PEER-BASED COMPUTER-SUPPORTED KNOWLEDGE REFINEMENT

AN EMPIRICAL INVESTIGATION

Nonexpert peer-based knowledge refinement, it turns out, is just as helpful as expert-centric knowledge refinement for improving the quality of results.

BY KWANGSU CHO, TINGTING RACHEL CHUNG, WILLIAM R. KING, AND CHRISTIAN SCHUNN

Knowledge management (KM) repository-based systems (such as those involving "best practices" and "lessons learned" applications) are generally costly to operate since they require expert judgment to determine which knowledge submissions are to be included in the system and refined to make them as efficacious as possible.

Empirical evidence in cognitive psychology suggests that experts may not be needed for such refinement when the knowledge consumers are nonexperts. The knowledge "distance" between experts and nonexperts may indeed weaken expert-centric knowledge refinement. In addition, peer judgments delivered to nonexpert end users may substitute well for expert judgments. Multiple peer judgments, especially when provided through computer-supported knowledge-refinement systems, may be much less costly and just as good or perhaps even better than expert judgments, since peers are likely to think more like nonexpert users.

A computer-support system is helpful for facilitating peer-based knowledge refinement, since more peers than experts are probably required for peer-based refinement

THE PROBABILITY OF A SERIOUS PROBLEM BEING DETECTED INCREASES AS THE NUMBER OF PEER REVIEWERS GROWS.

to be effective. Here, we present the results of an experimental study we conducted and a corporate application that confirm our hypothesized equality or superiority of peer-based knowledge refinement compared to expert-centric knowledge refinement. We make no attempt to compare the actual costs of using individual experts vs. multiple peers; we address only the issue of comparing the quality of the results of the two options.

Knowledge repositories are KM systems that play a crucial role in extending organizational memory, preventing organizational forgetfulness, and leveraging competitive advantage. When an organization solicits knowledge from its employees, it must evaluate the submitted knowledge and refine it into a form that is most useful for system users. A traditional assumption in the knowledge-refinement process suggests that experts "provide a framework for evaluation and incorporating new experiences and information" [5].

However, when the users of a repository-based KM system are nonexperts, the knowledge selected and refined by experts may not necessarily be most relevant or useful to those users. For example, technical-support staff members usually access knowledge-repository systems for topics that are unfamiliar to them. And scientists and engineers working on radical innovations often search beyond their own domain expertise for reusable ideas. They may be experts in their own domains but are nonexperts when searching for and evaluating knowledge created in other domains.

Research in cognitive psychology suggests that experts and nonexperts are fundamentally different in the way they represent knowledge and solve problems [3]. Experts have plentiful domain-specific knowledge that is highly organized, while nonexperts have loosely organized knowledge. Compared to nonexperts, experts are quicker at detecting problems; they need fewer cues, access their memory rather than focus on the task at hand, use heuristics rather than exhaustive search, recognize data patterns, and use compiled rules and response plans.

Paradoxically, experts may also have "spectacularly narrow" views on knowledge (from the nonexpert user's point of view). For instance, a study [1] showed that in reading and evaluating requirements documents, better results are achieved when software developers are able to first explore and absorb the perspectives of others (such as testers and users), in addition to their own.

One well-known expert behavior is the way experts tend to underestimate the difficulty of tasks to be performed by nonexperts. Laboratory experiments have shown that people who know a solution to a problem tend to underestimate the difficulty others face in solving it. Similarly, researchers have found that readers with extensive topic knowledge have trouble estimating how well readers with limited topic knowledge will understand the same topic.

Experts also tend to overestimate the potential performance of nonexperts, probably because they detect and diagnose problems on the basis of knowledge readily available in their memories. Experts were shown in [8] to overestimate nonexperts' performance by underestimating the time the nonexperts needed for task completion.

PEER-BASED KNOWLEDGE EVALUATION

When nonexperts code their knowledge into a form that is later refined by an expert, their knowledge differences may prevent them from comprehending the results, even when the refinement is of high quality. Since experts often anchor the knowledge-refinement process to knowledge that nonexperts may not be able to access, such feedback often leads to misunderstanding by nonexpert users. Moreover, explanations that are helpful to experts may not necessarily be helpful to nonexperts.

Contrary to intuition, similarities among peers can actually facilitate knowledge refinement. When the intended audience consists of nonexperts, a group of peers is often better than experts in providing evalua-

tions and feedback on peer-developed materials. Peers are more likely to share knowledgebases, experience, and problems; such socially shared cognition enables them to establish common ground that stimulates development of mutual knowledge, along with an error-correction mechanism. Because peers are closer to one another cognitively and behaviorally, they better understand what they need in a codified piece of knowledge than do experts.

The apparent superiority of peers is presumably due to nonexperts being better at detecting or diagnosing problems that are relevant to their peers. They are more accurate than experts at understanding other nonexperts' problems because they use more cues and more exhaustive search strategies and do not restrict themselves exclusively to private knowledge. These qualities are likely to make nonexperts more adept at coaching their peers in problem solving.

Peer benefits are augmented through the interaction of multiple peers. The value of multiple peers was recognized in [10], observing that large groups of average people can make better decisions than those made by a single expert when they contribute diverse unbiased opinions. When multiple peers participate in knowledge refinement, the aggregated benefits can be significant. One possible reason is that multiple peers create a larger search space for potential problems in the target knowledge; more reviewers find more problems. Also, multiple peers can make a serious problem in the material more salient for the contributor. The probability of a serious problem being detected increases with the number of peer reviewers. When more than one reviewer points out the same problem, the contributor has more reason to take the comment seriously. Moreover, multiple peers counterbalance idiosyncratic biases from individual reviewers, making the refinement process even more reliable.

Peer-Based Refinement System
Here, we describe peer-based knowledge refinement supported by a system described in [4]. Because knowledge is usually submitted in the form of a document, the system facilitates the processing of documents (such as those submitted to best-practice repositories). First, knowledge needs are solicited and potential contributors asked to submit knowledge in, say, the form of proposed "best practices" they have developed for a particular procedure, method, or problem. The first drafts authors submit to the system serve as raw material for the refinement process; the system identifies a diverse set of nonexpert peers to serve as reviewers for each submitted draft. The contributors' identities need not be revealed.

Peer reviewers make qualitative, then quantitative assessments of each submission on the basis of multiple criteria; for this study, the reviewers used flow, logic, and insight as the criteria. They submitted written comments on each of the dimensions. They also rated the quality of each draft on a seven-point scale, with 0 the lowest and seven the highest. After the system collected all evaluations from the reviewers, the system integrated them and computed quality measures for each knowledge submission, as well as the reliability of the individual reviewer evaluations based on built-in reliability-estimation algorithms. At this point, drafts were "stamped" with these initial quality measures based on aggregated ratings (weighted by the computed reliability). The system provided authors with reviewer comments, as well as quantitative quality measures, based on which submitters revised and resubmitted the knowledge document. Each contributor was also able to evaluate the helpfulness of each reviewer's feedback in terms of refining the document.

During the final stage, reviewers from the first-round review process read the revised drafts and generated another round of written comments and ratings the system again used to compute quality measures for the final drafts and reviews. In addition to evaluating the documents, the reviewers determined whether they qualified for inclusion in the repository. Again, contributors assessed the usefulness of these second-round reviews. Synthesizing reviewer recommendations, two rounds of reviews, and contributor evaluation of review quality, the system selected the most valuable documents to be stored in the repository in which contributors and reviewer identities are revealed, recognizing their contribution.

Experimental Study
Our empirical study, conducted at the University of Pittsburgh, simulated the knowledge-refinement process in an experiment that allowed us to observe the relative impact of experts vs. peers on the quality of codified knowledge intended for use by nonexperts. We compared quality improvement in technical reports refined by feedback from a subject-matter expert, a nonexpert peer, or multiple nonexpert peers. The subject-matter expert was an academic scholar with a Ph.D. and social-science expertise. The nonexpert participants were 28 undergraduate students with an average of 3.4 college years enrolled in a social science course not taught by the expert but covering a portion of the expert's domain knowledge.

Individual participants wrote and revised drafts that partially fulfilled the requirements of the course.

The same participants, serving as reviewers and refiners, evaluated six peers' first and revised drafts. Prior to the experiment, participants were tested on basic document-generation skills. Based on the scores, they were matched into blocks, then randomly assigned to one of three feedback source conditions: single expert (SE), single peer (SP), and multiple peers (MP). In the SE condition, the author received only the expert's review. In the SP condition, the best review of the six nonexpert reviews was selected and made available to the author. And in the MP condition, participants received all six peer reviews but not the expert review. To minimize bias, the participants were blind to reviewer identities and status, whether their documents were evaluated by nonexpert peers or by an expert, contributor identity, and who could receive reviewer feedback. The participants were all told that the anonymous reviewers alone would determine the quality of their submitted documents.

About three months after the experiment, the expert reevaluated 30% of the same documents (randomly selected). There was significant agreement between the expert's first and second evaluations, showing strong test-retest reliability. A second expert then independently evaluated the expert and nonexpert evaluations while being blind to the source of evaluations and showed a statistically significant correlation with the expert.

After being assigned to experimental "treatments," the course instructors introduced the participants to the system. All remaining procedures were managed online by the system. On the system's prompt, all participants submitted their first drafts online by the same date. Each participant then received six drafts (randomly selected by the system). At the same time, the expert reviewed all the drafts. The expert and nonexpert reviewers assessed the drafts on the three evaluation dimensions: flow (document readability, emphasizing organization and overall composition); logic (the degree to which the arguments presented in the document were supported by evidence); and insight (whether the document provided innovative ideas that indicate creativity).

Even though each document received one expert and six peer reviews, only some of the reviews were made available to each of the contributors, depending on the experimental condition. After receiving feedback, each participant submitted revised drafts, evaluated the helpfulness of the feedback, then received the second reviews on the revised drafts. Contributor evaluations of these reviews were not available to the reviewers.
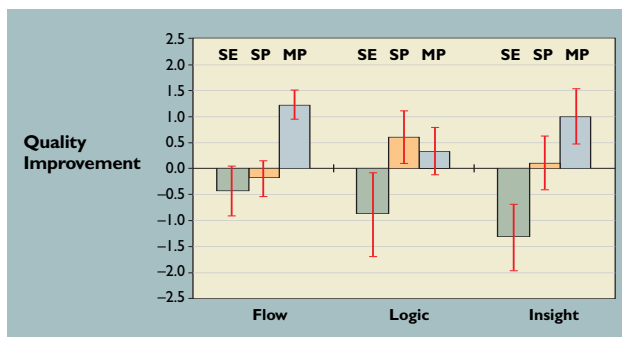

Figure 1. Quality improvement with standard error bars.

RESULTS OF THE EXPERIMENT

We assessed the differences in the impact of expert-based vs. peer-based knowledge refinement through a two-way mixed ANOVA on quality improvement from first to second drafts as measured by expert evaluations. Simple main-effect analyses revealed that flow in MP was significantly better than in SE and SP, that logic was the same across all conditions, and that insight in MP was significantly better than in SE (see Figure 1). Documents reviewed by multiple peers undergo significant quality improvement in both flow and in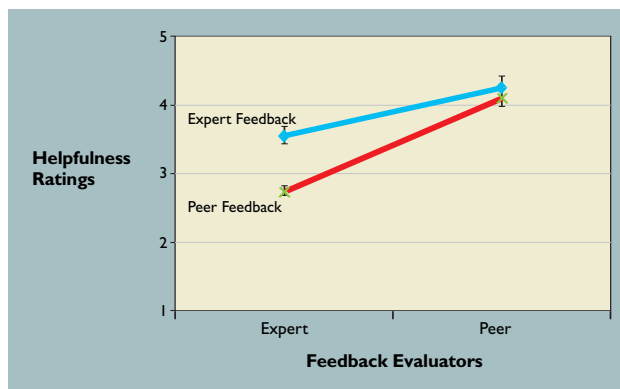sight. This result supports our hypothesis that knowledge distance between experts and nonexperts may hinder knowledge refinement and our hypothesis that knowledge refinement benefits from multiple peer feedback. The effect was particularly strong on the insight criterion, suggesting that multiple peer reviews are likely to stimulate creative ideas. In addition, peers judged expert and peer feedback as equally helpful, while experts judged expert feedback as more helpful than peer feedback, suggesting that when the target audience consists of nonexpert users, nonexpert refiners are just as helpful as expert evaluators (see Figure 2).

Figure 2. Expert and peer perception of expert and peer feedback.

WHEN THE TARGET AUDIENCE CONSISTS OF NONEXPERT USERS, NONEXPERT REFINERS ARE JUST AS HELPFUL AS EXPERT EVALUATORS.

REAL-WORLD APPLICATION

We verified these experimental findings with data we collected from an IT consulting firm in Korea that had implemented a procedure enabling peer-expert comparisons in its customer service process. It asked customers to evaluate answers ("thumbs up/satisfactory" or "thumbs down/unsatisfactory") provided by customer service representatives. Answers to customer questions in an existing database were provided by either expert designers or programmers. Out of 245 answers, 134 (55%) were refined by peers, and 111 (45%) were refined by experts before customers received responses. Among the 162 (66%) rated by customers, 116 (71%) were voted satisfactory and 28 (46%) unsatisfactory. Meanwhile, 79% of the peer-refined responses were satisfactory, and only 62% of the expert-refined responses were satisfactory (see Figure 3). In contrast, 21% of the peer-refined responses were unsatisfactory, whereas 38% of the expert-refined responses were unsatisfactory. These differences are statistically significant, giving further credibility to the generalizability of the results of our experiment.

Figure 3. Customer satisfaction ratings of answers refined by experts and by peers.

RESULTS

Our study represents one of the first attempts to empirically investigate properties of the knowledge-refinement process. The results show that, as we hypothesized, the knowledge "distance" between experts and nonexperts impaired expert-based knowledge-refinement, while the close knowledge distance among peers facilitated knowledge refinement.

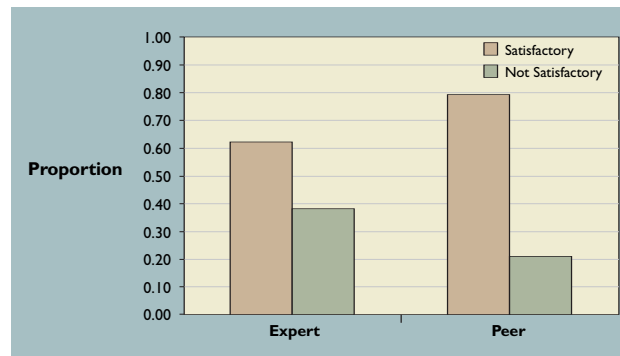Considering that the work of experts is rarely done without time constraints, expert-based knowledge refinement may be worse. Experts under time pressure often default to such typical expert behaviors as using fewer cues than nonexperts to build task representations. They exhibit these behaviors by activating existing task schemata in memory. By relying more on their expert instincts under time pressure, experts may be less likely to take the time to consider the nonexpert's perspective.

Computer-supported peer review systems offer what may be a more financially affordable option. Contributors and reviewers alike may be rewarded for their service by publishing their names when their knowledge is selected for inclusion in a repository. Thus, contributors and reviewers alike enjoy substantial social rewards (such as citation credit and peer approval). This was observed in an informal knowledge repository in a software development firm [6] and is a well-known characteristic of the Linux virtual development team [9]. Over time, such repositories enable development of reputation systems, helping users assess likely knowledge relevance and quality. A similar reputation system at the group level has directed user attention to high-quality content in an electronic document database [7].

A peer-based knowledge-refinement process designed to help users assess content quality should gradually build trust as a result of the creation of a virtual community. When a large number of documents is available in such a repository, the system becomes a knowledge market that competes for the reader's limited attention [7]. Systems allowing users to inspect feedback from other users may help maintain trust in the repository in a way that's similar to the way online merchants (such as Amazon.com and eBay.com) establish consumer trust through a peer-feedback mechanism.

Peer-based knowledge refinement makes it possible for organizations to establish reward systems for both

knowledge codification and refinement. Such systems may use extrinsic incentives, as well as social reputation rewards, to motivate employees to participate in knowledge refinement. Employees rewarded through this process may be more likely to use knowledge from the repository system for their own tasks. These benefits may explain the success of peer reviews (such as software pattern writers relying on peer discussions when writing up patterns). **C**

REFERENCES
1. Basili, V., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Sørumgård, S., and Zelkowitz, M. The empirical investigation of perspective-based reading. *Empirical Software Engineering 1,* 2 (1996), 133–164.
2. Besnard, D. and Bastien-Toniazzo, M. Expert error in troubleshooting: An exploratory study in electronics. *International Journal of Human-Computer Studies 50* (1999), 391–405.
3. Chi, M.T.H., Feltovich, P., and Glaser, R. Categorization and representation of physics problems by experts and novices. *Cognitive Science 5* (1981), 121–152.
4. Cho, K. and Schunn, C. Scaffolded writing and rewriting in the discipline: A Web-based reciprocal peer-review system. *Computers and Education 48,* 3 (2007), 409–426.
5. Davenport, T. and Prusak, L. *Working Knowledge: How Organizations Manage What They Know.* Harvard Business School Press, Cambridge, MA, 1997.
6. Dingsoyr, T. and Royrvik, E. An empirical study of an information knowledge repository in a medium-sized software consulting company. In *Proceedings of the International Conference on Software Engineering* (Portland, OR, 2003).
7. Hansen, M. and Haas, M. Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company. *Administrative Science Quarterly 46,* 1 (2001), 1–28.
8. Hinds, P. The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied 5,* 2 (1999), 205–221.
9. Lee, G. and Cole, R. From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organization Science: A Journal of the Institute of Management Sciences 14,* 6 (2003).
10. Surowiecki, J. *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations.* Doubleday, New York, 2004.

KWANGSU CHO (kwangsu.cho@gmail.com) is an assistant professor in the School of Information Science and Learning Technologies at the University of Missouri, Columbia.
TINGTING RACHEL CHUNG (ting@pitt.edu) is a Ph.D. candidate in the Katz Graduate School of Business of the University of Pittsburgh, Pittsburgh, PA.
WILLIAM R. KING (billking@katz.pitt.edu) is a University Professor in the Katz Graduate School of Business of the University of Pittsburgh, Pittsburgh, PA.
CHRISTIAN SCHUNN (schunn@pitt.edu) is an associate professor in the Learning Research and Development Center of the University of Pittsburgh, Pittsburgh, PA.