Blended training and collaborative learning for software users in the geosciences

Skordaki, E.M  
Royal Military College of Canada, Kingston, Ontario, Canada

Bainbridge, S.  
Centre for Distance Education – Athabasca University, Athabasca, Alberta, Canada

ABSTRACT
Geoscientists and geo-engineers use scientific software in applications that can directly affect public safety. Geological hazard risk assessments, water quality geochemical assessments, and rock stability evaluations, all involve the use of scientific software (i.e. numerical modeling programs). Although training on this type of software is necessary to ensure correct scientific decisions, it is currently conducted without the backing of supportive theory. This study examines the delivery of scientific software training, in blended learning settings, through the eyes of software users in their natural setting of practice. Preliminary data reveal that scientific software users perceive training on this type of software as the process wherein the users inform their practices by developing their conceptual skills. This depends on their participation in knowledge mobilization within an online/onsite community of practice, risk management strategies (i.e. checking their software assumptions) and the profile of the user (background/ strengths, motivation regarding distance/self-directed learning or willingness to openly share information).

RÉSUMÉ
Les géologues et les géo-ingénieurs utilisent des logiciels scientifiques pour analyser des situations pouvant directement affecter la sécurité publique. L'évaluation des risques liés aux dangers géologiques, de la stabilité rocheuse et l'évaluation géochimique de la qualité de l'eau impliquent l'utilisation de logiciels scientifiques (p. ex. programmes de modélisation numérique). La formation pour ces logiciels est nécessaire pour assurer de bonnes décisions scientifiques. Cependant, cela est actuellement mené sans l'apport de la compréhension théorique. Cette étude examine l'offre de formation sur les logiciels scientifiques et sur l'apprentissage des paramètres à travers les yeux des utilisateurs de logiciels dans le cadre naturel de leur pratique mixte. Les données préliminaires révèlent que les utilisateurs de logiciels scientifiques perçoivent leur formation comme un processus dans lequel ils partagent leurs pratiques en développant leurs compétences conceptuelles. Cela dépend de leur participation à la mobilisation des connaissances au sein d'une communauté en ligne et dans leur communauté pratique immédiate, de leurs stratégies de gestion des risques (c'est-à-dire vérifier leur hypothèse de logiciel) et du profil de l'utilisateur (connaissances / points forts, motivation au sujet de l'apprentissage à distance ou autodidacte ou la volonté de partager ouvertement l'information).

1 INTRODUCTION
Peer collaboration, albeit in an informal fashion outside of a theoretical framework, is a method used extensively by researchers in science and engineering in the quest to expand their background knowledge (Harp, Satzinger & Taylor, 1997; Fischer, 2011). As such, several researchers in current literature indicate the need for methodical teaching of collaborative (peer) learning skills in academic programs of science and engineering. However, Lingard (2010) points out that employers often report that new hires typically do not know how to communicate and that they have insufficient experience and preparation for working (and continuing to learn) as part of a team. This is a potential result of the ineffective teaching strategies and assessment tools of collaborative learning within teams in academic curricula: "Although many universities have recognized the need to assign group projects and have begun efforts to improve engineering and computer science curricula in this regard, students seldom receive any specific training on how to function collaboratively before such assignments are given, and little attention is given to how teams are formed" (Lingard, 2010). Similar concerns have been cited in literature since the 1990s (McGinnes, 1995; Green, 1999; Berry, Robert Lingard, 2001; Hernández & Ramirez, 2008; Purzer, 2009; Purzer, 2010, Borrego, Karlin, McNair & Beddooesc, 2013). According to Vygotsky (1978), students can perform at higher intellectual levels in collaborative situations than when working individually. Collaborative learning requires "working together toward a common goal" and that "students are responsible for one another's learning as well as their own" (Dooly, 2008). This collaboration entails the whole process of learning, not only the teacher teaching the students; it may involve students teaching one another or students teaching the teacher. Notably, the learners have attained their goal (to enrich their background) by helping each other understand the new concepts at hand (Dooly, 2008).

With technological advancements, tasks involving collaborative learning and completing projects in a team environment have largely moved online. This adds a new vector in the sum of the parameters that determine learning in science and engineering. Modern researchers are encouraged, and even expected by their peers, to share data in online resources and learn to incorporate distance technologies in their daily work routines. Although Olson and Olson (2000) determined that
“distance matters”, and that interactions over distance can never replace collocated interaction, the number of scientific papers published by international collaborations doubled in the last decades (Nentwich, 2008). Miller (2009) argues that online collaboration, as a way of doing scientific research, is becoming more and more common; the work of scientific research is becoming increasingly distributed and collaborative. For example, this tendency is indicated in the formation of collaboratories, i.e., organizations of researchers that, with the help of special technological systems, conduct science in a geographically distributed manner.

As the study of “complex, multidisciplinary, multiphenomena behaviours of large physical, biological, or social systems” often has researchers performing together in larger groups than those that traditionally make up a lab. For many of these projects, equipment and computing are distributed over large distances, new challenges are created for collaborative learning and technological equipment (Cummings, Finholt, Foster, Kesselman and Lawrence, 2008; Miller 2009). Further, this creates a new status quo, where large amounts of data with diverse characteristics are being shared and manipulated on a global scale (Hey and Trefethen, 2008).

Computer science, and more specifically, software developers are presented with a new demanding task: To design software that can address the needs of the scientist/engineer in the new, distributed working environment. As expected, effective training in the use of scientific software in a blended, collaborative learning environment is necessary in order to ensure correct scientific decisions.

Geoscientists and geengineers use scientific software in applications that can directly affect public safety. Water quality geochemical assessments, marine chemistry studies or geological hazard risk assessments. all involve the use of specialized software (including numerical modeling programs). Managing the risk in making errors in the use of scientific software is essential (Hannay, Langtangen, McLeod, Pfahl, Singer & Wilson, 2009; Howison & Hsrbsteb, 2011). Risk in this context refers to the likelihood of unintended, mistaken scientific and engineering decisions based on incorrect use and misinterpretation of data output from scientific software. Collaborative learning in blended settings has been employed by researchers in order to expand their knowledge (Csanyi, Reichl, & Jerlich, 2007; Hannay et al., 2009; Fischer, 2009; Fischer, 2011). However, despite the abundance of training literature, there is limited research that investigates how training is employed by geoprofessionals specifically in the use of scientific software. Thus, the very real problem is that, currently, there is a growing need for identifying successful practices for training in order to accurately apply scientific software. However, at the moment, there is a deficiency in the literature regarding this topic. This paper presents preliminary results from an ongoing research project that focuses on investigating software users’ experiences with respect to training on the reliable application of scientific software. The scope of this study is to create a scientific software training framework for users whose goal is the accurate application of the software. As such, the needs of scientific software users as learners in blended, collaborative environments are examined.

1.2 Training versus Learning

The term training has no generally accepted definition (Bramley, 1986; Mandefrot, 1997). According to Mandefrot (1997), a precise description for learning and training is not available. Training can be a means to bring about learning and create a learning environment where people acquire new knowledge, workers learn, and help each other learn. In literature, however, training is often considered a process that only requires that one learns a specific item by following exact directions (Bramley, 1986; Goldstein & Gessner, 1988). Often, training recognizes only formal instruction, though; it does not include the chance for people to learn through observation, direct experience and from each other. Dearden (1984) gave a more holistic definition of training and linked it to learning: “Training typically involves instruction and practice aimed at reaching a particular level of competence or operative efficiency... Often, training addresses itself to improving performance in direct dealing with things... Other sorts of training are more concerned with dealing with people... Yet other kinds of training are more indirectly concerned with changing or controlling people or things.” In every case what is aimed at is improved level of performance... brought about by learning” (Dearden, 1984, p. 58-59).

What makes Dearden’s definition of training relevant for this research undertaking is that his definition clearly emphasizes the link between training and learning. He indicates that the purpose of training is not the narrow focus of skill acquisition but that of behavioural change, which is a characteristic of comprehensive learning. This research adopts the definition of training by Dearden as it seeks to investigate how dealing with items (scientific software in this case), with people (learners-scientific software users), and with change (from traditional to blended, collaborative learning) can influence the educational processes within the community of scientists and engineers.

1.3 Blended Learning

In terms of defining “blended learning”, Garrison and Kanuka (2004) have pointed out that it is the thoughtful integration of classroom face-to-face learning experiences with online learning experiences. According to current studies, blended instruction has been used by an increasing number of post-secondary institutions to enhance science and engineering research training (Kyriazis, Psycharis & Korres, 2009; Graham, 2013). What makes blended learning particularly suitable for interactions within a community that shares scientific knowledge is its ability to facilitate a community of inquiry; it allows for comprehensive learning to occur within scientific communities as it fosters opportunities for reflection along with independence and increased control essential to developing critical thinking (Garrison & Kanuka, 2004; Graham, 2013). The social interrelations within a community of learners balance out the open communication and limitless access to information on the
Internet. Therefore, blended learning can provide a suitable environment for the learner to benefit from social presence, cognitive presence and teaching presence, cultivating a vigorous community of inquiry: “One of the characteristics of the community of inquiry is that members question one another, demand reasons for beliefs, and point out consequences of each other’s ideas—thus creating a self-judging community when adequate levels of social, cognitive, and teacher presence are evident” (in Garrison, Anderson & Archer, 2001, p. 6). Further, as blended learning environments can afford opportunities for multiple forms of communications, critical properties associated with reliable scientific knowledge and quality higher education are strengthened through free and open dialogue, critical debate, negotiation and agreement (Garrison & Kanuka, 2004; Graham, Henrie & Gibbons, 2014).

2 METHODOLOGY

The study examines the delivery of scientific software training (including numerical modelling programs), in conventional and distance learning settings, through the eyes of the participants in their natural setting of practice. It does not seek to test a particular hypothesis on this field of interest; new concepts on scientific software training will emerge from the themes that will be generated from the analysis of the data collected. In this study, a grounded theory approach is employed. Grounded Theory is a systematic methodology in the social sciences involving the generation of theory from data (Glaser & Strauss, 1999; Strauss & Corbin, 1998). Qualitative inquiry is preferred in this study due to its significant advantages. Strauss and Corbin (1998) indicate that qualitative research is necessary when the researcher seeks to better understand any phenomenon that has not been adequately investigated; little has been documented regarding scientific software user training and its relationship with collaborative, blended learning. Further, qualitative research acknowledges the importance of context, allows for reconsideration of issues which are considered unreliable and subjective in quantitative research. The use of a qualitative data gathering method is highly flexible, allowing for modifications of the research hypothesis as the study progresses (Berg, 1998; Cohen, Manion & Morrison, 2007). According to Cohen et al. (2007) qualitative research reports are typically rich with detail and insights into participants’ experiences around the world, providing the capacity to holistically describe a phenomenon, which can be more meaningful from the reader’s perspective (in this case, the scientific software users). However, qualitative research has shortcomings with respect to the quality of data and objectivity. Although qualitative research allows for a deep understanding of information, the knowledge gained might not generalize to other people or other settings (Patton, 2002). The information provided and the interpretation of the information is subjective due to the human element (Strauss & Corbin, 1998; Patton, 2002). As such, various techniques to enhance and ensure rigour and validity of the study have been employed and are discussed in subsequent sections within this paper.

This ongoing investigation employs an ethnographically-informed qualitative research methodology in order to explore scientific software training. It utilizes open-ended interviews as the primary data source so that research findings are delivered in the words of the participants. The participants are recruited within science and engineering laboratories in academia and industry. Secondary data sources include reviews of laboratory software manuals and pertinent publications (Hammersley & Atkinson, 1995; Robinson, Segal & Sharp, 2007). Extended immersion of the researcher in the participants’ naturalistic setting will not be feasible due to time considerations. The research design of this study is based on ethnography adaptations made by Robinson, Segal and Sharp (2007) in their series of qualitative software development studies. Further, it involves coding of information and constant comparative analysis until saturation of data is achieved in order to construct new theory from the information that will surface from the careful and detailed analysis of the data collected.

2.1 Reliability and Validity Strategies during Data Collection

Qualitative researchers have adhered to a list of five criteria towards evaluation for trustworthiness (a parallel concept to reliability and validity in qualitative research), i.e. i) credibility, ii) dependability, iii) transferability, iv) confirmability and v) authenticity. These criteria are used for the validation of a study after its completion (post hoc) as well as during its development (Lincoln and Guba, 1985; Guba & Lincoln, 1999; Morse, Barret, Mayan, Olson, & Spiers, 2002).

As such, the research design of this investigation incorporates strategies for reliability and validity checks, such as categorizing and comprehending (or “listening to”) the data during the data collection stage; this technique can influence the course of the investigation and enhance the quality of the research as well as its replication and confirmation (Glaser & Strauss, 1999). A second strategy involves including the interview of the researcher in the study in the data collection; in this manner the researcher confronts her own opinions and preconceptions and can compare them with the views of the actual participants. A third strategy to ensure the validity of the data is employing respondent validation; in this technique, a comparison between the accounts of multiple participants differently “placed” in the same lab or work environment is conducted and the emerging themes from the coding of the data is tested accordingly (Rajendran, 2001; Cohen, Manion, & Morrison; 2007).

Further, the ongoing development of sensitivity and flexibility of the researcher with respect to the emerging themes from the data collected is also an important parameter in the study, which can enhance the verification process during the investigation (Berg, 1998).
With a view to managing the prospective of researcher’s bias in this study, unstructured and open-ended interviews are conducted, allowing the interviewees’ own experiences to shape the course of the interviews. Prodding questions from the researcher included warm-up questions, past experience questions or lessons learned questions, such as the following: i) “Can you walk me through a typical day in your work that involves using scientific software (including numerical modeling software)?”, ii) “can you tell me about the last time you sought and obtained help with a problem regarding the software you are using?”, or iii) “can you describe your collaborations with colleagues regarding scientific software?”

This is an ongoing research study; field notes from observations of research teams in their workplace, review of relevant software training materials available online and in laboratory settings, as well as interviews have been utilized in the study, thus far. For this paper, the first author has conducted ten interviews with scientific software users working currently on research projects; the data collection is continuing currently. These ten interviews have been complemented by field observations and document analysis. All field notes, including transcripts and recordings of meetings and interviews, have been categorized, tabulated, and then analyzed via cycles of coding, initially using phrases to capture the meaning of the data; subsequently, the first coding cycle of the data was examined for trends/patterns (second coding cycle) in order to identify core categories or themes in the data about scientific software training processes.

Selected results stemming from the analysis of the interview data amassed have indicated that the issue of informal learning with peers online/onsite is an important ingredient in the training process on scientific software. Participant A, a graduate student, stated: “My supervisor sat beside me for a week, gave me an introduction to their software and then I continued on my own and figured things out myself.”. Participant B, a Master’s student, also indicated that:“...if you stumble on something, go ask a colleague or a friend, it is much faster... There is also an online community that shares ideas, we help each other, it speeds up the process”. Participant C, a PhD student commencing their program, pointed out: “With my lab mates I feel a lot more comfortable asking questions than if you work with a senior software developer, because they may not have time to answer questions at your level...”. Participant D, a Master’s student with some experience in computing, also added: “I have not really experienced formal training (industry seminars). It is expensive (…). After you learn the basics, there is an online community of users that you can go to”, Participant H, a faculty member, mentioned: “Group mentality, it actually produces some pretty good results”. Participant E, an industry expert, suggested: “A new software user? I would recommend they join an online group, they write to people, they ask”. Other Participants mentioned that their senior undergraduate and graduate students start their own online chat rooms in order to share ideas about their work: “They socialize online with a common issue; the exposure that they have to the pros (professors) is minimal to the one they get through online. But they need to know how to filter the information”.

Interviews that were conducted in geosciences and engineering laboratories revealed that if the learner interacts sufficiently with peers that maintain distinct roles within the onsite group as well as the online professional community, then the quality of the results is enhanced (Table 1). However, as Participant D acknowledged: “This environment is totally built on the attitude of the professors”. Consequently, the current training practices appear to be dependent on the good intentions of the faculty member who establish the research laboratory; they are not backed by theory on collaborative learning with the purpose of training young -or even experienced-professionals within scientific communities, and particularly with constantly changing technology.

Further, the data have shown several factors that are central in training a scientific software user; these factors are directly related to the motivation of the learner (Figure 1). Constant comparative analysis of the data confirms

<table>
<thead>
<tr>
<th>Table 1. Blended Software Training in Higher Education: Each Group Member has a Role to Play.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student</strong></td>
</tr>
<tr>
<td>• Take the time to review all available software</td>
</tr>
<tr>
<td>• resources, online/onsite</td>
</tr>
<tr>
<td>• Identify own interests, strengths and weaknesses</td>
</tr>
<tr>
<td>• Be motivated to communicate with peers</td>
</tr>
<tr>
<td>• online/onsite and seek feedback</td>
</tr>
<tr>
<td><strong>Mentor A (Senior Student or Industry Expert)</strong></td>
</tr>
<tr>
<td>• Identify what could be tacit (not documented)</td>
</tr>
<tr>
<td>• knowledge in the lab and take the time to</td>
</tr>
<tr>
<td>• explain to the student/trainee</td>
</tr>
<tr>
<td>• Encourage questions from new students, show</td>
</tr>
<tr>
<td>• respect</td>
</tr>
<tr>
<td>• Allow for sufficient one-on-one time with trainee</td>
</tr>
<tr>
<td><strong>Mentor B (Professor)</strong></td>
</tr>
<tr>
<td>• Provide as many resources as possible:</td>
</tr>
<tr>
<td>• Information technology, online access to</td>
</tr>
<tr>
<td>• journals, lab equipment etc</td>
</tr>
<tr>
<td>• Ensure that all students have adequate time</td>
</tr>
<tr>
<td>• to interact with their mentors</td>
</tr>
<tr>
<td>• Design ongoing training opportunities, not just</td>
</tr>
<tr>
<td>• at the beginning of the process (with a view to</td>
</tr>
<tr>
<td>• touching base with fundamentals as well as</td>
</tr>
<tr>
<td>• technology updates); encourage mentors and</td>
</tr>
<tr>
<td>• students to meet regularly in order to check</td>
</tr>
<tr>
<td>• students’ understanding</td>
</tr>
<tr>
<td><strong>Software Developer</strong></td>
</tr>
<tr>
<td>• Incorporate feedback in existing software</td>
</tr>
<tr>
<td>• design</td>
</tr>
<tr>
<td>• Maintain an on-going relationship with the user,</td>
</tr>
<tr>
<td>• at junior and senior level</td>
</tr>
<tr>
<td><strong>Online Community</strong></td>
</tr>
<tr>
<td>• Develop trust among peers by sharing reliable data</td>
</tr>
<tr>
<td>• Exchange meaningful feedback</td>
</tr>
<tr>
<td>• Identify and establish as much common ground with</td>
</tr>
<tr>
<td>• virtual peers as possible</td>
</tr>
<tr>
<td>• Communicate issues with data and equipment</td>
</tr>
<tr>
<td>• standardization among researchers in various</td>
</tr>
<tr>
<td>• countries.</td>
</tr>
</tbody>
</table>
that the usefulness of the training environment is impacted by the profiles of the users as learners. For instance, Participant B indicated that although she took many university-level courses, “these were not the places where I learned. I looked and found my internet sources and books”. Also, Participant H explained that: “Reading is important. People do not read as much as they should. They look for an answer, fast. I see that when students have a problem, it’s because they did not read enough, they develop this habit, they did not spend much time trying to search for a solution”.

In summary, the study has revealed that users are required to develop specific abilities or skills that percolate through their whole behaviour as learners and professionals. These abilities will be critical in attaining and maintaining a high level of understanding of the software and the research problem as well as creating new, reliable scientific information:

a. The ability of the user to manage and design own learning;

b. The motivation of the user to participate in collaborative learning activities within their community of practice in online and onsite settings;

c. Their level of comprehension of the research problem at hand;

d. Their ability to critically review scientific software interface design components (custom input of parameters versus set of default parameters); and,

e. Their ability to methodically test the software output and create own documentation with associated feedback, observations and reflections on their learning process.

These factors are summarized in the following diagram, which depicts an ongoing process with interrelated parameters (Figure 1).

![Figure 1. Specific parameters that characterize a scientific software user during the training process.]

4 DISCUSSION

Graduate students in geological science and engineering laboratories, among faculty and industry experts, have been interviewed as part of this study on scientific software training. Their specific concerns involved the undergraduate academic preparation that they had received and how it could be improved in order to train them adequately for the demanding collaborative learning environment in a research laboratory as well as their ability to interact effectively with peers locally as well as internationally (in online communities). All open interviews were conducted around the participants’ fields of research and their unique characteristics: Studying natural phenomena with high risk and having to use software input parameters that have not been tested in the lab satisfactorily (as large-scale experiments are expensive and difficult to reproduce).

Hannay et al. (2009) conducted a study on formal scientific software training especially for large research projects. They revealed that their research participants did not see the need for more formal training in scientific software in its current form, as general software courses at a computer science department usually have little relevance to their research questions. They identified an increasing awareness among scientists of the requirement for improved formal training on this type of software. Additionally, Joppa et al. (2013) discussed the lack of formal training in computational methods for scientists who graduate from natural science research programs. According to their research, an overwhelming majority of researchers in natural sciences wish for increased computational skills, as they need to have sufficient knowledge of what (or how) the software is undertaking and whether it is in fact doing what is expected. As societally important scientific decisions rely on accurate scientific software application, “the scientific community must ensure that the findings and recommendations put forth based on software models conform to the highest scientific expectation” (Joppa et al., 2013, p. 815). Further, Parnas (2010) has pointed out that teaching students how to work in disciplined ways and diligently test their software programs are critical elements in formal science and engineering education.

Onsite/online collaborative learning is essential for in-depth training on scientific software; this is an emerging theme in the data collected thus far. However, the methodology according to which this collaborative learning is expected to be realised among software users was not evident (neither was communicated by their supervisors) to any of the interviewees that contributed to this paper. Based on the participants’ feedback, it primarily falls on the faculty/researchers to design this team learning process, specifically for users of these specialised software programs. Further, as Lingard (2010) has pointed out –and it was mentioned earlier in this paper– formal team learning preparation of the students (through formal university courses) can be insufficient. As one of the interviewees pointed out: “It really depends on the person, background and personality”. This improvised situation becomes a deeper issue for young software
users and can impede their overall learning, especially if they do not feel comfortable asking their peers onsite (due to personality conflicts) and do not have sufficient background to search for reliable answers online, either. A Master’s student in his first year, stated: “It is not that helpful to search for answers online... There is such a wide range of uses so when I try to Google stuff I did not really find anything useful...I would prefer talking to someone...but they may see that I do not know enough yet.” Another Master’s student said: “Sometimes the information is not out there, it is just known by people”. Other graduate students admitted that, in their groups “some people might be shy...they do not feel they know enough to go and ask a question to industry experts”. This can affect interaction and participation levels, as well as the overall learning journey of the user. Consequently, it is of high importance to design useful collaborative learning environments that can assist a geologist/geoengineer with training on software programs that are essential in their work; yet there is not enough literature on how to conduct this effectively - either formally (through university courses) or informally (in the laboratory, as part of their research projects).

5 CONCLUSIONS

Further research in the scientific software training practices will focus on investigating the different training needs among scientific software users based on their years of experience. The study has identified, thus far, variations in the users’ preferences towards various learning situations (classroom seminars, online tutorials, webinars etc), as well as degrees of social presence in their online interactions with peers. Literature has indicated that adults clearly prefer to seek rather than receive knowledge and this tendency increases despite adult learners’ different learning styles or level of cognitive ability. They seek to learn at their own pace and to learn at the right time so that they can apply new knowledge and skills immediately (Albright & Post, 1993; Harp et al., 1997). Ongoing analysis of the data will provide information about these trends in scientific software training for researchers in the geological sciences and geological engineering.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the grant contributions by Athabasca University in order to help support this research. As well, the time and efforts of the participants is greatly appreciated.

REFERENCES


