T. Koschmann, D. Suthers, & T.W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years!*
Mahwah, NJ: Lawrence Erlbaum Associates.

The Effects of Electronic Communication Support on Presence Learning Scenarios

Andreas Harrer, Sam Zeini and Niels Pinkwart
Collide Research Group,Universität Duisburg-Essen
{harrer, zeini, pinkwart}@collide.info

Abstract. This paper investigates in the effects of using electronic communication forms in web-based environments. Following the idea of triangulation, we used qualitative methods, statistical analysis and Social Network Analyses to explore the patterns of communication within one selected case of a mixed presence/web-based university course. The results show that while an isolated perspective does not suffice to explain the complex processes, taking more perspectives into account in a combined and integrated way provides a better understanding of technology enabled communication and interaction.

INTRODUCTION

In current educational practice, web based environments are an established means to accompany learning scenarios. In contrast to other computer based support methods, web based tools have some inherent practical advantages: they normally do not require the user to install any software, and a significant number of today's learners already has some experience in browsing web pages and therefore is used to the underlying usage patterns.
In learning contexts there are numerous variants of how the WWW is used. Different functions include web pages serving as a more or less static information source, web-based intelligent tutoring systems where the main purpose of the system is to teach rather than to be a learning resource, and environments which take into account social perspectives of learning and offer means for communication or collaboration as a central element of the web based learning support (Madrazo, 2003; Scardamalia, 2004).

In addition to these differences in the *function* of existing web based environments for learning support, also the usage *context* of these systems varies considerably: one characteristic factor is the learning group size, which can vary between very small groups (or even isolated single users that to not interact with others) and large communities with their special needs (Gaudioso & Boticario, 2003). Further distinguishing criteria include the course type (e.g., lecture vs. seminar), and the age of students. In addition, some approaches are related primarily to distance learning scenarios, while others focus on the support of presence courses.

Today's support of presence lectures at university level via web based environments is a typical intermediate case between presence and distance learning situations, sometimes denoted with the term "blended learning" (Sauter, Sauter & Bender, 2004): often, the lecture is done physically, but a lot of supporting actions are delegated to a web-based environment due to lack of time, university staff, or other constraints. Students and teachers can make use of web-based support environments in various ways (like, e.g., communication facilities, resource collections, the management of exercises if appropriate, etc). One goal of this paper to find out whether this web based support *improves the learning results* of the students.

Pinkwart et al. (2005) present some investigations that analyze the interrelations between active usage of the forum embedded in the iPAL (internet Portal to Augment Learning) web portal, and the student's final grade in the examination. Indeed, a positive correlation was observable. One of the particular results showed that an above-average system usage correlates positively with a good grade. However, interesting questions remain: In order to thoroughly understand the relations between the usages of the web based environment by learning groups and the learning outcome, we redesigned iPAL to be able to conduct a more detailed analysis of system usage and its comparison to learning outcomes.

THE SETTING: LECTURE AND WEB-BASED COMMUNICATION ENVIRONMENT

The course that was investigated in this paper is the lecture "Software Engineering" held in the summer term 2004 at the University of Duisburg-Essen in Germany. The course is taken mainly by second year students of computer science. It consists of a lecture with accompanying exercises being part of the presence learning scenario.

The exams were conducted in a mixture of small group (3-5 students) projects of 4 weeks duration and oral exams taken after the project submission. To support the small project groups with a proper communication infrastructure, each group was given a small group discussion forum, a Wiki, and CVS server access:

The small group *discussion forum* was meant for communication within the project groups and with their assigned "customers", our student tutors taking the role of the customer of the software project to be developed. The *Wiki* was introduced to the students in the lecture as a means of co-constructively editing and refining living documents, which can be used to find common ground on specific terms by defining their interpretation. The *CVS server* supported the distributed software development by taking responsibilty of version managment and conflict management in case of concurrent modifications of source codes and project documents.

In the following sections of the paper we will analyse and discuss the usefulness of these communication facilities and their impact on group structure and dynamics as well as on the outcome of the exams. This is meant to shed light on our preliminary results (Pinkwart et al. 2005) that showed that a strong participation in the lecture's discussion forums correlated with the achieved grades. At this point we investigate more deeply in the use of a variety of support tools for project work. Our hypothesis is that using computer-based communication infrastructures facilitates the success of project work: here, we put a specific focus on relating the different communication means with each other. Especially the question if there is a key communication infrastructure crucial for success, or if synergy / balance of different tools proves to be effective is a focus in our study.

The methodology of the study can be characterised as a mixed method design, following the idea of triangulation (Denzin 1980). The decision was to use qualitative methods, statistical analysis and Social Network Analyses (Wassermann and Faust 1994). This research design allows to use the results from one applied methodological approach as interpretation context for the other methodological pathways.

Qualitative analyses of the forums and the Wiki: Qualitative methods are suitable for understanding new phenomena. In triangulation designs qualitative methods are usually used with the aim of building typologies and hypotheses. In our case, the building of hypothesis was guided by the question of differences between the typologies we found, and also by asking how these are affected by other factors.

Social Network Analysis (SNA): In contrast to quantitative methods which analyse structures indirectly through the operationalised properties of the analysed cases, SNA allows the reconstruction of social structures, e.g. communication paths. In our study we used the typology derived by the qualitative analyses for sampling the most interesting groups (in the sense of the highest variance) in the way of how they organised their project.

Statistics: Based on categorisation of groups with different types of Wiki and forum usage by qualitative analyses, statistical analysis serves us to explain differences between groups, done by formulating hypothesises.

Long term statistical analysis: Since iPAL had already been used to support a past course, we decided to compare the results of both courses. In addition to the result comparisons, we were also able to make some long term analysis, because 75% of the students from the actual course were also present at the past course.

The qualitative analysis was mainly done by long term observations through the teaching staff and by analysing the content of the Wiki and the forums. The data for the SNA and the statistical analysis was extracted from the database used by iPAL and the CVS log-files.

Qualitative and Quantitative Results from Wiki and Forum Usage

To understand how Wiki and the forums were used within the different project groups, we analysed the content and the creation process through its versions as well as the forums qualitatively. We found out that the Wiki usage varied widely in separate dimensions:

On the one hand the *interactivity of the construction*, i.e. number of different authors, number of versions and scope of changes between versions varied: some groups made small and frequent updates/modifications, some had few but rather big changes between versions. Additionally some Wiki pages seemed to have been the "property/responsibility" of one person, because they were edited exclusively/mainly by one person.

On the other hand the *content* and thus the *purpose of the usage* varied: we found and indexed four categories of usage of the Wiki: *project management, glossary construction, reference lists*, and/or *coding conventions*: In *project management*, the Wiki is used to coordinate team members' activities and document their planning. Updates are usually done when replanning, rescheduling and making counterproposals. The final version is (probably) the documentation of the project process as it happened in reality. For *Clarification of terms/Glossary Construction* the Wiki is used to find a common ground and understanding of central terms and concepts for the project work. Updates are usually done when introducing or defining new terms. The final version is a glossary of used concepts and terms of the project. *Reference List* usage provides a common index to outside resources. Updates occur when giving new references and links. If used for *coding conventions*, a style guide for programming and/or documenting code is created. Updates are usually done when conventions are proposed, changed or retracted. The final version represents the conventions to be used within the project.

We analysed the interactivity of construction and the usage type of the different project groups. 10 out of 20 project groups used the Wiki extensively, while 10 used it hardly or not at all. Some groups mainly communicated outside of our support environment, e.g. via ICQ. For the 10 groups using Wiki we manually indexed the type of usage with the following results:

Table 1 – Categories of Usage for the project groups' Wiki (N=73, Average Scores in Parentheses)

	Purpose of Usage			
Interactivity of Construction	*Glossary Construction*	*Project Management*	*Reference List*	*Coding Conventions*
Few versions, large differences	Group A (71.12) Group K (75.8)	Group A (71.12) Group G (65.5)	-	Group F (84.625)
Frequent versions, small differences	-	Group B (68.667) Group D (58.375) Group E (77.625) Group H (d.n.f) Group I (85.25)	Group C (63.0) Group I (85.25)	-

In the case of the forums' analysis (all 20 groups used this communication means), four different types of usage could be found. The first category shows a very structured behaviour of using the forums. We could usually find more threads than in other groups. The topics of threads were structured but the threads were short. The second category posted just a few but long threads. The third category posted there were both a high number of threads and some of the threads were also very long. In this case we could also observe a differentiated topic structure. The fourth category used the forums just for planning meeting dates. We classified each group according to these categories.

It is significant that category 3 has the highest average of postings (11.13) and also the best results with respect to the average score (87.43). This category produced also most total files (383) in the CVS and second most versions after category 1. There is another interesting result by looking at category 2. One of the project groups within category 2 decided to use the agile programming paradigm and another project group chose a modular approach based on the division of labour. Putting them to a subcategory, this subcategory reached a score average of 86 while the other two project groups within category 2 which had not followed a systematic approach reached an average of only 53 points. This bias has to be mentioned because the agile approach usually shows an extensive face to face communication structure and the modular approach shows a rationalized communication structure in favour of the division of labour concept. Category 4 (no use of the forum for content structuring) had the smallest average score (66.88) of the categories.

Counting both concepts together yields that the project group with the highest score average (93 points, $T = -6.29$ significant at $P < 0.001$) belongs to category 1 of the forum characterization and showed no extensive Wiki usage at all. All members of this group were also present at the course we analysed last term. The group with the fewest average score (40 points, $T = 3.51$ significant at $p = 0.001$) used the forum but not the Wiki. It is also interesting that there is no significant difference in the average scores by categorizing the project groups into categories which just used Wiki or forum or used Wiki and forum both.

Social Network Analysis

For our plans to investigate in the patterns of usage of the discussion forum and the resulting communication structures, we followed the method of Social Network Analysis (Wassermann and Faust 1994, for applications in CSCL: Reffay and Chanier 2003). For this study we decided to concentrate on "direct active communication", which manifests itself in a discussion forum by a direct answer of an actor to an actor's posting. For detailed analysis of the communication structures we had the general discussion forum open to every user of the iPAL system and additionally separate forums for each project group and their "customer" (cf. Section The Setting). We will focus on selected SNA features which are applied to the general forum and contrast/relate it with a few project groups with distinctive project processes, communication structures and project results. Among the SNA traits are the centrality of one actor, the centralization of the respective network, and the prestige of an actor, all of them computed based on the degree within the graph.

The general discussion forum had 64 persons creating 276 postings. The computed value for degree-based centralization is $C_D = 0.283$ (0 means a completely balanced network, 1 a completely centralized network). This shows that the network had some "keyplayers", but also that in general the network was not dominated by any actor. The average of individual actors' centrality was $Avg(C_D(n)) = 0.044$, which means that the general centrality of actors was quite low, so nobody would be called "hub" in this network.

For the project groups, which typically consisted of 3-5 students and one "customer", we were mainly interested in differences between the groups and relations between communication structure, project

organisation, and final outcomes. Driven by our qualitative categorization and the concept of maximal variance, we present 3 selected groups (see figure 1) that are distinctive with respect to the way they communicated and their general project organisation:

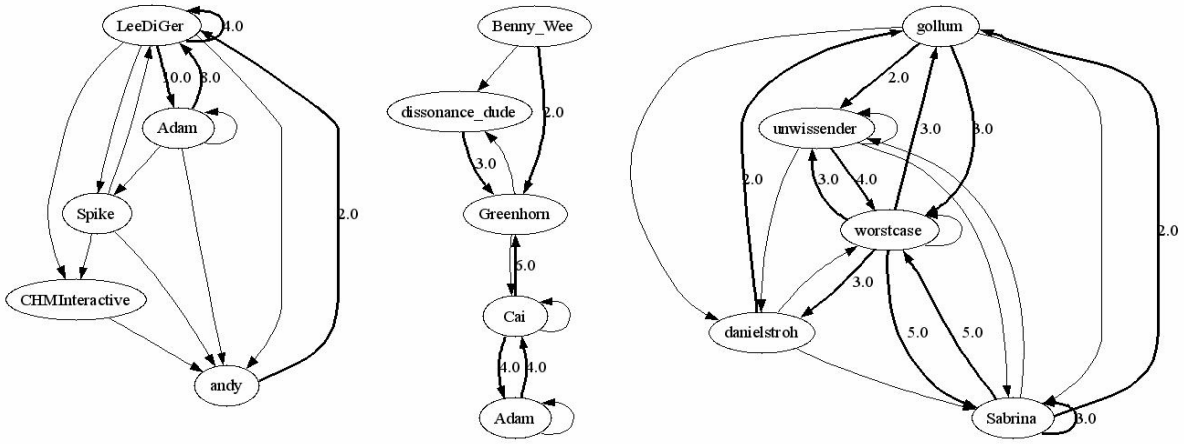

Figure 1 Sociograms of groups (left: group1, middle: group2, right: group 3)

Group 1 (no Wiki usage, little CVS, long threads with few topics) had a centralization $C_D = 0.5$ of the network with one student as central actor (centrality C_D (p) = 1.0 and prestige $P_D = 0.75$), the customer (Adam) with a small prestige of $P_D(c) = 0.25$ (in fact the smallest in this network) and other actors with centrality ranging from 0.25 to 0.75, prestige from 0.5 to 1.0. This group had indeed problems with internal communication (inside and outside the iPAL system), which led to a limited involvement of their customer, separate development of project subparts and integration problems for the project submission. This resulted in an inferior project outcome than the individual skills of the group members would suggest.

Group 2 (no Wiki use, highest CVS, differentiated topics with short threads) shown in the sparsely connected graph had a small centralization $C_D = 0.125$ of the network and a low individual centrality $C_D(n)$ of the members ranging from 0.25 to 0.5. The prestige $P_D(n)$ varied from 0 to 0.75, with the customer having 0.25 in both centrality and prestige. This can be explained because of the specific process and distribution of labour this group chose: One of the members (Cai) was assigned as "the Key Account Manager" and exclusively communicated to the customer (Adam), both in forum and personal meetings. Since the planned project process was followed consequently the project outcome resulted in the highest score of all the project groups. This group used other support facilities we provided extensively, especially the CVS with more than 140 files and 1400 versions.

Group 3 (extensive Wiki and CVS usage, differentiated topics with long threads) shown in the densely connected graph had also a small centralization $C_D = 0.1875$ but a consistently high individual centrality $C_D(n)$ ranging from 0.75 to 1.0 and prestige $P_D(n)$ between 0.75 and 1.0. The customer (Sabrina) was intensively involved with centrality $C_D(c) = 0.75$ and prestige $P_D(n) = 1.0$. All provided support facilities led to a well-coordinated project that scored second among all the project groups.

Reviewing the SNA results, we found that the exclusively structural analysis might not be sufficient to explain process and outcome of the group's work, but with the additional information we had as creators of the course, most of the phenomena could be explained utilizing both SNA and the process knowledge. This result indicates that especially in mixed presence/web-based scenarios, SNA can be helpful to understand and interpret communication structures.

Statistical Analysis

The dataset represents 20 project groups including the average score, average number of postings, the number of the files produced and the number of file revisions made by each group. The interesting outcomes are a) that there is a middle strength correlation (0.541 significant at 0.05 level, Spearman) between number of files each group produced and the average score each group reached in the course, and also b) a middle strength correlation (0.571 significant 0.05 level, Spearman) between CVS revisions made by each group and the average score each group reached. Another hypothesis was a correlation between the average number of postings for each group and the CVS usage behaviour. Yet, correlations between the average number of posts and the number of files or versions produced using the CVS system could not be observed.

As mentioned in the section about research design, the approach presented in this paper is particularly based on an evaluation study that was carried out last term. Thus it is self-evident to compare the current results with the past evaluation. We compared the results from Pinkwart et al.(2005) with our current study:

In this case we can observe that there was a stronger relationship between the number of postings and the average score (0.485 significant at 0.001 level, Spearman) than in the current study (0.320 significant at 0.01 level, Spearman). This result led us to assume a fortification of personal relationships, and thus more direct (for us non-observable) communication between the students, since 75% of the students in the recent course know each other from the last course. The hypothesis that this is caused by usage of Wiki could not be proved, since the students who used Wiki did not show a significantly different posting behaviour in the average than the students who didn't use the Wiki. This result leads us to look at the difference between the 75% of the students (N=55) which were present in both courses. In this case the students have received an average score of 71 points in comparison to the last course they reached an average of 61 points. This difference is significant ($T= 4.72$, $p < 0.001$) and there is correlation between the scores of the pairs (0.643, $p< 0.001$) that can be interpreted that in most of the cases (students) who received a high average in the past course received a high average in the current course, too. On the other hand we could not see a significant difference according to the posting behaviours within the compared courses.

CONCLUSION AND OUTLOOK

In this paper we used mixed method design to evaluate communication processes and structures within the web-based support system iPAL that was used for a presence university course. Following the idea of triangulation, we utilized qualitative methods, statistical analysis and Social Network Analysis. Qualitative methods were used to classify the usage types of the communication facilities Wiki and discussion forums. Based on these categories, we selected project groups with maximal variance of their communication behaviour and conducted Social Network Analysis to explore communication structures in detail. The SNA of the whole learning community produced a non centralized network, which complies to the large variety of communication facilities student subgroups used in the project work, according to their own choice. This degree of freedom was intended by the pedagogical approach. Indeed this is supported by the fact that no single communication form proved to be superior. In fact the combined usage showed to produce better results with respect to the final scores. These findings indicate that more aspects of the respective communication forms should be taken into account to be able to compare them properly. To reduce the complexity of data collection and aggregation of these multi-perspective analyses, we plan to explicate standard procedures for the follow-up studies, such as automated processing of forum postings as well as representation formats suitable for analysis.

REFERENCES

Denzin, N. (1989): The Research Act. A Theoretical Introduction to Sociological Methods, Prentece Hall, Englewood Cliffs, NJ.

Gaudioso, E., & Boticario, J. (2003). Towards web-based adaptive learning communities. In *Proc. of Artificial Intelligence in Education(AIED 2003)* (p. 237-244). Amsterdam, IOS Press.

Madrazo, L. (2003). SDR: networking. In *Designing for Change in Networked Learning Environments: Proc. of Computer Support for Collaborative Learning (CSCL 2003)* (p. 255-264). Dordrecht, Kluwer.

Moreno, J. L. (1951): Sociometry, Experimental Method and the Science of Society. Beacon house, Inc. New York.

Pinkwart, N., Harrer, A., Lohmann, S., & Vetter, S. (2005). Integrating Portal Based Support Tools to Foster Learning Communities in University Courses. In Uskov, V. (ed.): Proceedings of the *4th International Conference on Web-Based Education (p. 201-206). Anaheim, ACTA Press.*

Reffay, C. & Chanier, T. (2003): How Social Network Analysis can hep to measure cohesion in Collaborative Distance Learning. *Proc. of CSCL 2003 (p. 343-352)*. Kluwer, Academic Publishers, Dordrecht.

Sauter, W., Sauter, A., & Bender, H. (2004). *Blended learning*. Neuwied, Luchterhand.

Scardamalia, M. (2004). CSILE/KnowledgeForum. In *education and technology: an encyclopedia* (p. 183-192). Santa Barbara: ABC-CLIO.

Wassermann, S., & Faust, K. (1994), *Social Network Analysis: Methods and Application*, Cambridge, University Press.