The Challenges of International Computer-Supported Collaboration

Kathleen Swigger¹, Ferda Alpaslan², Robert Brazile³, Brian Harrington⁴, Xiaobo Peng⁵

Abstract - This paper discusses results of a study analyzing how cultural factors affect the performance of distributed collaborative learning teams. Participants in the study included computer science students from the University of North Texas and students from the Middle East Technical University in Ankara, Turkey. The results indicate that a team’s cultural attributes are a significant predictor of its performance on programming projects. Cultural attributes most strongly correlated to group performance were those associated with attitudes about organizational hierarchy, organizational harmony, trade-offs between future and current needs, and beliefs about the influence individuals have on their fate. The type of programming task affected the strength of the relationship between culture and performance. These results may provide distance-learning programs a way to identify at-risk work teams.

Index Terms – computer-supported collaborative work, collaborative learning, distance education, groupware.

INTRODUCTION

As more and more students enroll in distance learning courses, Universities are responding by initiating special projects that teach people how to work with others who might be located at distant sites [2]. The collaborative projects assigned in these courses tend to vary, ranging from simple email exchanges to more elaborate global interactions, supported by complex, synchronous communication tools. While these types of educational experiences help advance our knowledge of how technology can be used by geographically distributed students, it is still too early to know exactly what works and what doesn’t; theories are sparse, controlled studies are difficult to arrange, and the volume of data is overwhelming. One of the growing concerns among educators is the question of how to put together distributed teams, particularly if some of the students live in another country and/or another time zone. Whenever this occurs, there is the question of ‘culture’ and how it might affect the performance of the student teams [18]. Cultural values constitute the framework within which acceptable solutions are negotiated and decisions made [1]-[11]-[16]. While differences in view and approaches offer potential for multicultural teams to perform better, they also influence members’ preferences for social interactions, which make cooperative decision-making more difficult. To successfully incorporate group work within a distributed classroom, educators need guidelines on how to create teams, monitor team progress, and design appropriate assignments. Exploring these questions at the intersection of research on collaborative work and computer-supported collaborative learning can strengthen knowledge in all these areas.

This paper describes a two-year project that examined the factors that contributed to successful distributed programming interactions. Students enrolled at Middle East Technical University (METU), Ankara Turkey were paired with students at the University of North Texas, Denton Texas and asked to work on a number of consecutive programming projects. The research emphasized the use of computer-supported collaborative tools and how performance using these tools might be affected by cultural variables. For the purposes of this research, culture was defined by a student’s score on the Cultural Perspective Questionnaire [15]. The specific context for this work was a platform independent, computer-supported collaborative environment in which teams of programmers performed three tasks (i.e., design, code, and test software) using a variety of tools and multimedia technologies. Furthermore, this particular collaborative environment supports both synchronous and asynchronous communication among geographically, remote groups (i.e., same time, different place). In their current research, the PIs have developed a special collaborative system that supports a number of different synchronous applications such as group editing, browsing, talking, application sharing, drawing, emailing, etc. [5]-[6]. It also contains a database that logs the group’s activities that occur within the environment. The results of a two-year study indicate that a team’s cultural attributes, as defined by the Cultural Perspective Questionnaire, are a significant predictor of its performance on programming projects. The cultural attributes that were most strongly

¹Kathleen Swigger, Assoc. Dean, College of Engineering, University of North Texas, Kathy@cs.unt.edu
²Ferda Alpaslan, Associate Professor, Middle East Technical University, alpaslan@ceng.metu.edu.tr
³Robert Brazile, Associate Professor, University of North Texas, brazile@cs.unt.edu
⁴Brian Harrington, Phd student, University of North Texas, brh@cs.unt.edu
⁵Xiaobo Peng, PhD student, University of North Texas, peng@cs.unt.edu

October 20 – 23, 2004, Savannah, GA
34th ASEE/IEEE Frontiers in Education Conference
S2F-13
correlated to group performance were those associated with attitudes about organizational hierarchy, organizational harmony, trade-offs between future and current needs, and beliefs about the influence individuals have on their fate. Moreover, the type of programming task also affected the strength of the relationship between cultural attributes and performance. The results presented here have important implications for the formation of distributed collaborations and to engineering programs offering distance-learning courses requiring team projects. Below, we present both a description of the study, as well as a discussion of the pedagogical issues found to be most important in global distributed learning situations.

**RELEVANT RESEARCH**

Culture is a term that has been used in the social science and work team literature for some time [3]-[4]. The basic premise that drives culture studies is the idea that people are expected to act more favorably toward those with a higher degree of social similarity [4]. Typical determinants of culture are nationality, race, and religion. For these types of studies, the means of considering similarity and differences in the distance vector between individuals in their cultural space.

Cultural values constitute the framework within which acceptable solutions are negotiated and decisions made [10]-[14]. Results of studies in the workplace that examine the relationship between performance and a team’s cultural composition have been mixed. While some research has found that cultural differences have hindered performance [12], others have shown that multicultural teams can perform as well as homogenous teams, particularly if groups participate in formal training programs [14]-[16]. For example, [8] contends that the phenomenon of globalization has led to the development of a "culture free" hypothesis that promotes the idea that organizations are converging on a Western style, thus making cultural factors in the workplace, particularly for software developers, insignificant.

Similar differences can be found in the research literature on geographically dispersed collaborative learning. For example, Jarvenpaa and Leidner [13] found that culture was not a factor in their study of trust among globally distributed teams enrolled in a course at the University of Texas. They suggest that computer technology minimizes cultural differences among groups because it allows nonverbal communication that renders users relatively anonymous. On the other hand, Atkins and Cogburn [2] found that cultural factors have a significant effect on students’ ability to develop trust. The conflicting results may be due, in part, to the different ways culture is defined and how performance is evaluated. Unfortunately, no consistent definition of culture is used in studies on teamwork performance. For example, [12] classifies people along a number of "dimensions of culture" that reflect attitudes toward their jobs and employers, while [10] uses the terms organizational fit and demographic attributes to examine issues related to diversity and individual differences. Thus, there is some confusion over whether cultural values represent personality traits or values developed by a personal community as opposed to those associated with a specific country or national ethos [7]. As a result, we will use the terms “social” and “culture” interchangeably throughout this paper.

Along with cultural factors, the interaction of task type with time zone differences may also affect a team’s performance [9]. For example, O’Hara-Devereaux and Johansen [19] suggest that media-rich communication such as face-to-face or computer conferencing is important only during the early and late stages of project development, whereas email, faxes and workflow tools are sufficient for the middle stages. On the other hand, Meadows [17] found no significant differences in the use of media-rich tools throughout the global software development process.

In response to some of these issues, the authors began a project to investigate how cultural factors affect the performance of distributed collaborative learning teams. The study included computer science students from the U.S. and Turkey. Students were divided into culturally diverse work teams and assigned programming projects, based on the students’ responses to the Cultural Perspectives Questionnaire developed by Maznevski [15]. The results of this two-year study indicate that a team’s cultural composition is a significant predictor of its performance on programming projects. Since this study occurred within the context of a university course, the results have implications for educational institutions offering such experiences. A description of this study is found below.

**METHODS**

Participants in the study were computer science students enrolled in an Introduction to Artificial Intelligence course at the two universities (UNT and METU). It should be noted, that English is the spoken language in all courses at METU. Pedagogical differences were minimized because of the Turkish researcher’s previous experience at the University of North Texas. Each work team consisted of one student attending UNT and one student from METU. A total of 55 work teams were studied over a two-semester period – 34 work teams during fall semester 2001 and 21 during spring semester 2002. The same instructors (one at UNT and one at METU) taught the courses both semesters. Students were assigned work team partners based on their responses to the Cultural Perspective Questionnaire (described below) and their grade point averages (GPAs). The Questionnaire had been pre-tested for several semesters prior to the start of the experiment to insure that it was understandable by both Turkish and US students. The overall goal was to obtain as many work team representatives of cultural factors and GPA combinations as possible.
Project Descriptions

Work teams in each semester received similar assignments. The projects were designed to look at how cultural factors affected performance for each of the different phases of the software development life cycle. The first projects focused on the design phase. Student teams were asked to implement a heuristic search to solve a problem. Search problems ranged from guiding a robot through a grid to producing a path through a maze. Team members were instructed to work together to produce a common design document, and to use only the collaborative software to complete the project. The second projects were assigned towards the end of each semester, and required teams to use the collaborative software to design, code, and test their programs. The second projects consisted of designing, implementing and testing a program that could play a game such as Hexapawn, a game that is similar to checkers. As the final step in the game projects, team members were required to run a tournament in which the students’ programs would play each other. To successfully complete this part of the project, team members had to agree on a common interface that would allow them to pass information between the two programs (i.e., between the game program of the META student, and the game program of the UNT student).

Work Environment

As previously stated, all work teams were required to use the Computer Supported Collaborative Work (CSCW) software developed by the authors (see http://zeus.csci.unt.edu) for each assignment. The special cooperative interface relies on a shared, real-time window system (i.e., What You See is What I See) that allows groups that are geographically separated to execute controlled tools [5]. The software is designed to run over the Internet and support collaborative activities such as a shared whiteboard, editor, chat, browsing, e-mail, scribble, talk, and file/application sharing tools. The system is multi-platform and equipped with both management and record keeping capabilities that allow researchers to analyze the collaborative activities that occur during a group session.

Work Team Performance Evaluation

As previously mentioned, work teams were required to provide a single design document for the first project and code for the programs. For the second project, work teams were required to turn in a single design document, code for each of their games, and a text document discussing the outcome of the tournament and a rationale for their heuristics. Projects were evaluated and an overall grade was assigned based on four criteria – accuracy, efficiency, thoroughness, and style. A design or a program was considered accurate if it satisfied the user’s functional requirements and contained no errors. A project’s efficiency score was determined by examining the number and type of program modules, whether the design or program was concise and contained few or no extraneous elements, and was easy to understand. A program’s thoroughness was scored on whether the design or program included all the necessary elements (e.g., tests for out-of-bounds data and incorrect input). Finally, good programming style was judged by use of proper documentation style, appropriate variable naming conventions, proper indentation, and the quality of the interface.

A principal objective of the analysis was to determine work team characteristics that indicate potential for poor performance. So, for the logistic model discussed below, project grades were categorized either as poor – one standard deviation below the class average – or satisfactory – better than one Standard Deviation below the class average. Standardizing grades helped account for any differences in grading styles and class dynamics from one semester to the next. Twenty-seven percent (15 out of 55) of the work teams performed poorly on the first project and 38% (21 out of 55) performed poorly on the last project.

Cultural Attributes

Cultural attributes of students were measured with the Cultural Perspectives Questionnaire (CPQ) developed by Maznevski, [15]. The CPQ measures a variety of respondent attitudes, from making decisions based solely on immediate concerns to the degree of belief in one’s destiny. The CPQ consists of 92 questions and tests five cultural dimensions – Activity, Human Nature, Relationships, Relation to Nature, and Time. Responses for each question are scaled from 1 to 5, where 1 corresponds to strong disagreement and 5 corresponds to strong agreement. Each category is divided into subcategories that are tested by three to eight questions. Descriptions of the dimensions that are found significant are given in Table I. (For a complete list, see [15]). Each of the cultural dimensions has been given a “factor name” to help the reader better understand the discussion.

<table>
<thead>
<tr>
<th>Dimension Category</th>
<th>Factor Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationships</td>
<td>Hierarchical</td>
<td>This person believes in rigid power structure.</td>
</tr>
<tr>
<td>Relation to Nature</td>
<td>Subjugation</td>
<td>This person believes that his/her destiny has been predetermined.</td>
</tr>
<tr>
<td>Harmony</td>
<td>Harmony</td>
<td>This person believes that it is important to strike a balance in life and work.</td>
</tr>
<tr>
<td>Time</td>
<td>Future-Oriented</td>
<td>This person believes that we should be concerned primarily with future consequences.</td>
</tr>
</tbody>
</table>
The variables included in the analysis were the maximum and minimum of each work team’s GPAs and CPQ subcategory scores. The maximum and minimum work team GPA and CPQ scores provide a concise way to characterize work team CPQ-GPA configurations. For instance, if the minimum relation-to-hierarchy (i.e., hierarchy) score for the work team is high, then both team members believe that work groups function best with rigid power structures, while a high maximum relationship-to-hierarchy (i.e., hierarchy) and low minimum indicates a work team with very different attitudes about group power structures.

The means of most CPQ subcategories were close to the mid-response of 3 and standard deviations were around a ½ unit. The notable exception was relationship-to-nature (i.e., Destiny-Predetermined) – the average score was close to 2 and it had the largest standard deviation of .84. There were no practical differences between the distribution of most other CPQ subcategory scores between students from UNT and those from METU. Scores that were anticipated for ‘Turkish’ culture such as high hierarchical, high predetermined, and low individualism did not really materialize; although Turkish students had slightly higher scores in the Hierarchical and Destiny-Predetermined categories. However, as discussed below, there was a noticeable difference between one of the important model variables for predicting work team performance.

**ANALYSIS**

Logistic regression using SPSS statistical software was performed to relate CPQ and GPA variables to work team performance. Separate logistic models were built for the first and second team projects.

**Logistic Models**

Bivariate logistic models were constructed using the work team maximum and minimum for GPA and each CPQ subcategory. The GPA-CPQ subcategories for which either the work team minimum or maximum variable was at least marginally significant (P-value ≤ .25) were initially included in the multivariate analysis. This served as the standard univariate screening process for building the multi-variate model. Both the maximum and minimum GPA-CPQ variables were included to guard against missing associations between performance and dissimilar work team members – that is, work teams for which one member had a high CPQ subcategory score and the other had a low score. Based on a variable’s Wald statistic and likelihood ratio tests for the difference between models with and without a variable, variables were eliminated from the initial multi-variate model to arrive at the final models in Table II for the first project and in Table III for the second project. The significance of each pairwise interaction term was tested. For the first project, only the interaction between “Maximum Harmony” and “Maximum Hierarchical” scores were significant at < .1 level. There were not any significant interactions for the second project. Inclusion of a “Maximum Harmony” – “Maximum Hierarchical” interaction term produced an over-fitted model with numerically unstable parameter estimates.

Both the models presented here had model –2log-likelihood chi-square significance levels < .001 and all the model variables were significant at < .05 level. Assessing the models’ goodness of fit, 85.5% of all work teams were correctly classified by the first project model (66.7% of poor performing work teams and 92.5% of satisfactory), and 94.5% were correctly classified by the second project model.

**TABLE II**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (B)</th>
<th>S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA_MAX</td>
<td>2.65</td>
<td>1.05</td>
<td>.012</td>
</tr>
<tr>
<td>Destiny_MIN</td>
<td>2.55</td>
<td>1.07</td>
<td>.017</td>
</tr>
<tr>
<td>Destiny_MAX</td>
<td>-2.15</td>
<td>.10</td>
<td>.031</td>
</tr>
<tr>
<td>Hierarchy_MAX</td>
<td>-3.05</td>
<td>1.32</td>
<td>.021</td>
</tr>
<tr>
<td>Harmony_MAX</td>
<td>3.29</td>
<td>1.52</td>
<td>.031</td>
</tr>
</tbody>
</table>

**TABLE III**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (B)</th>
<th>S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA_MAX</td>
<td>4.16</td>
<td>1.47</td>
<td>.005</td>
</tr>
<tr>
<td>Hierarchy_MAX</td>
<td>-6.06</td>
<td>2.19</td>
<td>.006</td>
</tr>
<tr>
<td>Hierarchy_MIN</td>
<td>-2.60</td>
<td>1.18</td>
<td>.027</td>
</tr>
<tr>
<td>Destiny_MIN</td>
<td>3.33</td>
<td>1.29</td>
<td>.010</td>
</tr>
<tr>
<td>Future-Time_MAX</td>
<td>3.37</td>
<td>1.42</td>
<td>.018</td>
</tr>
</tbody>
</table>

(95% of poor performers and 94% of satisfactory) using a .5 classification cut-off. Both models do substantially better at correctly predicting performance for those work teams that performed poorly than models based solely on GPA – first and second project models based only on GPA correctly predicted only 33% of poor performing work teams.

Thirteen of the twenty-one work teams that performed poorly on the second project also performed poorly on the first project. So it would be expected that many of the same variables would be significant for both first and second project models. On the other hand, the “Maximum Future-Oriented” factor was not even moderately significant for the first project, indicating an important difference between the work team characteristics necessary for success with the two projects. Work team “Hierarchical” attitudes were also a more important factor in predicting performance for the second project than in the first.

**Interpreting the Models**

The model coefficients indicate that work teams with at least one member with a high GPA have greater odds of performing satisfactory. This is not surprising. However, it is interesting that once the Maximum GPA is given,
knowing the GPA of the low-GPA team member does not add to the predictive power of the model. Work teams with a member that strongly believes in rigid organizational hierarchy (high “Hierarchical” score) have greater odds of poor performance. This was particularly evident for the second project where all work teams for which both members had “Hierarchical” scores above 3 performed poorly and 52% performed poorly when both members scored above 2.5, compared to 27% when at least one scored below 2.5. Maximum GPA values for the poorly performing work teams with “Minimum Hierarchical” > 3 were distributed fairly uniformly from 2.6 through 3.6 and from 2.5 through 3.9 for poor performance teams with “Minimum Hierarchical” > 2.5. Thus, the relationship between “Hierarchical” scores and performance seems to hold regardless of GPA. Work teams that had a member with a low “Destiny-Predetermined” score had higher odds of performing poorly, and these odds were increased if the other team member had a high “Destiny-Predetermined” score (this held for both projects but the “Maximum Destiny-Predetermined” factor was not statistically significant for the second project). That is, work teams with noticeable differences in attitudes about their predetermined destiny had greater odds of performing poorly. Work teams for which there was a team member with a high harmony score had greater odds of performing satisfactory on the first project. Finally, for the second project, work teams for which at least one member had a high “Future-Oriented” score had greater odds of satisfactory performance.

**DISCUSSION**

Cultural attributes appear to be valuable predictors of work team performance. Moreover, the results indicate that the characteristics of programming tasks affect which cultural attributes are most strongly associated with performance. The models presented here are substantially better at identifying poorly performing student work teams than models based solely on GPA. Important CPQ attributes for predicting work team performance were “Hierarchical,” “Harmony,” “Destiny-Predetermined,” and “Future-Oriented.” Although there was no practical difference between the US and Turkish student scores on the CPQ in these categories, there is some indication that the “Hierarchical” scores were higher for Turkish students. The more general conclusions are the following:

- Work teams for which both members had high “Hierarchical” scores had greater odds of poor performance. This was especially true for the second project. A critical element for the success of the second project was close collaboration about how to exchange the information, and about how to repair the interface if it failed. Work teams with high hierarchical scores seemed more rigid and less able to adapt to changes and/or problems that occurred.
- Work teams with a low “Destiny-Predetermined” member have greater odds of performing poorly. The odds of poor performance are increased further if the other team member has a high score on this factor. A low “Destiny-Predetermined” score indicates strong disagreement with the idea that the outcomes of events are predetermined. So, work teams with divergent attitudes about their Destiny have fundamental differences over whether they have the ability to control events and how to manage unexpected problems.
- Not surprisingly, work teams with at least one member with a high “Harmony” score had greater odds of performing satisfactory (statistically significant for the first project).
- Members who scored low on future-oriented factors had greater odds of poor performance on the second project. This project was assigned towards the end of the semester, and demanded a high degree of synchronization between the team members. For example, team members had to wait to run the tournament until both members had completed the program. So, Future-Oriented factors provide another good indicator of the potential for success on time-critical projects.

The study also produced a number of results that relate to pedagogical issues surrounding the administration and teaching of geographically dispersed students. Important conclusions include:

- Students discovered quickly that team projects conducted within a computer-supported collaborative environment were very different than those set in traditional classrooms. In traditional class setting, students can schedule meetings more easily; put things off to the last minute; or re-allocate a task or program. In this study, none of the students knew each other previously; they had to learn to schedule meetings in advance (at least one day); and they had to learn to share workload responsibilities, since the experimental collaborative software could detect inequalities in team participation. To alleviate some of the scheduling problems, a web-based scheduling tool was designed to help students identify their availability on specific days. While the tool was somewhat helpful, students continued to have difficulties coordinating work and class schedules to accommodate meetings with their team members.
- US students’ knowledge of other countries is very poor. Misconceptions about other’s culture can lead to distrust between group members. To help establish trust between the students, a number of “soft” exercises were introduced to encourage initial communications.
- The administrative overhead of globally distributed learning is extremely high. Along with the obvious
problems, difficulties included different start times for semesters, different holidays, different test times, unequal class sizes, unhappy groups, and interrupted schedules caused by ice storms and/or other unanticipated events. To cope with these difficulties, a number of administrative tools were developed that enabled staff to communicate quickly with the different student groups.

- As claimed by other researchers (Atkins et al., 2000; McEllan 1997), we found that groups who communicated with each other during the early stages of the project and then continued this communication throughout the project performed better than those groups who tended to wait until a few days before the project was due. Frequent interaction provides information about others' actions that can help reduce uncertainty about future events. It also seems to increase awareness of others and the task.

**CONCLUSION**

The results reported in this study document the importance of the cultural composition of the group. Cultural attributes that were most strongly correlated to group performance were those associated with organizational hierarchy, organizational harmony, tradeoffs between future and current needs, and beliefs about the influence individuals have over their fate. The type of programming task also seems to have affected the relationship between culture and team performance. As the task switched from the design phase to the program and test phases, students who were less rigid about organizational structure and who were more “future oriented” outperformed those students who were more hierarchical and less future oriented.

This study also confirms previous research regarding the importance of developing early and frequent communication among cross-national collaborative learners [4]. It also reinforces studies that highlight the difficulties of scheduling work across several time zones [11]. While our web-based scheduler allowed users to see the appropriate clock, calendar, and person information for their team, many students complained about scheduling problems and delays.

Overall, the findings reported here, both empirically and qualitatively, show that it is possible to teach groups of computer science students who are geographically distributed how to use a suite of computer-supported collaborative tools to design, program, and test programming projects. The study also demonstrates that critical to the success of such an initiative are the cultural attributes of the group, technology factors such as adequate information and scheduling infrastructure, and human resources such as training, on-site coordination support, and administrative support. Results from this study are now being used to look at time zone differences and how to adapt the user interface to accommodate the effects of these factors. During the upcoming year, we will be evaluating these and other factors to determine how to construct more effective interfaces to support global distributed learning teams.

**ACKNOWLEDGEMENT**

This work was partially supported by NSF-TUBITAK Computation and Social Systems grant CSS-9988605.

**REFERENCES**


