

Usability, Sociability, and Learnability: A CSCL Design Evaluation Framework

Ravi Vatrapu, Dan Suthers, Richard Medina

Department of Information and Computer Sciences, University of Hawai'i at Mānoa, USA
{vatrapu,suthers,medina}@hawaii.edu

Abstract: In this paper, we propose a three component CSCL design evaluation framework of usability, sociability, and learnability. Usability refers to the ease of use and subjective learner satisfaction with CSCL systems. Sociability refers to the CSCL system support for social interactional processes such as conversation, cooperation, deliberation, and/or argumentation. Learnability refers to the evaluation of CSCL systems in terms of student learning processes and products. We apply this CSCL system design evaluation framework to results from a prior experimental study of three different CSCL design environments for asynchronous computer supported collaborative problem solving. Our results suggest a “preference vs. performance paradox” in CSCL across the usability, sociability, and learnability components of the design evaluation framework.

Keywords: usability, sociability, learnability, preference vs. performance paradox, human-computer interaction, computer supported collaborative learning.

Introduction

In this paper, we propose a three-component design evaluation framework for computer supported collaborative learning (CSCL). This framework consolidates existing evaluation criteria—usability, sociability, and learnability—from different research and design literatures. *Usability*, taken from the human-computer interaction literature [3], refers to the ease of use as well as subjective learner satisfaction with the CSCL system. *Sociability*, taken from the online communities and learning sciences literatures [2, 7], refers to the support for social interactional processes such as collaboration, conversation, cooperation, deliberation, and/or argumentation, as well as for the development of a group's identity and norms. *Learnability* assesses CSCL systems in terms of actual student learning processes and products.

There are two motivations for this design evaluation framework. The first motivation is to provide a conceptual framework that can help inform the systematic design and assessment of CSCL environments. The three dimensions of (1) ease of use and subjective learner satisfaction, (2) ability to support the processes of social interaction, group identity, and intersubjectivity, and (3) ability to effectively and efficiently enhance learning directly map on the software, social, and learning aspects of CSCL. The second motivation is to use the framework to make sense of the results of empirical studies. After further discussion of the framework, the present paper focuses on this second motivation by providing a case study. We interpreted the cumulative results of an experimental study of quasi-asynchronous computer supported collaborative problem solving with the proposed framework. In particular, we use the framework to interpret a “preference vs. performance paradox” in the empirical results of an experimental study of three different designs of CSCL environments.

1. Framework

1.1 Usability

According to the International Standards Organization's standard ISO 9241-11, usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [3, p.536]. In the realm of CSCL, the "specified users" are learners. The "specified context of use" could be formal, informal, blended or online learning settings. The "specified goals" are the extrinsic pedagogical goals set by an instructor in formal learning settings and also the intrinsic learning goals set by learners themselves. The "specified goals" from the perspective of CSCL system designers, developers, and evaluators are to enhance and enrich collaborative interaction processes that lead to improved assessment outcomes while ensuring high learner satisfaction. Effectiveness and efficiency of the CSCL systems are measured from the pedagogical perspective whereas learner satisfaction is to be evaluated by subjective user satisfaction measures. Learner satisfaction with the CSCL system is an indirect indicator of learner engagement and motivation. Therefore, designing for and evaluating subjective satisfaction is not only a good pragmatic prescription but also a good research strategy.

1.2 Sociability

Sociability is taken from the field of online communities, where it "is concerned with planning and developing social policies and supporting social interaction" [7]. Prior work has already identified usability and sociability as mutually interdependent entities [2, 7]. Distinguishing sociability from usability, Preece and Maloney-Krichmar [7] say:

The key components of sociability in this framework are the community's purpose, its people and the policies that help to guide online behavior. The key components of usability are dialog and social support, information design, navigation and access.

In CSCL, there are a plurality of design concerns regarding "purpose, people, and policies" for sociability. For example, scripting in CSCL can be construed as explicitly dealing with the "policies" component of the notion of sociability. "Purposes" of CSCL systems vary from collaborative argumentation to science problem solving to mathematical reasoning. In the era of globalization, online and offline classrooms increasingly incorporate diverse learners and teachers. Culture can influence how students socially interact with each other in *intra-* and *inter-*cultural CSCL settings [11, 12]. In CSCL, pedagogy is actively de-centered from the teacher-learner realm and distributed across the learner-learner intersubjective realm. Each learner is both a user and a resource in CSCL as well as ALN environments. Therefore, sociability of CSCL systems should be an explicit design evaluation concern.

1.3 Learnability

In the field of usability, learnability is often treated as a component of usability and refers to the ease of learning the software. In our framework, learnability embodies a double sense: (a) learning the software and (b) learning the domain using the software. In this paper, we refer to learnability in the second sense of domain learning. CSCL researchers and practitioners assess learnability according to their respective theoretical and methodological perspectives. From a theoretical perspective, learnability can be evaluated at the individual, small group, and/or community level as well from the socio-cultural and/or socio-cognitive theoretical traditions. From a methodological perspective, field studies, design studies, experimental studies, and case studies are used to evaluate learnability by appropriate quantitative and/or qualitative methods. Historically, unlike usability and sociability,

learnability has been the central concern in the design, analysis, and evaluation of CSCL systems.

2. Application of the Framework to Experimental Results: A Case Study

In a prior study, three software environments were designed and developed in order to test competing hypotheses about the relationship between collaborative discourse and knowledge construction [10]. This paper reports additional results from a usability questionnaire that was administered after study sessions, and applies the framework to organize the results for interpretation, resulting in identification of a preference vs. performance split. Below we summarize the study design, organize prior results reported in [10] according to the learnability and sociability constructs, report the additional usability results, and interpret the three sets of results in relation to each other.

2.1 Study Design

The prior studies compared three software conditions. The shared workspace in the “Text” condition was a conventional threaded discussion tool. The shared workspace for the “Graph” condition included an evidence-mapping tool with embedded notes that supported simple linear (unthreaded) discussion. The shared workspace of the “Mixed” condition also had an evidence-mapping tool for representing conceptual structure, but provided a threaded discussion tool on the side instead of embedded notes. For the point of this comparison, see [10].

The study challenged participants to identify the cause of a disease on the island of Guam known as ALS-PD. The materials incorporated a “hidden profile” in which information is distributed across participants such that a participant relying only on information he or she directly received would come to a suboptimal conclusion.

Pairs of participants were recruited from introductory courses in the College of Natural Sciences at the University of Hawai‘i, and were paid for participating in the study. There were 10 gender-balanced pairs of participants for each of three treatment groups.

Participants worked for up to 30 minutes on a “warm-up” problem to familiarize themselves with the software. Participants were then given up to 120 minutes to work on the main problem, Guam ALS-PD. At the conclusion of their final study session, each participant working alone was given up to 30 minutes to write an essay on the hypotheses that were considered, the evidence for and against these hypotheses, and the conclusion reached. One week after the experimental session, each participant was required to complete an online post-test before payment was sent.

2.2 Learnability Results

In the prior study, essay contents and the post test were used to compare learning outcomes between the three groups [10]. The three conditions did not differ in optimality of conclusions in the essays: relatively few participants in all conditions identified the optimal explanation of the epidemiological facts. Pairs in the Graph condition were more likely to converge on the same (not necessarily optimal) conclusion than pairs in the other conditions ($\chi^2(2, N=30)=7.5, p=0.025$). Finally, Graph users performed significantly better than Mixed users on questions on the post-test requiring high integration of information ($F(2,57)=4.40, p=0.0167$), suggesting that they were able to more effectively bring relevant and distributed information together. Moreover, Graph and Mixed users expressed hypotheses significantly earlier in the sessions than Text users ($F(2,57)=10.14, p=0.0002$), and Graph users expressed more hypotheses than Text users ($F(2, 57)=4.73, p=0.0126$). Graph and Mixed

users elaborated on hypotheses significantly more than participants in the Text condition ($F(2, 57)=6.86, p<0.0021$). Overall, Graph appears advantageous for learning.

2.3 Sociability Results as Measured by Round Trips of Interaction

Further analyses in our prior study focused participants' interactional use of the shared workspace in order to explain the above outcomes and to better understand how they use the resources of the software environments. Those analyses deployed two distinct measures of interaction: *one-way information sharing* and *interactional round trips* [9]. These were measured by tracing the movement of information that was uniquely provided to one or the other participant in the source materials. Our information-sharing construct is not a suitable measure of sociability because persons could have placed information in the workspace for their own later retrieval. Therefore we focus on interactional round-trips, in which information shared by one person was taken up by the other person in an inscription that was then accessed by the originating person. Details on how the notion of "round trips" was devised and measured are in [9]. A one-way analysis of variance (ANOVA) of number of round trips between the three treatment groups suggests marginally non-random distribution of interactivity ($F(2, 27)=3.03, p=0.07$), with Graph having the highest average. Yet the pair-wise differences did not fall within the Bonferroni 90% confidence interval, so we cannot say that the environments differed in terms of round trips.

2.4 Usability Results as Measured by User Satisfaction

In the prior study, at the conclusion of each session we administered the Questionnaire for User Interface Satisfaction (QUIS) (<http://lap.umd.edu/quis/>) to measure the participants' satisfaction with the instructions and the software environment they used. There was no significant difference across groups in participants' satisfaction with the instructions and software demonstration given by the experimenter, providing a check against experimenter bias in these presentations. A one-way ANOVA of the screen section of the QUIS instrument showed significantly higher subjective user satisfaction with the Text condition ($F(2, 27)=5.20, p=0.008$). On average, subjective user satisfaction scores were higher for the Text condition. Graph and Mixed conditions received more negative comments, particularly with respect to screen clutter. In the Graph and Mixed conditions, "undo" was the most requested feature. In the Text conditions, participants frequently asked why they could not interact in real-time with synchronous messaging.

2.5 Discussion

With respect to *learnability* as measured by essays and post-tests, the design of Graph environment that embedded notes within a shared evidence-map was significantly better than the designs of the Mixed and Text environments. With respect to *sociability* as measured by "round-trips," there is no clear difference in interactivity between the three software design conditions in spite of higher counts for Graph. With respect to *usability* as measured by user satisfaction, participants prefer the Text condition. So even though learnability measures were better for the Graph condition, and the sociability measure of "round-trips" did not significantly vary between the three conditions, learners' subjective preference was for the Text environment. This might be partly due to the close resemblance of the Text environment's design to conventional online learning discussion boards used in environments such as WebCT, Blackboard, and Moodle that may be more familiar to students. These findings indicate a "preference vs. performance paradox" that has also surfaced elsewhere in HCI [1, 5], in CSCL [4, 8], and in online learning [6]. Whether the "preference vs. performance paradox" is prevalent in CSCL or not is an open empirical

research question; however its identification in our analyses is an example of the interpretive potential supported by the framework.

3. Conclusion

Since CSCL systems are socio-technical systems with learning criteria, design evaluation should incorporate (a) usability, including subjective user interface satisfaction; (b) sociability or support for social interactional processes, and (c) learnability, learning assessment with respect to the theoretical perspective of the researcher and/or the pedagogical perspective of the practitioner. As the field of CSCL continues to grow both intellectually and institutionally, such an explicit concern with systematic design evaluation should help in informing policy decisions and practice prescriptions.

Acknowledgements

Sam Joseph and Nathan Dwyer contributed to the study that produced the data revisited in this paper. This work was supported by the National Science Foundation under award 0093505 to the second author. The first author was partially supported in the writing of this article by NSF award 0535036 to Scott Robertson.

References

- [1] Bailey, R. W. (1993). Performance Vs. Preference. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 5, 282-286.
- [2] Barab, S. A., Makinster, J. G., Moore, J. A., Cunningham, D. J., & Team, T. I. D. (2001). Designing and building an on-line community: The struggle to support sociability in the inquiry learning forum. *Educational Technology Research and Development*, 49(4), 71-96.
- [3] Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human-Computer Studies*, 55(4), 533-552.
- [4] Khandaker, N., Soh, L. K., & Jiang, H. (2006). Student Learning and Team Formation in a Structured CSCL Environment. *Learning by Effective Utilization of Technologies: Facilitating Intercultural Understanding*, 185-192.
- [5] Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. *Communications of the ACM*, 37(4), 66-75.
- [6] Picciano, A. G. (2002). Beyond Student Perceptions: Issues of Interaction, Presence, and Performance in an Online Course. *Journal of Asynchronous Learning Networks*, 6(1),
- [7] http://www.sloan-c.org/publications/jaln/v6n1/v6n1_picciano.asp.
- [8] Preece, J., & Maloney-Krichmar, D. (2003). Online Communities: Focusing on Sociability and Usability. In J. Jacko & A. Sears (Eds.), *Handbook of Human-Computer Interaction* (pp. 596–620). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- [9] Prins, F. J., Sluijsmans, D. M. A., Kirschner, P. A., & Strijbos, J. W. (2005). Formative peer assessment in a CSCL environment: a case study. *Assessment & Evaluation in Higher Education*, 30(4), 417-444.
- [10] Suthers, D., Vatrappu, R., Medina, R., & Dwyer, N. (2007). *Information Sharing and Interaction in Collaborative Convergence*. Paper presented at the 15th International Conference on Computers in Education (ICCE2007), Hiroshima, Japan.
- [11] Suthers, D., Vatrappu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond Threaded Discussion: Representational Guidance in Asynchronous Collaborative Learning Environments. *Computers and Education*, 50(4), 1103-1127.
- [12] Vatrappu, R. (2008). Cultural Considerations in Computer Supported Collaborative Learning. *Research and Practice in Technology Enhanced Learning*, 3(2), 159-201.
- [13] Vatrappu, R., & Suthers, D. (2007). Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning. In T. Ishida, S. R. Fussell & P. T. J. M. Vossen (Eds.), *Intercultural Collaboration I : Lecture Notes in Computer Science* (pp. 260-275): Springer-Verlag