

Effects of Cohesion-Based Feedback on the Collaborations in Global Software Development Teams

Alberto Castro-Hernández
Computer Science
University of North Texas
albertocastrohernandez@my.unt.edu

Kathleen Swigger
Computer Science
University of North Texas
kathy@cs.unt.edu

Mirna P. Ponce-Flores
Ingeniería en Tecnologías de la Información
Universidad Politécnica de Altamira
mirna.ponce@upalt.edu.mx

Abstract—This paper describes a study that examines the effect of cohesion-based feedback on a team members' behaviors in a global software development project. Chat messages and forum posts were collected from a software development project involving students living in the US and Mexico. Half of the teams in the project received feedback in the form of a graphical representation that displayed the group's cohesion level, while the other teams received no feedback. The nature of the group interactions as well as the linguistic content of such interactions was then analyzed and compared. Results from this analysis show statistically significant differences between the feedback and non-feedback conditions. More specifically, cohesion-based feedback had a positive relation to a team's total message count, response rate, and individual cohesion score. In addition, the analysis of linguistic categories showed that the most salient categories observed were related to words about time and work. Although the feedback system did not appear to affect individual performance, the findings suggest that the cohesion measure defined in the study is positively correlated to the task cohesion construct and is also related to individual and team performance.

Keywords—Feedback, cohesion, teamwork, collaboration, global software development, virtual teams

I. INTRODUCTION

Because of the growing trend toward globalization in industries, there has been a rise in the use of global teams in organizations [1]. These types of teams generally consist of people who have diverse skills, are remotely located, operate within a global organization, and collaborate on tasks by using telecommunication technologies [2]. The reasons why global teams have flourished over the past few years is because they provide industry with many benefits such as reduced costs, access to people with different skills, a flatter organizational structure, and closer proximity to local markets [3].

On the other hand, managing global teams has its own special challenges. Dealing with cultural diversity [2] and coping with different perceptions of time and relationships [4] can have significant effects on team performance. The lack of trust within global teams is also a major problem [5], largely because of the lack of face-to-face contact [6]. Another challenge for global teams is the difficulty in establishing effective communication channels among team members. Because participants on global teams generally use some type of telecommunication technology to share information and

achieve their goals, members often find themselves using cumbersome software tools to managed their distributed communications [7].

Research has recently found that many of the problems found in virtual teams are often caused by a lack of cohesiveness among group members. For example, researchers report that the use of communication technologies often hinders the development of the cohesion construct [2]. Global teams have much lower cohesion levels than co-located groups, largely because they collaborate via technology rather than face-to-face. On the other hand, research suggests that the lack of group cohesiveness can be overcome through the exchange of more social communications among group members [8]. Social interactions allow team members to create stronger links, which has the effect of increasing the levels of group cohesiveness. This particular factor appears to be extremely important because it is also linked to group performance [9].

Thus the question arises about whether one can improve the performance of global teams by providing groups with information about their overall cohesiveness levels. For example, individuals within a group might change their behavior if they receive some type of feedback about how well their team is communicating with one another. Feedback has been found to be related to performance when it is adequately provided [10]. It has also been shown to increase the engagement level among team members [11]. Therefore, it should be possible to increase cohesiveness and performance within the group by providing effective feedback about how closely team members are working together.

In this paper we describe a feedback system that is intended to show individual cohesion levels within a global software development team. Our main objective is to determine whether such displays can modify the communication behavior of team members who are participating in the project. Moreover, it is proposed that the change in the communication behaviors as a result of this feedback will have an effect on team performance. Additionally, we propose that an interaction-based measure is adequate to measure task cohesion within a team.

II. TEAM COHESION LITERATURE

Cohesion is an important emergent state that is usually defined as "a dynamic process which is reflected in the tendency for a group to stick together and remain united in

the pursuit of its goals and objectives” [12]. This construct has been studied at both the individual and group levels [13] and has been linked to group performance [14]. Moreover, the strength of this relationship seems to be affected by the group’s task [15].

In comparison with co-located teams, virtual teams tend to be less cohesive [5], although the performance in both types of teams is essentially the same, even when different tasks are considered. Group cohesiveness in any type of team seems to increase over time, particularly when there is a leader in the group [16]. Other elements that affect group cohesiveness include team size, degree of democratic behavior within a group, participation, and satisfaction [17].

However, strong group cohesion may not always be a positive thing for a team. For example, [17] shows that high social cohesion within a group can sometimes lead to poor performance. One explanation for this negative relationship is that a team with high social cohesion may actually lead to high levels of group conformity and a reluctance to criticize a teammate’s performance. If such a condition persists, then high social cohesion values will eventually lead to lower performance.

Although researchers use similar words to describe the cohesion construct, they generally use different techniques to measure levels of cohesiveness among groups. One highly cited work [18] describes using the Group Environment Questionnaire (GEQ) to measure different levels of cohesion within groups. In this particular study, the authors propose measuring cohesion levels across four different dimensions: Group Interaction - Task (GIT), Group Interaction - Social (GIS), Individual Attraction to the Group - Social (ATGS), and Individual Attraction to the Group - Task (ATGT). Other researches propose similar dimensions and use surveys to measure their different group cohesiveness constructs [4] [19] [20] [21].

A. Related Work on Cohesion Measures

In addition to measuring group cohesiveness through surveys and self-reports, researchers have also developed techniques for measuring the quantitative aspects of a team’s interactions. More specifically, [22] calculates group cohesion using a Social Network Analysis technique that creates weighted links between participants based on the number of messages exchanged. Once these adjacency matrices are computed, the authors establish a group cohesion score by looking at only the links that have weights higher or equal to a pre-defined number. They argue that this particular measure is able to detect the position of the agents for a specific level of communication.

Another cohesion measure, called Linguistic Style Matching (LSM), was developed by [20]. This particular measure is based on the similarity of the use of function words between two individuals. Once all paired similarities among group members are computed, the paired values in a group are then averaged, and this number becomes the group cohesiveness score. Using this technique, the researchers found a correlation between LSM and the cohesion construct, and a limited relation between LSM and performance. This particular study tested the LSM measure using chat communications generated during a one-hour session from single gender teams. However, researchers who have applied LSM to the analysis of email

messages among team members over an extended period of time were unable to duplicate the significant relationships between cohesion and performance [23].

In this paper, we use survey data and a form of LSM to measure different aspects of group cohesiveness. The survey described in the paper was developed by [8] and was based on the GEQ survey. This particular survey measures cohesiveness among work groups along three different constructs: GIT, GIS and ATGS. In this paper we report on only the GIT results. The four items related to the GIT were measured using a 9-point scale. In addition to the survey data, we used a form of LSM to measure (and display) cohesion among group members. A more detailed description of how this is incorporated in our study can be found in Section IV-D of this paper. Related literature describing various feedback characteristics now follows.

III. RELATED FEEDBACK LITERATURE

As explained in [24], feedback can take on many different forms and uses. Thus, it is important to look at the literature on feedback to determine which characteristics are most important when developing a display that is intended to increase team interactions. Below is a description of a number of characteristics that have been suggested by researchers in the feedback literature. In our opinion, these characteristics represent some of the best formal research thinking about developing effective feedback systems.

A. Characteristics of Effective Feedback

[10] recommends providing feedback about how well the user is progressing toward the accomplishment of a specific task. On the other hand, [25] suggests giving users two types of feedback: *outcome* feedback that relates to how well the person is doing on the task; and *process* feedback that captures how well the individual is developing. Both types of feedback seem to be related to performance. Although the literature recommends that both outcome and process feedback should be offered, it is not always practical to display advice about a project’s outcomes, particular if the feedback must be presented in real-time. Thus, providing process feedback becomes a more realistic option for a real-time system.

According to [26], feedback should be *simple* and clear. Large amounts of feedback information can often be confusing, because it may seem irrelevant to a specific task or event. According to researchers, it is especially important that automatically-generated feedback systems provide clear explanations about how the display information is calculated, and how it changes over time [24].

[27] recommends avoiding any feedback that might be construed as a *normative* comparison, because this type of information may discourage participants from pursuing the task. On the other hand, [24] argues that sharing individual information about others’ performances can support adherence to social group norms.

The research literature also suggests that feedback can be displayed either *immediately* or delivered at a *later* time. If something requires a user’s attention in real-time, then immediate feedback is preferred [28]. However, immediate feedback

can sometimes make individuals focus on the feedback system rather than on the task. Moreover, delayed feedback can be effective if it is provided in a timely manner, such as just before a milestone or a specific date [29]. For example, software development teams often hold a daily Scrum in order to improve the communications among team members [30]. This type of daily event provides team members with not only necessary feedback but also a routine that helps set the pace for task completion.

In a formative feedback system, the intervention usually occurs at a specific time (either immediate or delayed) [27]. The *consistency* in the timing of the feedback allows the user to remain focused on the task. However, in an automatic real-time system, the intervention is continuous, which means that the user's attention may be divided between the task and the feedback display. Thus, a real-time feedback system must somehow provide the desired information without interfering with the user's engagement in the collaborative exercise [31].

One of the obvious characteristics of an effective feedback system is a good *design*. An example of this type of exacting development process is presented in [24], which describes the process that was used to develop a feedback tool called GroupMeter. The paper highlights the different design phases and relates the reasons why particular decisions were made. In the end, the authors found that a simple, but playful, display was the best design element to use in their feedback system.

B. Related Work on Feedback and Groups

Given the amount of research that has examined the efficacy of providing feedback to individuals, it is not surprising to find that there are studies that have looked at the specific relationship between feedback and groups. In [11], the researchers report on using Social Network Analysis (SNA) to provide feedback concerning a team's communication activities. Teams working on a specific task were provided information about the group's centrality, reciprocity, density (cohesion) and centralization after each session of work (a total of four). The researchers found that all forms of feedback tended to increase the number of messages that were generated by group members.

Another example of a feedback system that is aimed at supporting group activities is described in [31]. This particular system was designed to provide feedback concerning a group's positivity, engagement, information exchange and participation levels. Data for these categories were obtained through the use of the Linguistic Inquiry and Word Count (LIWC) tool [32]. This particular software analyzes each word in the text, and a total count of all values is kept for a particular category. From these counts, one can determine trends in conversations and/or discussions for individuals and groups. As can be seen from the study, analyzing word usage in a group's communications can lead to a better understanding of group cohesiveness. Once the communications were analyzed, the authors provided feedback to groups through a simple text display that explained how well a group was performing on each of the four cohesion constructs. After several studies, the authors concluded that providing groups with feedback concerning their levels of engagement and information exchange had positive effects. Both types of feedback had a significant relationship to the type of words that were used during the group discussion.

Thus, the literature on feedback systems and group cohesiveness suggests that providing groups with information about their interactions can lead to changes in group behavior. However, there remain many questions about which elements of a feedback display are most effective, and what specific group behaviors are most impacted by the feedback? For example, [31] reports that students were unhappy with the immediate feedback pop-up displays, yet [28] recommends presenting feedback in a timely manner. Despite the many studies on feedback, formal research on the relationship between feedback and group cohesion is still small. In order to address this need, we present a study that looks at the effects of feedback on global software development teams.

IV. RESEARCH METHODOLOGY

A. Teams

The students who participated in this study were drawn from two remotely-located universities; one group came from the University of North Texas (UNT) in the US, and the second group came from the Universidad Politecnica de Altamira (UPA) in Mexico. US participants were all enrolled in a Human-Computer Interface course, and the Mexican participants were enrolled in a Database design course.

Initially, 75 students participated in this project: 35 students from UNT and 40 from UPA. However, two of the Mexican students dropped the database course after the first week of the project, leaving a total of 73 participants.

Prior to the start of the project, students from both universities completed a Software Development Skill Survey in which subjects were asked to rate their competency in the different areas of software development such as Java proficiency and knowledge of database design. Responses on this survey were averaged, and an Individual Skill Level score (ISL) was assigned to each student who participated in the study. These scores were then used to place the students into the various work groups. After sorting each school's students by ISL score, 1-2 students from the top of the list were paired with a student from the bottom of the list. Once each country's teams were assembled, a Gini index was calculated for each country-team on the two lists. (A Gini index is a measure of inequality [33] which we used to operationalize member diversity). Each university's country-team list was then sorted by their Gini index, and country-teams with similar Gini indexes were then combined into a single team. Using this procedure we were able to create 15 teams. A more detailed description of the composition of each team can be found in Section V.

B. Software Development Project

Each team was asked to complete the same software development project. The assignment consisted of a redesign of an existing non-profit website. More specifically, students were asked to redesign three sections of the website (i.e., the home page, the events page, and the contribution page) and implement a database that could support the various operations that were needed to maintain the pages. The responsibilities for completing this task were divided among the two teams: UNT members developed the website front-end, whereas UPA members designed and implemented the database. Moreover,

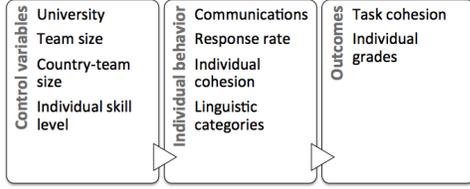


Fig. 1. Control, individual behavior and outcome variables at individual level.

the project deliverables were further divided into four separate milestones as indicated below:

- 1) Team members introduced themselves to each other; exchanged information about requirements (UNT students); created an Entity-Relationship diagram for the database (UPA students). Duration: 6 days.
- 2) UNT members re-designed and created the new home page, and UPA members implemented the database. Duration: 8 days.
- 3) UNT members created remaining website modules and UPA members provided database queries. Duration: 7 days.
- 4) All members collaborated to finish the project. Duration: 7 days.

Members were reminded to complete their milestones on time, and checks were made to ensure that most members complied with these instructions.

C. Collaboration Tools

Students who participated in the project were asked to communicate with one another using a project management web application based on the Redmine platform. The Redmine application platform supports several collaborative tools including chat, forums, wikis, document sharing, etc. Moreover, this particular Redmine application has been enhanced so that it records and timestamps all interactions among group members and transfers this information to a centralized database. Additional refinements were made to the Redmine tool in order to provide feedback to some groups concerning their collaborations and participation levels. A description of this special module can be found in Section IV-E.

In addition to the Redmine software, each team had access to a MySQL server, which allowed them to create and populate a database, as well as a web server, which allowed teams to publish their web pages.

D. Measures

In order to determine the effects of feedback on global software development teams, we examined a number of variables that capture different characteristics of a team's interactions. The list of variables are shown in Figure 1.

1) *Control Variables*: The Control variables used in this study were *University*, *Team size*, *Country-team size*, *Individual skill level*. Section V provides more details concerning each of these variables.

TABLE I. LEVEL OF ANALYSIS OF COMMUNICATIONS, RESPONSE RATE AND INDIVIDUAL COHESION. WHERE C MEANS CHAT, F MEANS FORUM, NS MEANS NON-SOCIAL AND S MEANS SOCIAL

| Level | Communications | | | | Response rate | | | | Indiv. cohesion | | | |
|---------|----------------|---|----|---|---------------|---|---|--|-----------------|---|---|--|
| Tool | C | | F | | C | | F | | C | | F | |
| Content | NS | S | NS | S | NS | S | | | NS | S | | |

2) *Individual Behaviors*: The Individual Behaviors of interest to this study were: *Communications*, *Response rate*, *Individual cohesion*, and the *Linguistic-categories*. For example, the total number of messages generated by a group has been used as a dimension of team performance [4]. Thus, we defined a measure called *Communications* to represent the total number of messages generated by team members.

The *Response rate* measure was operationalized by counting the number of replies in the chat and forum tools for each individual. A *reply* was defined as either a chat message or forum post that was sent by an individual who was different from the sender of the previous chat-message or forum-post (i.e., a sequence of chat-messages or forum-posts from the same author were counted as 1).

To show how much (or how similar) teammates were responding to one another's communications, we created an *Individual cohesion* measure. This measure was calculated by summing messages between each pair of individuals on a team, as follows:

$$cohesion_{ij} = 1 - \frac{abs(r_{ij} - r_{ji})}{r_{ij} + r_{ji}} \quad (1)$$

Where r_{ij} are the replies sent from member i to member j . An *Individual's cohesion* was then obtained by averaging all the paired cohesion values, as show in equation 2.

$$cohesion_i = \frac{\sum_{j \in M, j \neq i} cohesion_{ij}}{|M| - 1} \quad (2)$$

Where j are the teammates of i in team M . For a group-level measure, all team member's *Individual cohesion* values were averaged (see equation 3).

$$group_cohesion = \frac{\sum_{i \in M} cohesion_i}{|M|} \quad (3)$$

These measures were based on the similarity measure proposed in [20]. The scores on each of these measures range between 0 and 1, with a 1 representing perfect cohesion.

Communications, *Response rates* and *Individual cohesion* measures were also analyzed at the individual tool level [i.e., messages using either the (C)hat or (F)orum tool]. Moreover, these three measures were divided even further into those chat and forum communications that referred to Non-Social (NS) and (S)ocial¹ content (see table I). For example, we separated the chat room messages that were non-social and identified them as *Non-social-chat-communications*. The Non-social/Social Forum messages for the Response rates and Individual Cohesion measures were not included in this study because the levels of participation of Feedback groups in the forums was too low (as reported in the next section).

¹ A message was counted as Social, if according to its LIWC Social category value, it obtained 50% or higher. Otherwise, it was counted as Non-social.

The *Linguistic word category* measures were obtained by applying the sentiment analysis software (LIWC) [32]) to the group transcripts generated during this project. LIWC was utilized for this project because it looks at different linguistic features that characterize different individuals and groups. Through such a procedure, we were able to extract both an individual and a group’s use of different linguistic categories.

3) *Outcome Measures:* The two Outcome measures used in this study were *Task cohesion* and *Performance*. *Task cohesion* is a measure of a team’s perception of their level of commitment to complete the task. At the end of the project, *Task cohesion* scores for this project were calculated by averaging an individual’s responses on a 4-item Task Cohesion survey, as described in [8]. We also calculated *Group task cohesion* scores by aggregating the individual surveys by team, as described in Section V-D.

Performance scores were obtained by examining the individual grades assigned to students after the completion of the project. Each country team (i.e., UNT and UPA) was assessed separately by its corresponding course professor. That is, UNT students were graded by the US instructor, and UPA students were graded by the Mexican instructor. To avoid any inconsistencies between the two evaluation procedures, grades for all students were normalized by their university values. Moreover, *Team grades* were obtained by averaging student’s individual grades for each group, as described in Section V-D.

E. The Feedback Display

In order to determine the effect of feedback on teams, we created a special graphic that appeared in the chat module that resides within our collaborative tool. We chose the chat module for the feedback display because previous experience showed that this is the most frequently used tool for collaborations [34]. This particular module was designed to provide information about how a team was collaborating throughout the project. Following the recommendations suggested in [24], the special feedback visual was designed to consist of four elements (see Figure 2):

- 1) A piece of text showing the group’s current cohesion level. This number is computed using the *Group cohesion* formula.
- 2) A piece of text showing an individual student’s percentage of *Individual participation*. This number is computed as a proportion of the number of replies sent by the individual student over the total number replies sent by the team.
- 3) A graph that displays the group’s current cohesion level. The graphic consists of a central node (in a light gray color) surrounded by other nodes (in dark gray colors) which represent the teammates of the current viewer. The central node represents the position where, ideally, all nodes should be when a team has perfect cohesion (*Group cohesion* measure = 1). Edges represent the inverse *Individual cohesion* measure of a team member associated with its teammates (i.e. when an individual’s cohesion measure is close to 0, its node is far from the central node). The cohesion level of current viewer of the feedback display is represented as a light blue node.

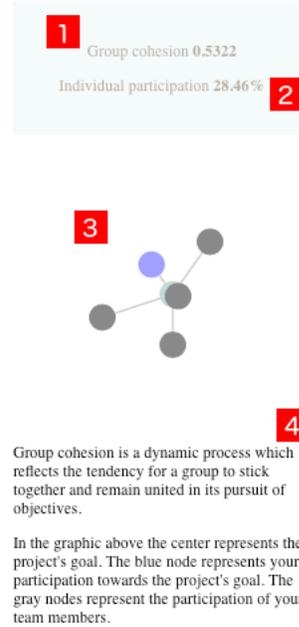


Fig. 2. Cohesion-based feedback display.

- 4) A text describing the above graph.

The feedback display was designed to provide students with a more focused understanding of their group’s communication activities. For example, the graph shows how closely the user (represented as the blue node) and their group (represented as the dark gray nodes) relate to ideal cohesiveness (represented as the center, light gray node). While such information may discourage some students, as suggested in [24], we believe that showing teams opportunities for improvement can motivate them to move in those directions.

In addition to the graph, the display also reports on a student’s individual participation rates in comparison to the total participation of the team. This information is not only helpful to the individual, but it may help offset any negative effects brought about by the group cohesiveness graph, since an individual can always control their own behavior, but perhaps not their team’s behavior. These types of mixed reviews concerning the effects of feedback are reported in [25] and [35].

The cohesion-based feedback display was provided to eight of the fifteen teams who participated in the study. These eight teams were randomly selected and designated as the Feedback teams. Seven teams received no feedback through Redmine and were, therefore, designated as the No-Feedback teams. Teams assigned to the Feedback group received an email describing the information contained in the feedback display at the beginning of the project. Once the project began, the feedback displays were updated every two hours.

V. DATA ANALYSIS

A total of 2831 messages were sent during this project: 2636 chat messages, 195 forum posts and 0 wiki pages. A total of 71, out of a possible 73, *Task Cohesion* surveys were received, with an internal consistency of 0.807 (Cronbach’s alpha).

TABLE II. CONTROL VARIABLES

| University | | Team size | | Country-team size | | Feedback | |
|------------|----|-----------|----|-------------------|----|----------|----|
| UNT | 35 | 4-members | 8 | 2-members | 34 | No | 34 |
| UPA | 38 | 5-members | 65 | 3-members | 39 | Yes | 39 |

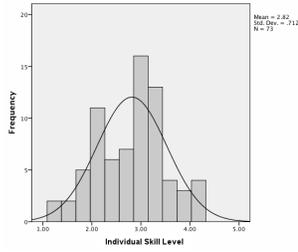


Fig. 3. Individual Skill Level distribution.

The total number of participants for each of the control variables is listed in Table II. Moreover, Figure 3 reports the overall distribution levels of student scores on the Individual Skill Level survey. The mean for subjects completing the Individual Skill Level survey was 2.82 ($SD = 0.712$).

Unless otherwise specified, the following correlations that are reported were controlled by *University*, *Team size*, *Country-team size* and *Individual skill level*. In addition, the presence or absence of *Feedback* was operationalized with a dummy variable (1 or 0).

A. Feedback Effect on Individual Behaviors

We anticipated that overall *Communications* would be higher in Feedback teams than in No-Feedback teams. The results from a partial correlation analysis show a marginal and positive relationship between Feedback and *Communications* ($r = 0.187$, $p = 0.062$). When the communication variables were analyzed at the tool-level (i.e., Chat and Forum communications), we found a positive correlation ($r = 0.206$, $p < 0.05$) between Feedback teams and Chat communications. The correlation between Feedback and Forum communications, however, was negative ($r = -0.204$, $p < 0.05$). Feedback seems to have affected the total number of Communications as well as the Chat communications generated by Feedback teams. The negative relationship between Feedback and Forum communications can be explained by the fact that Feedback teams had a much greater preference for the Chat tool as opposed to the Forum tool. Since our Feedback graphic was displayed within the Chat tool module, it seems reasonable to assume that the Feedback teams were more likely to keep returning to the Chat tool for their group information rather than use the Forum tool.

When *Chat and Forum Communications* were analyzed at the linguistic level, we found that there were no correlations between Feedback and *Social-Chat-Communications* or between Feedback and *Social-Forum-Communications*. On the other hand, there was a significant relationship between Feedback teams and *Non-social-Chat-Communications* ($r = 0.214$, $p < 0.05$), and between No-feedback teams and *Non-social-Forum-Communications* ($r = -0.196$, $p = 0.053$). As [2] reports, the low level of social communication is expected in virtual teams, since they tend to have weaker

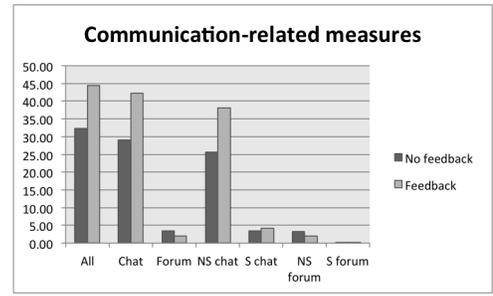


Fig. 4. Comparison of Feedback and No-Feedback teams on Communication-related measures.

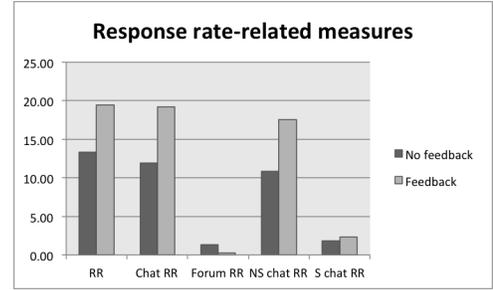


Fig. 5. Comparison of Feedback and No-Feedback teams on Response Rate (RR)-related measures.

social cohesiveness than face-to-face groups. However, both Feedback and No-Feedback groups seem to use non-social messages to communicate about the task. A more graphic illustration of the relationships that we found can be seen in Figure 4, which shows a comparison of the means of Feedback and No-Feedback groups on all the communication-related measures.

It was also hypothesized that the feedback display would affect a team's *Response rates*. Results of partial correlations indicate that Feedback was significantly and positively related to overall *Response Rates* ($r = 0.262$, $p < 0.05$), *Chat-Response-Rates* ($r = 0.296$, $p < 0.01$), and *Non-social-Chat-Response-Rates* ($r = 0.306$, $p < 0.01$). As anticipated, feedback teams had fewer responses within the forums, and so the relationship between Feedback and *Forum-Response-Rates* was negative ($r = -0.298$, $p < 0.01$). Lastly, there was no correlation between Feedback teams and *Social-Chat-Response-Rates* ($r = 0.105$, $p = 0.194$). Again, the mean comparisons of Feedback and No-Feedback teams on each of the Response-Rate measures can be seen in Figure 5.

Results also indicate that Feedback teams were much more cohesive than Non-Feedback teams. A partial correlation analysis of Feedback teams and *Individual Cohesion* rates indicate that individuals in Feedback teams were significantly more cohesive than individuals in No-Feedback teams ($r = 0.183$, $p = 0.066$). This higher cohesiveness extended to all chat communications ($r = 0.260$, $p < 0.05$), and to those chat communications related to non-social messages ($r = 0.330$, $p < 0.01$). Not surprising, *Forum-Individual-Cohesion* was negatively correlated with Feedback teams ($r = -0.369$, $p < 0.01$), and *Social-Chat-Individual-Cohesion* was not correlated with Feedback ($r = 0.014$, $p = 0.455$). The mean

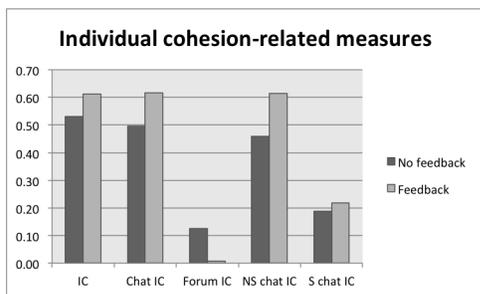


Fig. 6. Comparison of Feedback and No-Feedback teams on Individual Cohesion (IC) -related measures.

comparisons between Feedback and No-Feedback teams on the Individual Cohesion measures are presented in Figure 6.

To further understand the effects of feedback on global software teams, we examined the content of the messages that were sent between team members. As noted earlier, we used the LIWC software to determine the frequency of different *Linguistic categories* within Feedback and No-Feedback teams. Correlations between the different word categories and Feedback teams are presented in Table III. As the table shows, there were significant correlations between nine word categories and Feedback teams.

Interestingly, the words used most frequently in this project were related to the *inclusive* category. This category contains words such as *add*, *and*, *we* [36]. Moreover, the single word that was most frequently used in this category was the *we* word – 396 times in Feedback teams, and 260 times in No-Feedback teams. The heavy use of the *we* word suggests that groups in both Feedback and No-Feedback were starting to act (and think) as a team. Also, the use of *we* has been shown to have a relationship to hard work in organizations [37].

Conjunction words were also used quite frequently within the global teams. This category shares a number of words with the *inclusive* word-type such as *and*, *with*, *together*. There was a high correlation between both categories ($r = 0.945$, $p < 0.001$). This word category is related to narrative thinking, which is a trait often associated with people who have good social skills.

The *Time* category contains words such as *today*, *tomorrow*, *now*, *yesterday*, *date*. Thus, the substantial use of these types of words suggests that Feedback teams were very aware of schedules and the passing of time, which was probably related to the group’s efforts to complete the task.

The *Work* and *Achievement* categories have a number of words in common such as *accomplish*, *work*, *success*, *team*. As a result of these shared words, it is not surprising that there was a strong correlation between these two categories ($r = 0.854$, $p < 0.001$).

The *Assent* category includes words such as *agree*, *ok*, *yes*, *yeah*, *haha*, *cool*. The high correlation between Feedback teams and the *assent* category suggests that members within Feedback teams had similar opinions, and they tended to agree with one another more than members in Non-Feedback teams. However, frequent agreement has also been associated with passivity, as reported in [38].

TABLE III. LIWC CATEGORIES WITH SIGNIFICANT DIFFERENCES. * $p < 0.1$, ** $p < 0.05$

| Word category | Mean | Correlation with Feedback |
|---------------|-------|---------------------------|
| Inclusive | 27.48 | 0.202* |
| Conjunctions | 26.64 | 0.215* |
| Time | 20.76 | 0.298** |
| Adverbs | 17.57 | 0.201* |
| Work | 14.74 | 0.214* |
| Achieve | 10.58 | 0.206* |
| Assent | 5.33 | 0.273** |
| Leisure | 4.74 | 0.246** |
| Home | 2.00 | 0.242** |

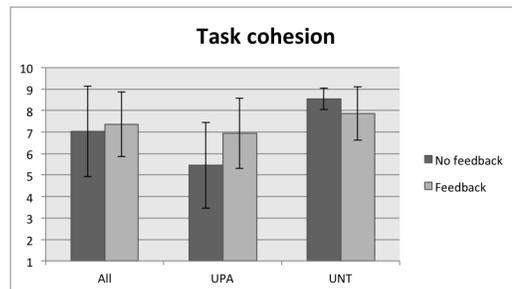


Fig. 7. Comparison of Feedback and No-Feedback teams on Task Cohesion values.

Categories of *leisure* and *home* were also strongly related to Feedback teams. Words related to these two categories generally indicate that the group is engaged in non-work related communications. The presence of words from these two categories suggests that members within Feedback teams were more likely to share social or personal information with one another than those members within No-Feedback teams.

B. Effects on Outcome Variables

We also analyzed the relationships between Feedback and the two outcome variables of *Task cohesion* and *Individual grades*. Results showed that Feedback had a positive correlation with *Task cohesion*, but it was not significant ($r = 0.118$, $p = 0.110$). Overall mean values on the Task Cohesion survey were much higher for Feedback groups than No-Feedback groups (mean difference = 0.3295). Moreover, a Levene’s Test of the data revealed that there was a significant difference between the variances of Task Cohesion values of No-Feedback teams and Feedback teams ($p < 0.05$) (see Figure 7, column All). Given this high variance, we analyzed the *Task Cohesion* responses by university and feedback type and found that the UNT students assigned to the No-Feedback teams had much higher perceived values of task cohesion than UPA students in No-Feedback teams, while the differences between the Task cohesion scores in the two universities was much lower in Feedback teams. Therefore, this result suggests that task cohesion’s perception was shaped by the feedback display.

It was suspected that students who received feedback would also receive higher grades on their global software development projects. However, a partial correlation analysis found that *Individual Grades* were unrelated to Feedback ($r = 0.026$, $p = 0.415$), and that the variance between the Feedback and No-Feedback groups was not significant.

TABLE IV. CORRELATIONS BETWEEN COMMUNICATION- AND RESPONSE-RATE- MEASURES AND INDIVIDUAL GRADES. * $p < 0.01$, ** $p < 0.001$

| | Individual grades | |
|------------------|-------------------|---------------|
| | Communications | Response rate |
| All | 0.345* | 0.336* |
| Chat | 0.337* | 0.350* |
| Forum | 0.122 | -0.124 |
| Social chat | 0.363* | 0.464** |
| Non-social chat | 0.319* | 0.003* |
| Social forum | 0.148 | |
| Non-social forum | 0.107 | |

TABLE V. CORRELATIONS BETWEEN INDIVIDUAL COHESION MEASURES AND TASK COHESION. * $p < 0.1$, ** $p < 0.05$

| Task cohesion | |
|-------------------------------------|---------|
| Individual Cohesion | 0.214* |
| Chat Individual Cohesion | 0.215* |
| Forum Individual Cohesion | 0.005 |
| Social chat Individual Cohesion | 0.135 |
| Non-social chat Individual Cohesion | 0.265** |

C. Effects of Other Measures on Outcome Variables

Since this study found significant relationships between Feedback and the various Communication and Response rate variables (see Section II), we anticipated that there might be similar correlations between the communication and response-rate measures and our outcomes measures (i.e., Task Cohesion and Performance). Partial correlations (controlled by *University*, *Team size*, *Country-team size*, *Individual Skill Level* and *Feedback*) between the *Communication* variables and *Task cohesion*, and between the *Response rate* measures and *Task cohesion* produced no significant relationships. However, as indicated in Table IV, all *Communication* and *Response rate* variables (except forum communications) had a positive and significant relationship to *Individual grades*. While individual communication behaviors do not appear to affect how an individual "feels" about their group and the task, they do seem to affect how well an individual "performs" the task.

The *Individual cohesion* measures used in this study capture the response rate similarity between an individual and his/her teammates. As such, it was assumed that there would be strong correlations between the Individual cohesion measures and the Task cohesion construct. Thus, we calculated partial correlations between *Individual cohesion* and *Task cohesion*, while controlling for *University*, *Team size*, *Country-team size*, *Individual Skill Level* and *Feedback*. Results indicate that *Individual cohesion* rates (except forum and social chat cohesion levels) were strongly related to a group's perception of how well they were committed to completing the task. Evidence for this relationship is strongest among individual cohesion and chat cohesion levels, particularly as they pertain to non-social activities (see Table V).

We also related the *Individual cohesion* measures to individual grades while controlling for *University*, *Team size*, *Country-team size*, *Individual Skill Level* and *Feedback* (see Table VI). Correlations show that all of the *Individual cohesion* measures predict individual performance, except forum messages.

TABLE VI. CORRELATIONS BETWEEN INDIVIDUAL COHESION MEASURES AND INDIVIDUAL GRADES. * $p < 0.001$

| Individual grades | |
|-------------------------------------|--------|
| Individual Cohesion | 0.412* |
| Chat Individual Cohesion | 0.432* |
| Forum Individual Cohesion | -0.079 |
| Social chat Individual Cohesion | 0.468* |
| Non-social chat Individual Cohesion | 0.426* |

TABLE VII. CORRELATIONS BETWEEN GROUP COHESION, CHAT GROUP COHESION, WORK-CHAT GROUP COHESION, GROUP TASK COHESION AND TEAM GRADES. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

| | Group Task Cohesion | Team Grades |
|--------------------------------|---------------------|-------------|
| Group Cohesion | -0.181 | 0.425** |
| Chat Group Cohesion | -0.087 | 0.466*** |
| Non-social Chat Group Cohesion | 0.307* | 0.587*** |

The relationships between the *Individual cohesion* measures and *Task Cohesion* supports the notion that a group's perception of how well they are doing on a task is shaped by a person's response activities [8]. In addition, the strong positive relation between the cohesion measures and *Individual grades* demonstrates the link between response levels within a group and individual performance.

D. Group Cohesion and Outcome Variables

Since the results showed that students who had more interactions with their teammates not only performed better but also had better perceptions of their team's ability to accomplish the task, it was anticipated that a group-level analyses of these variables would achieve similar results. Prior to running the analysis, however, it was necessary to determine how to aggregate the data and, at the same time, adjust for the large variation between UNT and UPA responses on the Task Cohesion survey. Techniques suggested by [39] indicate that the variance problem can be overcome by treating each country team separately. Therefore, we aggregated Task cohesion, Individual cohesion, Chat-Individual-Cohesion, Non-social-Chat-Individual-Cohesion and Individual grades for each team at each university and labeled these new variables as *Group Task Cohesion*, *Group Cohesion*, *Group Chat-Cohesion*, *Group Non-social-Chat-Cohesion* and *Team Grade*, respectively.

Correlations, as shown in Table VII, were controlled by *Team size*, *Country-team size*, *Feedback* and average of *Individual Skill Level*. Results show that only *Group Non-social-Chat-Cohesion* predicts *Group Task Cohesion*. On the other hand, all group cohesion measures were significantly correlated with *Team grades*.

It is important to note that *Group Task Cohesion* and *Team grades* were not correlated ($r = 0.168$, $p = 0.187$). A possible explanation for this result is that student responses on the Task Cohesion surveys represent an individual's perceptions of the "global" team, whereas Team grades represent student performance at the university level. While teams may have perceived that the global team was capable of completing the task, a specific country team may not have performed as well as other teams within the same country.

VI. CONCLUSION

Research on group dynamics has long established the relationship between a group's cohesiveness and team performance [40]. While early studies focused on investigations of group cohesion within face-to-face teams, more recent research has shifted toward analyses of the cohesion construct within technology-supported teams that are distributed in different cities and even countries [4]. Due to time pressures, geographical constraints, or organizational structures, global virtual teams do not always have the opportunity to come together and establish a sense of group cohesiveness. Therefore, it seems appropriate to explore ways in which distributed technology can be used to provide feedback to global teams that will help them establish group cohesiveness.

In this study we examined the relationships between cohesion-based feedback and group behaviors and outcomes. A cohesion-based feedback module was developed and inserted into a distributed collaborative software system. This cohesion-based feedback module provided a graphic display of a groups response rate similarities. The cohesion-based feedback display was then tested on students from the US and Mexico who were collaborating on a global software development project. Half of the teams received the feedback throughout the project, and half the teams received no feedback. Data obtained from the groups' communication activities, Task Cohesion surveys, and individual/group grades were then analyzed and compared. Results indicate that students in the Feedback teams had significantly more overall communications, higher response rates, and more similar response behaviors (as measured by the Individual Cohesion variable) than No-Feedback teams. These Feedback effects could be seen in all types of chat messages, particularly those non-social text messages.

Feedback was also found to affect the linguistic characteristics of a group. For example, Feedback teams were more likely to use words related to the work, time, and achieve categories than No-Feedback teams. At the same time, Feedback teams were more apt to use words linked to personal categories such as leisure and home as opposed to No-Feedback teams. Thus, while Feedback teams seemed more focused on the task, they also seemed more willing to engage in personal communications.

An analysis of the effects of Feedback on our outcome measures (i.e., Task Cohesion and Performance), however, was not significant. Students who viewed the feedback module, did not score higher on the Task Cohesion survey, nor did they receive higher grades on the projects. This result seems to suggest that increasing interaction among team members does not necessarily lead to better performance. Perhaps a more robust feedback system that included a link to different performance measures would result in improved group performance.

Although the study found no relationship between feedback and the outcome measures, it did find relationships between some of our other measures and the outcome variables. While we found no significant relationship between either Communications and Task cohesion or between Response Rates and Task cohesion, both Communications and Response Rates affected Individual grades. We also looked for relationships between our cohesion-based measures and Task cohesion and Individual grades and found that ALL the cohesion-based measures

(except forum cohesion) were positively correlated to Task cohesion and Individual grades. The relationship was highest for cohesion-based scores that were related to non-social chat messages. This result suggests that the cohesion-based measure is an accurate predictor of both task cohesion and performance; thus reinforcing the notion that effective team collaborations will lead to more effective team performance and a more positive perception of group members.

Finally, the individual cohesion-based measures were aggregated to the group level to determine whether there were any relationships between the group-cohesion measures and the outcome variables. The large variation between UNT and UPA responses on the Task Cohesion survey indicated that the group-level measures could be combined at only the country (i.e., university) level. Analysis of the University-level variables indicated that only the *Group Non-social-Chat-Cohesion* scores were significantly correlated to Task Cohesion, although all the Group-Cohesion scores were significantly related to *Team Grades*. Although these findings differ from the Individual cohesion results, they tend to confirm the relationship between the cohesion measure and performance. The relationship between a group's overall response similarity and task cohesion remains an open question. There is also some question about how instructors evaluate team projects and how this might affect the results reported in studies of this type.

Although the results presented in this study occurred within an academic setting, we believe that many of the characteristics of the project are similar to real-world environments [41], [42]. For example, many industry-related group projects must cope with rational task distributions, shared milestones, collaborative environments with team-specific information, and distributed leadership. The presence of these characteristics in a team will, surely, affect the cohesiveness within a group. The cohesion measure described in this study, as well as the feedback display, could potentially help groups better understand what communication is most effective.

Finding a feedback display that will affect group performance has always been a concern for researchers in human-computer interactions. This study examines the effects of a cohesion-based display on group behaviors. The display was intended to provide group members with a clear vision of not only how their group was interacting, but also how they, as individuals, were contributing to the group, and how this contribution was being measured. A clear understanding among team members of how their group is interacting appears to facilitate a group's communications, response levels, and individual cohesion scores. While our feedback module did not appear to significantly affect outcome measures, the cohesion-based measure that was used in the feedback display appears to be a good predictor of both Task Cohesion and Performance.

ACKNOWLEDGMENT

The first author gratefully acknowledges financial support from a CONACYT scholarship and from the Support for Graduate Studies Program of SEP. This material was also based upon work supported by the National Science Foundation under Grant No. 0705638.

REFERENCES

- [1] J. D. Herbsleb and D. Moitra, "Global software development," *Software, IEEE*, vol. 18, no. 2, pp. 16–20, 2001.
- [2] A. Powell, G. Piccoli, and B. Ives, "Virtual teams: a review of current literature and directions for future research," *ACM Sigmis Database*, vol. 35, no. 1, pp. 6–36, 2004.
- [3] P. J. Agerfalk, B. Fitzgerald, H. H. Olsson, and E. . Conchir, "Benefits of global software development: the known and unknown," in *Making Globally Distributed Software Development a Success Story*. Springer, 2008, pp. 1–9.
- [4] K. Swigger, F. Nur Aplasan, V. Lopez, R. Brazile, G. Dafoulas, and F. C. Serce, "Structural factors that affect global software development learning team performance," in *Proceedings of the Special Interest Group on Management Information System's 47th Annual Conference on Computer Personnel Research*, ser. SIGMIS CPR '09. New York, NY, USA: ACM, 2009, pp. 187–196.
- [5] K. Furumo and J. Pearson, "An empirical investigation of how trust, cohesion, and performance vary in virtual and face-to-face teams," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences, 2006. HICSS '06*, vol. 1, Jan. 2006, pp. 26c–26c.
- [6] E. F. McDonough, K. B. Kahn, and G. Barczaka, "An investigation of the use of global, virtual, and colocated new product development teams," *Journal of Product Innovation Management*, vol. 18, no. 2, pp. 110–120, 2001.
- [7] A. M. Townsend, S. M. DeMarie, and A. R. Hendrickson, "Virtual teams: Technology and the workplace of the future," *The Academy of Management Executive*, vol. 12, no. 3, pp. 17–29, 1998.
- [8] S. A. Carless and C. D. Paola, "The measurement of cohesion in work teams," *Small Group Research*, vol. 31, no. 1, pp. 71–88, Feb. 2000.
- [9] P. Goodman, E. Ravlin, and M. Schminke, "Understanding groups in organizations," *Tepper School of Business*, Jan. 1987.
- [10] A. N. Kluger and A. DeNisi, "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory," *Psychological Bulletin*, vol. 119, no. 2, pp. 254–284, 1996.
- [11] L. Gamberini, F. Martino, A. Spagnolli, R. Ba, and M. Ferron, "your team cohesion is low: A systematic study of the effects of social network feedback on mediated activity," in *Online Communities and Social Computing*, D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, A. A. Ozok, and P. Zaphiris, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, vol. 6778, pp. 172–181.
- [12] A. V. Carron, "Cohesiveness in sport groups: Interpretations and considerations." *Journal of Sport psychology*, 1982.
- [13] S. M. Gully, D. J. Devine, and D. J. Whitney, "A meta-analysis of cohesion and performance effects of level of analysis and task interdependence," *Small Group Research*, vol. 26, no. 4, pp. 497–520, 1995.
- [14] S. S. Webber and L. M. Donahue, "Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis," *Journal of management*, vol. 27, no. 2, pp. 141–162, 2001.
- [15] S. W. Kozlowski and D. R. Ilgen, "Enhancing the effectiveness of work groups and teams," *Psychological Science in the Public Interest*, vol. 7, no. 3, pp. 77–124, Dec. 2006.
- [16] V. L. Schwanda, K. Barron, J. Lien, G. Schroeder, A. Vernon, and J. T. Hancock, "Temporal patterns of cohesiveness in virtual groups," in *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, ser. CSCW '11. New York, NY, USA: ACM, 2011, pp. 709–712.
- [17] E. Rovio, J. Eskola, S. A. Kozub, J. L. Duda, and T. Lintunen, "Can high group cohesion be harmful? a case study of a junior ice-hockey team," *Small Group Research*, vol. 40, no. 4, pp. 421–435, Aug. 2009.
- [18] A. V. Carron, W. Widmeyer, and L. R. Brawley, "The development of an instrument to assess cohesion in sport teams: The group environment questionnaire." *Journal of sport psychology*, 1985.
- [19] W. D. Salisbury, T. A. Carte, and L. Chidambaram, "Cohesion in virtual teams: Validating the perceived cohesion scale in a distributed setting," *SIGMIS Database*, vol. 37, no. 2-3, pp. 147–155, Sep. 2006.
- [20] A. L. Gonzales, J. T. Hancock, and J. W. Pennebaker, "Language style matching as a predictor of social dynamics in small groups," *Communication Research*, vol. 37, no. 1, pp. 3–19, Feb. 2010.
- [21] P. J. Carrington, J. Scott, and S. Wasserman, *Models and methods in social network analysis*. Cambridge university press, 2005, vol. 28.
- [22] C. Reffay and T. Chanier, "How social network analysis can help to measure cohesion in collaborative distance-learning," in *Designing for Change in Networked Learning Environments*, ser. Computer-Supported Collaborative Learning, B. Wasson, S. Ludvigsen, and U. Hoppe, Eds. Springer Netherlands, Jan. 2003, no. 2, pp. 343–352.
- [23] S. A. Munson, K. Kervin, and L. P. Robert Jr, "Monitoring email to indicate project team performance and mutual attraction," 2014.
- [24] G. Leshed, D. Perez, J. T. Hancock, D. Cosley, J. Birnholtz, S. Lee, P. L. McLeod, and G. Gay, "Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 537–546.
- [25] P. C. Earley, G. B. Northcraft, C. Lee, and T. R. Lituchy, "Impact of process and outcome feedback on the relation of goal setting to task performance," *Academy of Management Journal*, vol. 33, no. 1, pp. 87–105, Mar. 1990.
- [26] R. W. Kulhavy, "Feedback in written instruction," *Review of Educational Research*, vol. 47, no. 2, pp. 211–232, Jan. 1977.
- [27] V. J. Shute, "Focus on formative feedback," *Review of Educational Research*, vol. 78, no. 1, pp. 153–189, Mar. 2008.
- [28] B. J. Mason and R. Bruning, "Providing feedback in computer-based instruction: What the research tells us," *Retrieved February*, vol. 15, p. 2007, 2001.
- [29] C. Scharff and R. Verma, "Scrum to support mobile application development projects in a just-in-time learning context," in *Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering*, ser. CHASE '10. New York, NY, USA: ACM, 2010, pp. 25–31.
- [30] C. Scharff, "Guiding global software development projects using scrum and agile with quality assurance," in *2011 24th IEEE-CS Conference on Software Engineering Education and Training (CSEE T)*, May 2011, pp. 274–283.
- [31] Y. R. Tausczik and J. W. Pennebaker, "Improving teamwork using real-time language feedback," 2013.
- [32] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: LIWC 2001," *Mahway: Lawrence Erlbaum Associates*, 2001.
- [33] T. Ogowang, "A convenient method of computing the gini index and its standard error," *Oxford Bulletin of Economics and Statistics*, vol. 62, no. 1, pp. 123–129, 2000.
- [34] F. C. Sere, K. Swigger, F. N. Alpaslan, R. Brazile, G. Dafoulas, and V. Lopez, "Online collaboration: Collaborative behavior patterns and factors affecting globally distributed team performance," *Computers in Human Behavior*, vol. 27, no. 1, pp. 490–503, Jan. 2011.
- [35] Y. R. Tausczik, "Changing group dynamics through computerized language feedback," 2012.
- [36] J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzales, and R. J. Booth, "The development and psychometric properties of LIWC2007," *Austin, TX, LIWC. Net*, 2007.
- [37] J. Pennebaker, *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury USA, 2013.
- [38] G. Leshed, "Automated language-based feedback for teamwork behaviors," Ph.D. dissertation, Cornell University, 2009.
- [39] J. M. LeBreton and J. L. Senter, "Answers to 20 questions about interrater reliability and interrater agreement," *Organizational Research Methods*, 2007.
- [40] J. Mathieu, M. T. Maynard, T. Rapp, and L. Gilson, "Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future," *Journal of Management*, vol. 34, no. 3, pp. 410–476, Jun. 2008.
- [41] R. Battin, R. Crocker, J. Kreidler, and K. Subramanian, "Leveraging resources in global software development," *IEEE Software*, vol. 18, no. 2, pp. 70–77, Mar. 2001.
- [42] C. Ebert and P. De Neve, "Surviving global software development," *IEEE Software*, vol. 18, no. 2, pp. 62–69, Mar. 2001.