answer the author or validator questionnaire with respect to the nominated tip.

Questionnaire Instrument

Validated measurement items were adapted from existing literature and a small number of original items were created for the survey. Interviews with Eureka validators and users confirmed that the survey items were appropriate for the Eureka context. Table 1 lists the measurement items used in the survey.

Effectiveness of the knowledge refinement process was examined by measuring the quality of initial knowledge contributions and their refined versions. Consistent with prior studies that empirically assessed knowledge quality (Clay et al. 2005; Kulkarni et al. 2006; Qian et al. 2005; Wu et al. 2006), six validated measures of knowledge quality were adapted to capture the currency, accuracy, completeness, relevance and reliability of information content (Nicolaou et al. 2006).

<table>
<thead>
<tr>
<th>Construct ID</th>
<th>Measurement Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ1</td>
<td>This tip provides the precise information I need.</td>
</tr>
</tbody>
</table>

3 Survey instruments are available upon request. Please contact the first author.

4 To minimize the effort of respondents, a Eureka participant was only sent one invitation for participation. If an individual was involved in more than one nominated tip, either as an author or as a validator, a decision was made to invite for the individual for only one of the tips that he or she was involved in. The other tips were excluded from the study.
The user questionnaire focused on the assessment of Refined Knowledge Quality (KQ1). The author questionnaire focused on assessing the author's perception of the procedural justice of the refinement process (Perceived Procedural Justice (PJ) items adapted from Blader et al. 2003). Both author and the validator were direct participants in the refinement procedure, where the validator served as the decision maker, Therefore, the author was the appropriate person to evaluate the fairness of the process. In addition, the author rated the extent of shared understanding with the validator during the refinement process (Shared Understanding (SU), items adapted from Ko et al. 2005), the frequency of communication with the validator (Communication Frequency (CF)), and his or her own level of expertise. Self-reported levels of expertise were then used to calculate the expertise gap, to be discussed below.

The validator questionnaire focused on assessing the validator’s assessment of the quality of the author's initial contribution (Initial Knowledge Quality (KQ0)). In addition, the validator rated the extent of shared understanding with the author during the refinement process (Shared Understanding (SU)), the frequency of communication with the author (Communication Frequency (CF)), and his or her own level of expertise. Self-reported levels of expertise were used to calculate the expertise gap.

Expertise Gap (EG) was measured with four items in the author and the validator surveys and operationalized as the difference scores between the author and the validator’s self-reported levels of expertise. Similar to difference scores used in prior IS research (e.g., Barki et al. 2001; Jiang et al. 2002), EG was calculated as:

<table>
<thead>
<tr>
<th>(Rai et al. 2002)</th>
<th>IQ2</th>
<th>This tip provides sufficient information for me to use it</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ3*</td>
<td></td>
<td>This tip has errors that I must work around</td>
</tr>
<tr>
<td>IQ4</td>
<td></td>
<td>I am satisfied with the accuracy of this tip</td>
</tr>
<tr>
<td>IQ5</td>
<td></td>
<td>Information contained in this tip is helpful regarding my questions or problems</td>
</tr>
<tr>
<td>IQ6</td>
<td></td>
<td>This tip is presented in formats (e.g., text, visual, audio) that are sufficient for my use</td>
</tr>
<tr>
<td>Shared Understanding (Ko et al. 2005)</td>
<td>SU1</td>
<td>The tip validator and I solved problems the same way</td>
</tr>
<tr>
<td></td>
<td>SU2</td>
<td>The tip validator and I understood each other when we talked</td>
</tr>
<tr>
<td></td>
<td>SU3</td>
<td>The tip validator and I had no problem understanding each other</td>
</tr>
<tr>
<td>Comm. Frequency</td>
<td>CF1</td>
<td>The tip validator and I communicated over the phone</td>
</tr>
<tr>
<td></td>
<td>CF2</td>
<td>The tip validator and I communicated over email</td>
</tr>
<tr>
<td></td>
<td>CF3</td>
<td>The tip validator and I communicated in person</td>
</tr>
<tr>
<td>Perceived Procedural Justice (Blader et al. 2003)</td>
<td>PJ1</td>
<td>How often did you feel that validation decisions were made in fair ways?</td>
</tr>
<tr>
<td></td>
<td>PJ2</td>
<td>Overall, how fair would you say validation decisions and processes were in Eureka?</td>
</tr>
<tr>
<td></td>
<td>PJ3</td>
<td>How would you rate the overall fairness with which issues and decisions that came up during validation were handled?</td>
</tr>
<tr>
<td></td>
<td>PJ4</td>
<td>Was there a general sense among employees that submissions were handled in fair ways during validation?</td>
</tr>
<tr>
<td></td>
<td>PJ5**</td>
<td>How much of an effort was made to be fair to employees when validation decisions were being made?</td>
</tr>
<tr>
<td>Expertise</td>
<td>E1</td>
<td>At the time of submitting this tip, how many years have you worked with the product for which this tip supports?</td>
</tr>
<tr>
<td></td>
<td>E2</td>
<td>Please rate your expertise on this product at the time when you validated the tip: Novice ..... Expert</td>
</tr>
<tr>
<td></td>
<td>E3</td>
<td>How informed were you on this product at the time of the tip submission?</td>
</tr>
<tr>
<td></td>
<td>E4</td>
<td>To what extent were you an expert on this product?</td>
</tr>
</tbody>
</table>

*Item was reverse coded. **Item was dropped.
EG = | Validator’s Expertise – Author’s Expertise |

Guidelines provided by Petter et al. (2007) were followed to determine if a construct was reflective or formative.

Both the questionnaire instrument and the survey design was pilot-tested on 20 matched-triads. Results from the pilot displayed satisfactory levels of reliability, or internal consistency (as indicated by Cronbach’s α). They also displayed satisfactory averaged variance extracted (AVE), indicating appropriate convergent validity. Confirmatory factor analysis (CFA) using Structural Equation Modeling (SEM) procedure with the software package EQS (Bentler et al. 1995) revealed acceptable levels of model fit for the reflective constructs.

Subjects

The 145 user survey responses with a valid and unique tip ID, a unique author and a unique validator, formed the tip sample base for the present analysis. Email invitations to participate in the survey study were sent to authors and validators of these tips with the following restrictions: The authors and validators had a currently working email address with the Xerox company, and each participant only received one invitation (for either the author or the validator survey). Twenty-two participants were excluded from the invitation process because they had retired from Xerox at the time of the study. The “one-invitation” strategy was adopted to minimize the negative effects of overwhelming the participants and to maximize participation rate. If a participant was author or validator for more than one tip nominated for inclusion in the study, the study invitation referred to only one of the tips with which the participant was affiliated. The other tips with which the participant was affiliated were then excluded from the study. Decisions about tip exclusion were made such that a maximum number of tips were retained in the study. When the study was concluded, valid responses were received from 106 author and validator pairs.

With 106 cases, a total of 318 Eureka members participated in the study by filling out the questionnaires online. 106 participants answered the user questionnaire on a tip of their choice, 106 answered the author questionnaire about a tip they authored, and the remaining 106 answered the validator questionnaire about a tip they validated. Descriptive statistics of these pilot study participants are presented in Table 1.

| Table 2. Full Survey Study Participant Descriptive Statistics |
|-----------------|-----------------|-----------------|-----------------|
| Role            | Gender          | Median Age Range | Job Tenure (S.D.) |
| User            | 9 female/98 male | 49-58            | 8.71 (9.23)      |
| Author          | 2 female /105 male | 49-58            | 12.43 (11.31)    |
| Validator       | 8 female/99 male | 49-58            | 11.07 (5.70)     |

Control Variables

Three variables compiled from systems logs were included as control variables in the testing of the research model:

Author Experience: Author experience is defined as the participant’s experience as an author contributing to the knowledge management system, and is operationalized as the number of tips that an author has submitted for validation. This measure includes all submission attempts, without regard to whether a tip was eventually validated or rejected. Author experience of participants included in the full survey study ranged from 1 to 17, with a mean of 3.7 (S.D. = 2.13).

Validator Experience: Validator experience is defined as the participant’s experience as a validator participating in the knowledge management system’s refinement process, and is operationalized as the number of tips that a validator has reviewed. This measure includes all review effort, without regard to

\[\text{|} \text{ denotes the calculation of absolute value.}\]

Although we know for sure that authors and validators were unique participants, we are not certain if the users who answered the survey were not also authors or validators participating in this study. In other words, the total number of unique participants could potentially be fewer than 318.
whether a tip was eventually validated or rejected. Validator experience of participants included in the full survey study ranged from 1 to 23, with a mean of 4.6 (S.D. = 3.09).

Refinement Duration: Refinement duration is defined as the amount of time taken by the refinement process, and is operationalized as the number of days passing from tip submission to the day when a decision was made about the tip’s publication status. This measure includes all submission validation attempts, without regard to whether a tip was eventually validated or rejected. Refinement duration of all tips included in the full survey study ranged from 1 to 296, with a mean of 42.1 (S.D. = 39.72).

Results

Measurement Models

Table 2 below presents the means, standard deviations and other descriptive and correlational results of the four reflective constructs – Perceived Procedural Justice\(^7\), Shared Understanding\(^8\), Initial Knowledge Quality (KQ0) and Refined Knowledge Quality (KQ1) – used in the research model. None of the correlations between independent constructs was greater than .80, the critical level for concern about multicollinearity (Billings et al. 1978). The constructs display satisfactory levels of reliability, or internal consistency. Cronbach’s α exceeded the recommended minimum value of 0.7 for all reflective constructs.

<table>
<thead>
<tr>
<th>CONSTRUCT</th>
<th># Items</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>α</th>
<th>CR(^9)</th>
<th>AVE</th>
<th>PJ</th>
<th>SU</th>
<th>KQ0</th>
<th>KQ1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Procedural Justice (PJ)</td>
<td>4</td>
<td>99</td>
<td>5.84</td>
<td>1.33</td>
<td>.91</td>
<td>.94</td>
<td>.85</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shared Understanding (SU)</td>
<td>3</td>
<td>107</td>
<td>5.13</td>
<td>1.49</td>
<td>.94</td>
<td>.96</td>
<td>.86</td>
<td>.39</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Knowledge Quality (KQ0)</td>
<td>6</td>
<td>101</td>
<td>5.89</td>
<td>.60</td>
<td>.74</td>
<td>.82</td>
<td>.46</td>
<td>.33</td>
<td>.23</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>Refined Knowledge Quality (KQ1)</td>
<td>6</td>
<td>106</td>
<td>6.05</td>
<td>1.14</td>
<td>.94</td>
<td>95</td>
<td>.77</td>
<td>.37</td>
<td>.17</td>
<td>.22</td>
<td>.88</td>
</tr>
</tbody>
</table>

Note: The square root of average variance extracted (AVE) is on the diagonal in bold font.

Validity of reflective constructs was assessed in three ways. First, dimensionality of the four reflective constructs was assessed with confirmatory factor analysis (CFA) using EQS software (Bentler et al. 1995). As recommended by Gerbing and Hamilton (1996), this CFA was conducted with maximum likelihood estimation with a four-factor model. To evaluate how the model represented the data, we used both absolute fit indices, including the goodness-of-fit index (GFI) and the χ\(^2\) statistic, and incremental fit statistics, including the root square error of approximation (RMSEA) and the comparative fit index (CFI). Following Medsket et al.’s (1994) recommendations, we looked for good fit with values greater than .95 and acceptable fit with values greater than .90 when assessing CFI and GFI. With RMSEA, we followed these generally accepted guidelines: values less than .05 constitute good fit, values ranging from .05 to .08 constitute acceptable fit, values ranging from .08 to .10 constitute marginal fit, and values greater than .10 constitute poor fit (Browne et al. 1992). CFA results showed that the model produced an acceptable fit to the data, χ\(^2\) = 135.79, GFI = .93, CFI = .91, and RMSEA = .077.

Next, average variance extracted (AVE), or the percentage of the total variance of a measure represented or extracted by the variance due to the construct, was used to assess the constructs’ convergent validity. The AVE scores, shown in Table 3, were mostly above .5, the minimum level for adequate convergent validity suggested by Fornell and Larcker (1981). The only exception was Initial Knowledge Quality, the

---

\(^7\) Data collected from both the author and validator were averaged for evaluating the measurement model of the Procedural Justice construct.

\(^8\) Data collected from both the author and validator were averaged for evaluating the measurement model of the Shared Understanding construct.

\(^9\) CR stands for composite reliability.
AVE value of .46. Discriminant validity, on the other hand, was established by inspecting the square root of AVE scores (available on the diagonal of Table 2). All values were greater than the off-diagonal correlations, indicating proper discriminant validity for all constructs. In summary, measurement model validity has been established for reflective constructs.

Measures of two constructs - Communication Frequency and Expertise Gap - will not necessarily need to covary and therefore were modeled as formative indicators. Descriptive statistics of these formative constructs are summarized in Table 4.

| Table 4. Descriptive Results of Formative Constructs. |
|---------------------------------|----------|------|------|
| Dimension                       | # Items | N    | MEAN | SD    |
| Expertise Gap (EG)              | 4       | 105  | 1.81 | 1.06  |
| Communication Frequency (CF)    | 3       | 107  | 4.72 | 1.36  |

Validity of these measures of formative constructs was assessed using exploratory factor analyses with principle component analysis and Varimax rotation. Because data collected for the Communication Frequency (CF) construct from both authors and validators were determined to be valid, the average scores of author and validator responses were used for this analysis. The results, again, converged well, yielding components with eigen values greater than one that corresponded with the intended measures. These results confirm the validity of both formative measures.

**Hypothesis Testing**

Hypotheses 1a and 1b were tested using Pearson correlation and t tests. The remaining statistical analysis was performed using partial least squares (PLS) analysis for several reasons. First, this work is relatively exploratory as this work is among the first to examine knowledge refinement empirically. Second, the reflective and formative constructs can be considered within the same model (Chin, Marcolin, & Newsted, 2003). Finally, PLS is more appropriate than covariance-based approaches when the sample size is relatively small (Hair, Ringle & Sarstedt, 2011). To test Hypothesis 1a, Pearson correlation was obtained between KQ$_0$ and KQ$_1$. These two constructs demonstrated a modest but significant level of correlation ($r = .22$). Therefore, Hypothesis 1a is supported. To test Hypothesis 1b, paired sample t tests were performed between the six measures of KQ$_0$ and KQ$_1$. This analysis was performed to examine if KQ$_1$ values were significantly greater than the values of KQ$_0$. Five out of the six t tests were significant at the .05 level, suggesting that Hypothesis 1b was largely supported.

The structural aspect of the research model was tested using the software package SmartPLS version 2.0 (Ringle et al. 2005). Results of model testing revealed that the antecedent constructs explained 22.7 percent of the variance in the level of refined knowledge quality. Statistical tests of the model paths were conducted using the bootstrapping technique with 1000 resamples. This bootstrapping procedure allowed us to examine the convergent validity of measurement items by examining the t values of outer model loadings (Gefen et al. 2005). All of the outer model loadings of reflective constructs were statistically significant at the .01 level. These results confirmed the strong convergent validity for reflective constructs with multiple indicators.

The bootstrapping procedure also allowed us to examine hypotheses 2-6. Findings are summarized in Figure 3. The majority of the structural paths in the research model were significant statistically.

Perceived procedural justice significantly explained the level of initial knowledge quality ($\beta = .324, p < .01$). Perceived procedural justice also significantly explained the level of refined knowledge quality ($\beta = .297, p < .01$). Communication frequency significantly explained the level of refined knowledge quality ($\beta = .187, p < .05$). Expertise gap significantly explained the level of refined knowledge quality ($\beta = -.175, p < .05$). The path from shared understanding to refined knowledge quality was the only one significant at a marginal level. None of the control variables exerted statistically significant effects on the dependent variable. These findings are summarized visually in Figure 3.
Discussion of Research Findings

This study examined the knowledge refinement process in KM implementation, which has received relatively little attention from researchers. Empirical findings were reported from studying Eureka, one of the most successful and long-lived repository systems that have been discussed widely in the KM literature. We evaluated the knowledge refinement process by observing the evolution in knowledge quality (i.e. the input to refinement) from submission to publication (i.e. the outcome of refinement). While the refined knowledge quality is correlated with the quality of initial submissions, other factors also contribute to the effectiveness of the refinement process. We proposed a framework to understand organizational as well as dyadic factors that contributed to effective knowledge refinement. Here we discuss findings of our empirical studies, and their implications for future research as well as managerial practices.

At the organizational level, we theorized that a fair refinement process would be effective in improving knowledge quality. Because prior research suggests that procedural justice promotes organizational citizenship behavior, we hypothesized that perceived procedural justice would enhance the quality of both initial and refined knowledge contributions. Our results provided support for both hypotheses. This finding suggests that a fair refinement process would produce benefits to the users, and not just the authors or validators who participated in the refinement work. Prior research investigated the role of fairness in the quantity of knowledge contribution (Bock et al. 2005). Our finding demonstrates that, consistent with Kim et al.’s (1997) emphasis on fairness in KM, the impact of fair processes extends beyond sheer quantities of knowledge sharing and appears in other domains of KM.

At the dyadic level, we found three factors that were critical in determining the quality of refined knowledge contributions. First, communication frequency positively influenced the quality of refined knowledge, contributing to the effectiveness of the refinement process. This finding supports the view that increased interaction between the author and the validator helps strengthen the validator’s understanding of the author’s intended goal in sharing a tip, which in turn improves the quality of the
refinement outcome. In open-ended comments, some participants pointed out the relatively low level of interaction between the author and the validator during the refinement process. The positive relationship between communication frequency and refined knowledge quality suggests that more interaction between the author and the validator could be encouraged.

The size of the expertise gap, as predicted, is negatively related to the quality of refined knowledge. In other words, the greater the difference between the author’s expertise level and the refiner’s expertise level, the less effective the refinement process was. In contrast, when the author and the refiner were close to each other in terms of expertise level, the refinement process was more effective in improving the quality of refined knowledge. This finding is consistent with existing theory (Hinds et al. 2003) and experimental research findings (Cho et al. 2008). This study is the first time this effect is demonstrated in the context of a corporate KM system, lending much needed external validity to the field of expertise research.

Shared understanding significantly determined the quality of refined knowledge, although the relationship was significant only at a marginal level. This relatively weak finding may be explained by the relationship between shared understanding and the other antecedents included in the research model. Frequency communication would increase shared understanding. The more often the validator interacted with the author, the more likely they would achieve a common ground in their understanding of the common task and approach to problem solving. Similarly, when the author and the validator are close in terms of expertise levels, they would examine the tip with more similar perspectives, which would lead them to a more common understanding of the validation task. In both cases, the conceptual overlap suggests that the explanatory power of shared understanding may have been largely accounted for by the other factor – communication frequency or expertise gap, leaving little variance in the outcome variable for which shared understanding could explain. If this is true, then this finding by no means suggests that shared understanding is unimportant for refinement. Quite the contrary, shared understanding may have mediated the effect of expertise gap and communication frequency on refinement outcomes.

Implications for Future Research

This research focuses exclusively on the role of perceived procedural justice in knowledge refinement. The organizational justice literature, however, provides rich discussions of other forms of justice – interactional justice, informational justice and distributive justice. Some of these justice concepts have been discussed in the MIS literature (Joshi 1989), but their distinctions are largely non-existent in the KM literature. Understanding the antecedents and consequences of these justice perceptions in the KM context can provide important theoretical and managerial insights. At the same time, the organizational justice perspective can enrich the understanding of other KM procedures and practices, such as knowledge sharing, seeking, and transfer.

It would be particularly interesting to explore how distributive justice should be conceptualized in the context of repository contribution. Unlike other forms of work rewards, repository inclusion is not a zero-sum or constant-sum game. In other words, having a submission accepted for inclusion in a repository does not take away the opportunity for another submission to be accepted. Whether distributive justice is relevant in this context and how it affects contribution behavior should be examined with additional research.

Future research should also explore organizational factors besides justice in knowledge refinement. For instance, what organizational design – centralized, decentralized, etc. – makes refinement more effective? What organizational culture is more conducive to effective knowledge refinement? What organizational reward systems would effectively promote the quality of refined knowledge? Additional examination of the relationships among the antecedents would also be useful, in particular because of the correlations that were found.

Refinement practices are now increasingly common in contexts outside of knowledge repositories. The open-source environment for software development, for instance, has adopted a variety of refinement practices (Halloran et al. 2002). How concepts and relationships discussed here can be extended or modified to predict knowledge quality and usage in other contexts should certainly be explored in the future.
Implications for Managerial Practices

This study is one of the few that focus on how knowledge stored in repositories is refined to increase its value to an organization. Without refinement, the overall quality of the knowledge in a repository will suffer, and an organization will find it increasingly difficult to accomplish its objectives. While these objectives might appeal to common sense, it is striking how little research has focused on this process.

Refinement research has implications for the management of communities that support the collaborative creation of knowledge repositories. This research highlights the importance of procedural justice in successfully motivating quality knowledge contribution and refinement. As these organizations grow in size, disputes over the refinement process or outcomes would surely become more commonplace (Butler et al. 2008). Promoting and maintaining procedural justice might be a key element in the successful operations of these collective commons.

This research also has implications for the management of innovation. A formal review process is often implemented when firms evaluate innovative ideas for further development (Lafley et al. 2008). Results of this research could shed light on the extent to which the perceived level of procedural justice affects the quality and quantity of innovative ideas that are contributed and shared by a firm’s intellectual capital. Moreover, the present research could help managers develop organizational processes that review and refine innovative ideas more effectively.

Limitations

As with all studies, this research is subject to a number of limitations. First, although instructions given to the participants were presented in a neutral fashion, it is likely that participants nominated tips that they recalled particularly favorably or unfavorably. This potential selection bias towards more memorable tips could have reduced the representativeness of the tip sample included in the study.

Moreover, the sampling procedure used in the study excluded tips that were rejected for publication after the validation process. Because tips rejected for Eureka inclusion were never published for user access, they could not possibly have been nominated for the present study. Sampling only tips that were approved by validators has likely limited the range and variation of quality ratings, and skewed the ratings towards higher scores. Future research should strive to overcome these methodological limitations using more creative sampling strategies.

More research is needed to overcome the limitation of the single-system, single-company research context. It should be noted that Xerox is a product-based organization in a high volatility context (Kankanahalli et al. 2003). Findings from Eureka may not generalize to the management of repository systems for other types of organizations or other contexts without caution. This case-study design allows greater control for systems-level variations that are beyond the focal interest of the present study. At the same time, however, generalization of the study’s findings to other contexts must be made with consideration of the specific context. It is possible that the findings are dependent upon the unique organizational setting of Xerox, and the nature of technical tasks that Eureka is designed to support. Future research should validate the research model in multiple organizations.

Conclusion

This study contributes to the literature by defining, theorizing and empirically examining the process of knowledge refinement. For many years to come, researchers will continue to be concerned with the manner in which organizations manage the knowledge that is often crucial to their success. This research has identified and studied several relevant constructs which will enable researchers to offer broad guidance to practitioners and also to provide a strong foundation for future studies that include additional constructs and relationships to inform the future scholars of knowledge refinement. We hope these initial findings about knowledge refinement will stimulate more innovations in this important area of research.
Acknowledgement

The authors would like to thank the International Business Center at the University of Pittsburgh, and Michel Boucher, Database Quality Coordinator of the Xerox Eureka system, for their support of this research.

References

Boucher, M. "Growing community knowledge through a socio-technical tip sharing system," Xerox, Montreal, Canada.


Joshi, K. "The measurement of fairness or equity perceptions of management information users," *MIS Quarterly* (13:3) 1989, pp 343-358.


