Experts Learn More (than Newcomers): An Exploratory Study of Argumentation in an Online Help Forum

Hon Jie Teo, Aditya Johri, Virginia Tech, USA
Email: hjteo@vt.edu, ajohri@vt.edu

Abstract: Online help forums have emerged as a powerful form of knowledge sharing with significant implications for CSCL. To understand knowledge construction and learning within online help forums, this study examined the interactional dynamics amongst help-givers (experts) and help-seekers (newcomers) in an open, voluntary help forum for Java computer programming language. A purposive sample of active threads was analyzed using a coding scheme adapted from the argumentative knowledge construction (AKC) framework. The study found that while help-givers actively engaged in advanced argumentative and social moves, help-seekers seldom did. Furthermore, the number of advanced epistemic messages was lower than expected and both help-givers and help-seekers often veered off-topic towards socialization and conflict-oriented interactions. The study questions the feasibility of online help forums for facilitating acquisition of domain-specific knowledge by help-seekers but demonstrates the viability of the platform as a mechanism for experts to expand their knowledge. Our findings suggest that those who know more are able to learn more given their ability to engage in argumentation.

Introduction

With the increase in access and use of digital media among youth (Johri, Teo, Lo, Dufour, & Schram, 2013), research on online learning has gained increased significance. Within CSCL, contexts for research have included online virtual math chat groups (Stahl, 2006, 2011), online mathematics help forums (van de Sande, 2010; van der Sande & Leinhardt, 2007), forum for mathematics teachers (Remniger & Shumar, 2002), online distance learning classes (Johri, 2005), free non-course-related mathematics forum (Chen, Chiu & Wang, 2010), and K12 programming communities such as Scratch (Resnick et al., 2009) and MOOSE Perl (Bruckmann, 2006). Researchers are interested in online contexts because of their ability to augment formal learning by supporting knowledge construction (Scardamalia & Bereiter, 2006; Law, Yuen, Wong & Leng, 2011), sense-making (Kirschncher, Buckingham Shum & Carr, 2003) and facilitating complex problem solving (Munneke, Andriessen, Kanselaar & Kirschner, 2007). Furthermore, the online discussion can be enhanced through the use of mechanisms such as ‘scripts’ which allow learners to engage in discourse of high epistemic quality (Stegman, Wecker, Weinberger & Fischer, 2007). The viability of asynchronous online discussions in fostering knowledge construction and acquisition has been well-documented. Analytical frameworks (Suthers et al., 2010; Weinberger & Fischer, 2006) and content analysis instruments (Clark & Sampson, 2007; Gunawardena et al., 1997; Pena-Shaff & Nicholls, 2004) have been used to understand learning in online settings. In this study, we adopt the argumentative knowledge construction (AKC) framework proposed by Weinberger and Fischer (2006) to obtain an in-depth picture of the interactions between help-givers and help-seekers in an unscripted online forum (1). Our interest in this study and the framework emerged from two primary research questions about online help forums – what does learning look like and who learns by participating in these communities? As we discuss in detail below, our use of the Weinberger and Fischer (2006) framework is that it lends itself well to our goal – an examination of the quality of arguments and the degree of social interaction in an unscripted environment where contributions are voluntary and expertise levels may vary amongst learners.

Argumentative Knowledge Construction

Knowledge construction is a useful framework for analyzing collaborative online activity as it emphasizes the role of social activity in negotiating meanings relevant to the learning task (Dillenburg, 1999; Stahl, Koschmann & Suthers, 2006). Through dialogue, participants are able to create a coordinated activity where they can maintain a shared conception of a problem (Dillenburg, 1999; Scardamalia & Bereiter, 2006; Pena-shaff & Nicholls, 2004; Hmelo-Silver, 2003). Within the knowledge construction framework, researchers interested in the design of learning environments have examined learner interaction and discourse to identify mechanisms that make knowledge construction viable and have found that argumentation activities support help students learn about argumentative structures and constructing explanations (Andriessen, 2007; Berland & Resier, 2008; Toulmin, 1958). Argumentative knowledge construction (AKC) has emerged as a useful construct for framing and understanding learning tasks. It emphasizes problem discussion by making argumentative elaborations with the goal to contribute towards multiple perspectives, produce joint-solutions, and acquire domain-specific knowledge (Astleitner, Brunken & Leutner, 2003; Jamaludin, Chee & Ho, 2009; Stegmann, Weinberger & Fischer 2006). To achieve AKC, learners collaborate and exchange their knowledge through discourse on an
open and complex learning task through the construction of arguments, balance of arguments and supporting their arguments with evidences, and counter arguments (Weinberger, 2003; Weinberger & Fischer 2006). The examination of argumentative discourse produces further insights into how learners engage in negotiation of different perspectives through theoretical conceptualization and application (Weinberger & Fischer, 2006). For instance, Stegmann and colleagues (2007) studied a problem-based learning in a higher education setting and found that argumentative collaboration scripts can foster the quality of argumentations, such that there are more arguments contain verifiable claims and qualifying statements, and help learners acquire knowledge about argumentation. However, they reported that there are no learning gains when it comes to the acquisition of content-specific knowledge. Overall, the ability of the knowledge construction framework to illuminate joint activity where learners improve upon initial ideas through dialogue makes is pertinent for studying online help forums. In these forums interaction takes place through text based dialogue which is stored and available for analysis and this is the only form of interaction that takes place among the participants thereby increasing the usefulness of the data in terms of interpretation.

Research Setting – Java Newcomer Forum

Online help forums are as old as electronic communication itself and their continued viability and growth over the years is an indicator of their usefulness to users. In recent years, one form of online communities – help forums – have flourished in membership and have become a powerful complement to informal learning opportunities available to people and both the level of participation as well as the diversity of topics covered by help forums is astounding (Singh, 2008). In the seven websites she sampled, Singh (2008) found approximately 3.8 million registered users excluding the guests who post and the visitors who lurk on the forums as invisible visitors and never post or register. One arena in which help forums have had a persistence and successful presence is computer programming. The usefulness of online help-forums for learning to program is easy to see. It is easy to represent software code online and relatively straightforward for the help giver to ‘run’ the code and see if it works. For our study, we targeted a popular programming language with millions of user worldwide and significant online support – Java™. The popularity of Java comes from its usefulness for developing Internet related applications and the fact that it is open-sourced. Open sourcing of any software artifact – language, program or product – results in significant growth in the supporting online community (Crowston & Howison, 2003) and Java is no exception (Johri, Nov & Mitra, 2011a, 2011b). Studies have found that users in online communities that are formed around open source software actively engage in collaboration, learning and socialization (Lakhani & Wolf, 2005) and leverage tools such as code repositories, forum discussion, blogs and chat clients to facilitate interaction (Crowston, Wei, Howison & Wiggins, 2012). In these communities help-seeking and help-giving are common and Lakhani and Von Hippel (2003) found that help-givers spent 2% of the total spent time on the online community answering questions and 98% of their time reading questions as well as crafting potential answers to the questions. Singh (2008) found that the process of help-giving in these communities is many-to-many (many help-seekers and many help-givers are involved) and that the process of help-giving is seldom dyadic and often iterative (Singh & Twidale, 2008; Singh, Kathuria & Johri, 2012).

In contrast to formal and highly structured instruction, the site we selected for this study (the Java community) has an informal structure and both participation and collaboration is voluntary and not mandated by coursework. Help is mainly provided by a core group of volunteers and the discussion task structure deviates from the common set-ups of common wrapper/starter roles and open-ended class discussions without pre-designated roles. This setting is problem driven as help-seeker starts a discussion soliciting help from voluntary help-givers to assist them with their learning needs and that challenges emerge through the engagement of newcomers with a programming task. Figure 1 shows a screenshot of the forum and Figure 2 represents the structure of a sample discussion. The forum explicitly prohibited off-task discussion – such as discussion of Java as a language or its open sourcing – in the forum. The terms of participation made this explicit.

![Figure 1](image_url). A page listing discussion topics in the “New to Java” discussion forum
Methodology, Data Collection, and Data Analysis

The discussion forum data that we collected consisted of a total of 37,472 discussion topics. Our data collection complied with both internal institutional IRB protocol and the 'terms of services' of the forum. We limited data collection to public data that did not require creating an account and logging in to the site to be able to read the forum messages. The data ranged from a time period between 2006 and 2010. User post counts (see Table 1) indicates that most users have less than 10 posts in this discussion forum with only 326 users having post counts more than 500. Therefore, similar to other help forums, the majority of users participate infrequently in contrast to a much smaller number of expert helpers who contribute much more frequently to the forums. For the study reported in this paper, data were collected from the most active discussion topics created in the month of September in year 2009 and which accounted for 9.4% of all topics started in the month. By only examining the most active discussion topics, we are able to focus on the entirety of discussion to account for a wide range of domain-specific content, argumentative and interactional moves. In total, 1119 messages from 47 discussion threads amounting to a total of 108 pages of conversations were analyzed. Further details of the research site and data collection are available in Mitra (2011).

<table>
<thead>
<tr>
<th>Post Count</th>
<th>Number of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 10</td>
<td>14114</td>
</tr>
<tr>
<td>11 to 50</td>
<td>5076</td>
</tr>
<tr>
<td>51 to 500</td>
<td>1922</td>
</tr>
<tr>
<td>&gt; 500</td>
<td>326</td>
</tr>
</tbody>
</table>

Our coding scheme is based on the three independent dimensions: epistemic, argumentative and social modes (See Table 2) as described by Weinberger & Fischer (2006). The epistemic dimension allows us to investigate if learners are on-task and if they are able to convey theoretical concepts to explain concepts or apply concepts to provide a solution to the case at hand. The argumentative dimension is focused on the argumentation sequences and whether learners make arguments, counter arguments, and interactive arguments in addition to drawing from prior knowledge. The social modes dimension is concerned with the extent to which the learners externalize their thoughts, question others, accept the contributions of others and disagree or agree with the perspectives of others. While the analytical framework (Weinberger & Fischer, 2006) is useful for guiding analysis at both the micro and macro levels, a macro-approach was chosen such that the complete message was chosen as the unit of analysis for the coding. This is in line with the suggestion by Rourke et al. (2001), who
argued that taking the complete message as the unit of analysis is the most objective approach as data analysis is performed under the entirety of interaction. The inter-rater agreement for this study is 0.78 (Cohen, 1960).

Table 2: Coding scheme with categories in epistemic, argumentative and social dimensions (adapted from Weinberger & Fischer, 2006)

<table>
<thead>
<tr>
<th>Code</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPI1</td>
<td>Non-epistemic activities</td>
<td>Learners discussing off-topic or digressing off-topic</td>
</tr>
<tr>
<td>EPI2</td>
<td>Construction of problem space</td>
<td>Learners convey contextual information within the argumentation space with the aim of fostering understanding of the topic</td>
</tr>
<tr>
<td>EPI3</td>
<td>Construction of conceptual space</td>
<td>Learners convey theoretical concepts and explain theoretical principles to foster theoretical understanding of the topic</td>
</tr>
<tr>
<td>EPI4</td>
<td>Construction of adequate relations between conceptual and problem space</td>
<td>Learners convey theoretical concepts of case information and apply relevant theoretical concepts adequately to provide a solution for the topic problem</td>
</tr>
<tr>
<td>EPI5</td>
<td>Construction of relations between prior knowledge and problem space</td>
<td>Applying concepts that stem from prior knowledge rather than the new theoretical concepts that are to be learned</td>
</tr>
<tr>
<td>ARG1</td>
<td>Non-argumentative moves</td>
<td>Questions, coordinating moves, and meta-statements on argumentation</td>
</tr>
<tr>
<td>ARG2</td>
<td>Argument</td>
<td>Statement put forward in favor of a specific proposition</td>
</tr>
<tr>
<td>ARG3</td>
<td>Counterargument</td>
<td>An argument opposing a preceding argument, favoring an opposite proposition</td>
</tr>
<tr>
<td>ARG4</td>
<td>Integration</td>
<td>Statement that aims to balance and to advance a preceding argument and counterargument</td>
</tr>
<tr>
<td>SOC1</td>
<td>Externalization</td>
<td>Externalizing or articulating thoughts to the group</td>
</tr>
<tr>
<td>SOC2</td>
<td>Elicitation</td>
<td>Eliciting a response, questioning the learning partner, or provoking a reaction from group</td>
</tr>
<tr>
<td>SOC3</td>
<td>Quick consensus building</td>
<td>Accepting the contributions of group members in order to move on with the task</td>
</tr>
<tr>
<td>SOC4</td>
<td>Integration-oriented consensus building</td>
<td>Taking over, integrating and applying the perspectives of group members</td>
</tr>
<tr>
<td>SOC5</td>
<td>Conflict-oriented consensus building</td>
<td>Disagreeing, modifying or replacing the perspectives of group members</td>
</tr>
</tbody>
</table>

Findings

Our analysis (see left most bar in Fig. 2) indicates that non-epistemic (EPI1) messages comprised 42.2% of all messages corresponding to the epistemic dimension followed by epistemic activities of higher order such as conveying of contextual information (EPI2) at 32.6% and conveying of theoretical concepts (EPI4) at 17.6%. We, however, also found that both help-givers and help-seekers seldom explicitly drew from any prior knowledge (EPI5) at 1.6%. Chi-square analysis showed that help-givers are more likely to engage in non-epistemic messages ($\chi^2 (18, N = 94) = 43.5, p < 0.001$). Together with the previous finding, that there are a large number of non-epistemic messages, this significant difference suggests that as help-seekers engage in help-oriented discussions they do not at the same time actively participate in any off-task activities, such as engaging in socialization moves. On the other hand, experts are more likely to engage in higher level epistemic activities from explanation of theoretical concepts ($\chi^2 (5, N = 94) = 38.5, p < 0.001$) to application of theoretical concepts ($\chi^2 (10, N = 94) = 60.6, p < 0.001$). This finding reflects the nature of help-oriented discussion where help-givers have more expertise and more likely to share their theoretical and conceptual knowledge to provide a solution to the case at hand or to explain theoretical concepts.

In the argumentative dimension, we found that non-argumentative moves (ARG1) constitute 90.6% of all argumentative messages and made up the majority of messages in the argumentative dimension (see middle bar in Fig. 2). This is not surprising since topics started by help-seekers may not be entirely conductive for AKC. Amongst the argumentative moves, the occurrences of ARG2, ARG3 and ARG4 make up 37.5, 41.3 and 21.2% of the total moves. Our analysis also highlight that help-seekers seldom engaged in any of these argumentative moves at 99%. Chi-square analysis showed that that help-givers are more likely to engage in argumentative moves ARG3 and ARG4 which are putting forth arguments ($\chi^2 (4, N = 94) = 16.1, p = 0.0013$) and counterarguments ($\chi^2 (5, N = 94) = 13.8, p = 0.0058$) and integrative arguments ($\chi^2 (3, N = 94) = 16.5, p < 0.001$). This discussion platform, being a help-oriented online discussion forum, is made up of help-givers dominantly engaged in putting forth their arguments when the topic at hand requires it. This seems to suggest that help-seekers are not actively participating in the argumentative processes as they may lack the knowledge needed to engage in any arguments. It is therefore suspect that the multiple perspectives that originate from help-givers’ arguments have any positive impact on help-seekers’ learning experiences, especially when experts
touch on advanced domain-specific content without explicit explanation of the content. The limitation of this finding is that we are not able to detect any form of lurking behavior and therefore cannot account for help-seekers who are observing yet do not make a contribution to the discussion.

In the social dimension, the use of questions or provocation to elicit responses (SOC2) constitutes 32.0% of all social messages whereas externalization of thoughts occurred at a frequency of 22.7%. On the other hand, quick consensus building and conflict-oriented consensus building messages (SOC3 and SOC4) constitutes 14.0% and 26.4% of all social messages respectively (see right most bar in Fig. 2). The relatively higher number of SOC2 moves can be explained by the nature of the help-oriented discussions where information or prompts, rather than solutions, was provided to enable help-seekers to meeting their learning needs (Lakhani & von Hippel, 2003). It is also observed that both help-givers and helps-seekers seldom participate in moves that show integration of others’ perspectives into their viewpoints (SOC5) and these moves only feature in 4.9% of all social messages. Through chi-square analysis, our analysis suggests while help-seekers are more likely to use questions and provocation to solicit responses, help-givers are more likely to take over the perspectives of others ($\chi^2 (4, N = 94) = 9.5, p < 0.028$) and are more likely to engage in conflict-oriented consensus building social moves ($\chi^2 (13, N = 94) = 44.7, p < 0.001$). There are no significant differences in our examination of the frequencies of SOC2 ($\chi^2 (7, N = 94) = 8.6, p < 0.001$) and this suggests that both help-seekers and help-givers are also engaged in questioning and provocation, which is a positive aspect of the online help discussion. To complement the discussion of our findings, we also present Table 3 which shows an excerpt for a help-discussion featuring acts from the three dimensions.

![Figure 3. Frequencies of Occurrence of Epistemic, Argumentative and Social Acts](image)

Table 3: Excerpt from a help discussion on the computational concept of “inheritance”

<table>
<thead>
<tr>
<th>Actor (Role)</th>
<th>Message Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Help-seeker (HS)</td>
<td>Hello again, Guys I've read this tutorial: <a href="http://java.sun.com/docs/books/frietutorial/java/landisubclasses.html">http://java.sun.com/docs/books/frietutorial/java/landisubclasses.html</a> As one of you suggested in one of your replies to my question on topic about inheritance and that's fine. Unfortunately in this tutorial two sentences contradict each other: 1. A subclass inherits all the members (fields, methods, and nested classes) from its superclass. Constructors are not members, so they are not inherited by subclasses, but the constructor of the superclass can be invoked from the subclass. 2. A subclass does not inherit the private members of its parent class. Moreover in netbeans while debugging it shows that private members of superclass are inherited by child class. So what's going on? Which sentence is true? And really, but on sun's official website this kind of error shouldn't have place. Looking forward to your reply</td>
</tr>
<tr>
<td>Help-giver (HG) #1</td>
<td>It's really semantics. The only difference it makes if they're inherited or not is what the definition of inheritance must be. However, according to the Java Language Specification, I would say that private members are not inherited. Quote: 6.4.3 The Members of a Class Type says: Members are either declared in the type, or inherited because they are accessible members of a superclass or superinterface which are neither private nor hidden nor overridden. Although it's entirely possible that the JLS contradicts itself elsewhere in that regard. As I said though, it makes no difference.</td>
</tr>
</tbody>
</table>
Online communities represent a significant avenue to understand many learning and education related issues such as newcomer socialization (Ducheneaut, 2005), free user-to-user help (Lakhani & von Hippel, 2003), and distribution of tasks (von Krogh, Spaeth & Lakhani, 2003; Mockus et al., 2002). Through examination of the epistemic, argumentative and social tendencies of help-givers and help-seekers in an open unscripted learning environment, we gained deeper insights into how help discussions can be characterized and how learning is taking place in unscripted online discussion. The high number of off-topic messages raises a concern whether the forums are a productive avenue for significant acquisition of domain-specific knowledge and echoes the concerns raised by Stegmann and colleagues (2007). Both help-givers and help-seekers were likely to veer off-topic towards socialization and conflict-oriented social interactions. In our analysis, we often find that the quality of help is highly dependent on the social interactions between the help-seeker and the help-giver. Taking into account that help-givers are volunteering their time, effort and offering their expertise, it is not surprising that specific social and community practices have been established to allow them to maximize their voluntary efforts. As such, help-seekers who are accustomed to these practices are more likely to attain highly productive learning experiences. In our data analysis, we observed that help-seekers are more likely to engage in argumentative moves and engage in integration-oriented consensus building moves. These findings highlight an opportunity for community leaders to address the design of tasks and roles to address the expertise gap between the experts and newcomers as well as to foster more engaged learning experiences. Online communities have much to gain by encouraging participation from the larger pool of newcomer learners as it will in turn encourage higher levels of collective contributions and learner diversity. Finally, consistent with prior research on expertise we found that due to their ability to organize knowledge and connect concepts – metacognitive proclivity – experts were able to engage more centrally with the discussion, argue strongly, and thereby also construct knowledge and learn (NRC, 2000). Our findings hint at a unique problem with informal learning in that those who know more are able to learn more as they can argue more both due to prior knowledge, and in some instances, the culture of their profession or community.
Acknowledgement(s)
This material is based upon work supported, in part, by the U.S. National Science Foundation under Grants EEC-0954034 & EEC-0935143. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agency.

Endnotes
(1) As other scholars have suggested, the role of ‘help-givers’ and ‘help-seekers’ in online communities is correlated with ‘experts’ and ‘newcomers’ respectively and we use the terms interchangeably.

References

© ISLS


