

Using Innovation Diffusion Theory to Guide Collaboration Technology Evaluation: Work in Progress

Diane H. Sonnenwald Kelly L. Maglaughlin
CB #3360
University of North Carolina
Chapel Hill, NC 27599-3360 USA
sonnenwald@ils.unc.edu maglk@ils.unc.edu

Mary C. Whitton
CB #3175
University of North Carolina
Chapel Hill, NC 27599-3175 USA
whitton@cs.unc.edu

Abstract

Cost effective development of collaboration technology requires evaluation methods that consider group practices and can be used early in a system's life-cycle. To address this challenge we developed a survey to evaluate collaboration technology based on innovation diffusion theory[23]. The theory proposes five attributes of innovations that influence technology adoption: relative advantage, compatibility, complexity, trialability and observability. Selecting items from existing surveys related to these attributes, we developed a prototype multi-scale survey to help evaluate whether using a system face-to-face or distributively influences study participants' attitudes towards system adoption. We have begun refining the survey instrument and report on this process, the proposed survey questions, and the reliability and validity of the survey instrument.

1. Introduction

Increasingly organizations are relying on technology to support collaboration among team members in different locations. Collaboration technology is not, however, universally adopted by potential users [e.g., 20, 21]. An evaluation method to help predict users' adoption of collaboration technology before final product development and deployment could help maximize success and minimize development and costs.

Traditional technology evaluation methods are problematic for collaboration technology [16, 22]. Usability engineering typically optimizes for a single user and does not address whether a system will be compatible with work norms, culture and practices of a collaborative group. Case studies can help determine whether a system is compatible with a group's work norms, culture and practices but they are resource intensive to implement for large distributed groups and are typically conducted after a system is developed and deployed.

We have begun developing a multi-scale evaluation instrument based on Rogers' innovation diffusion theory

[23]. The theory proposes that five attributes of an innovation influences its adoption. The five attributes have concepts in common with CSCW research and have been operationalized in previous surveys used to evaluate a variety of communications and information technology. We used these previous surveys as a starting point for our survey instrument. We have used the instrument to evaluate a collaborative system to support scientific research activities. In addition to the survey, we used task outcome measures and interviews to allow a comparative evaluation. Data from the three sources converged, suggesting the survey instrument should be further evaluated to judge its broader applicability and utility.

2. Innovation diffusion theory and survey design

Synthesizing five decades of diffusion research, Rogers [23] has identified five attributes of innovations that are correlated with the adoption of innovations. He and others [e.g., 28] cite research validating these attributes in domains such as medicine, engineering, and airline reservation information systems. Research [e.g., 16, 20, 21] has also shown the importance of several attributes in CSCW contexts. The five innovation attributes are: relative advantage, compatibility, complexity, trialability and observability.

We are using these attributes as a theoretical foundation to evaluate collaboration technology. Since use of a theoretical framework helps insure instrument validity [1], we used the theory and related research to design a survey instrument. The instrument design contains scales based on the attributes.

The attribute, *relative advantage*, is the degree to which a new innovation surpasses current practices. Grudin [16] has observed that CSCW applications often require additional work without providing obvious benefits. Relative advantage can be operationalized, or measured, in terms of variables such as usefulness in accomplishing work goals, quality of work outcomes, added convenience and social prestige provided by the innovation. Combining this theory with our focus on collaborative work, we selected questions for this scale

from instruments measuring perceived system usefulness [9], credibility and use of information provided by group members [10], group decision satisfaction [15], solution satisfaction [15], and prestige/image [18].

Compatibility is the degree to which an innovation is perceived to be consistent with adopters' existing values, past experiences and needs. It includes individual, group and organizational goals, needs, culture and structure. It is concerned with the agreement/differences between a group's traditional work patterns and the work patterns required by the innovation. CSCW research, e.g., [16, 20, 21] also stress the importance of compatibility for CSCW applications. Although specific values, needs and work practices vary with individual preferences and context, collaboration technology should, ideally, be compatible with fundamental or general values, needs and work practices. To measure compatibility we selected questions derived from small group and usability theories that propose that satisfaction, with respect to participation in a group [15], communication in a group [10,11], experiences with communications technology in a group setting [3,11,25], and user needs related to general system qualities, such as reliability and response time [4], are fundamental to successful computer-mediated group work.

Complexity refers to the perceived difficulty of learning to use and understand a new system or technology. When a system is perceived as difficult to understand, learn and use, it will not be adopted. Research in usability engineering [e.g.,19,24] emphasizes the importance of reducing complexity in human-computer interaction. Therefore we selected questions from usability instruments, including questions that measure perceived ease of use [9,14,18] and ease of learning a system [4].

Trialability refers to the ease of experimenting with an innovation. It includes the level of effort needed and risk involved in observing and participating in small scale demonstrations of the system, including easily recovering from, or "undoing," operations using the systems and the costs involved in reversing the decision to adopt. Experimenting with and exploring system features is also a component of usability engineering. Thus we selected questions from usability questionnaires that focus on exploring and trying system features in general [4] and developed questions that focus on features specific to our system.

Observability is the degree to which the results of the innovation are easily seen and understood. Grudin [16] cautions that users need to develop a clear understanding of collaboration technology before they will adopt it. Observability has been operationalized as "results demonstrability," i.e., the ease of telling others the consequences or results of using information technology [18], and thus we selected questions from

this scale. Observability also includes visibility, i.e., the degree to which the results of an innovation are visible to others. Observability questions reported in the literature apply to situations where the innovation is observed (over time) within or outside an organization. Because the evaluation method reported in this paper controls study participants' exposure to the technology, we developed questions focusing on observability between collaborating users. For example, the question "I learned new ways of using the technology from my partner" investigates the degree to which a user can observe and learn about the technology from his or her partner.

The wording of some questions was customized for our context. For example, the question "Having a <study-specific instrument> is a status symbol in my organization" [18, p. 216] was changed to "Having this technology is a status symbol for [the university]." This yielded a survey instrument with 109 questions. Each question used a 5-point semantic differential scale with endpoints "strongly agree" and "strongly disagree" or with opposing adjectives, such as "easy" and "hard."

3. Survey instrument development

3.1 Initial item refinement: Pilot test

To refine the instrument we conducted a pilot test that included 3 pairs of study participants who participated in both face-to-face and distributed sessions and completed the survey after each session. Participants reported the survey was too long; so we eliminated similarly worded, redundant questions. For example, the original survey had the following questions: "I was very satisfied with the interaction" and "I was very dissatisfied with the conversation." The second question was omitted to reduce the survey's length. The resulting survey contained 93 questions distributed as follows: relative advantage (15), compatibility (the most complex attribute, 40), complexity (17), trialability (15), and observability (6).

3.2 Survey administration and data collection

We administered the survey as part of an evaluation study of a scientific collaborative system called the nanoManipulator [26,27]. The study was a repeated measures, or within-subjects, controlled experiment with the order of conditions, working face-to-face and working distributively, counterbalanced. A total of twenty pairs of study participants conducted scientific research activities working face-to-face using the system in stand-alone mode, and also working in separate rooms using the system in a networked, distributed mode. The study participants, 19 males and 21 females, were junior and senior undergraduate science majors at research I universities. The participants were randomly assigned to pairs; 9 pairs were of mixed gender, 6 pairs

were female, and 5 were male. Twenty-five percent of the pairs knew their partner previously. Participants were given up to 5 hours to complete the scientific research activities, which included producing a scientific lab report. After each condition, every participant completed the survey instrument and participated in an interview.¹ Participants indicated they were intensely absorbed in the work activity (4.11 on a 5-point scale), concentrated fully on the activity (4.31 on a 5 point scale), and were provided ample training (4.5 on a 5-point scale). There were no correlations between gender, (self-reported) grade point average, major, survey responses and/or lab report grades.

We used the 80 survey responses to explore the internal reliability and construct validity of the survey instrument. Eighty responses satisfies the heuristic rule that advise 5 to 15 responses per potential factor.

3.3 Construct validity

To assess the construct validity of the instrument we conducted a principal components factor analysis. The analysis was conducted iteratively, adding and deleting questions until factors emerged whose eigenvalues are greater than one based on Kaiser's stopping rule [2]. Together the factors explain 66% of the total variance in participant responses, and correspond to the innovation adoption attributes discussed earlier.

Relative Advantage. This factor contains five questions that probe participants' perceptions regarding the quality of the group problem solving process and outcomes. It includes questions such as "the group's problem solving process was efficient/inefficient" [15] and "how satisfied or dissatisfied are you with the quality of your group's results (satisfied/dissatisfied)" [15]. The factor analysis indicates that participants perceive the quality of the group problem-solving process and its outcome as an important criteria when evaluating collaboration technology. The data further suggests that other potential components of relative advantage, such as perceived system usefulness, credibility and use of information provided by group members and prestige, do not appear to significantly contribute to participants' perceptions of relative advantages when evaluating collaboration technology.

Compatibility. The five questions in this factor focus on participants' perceptions regarding the quality of interaction among collaborators. It includes questions such as "How do you feel about the group's discussions? (satisfied/dissatisfied)" [10] and "Your interactions with your partner were: free/constrained"

¹ A complete description of this evaluation and system is outside the scope of this paper.

[3,25]. These questions suggest that the compatibility of collaboration technology with traditional work practices can be measured in terms of the interaction during work, or task, process.

A second component, or aspect, of compatibility may focus on overall system quality. In our analysis, four questions:

"the system is reliable: always/never" [4]

"operations are: dependable/undependable" [4]

"response time for most operations is: fast enough/too slow" [4]

"the system often behaves in unexpected ways (agree/disagree)" [9]

were grouped as a factor but this factor only accounted for an additional 2.8% to the total variance. These questions appear to address the degree to which the system is compatible with participants' expectations of system quality. Our data do not support the inclusion of these questions in the survey, however, the data suggest that future research should consider system quality as a potential component, or aspect, of compatibility

Complexity. As predicted by innovation diffusion theory and usability engineering, one factor focuses on complexity, i.e., the perceived difficulty of using and understanding a system. This factor contains five questions, including "Interacting with the system is frustrating (disagree/agree)" [9] and "Overall, I find the system easy to use." [9,14,18].

Surprisingly, ease of learning to use a system did not emerge as a component of complexity. This may be influenced by the system used in the evaluation. The nanoManipulator system provides over 100 features to support specialized scientific experimentation in both the stand-alone and distributed mode. Therefore, ease of learning may not be a contributing factor when comparing system complexity working face-to-face and distributedly because the system is complex in both modes of working. Additional research is needed to investigate whether this result is generalizable for other types of collaboration technology.

Trialability. Trialability refers to the ease of experimenting with a system, and four questions emerged in this factor. The questions include "It is easy to recover from mistakes (agree/disagree)" [4] and "Exploration of features is: safe/risky" [4]. The questions are in concurrence with the definition of trialability in both innovation diffusion theory and usability engineering.

Observability. Three questions, including "I learned new ways of using the technology from my partner," "I changed my way of using the technology based on what I learned from my partner" and "I learned new ways of doing science from working with my partner," grouped together in the factor analysis to

comprise the observability factor. These questions investigate the degree to which the results of the technology are visible to others. As discussed previously, these questions were developed specifically for our evaluation because our experimental design did not allow for observability over time within an organization as is the case for other innovation adoption studies reported in the literature. These results indicate that for comparison between face-to-face and distributed conditions, observability can be measured by asking users if the technology enables them to learn from each other. When team members are separated by distance, it may be more difficult for them to learn from each other because the technology may reduce opportunities to learn from observation.

Results demonstrability, or the ease of telling others the consequences or results of using a system, did not emerge as a component, or aspect, of observability in this survey.

3.4 Internal reliability of attribute scales

Cronbach's alpha test [6] was used to assess the internal reliability of each attribute scale in the refined survey instrument. Accepted minimum reliability is .70 [18]. The analysis yielded alpha values ranging from .72 to .88 (see Appendix A). Therefore, the scales meet the criteria for internal reliability.

4. Discussion

In order to compare the participants' perceptions of working with the distributed, collaborative system with their perceptions of working face-to-face, we administered the survey to each participant twice, once after collaborating face-to-face and once after collaborating distributively. An analysis of variance (ANOVA) provided no evidence that there was a significant difference in participants' scale scores attributable to condition (face-to-face or distributed) for any of the scales or that there was an interaction effect due to order. Relative advantage was the only scale score that was significantly higher ($F(1,76)=8.11, p<.05, R^2=.100$) after the second research activity, regardless of condition.

To investigate whether the lack of difference is due to a poorly constructed survey instrument, triangulation of survey data with additional types of evaluation data is ongoing. We compared science lab report grades and interview data with the survey results. Each team completed a lab report both times they used the system and these reports were graded blindly with no reference to condition. Analysis showed there was no statistical difference between scores on reports written working face-to-face and distributively. Furthermore, interview data appears to indicate that study participants had equivocal perceptions of the system under both

conditions. For example, when comparing their experiences working face-to-face and distributively participants reported:

It really didn't make too much of a difference to me...I don't believe [working distributively] made [the tasks] any harder to do.

I think I liked working at a distance better.

[Working face-to-face] was probably a little bit easier to brainstorm..., but [working distributively] seemed easier to discover things on my own and figure things out.

Thus initial analysis of lab report grades and interview data show a convergence with survey results, illustrating the reliability and potential utility of the survey instrument and its use of innovation diffusion attributes.

Additional insight into the validity and reliability of the survey instrument may emerge from an ongoing longitudinal field evaluation of the system. The collaboration system has been deployed and scientists have the option of using the system to conduct collaborative scientific experiments or traveling (approximately .5 miles) to their collaborator's location to use the system. We will survey, observe and interview the scientists over the upcoming months to investigate whether they adopt the system, and why or why not.

Wording of items may also benefit from additional refinement. For example, the questions about "your group's problem-solving process" could perhaps be generalized to "your group's work process" or "your group's task process". Other questions should be customized for specific evaluation contexts. For example, in our context only two study participants used the system at any one time, and hence the term, *partner*, was appropriate. The term, *team members*, may be more applicable in other contexts.

The final scales appear to make sense when used in a repeated measures context where scores could be compared between face-to-face and distributed conditions, or during an iterative design process. However it may be difficult to interpret responses from a single administration unless desired values are established a priori.

Our analysis is also limited by the number of survey responses (80) and the single context in which the instrument was used. Testing the survey instrument in additional contexts with larger numbers of respondents and independent analyses of face-to-face and distributed responses may increase its construct validity and generalizability.

Acknowledgements. *Our thanks to the study participants; the team who program the nanoManipulator system, including Frederick P. Brooks, Jr., Aron Helser, Tom Hudson, Kevin Jeffay, Don Smith, Russell M. Taylor II; Martin Guthold and Rich Superfine, scientists who assisted in the design the natural science content of the evaluation experiment; Ron Bergquist, Bin Li, Atsuko Negishi, and Leila Plummer who assisted in running the evaluation; and Bob Losee and the anonymous reviewers for their comments. This research has been funded by the NIH National Center for Research Resources, NCRR 5-P41-RR02170. The nanoManipulator project is part of the GRIP Research Resource at the University of North Carolina at Chapel Hill.*

References

- [1] A. Anastasi, "Evolving concepts to test validation", *Annual Review of Psychology*, 37, 1986, pp. 1-15.
- [2] F.B. Bryant and P.R. Yarnald, "Principal-component analysis and exploratory and confirmatory factor analysis". *Reading and Understanding Multivariate Statistics*, APA, Washington, DC, 1995, pp.99-136.
- [3] L. Chidambaram and B. Jones, "Impact of communication medium and computer support on group perceptions and performance: A comparison of face-to-face and dispersed meetings", *MIS Quarterly*, Dec. 1993, pp. 465-491.
- [4] J.P. Chin, V. Diel, and K. Norman, "Development of an instrument measuring user satisfaction of the human-computer interface", *CHI '88 Conference Proceedings*, ACM, NY, 1988, pp. 213-218.
- [5] D. Copeland and J.L. McKenny, "Airline reservations systems: Lessons from history", *MIS Quarterly*, 16, 1988, pp. 353-370.
- [6] L.J. Cronbach, *Essentials of Psychological Testing*, Harper and Row, NY, 1980.
- [7] R.L. Daft and R.H. Lengel, "Organizational information requirements, media richness and structural design", *Management Science*, 32(5), 1986, pp. 554-571.
- [8] R.L. Daft and N.B. Macintosh, "A tentative exploration into the amount and equivocality of information processing in organizational work units", *Administrative Science Quarterly*, 26, 1981, pp. 207-224.
- [9] F.D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, 1989, pp. 319-339.
- [10] A.R. Dennis, "Information exchange and use in group decisions making: You can lead a group to information, but you can't make it think", *MIS Quarterly*, 1996, pp. 433-457.
- [11] A.R. Dennis and S.T. Kenney, "Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality", *Information Systems Research*, 9(3), 1998, pp. 256-274.
- [12] T. Finholt, "Collaboratories", *Annual Review of Information Science and Technology*, ASIST, Washington, DC, 2001.
- [13] J.A. Ghani, R. Supnick, P. Ronney, "The experience of flow in computer-mediated and in face-to-face groups", *MIS Quarterly*, pp. 229-237.
- [14] J.A. Gowan and M. Downs, "Video Conferencing human-machine interface: A field study", *Information and Management*, 27(6), 1994, pp. 341-356.
- [15] S.G. Green and T.D. Taber, "The effects of three social decision schemes on decision group process", *Organizational Behavior and Human Performance*, 25, 1980, pp. 97-106.
- [16] J. Grudin, "Eight challenges for developers," *Communications of the ACM*, 37(1), pp. 92-105.
- [17] M.L. Hecht, "The conceptualization and measurement of interpersonal communication satisfaction", *Human Communication Research*, 4(3), 1978, pp. 253-264.
- [18] G.C. Moore and I. Benbasat, "Development of an instrument to measure the perceptions of adopting an information technology innovation", *Information Systems Research*, 2(3), 1991, pp. 192-222.
- [19] D. Norman, "Cognitive engineering", *User Centered System Design*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1986, pp. 31-61.
- [20] J.S. Olson and S. Teasley, "Groupware in the wild: Lessons learned from a year of virtual collocation", *Proceedings of the ACM 1996 conference on computer supported Cooperative work*, ACM, 1996, pp. 419-427.
- [21] W. Orłowski, "Learning from Notes: Organizational issues in groupware implementation", *The Information Society*, 9(3), 1993, pp. 237-252.
- [22] D. Pinelle and C. Gutwin, "A review of groupware evaluations", *Proceedings of WET ICE 2000*, IEEE Computer society, Washington, 2000, pp. 86-91.
- [23] E. Rogers, *Diffusion of Innovations*, The Free Press, NY, 1995.
- [24] B. Schneiderman, *Designing the User Interface*. Addison Wesley, MA, 1997.
- [25] J. Short, E. Williams and B. Christie, *The Social Psychology of Telecommunications*, Wiley, NY, 1976.
- [26] D.H. Sonnenwald, R. Bergquist, K.L. Maglaughlin, E. Kupstas-Soo, and M.C. Whitton, "Designing to support collaborative scientific research across distances: The nanoManipulator example", *Collaborative Virtual Environments*, Springer Verlag, London, 2001.
- [27] R.M. Taylor II and R. Superfine, "Advanced interfaces to scanning probe microscopes", *Handbook of Nanostructured Materials and Nanotechnology*, Academic Press, NY, 1999.
- [28] L.G. Tornatzky, and M. Fleischer, *The Process of Technological Innovation*, Pergamon Press, NY, 1990.

Appendix A: Revised Survey

	Alpha	Question	Anchors	Source	Factors				
					1	2	3	4	5
Relative Advantage	.88	How satisfied or dissatisfied are you with the quality of your group's results?	Satisfied-Dissatisfied	[15]	0.731	0.184	0.140	0.136	0.180
		Group's problem solving process	Efficient-Inefficient	[15]	0.716	0.322	0.270	0.145	0.010
		Group's problem solving process	Coordinated-Uncoordinated	[15]	0.800	0.310	0.099	-0.044	0.076
		Group's problem solving process	Understandable-Confusing	[15]	0.608	0.351	0.155	0.068	0.197
		Group's problem solving process	Satisfying-Dissatisfying	[15]	0.728	0.385	0.046	0.105	0.172
Compatibility	.84	How do you feel about the group's discussions?	Satisfied-Dissatisfied	[10]	0.289	0.709	0.025	0.069	0.206
		Your interactions with your partner	Free-Constrained	[3,25]	0.276	0.712	-0.134	-0.020	-0.102
		Your interactions with your partner	Good-Bad	[3,25]	0.176	0.763	0.109	0.018	0.136
		Your interactions with your partner	Accurate-Distorted	[3,25]	0.211	0.696	0.149	-0.247	-0.093
		Your interactions with your partner	Easy-Difficult	[3,25]	0.294	0.713	0.244	0.068	0.060
Complexity	.72	Interacting with the system is frustrating.	Disagree-Agree	[9]	0.178	0.327	0.679	-0.055	-0.120
		The system is rigid and inflexible to interact with.	Disagree-Agree	[9]	0.153	-0.128	0.618	0.247	0.275
		I find it easy to get the system to do what I want it to do.	Agree-Disagree	[9]	0.425	0.056	0.630	0.190	0.133
		The system often behaves in unexpected ways.	Disagree-Agree	[9]	0.205	-0.127	0.638	0.070	-0.339
		Overall, I find the system easy to use	Agree-Disagree	[9,14,18]	-0.125	0.317	0.682	0.130	0.173
Trailability	.79	It is easy to recover from mistakes when using the technology	Agree-Disagree	[4]	0.168	0.127	0.008	0.845	0.114
		Exploration of features	Safe-Risky	[4]	-0.281	0.166	0.243	0.559	-0.013
		Correcting my mistakes	Easy-Difficult	[4]	0.191	-0.117	0.111	0.850	0.018
		Ability to undo operations	Adequate-Inadequate	[4]	0.123	-0.196	0.118	0.777	-0.186
Observability	.77	I learned new ways of using the technology from my partner	Agree-Disagree		0.158	0.020	0.020	-0.047	0.840
		I changed my way of using the technology based on what I learned from my partner	Agree-Disagree		0.314	-0.145	0.078	-0.209	0.751
		I learned new ways of doing science from working with my partner	Agree-Disagree		0.041	0.264	0.002	0.176	0.735