CONTENT AND TEMPORAL ANALYSIS OF COMMUNICATIONS TO PREDICT
TASK COHESION IN SOFTWARE DEVELOPMENT GLOBAL TEAMS

Alberto Castro Hernandez

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2017

APPROVED:

Kathleen Swigger, Major Professor
Yan Huang, Committee Member
Cornelia Caragea, Committee Member
Rodney D. Nielsen, Committee Member
Barrett Bryant, Chair of the Department of
Computer Science and Engineering
Costas Tsatsoulis, Dean of the College of
Engineering
Victor Prybutok, Vice Provost of the
Toulouse Graduate School

Castro Hernandez, Alberto. *Content and Temporal Analysis of Communications to Predict Task Cohesion in Software Development Global Teams*. Doctor of Philosophy (Computer Science and Engineering), May 2017, 53 pp., 25 tables, 5 illustrations, 51 numbered references.

Messages were collected from three software development projects that involved students from two different countries. The similarities and quantities of these interactions were computed and analyzed at individual and group levels. Results of interaction-based metrics showed that the collaboration variables most related to Task Cohesion were Linguistic Style Matching and Information Exchange. The study also found that Information Exchange rate and Reply rate have a significant and positive correlation to Task Cohesion, a factor used to describe participants' engagement in the global software development process. All these results suggest that metrics based on rate can be very useful for predicting cohesion in virtual groups. Similarly, content features based on communication categories were used to improve the identification of Task Cohesion levels. Also, at a group-level, all models were found correlated to Task Cohesion, specifically, Similarity+Rate, which suggests that models that include social and work communication categories are also good predictors of team cohesiveness. Finally, temporal interaction similarity measures were calculated to assess their prediction capabilities in a global setting. Results showed a significant negative correlation between the Pacing Rate and Task Cohesion, which suggests that frequent communications increases the cohesion between team members.. The contributions in this dissertation are three fold. 1) Novel use of Temporal measures to describe a team's rhythmic interactions, 2) Development of new, quantifiable factors for analyzing different characteristics of a team's communications, 3) Identification of interesting factors for predicting Task Cohesion levels among global teams.

Copyright 2017
by
Alberto Castro Hernandez

ACKNOWLEDGMENTS

I would like to acknowledge all the people who encouraged me to pursue my PhD studies. I would not have been able to reach this personal goal without their support, patience, love and guidance during all these years.

First, I am deeply grateful to my advisor, Kathleen Swigger, for her patience, help and understanding through my research. She always provided me with her support and valuable advice even in hard times. I have been really blessed of having the opportunity to work with an excellent person.

I also want to thank my dissertation committee, professors Yan Huang, Cornelia Caragea, and Rodney D. Nielsen for their time and feedback during the development of this work. I greatly appreciate their assistance.

I also would like to express my profound gratitude to Veronica for her support, guidance and inspiration, because her encouragement helped me to start and finish this beautiful dream.

I gratefully acknowledge the funding sources that made my PhD studies possible. I was funded by the National Council of Science and Technology of Mexico, and by the Mexican Secretariat of Public Education.

Last but not least, I want to thank my family for always encouraging me and believing in me. This dissertation is also a result of their unconditional love and support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	viii
CHAPTER 1. INTRODUCTION	1
1.1. Motivation	2
1.2. Problem Definition	3
1.3. Objective	3
1.4. Contributions	4
1.5. Dissertation Outline	5
CHAPTER 2. BACKGROUND	6
2.1. Cohesion Definition	6
2.2. Cohesiveness and Other Group Variables	7
2.3. Team Characteristics and Behavior	7
2.4. Work Related to Measuring Cohesion	8
2.5. Group Communication Types	9
2.6. Temporal Communication Patterns	11
2.7. Summary	12
CHAPTER 3. METHODOLOGY	13
3.1. Global Software Development Projects	13
3.2. Software	14
3.3. Measures	14
3.3.1. Cohesion Captured at the Individual vs Group Level	15
3.3.2. Collaboration Measures	15
3.3.3. Content Measures	20
3.3.4. Temporal Measures	21

3.4. Models	23
CHAPTER 4. RESULTS	27
4.1. Sample Characteristics	27
4.1.1. The Culture Effect	27
4.1.2. Collaboration Measures	28
4.1.3. Relation of Quantity-Based Measures	29
4.1.4. Effect of Group Collaboration Measures	30
4.2. Summary of Collaboration-based Measures	31
4.3. Content-based Features as Cohesiveness Predictors	32
4.4. Temporal Features for Predicting Task Cohesion	38
4.5. Summary of Results	40
CHAPTER 5. DISCUSSION AND CONCLUSIONS	42
5.1. Findings	43
5.2. Contributions	45
5.3. Future Work	46
BIBLIOGRAPHY	48

LIST OF TABLES

3.1	Description of projects	14
3.2	Example of communication among three team members.	16
3.3	Message interaction matrix.	17
3.4	Message reply matrix.	17
3.5	Word count totals.	17
3.6	Message classification.	21
3.7	Interaction models	24
3.8	Group Interaction models	25
4.1	Task cohesion values by culture. $*p < 0.05$.	28
4.2	Partial correlations of collaboration measures and density to Task Cohesion, controlled by culture. $*p < 0.10$, $**p < 0.05$.	28
4.3	Partial correlations of quantity-based measures to Task Cohesion, controlled by culture. $*p < 0.10$, $**p < 0.05$.	30
4.4	Partial correlations of group measures to Task Cohesion, controlled by culture. $*p < 0.10$.	30
4.5	Initial dataset	33
4.6	3-classes dataset	33
4.7	F-score values for classification algorithms.	34
4.8	3-classes distribution in the three projects.	34
4.9	Relevant words by category, according to classifier.	35
4.10	Correlations of content-based variables with Task Cohesion. $*p < 0.1$, $**p < 0.05$	36
4.11	Average and sum of similarity and rate content measures	37
4.12	Comparison of similarity measures and metric measures models. $*p < 0.05$, $**p < 0.01$	37
4.13	Correlations of content-based variables at group-level with Task Cohesion.	38
4.14	Temporal metric statistics	38

4.15	Pacing correlation with Task Cohesion. $*p < 0.1$	39
4.16	Effect of coherence similarity to Task Cohesion. $**p < 0.05$	39
4.17	Linear temporal models to predict Task Cohesion. $***p < 0.01$	40

LIST OF FIGURES

3.1	Frequency graph of 4 team members.	22
3.2	Periodogram of 4 team members.	23
3.3	Coherence graph of each member pair.	24
3.4	Content model.	25
3.5	Temporal model.	26

CHAPTER 1

INTRODUCTION

The growing trend towards globalization in recent years has given rise to the need for greater collaboration among workers living in different countries. Groups of individuals within an organization who are remotely located and who must accomplish tasks using telecommunications technologies are often known as global or virtual teams.

There are numerous benefits for companies who decide to create teams with members from different regions. These benefits include reduced costs, access to people with different skills, and closer proximity to local markets. At the same time, there are numerous drawbacks to using global teams. For example, cultural and language differences among team members can affect how groups communicate with one another; often resulting in a decrease in the amount of collaboration within the group. Also, different perceptions of time and its relationship to the task can affect overall group performance. Finally, many problems can occur as the result of using different types of communication technology, because these tools may not support quality interactions among team members. For example, teams who share information through electronic means, as opposed to having face-to-face exchanges, often have uncertainties about what their distant colleague is doing at any particular time. Similarly, task coordination is generally much more difficult in distributed teams because current scheduling software may not be able to capture the overlap time among team members who are working on the same project in different time zones.

Another important team process that is affected by conditions within a remote collaboration setting is group cohesion. Previous research has found that communication technology (such as chat and email) often delays the development of the group cohesion construct [38]. As a result, global teams tend to have much lower group cohesion levels than co-located groups. One of the reasons why it is important to examine group cohesion levels within global virtual teams is that research has shown that this construct seems to affect how a team deals with different obstacles during the development of a project. In addition, the relation between group cohesion and other constructs (e.g. trust) has been shown to have an indirect relationship to team performance [21].

Recognizing the importance of the group cohesion construct, researchers have developed a number of different methods to measure the extent of group cohesiveness within teams; everything from the use of paper-based individual surveys to the more automatic process of following a team's communication trails. Both methods have shown the importance of looking at the cohesion construct at the team process level.

Following this line of work, this researcher sought to explore the collaboration, content, and temporal features of a team's communication as a way to predict cohesion levels within a global virtual team. In particular, this researcher decided to focus on the analysis of collaboration, content and temporal features of communications within medium-sized global software development teams and use these features to predict the cohesion levels within a group. Subjects for this study were students enrolled in courses that contained a global software development project. Thus, the data gathered for the analysis consisted of team members' interactions about various aspects of these software projects. It was the intention of this research to see which collaboration, content, and temporal feature led to greater perception of cohesion within a group.

1.1. Motivation

The wide use of computer-based communication technology to support team collaborations, even co-located work teams, has given researchers the ability to track the interactions among participants. These interactions can be analyzed and sorted by quantity, content and time; and can provide a rich data source for describing how a team's interaction can affect group work. Previous researchers have used this type of data to predict or describe a number of different group constructs such as leadership, cohesiveness, and group potency [3, 42, 10]. Moreover, some of the data has been used to describe behaviors that were related to effective team performance [45].

Since most work-related interactions now use some type of electronic communication, it should be possible for a computer-based communication system to examine these interactions and provide some useful insights into which variables might affect cohesion levels within groups. Moreover, new methods for analyzing group communications with respect to their collaboration levels, content and temporal meanings should lead to more timely analyses of team processes.

1.2. Problem Definition

In this work we propose a method for integrating collaboration, content-based features and temporal data into individual and group measures that can predict a team's cohesiveness within a global work unit. In previous group behavior research, the quantity of messages has been linked to positive actions within a team and/or effective performance [45]. Moreover, the use of content features has been used to describe other types of group behaviors such as trust, leadership, and agreement [3, 32, 46]. To a lesser degree, the use of temporal analysis has been shown to be an effective measure for evaluating team performance [12].

Using the above studies as a starting point, we studied the integration of metrics based on interaction type and amount, content features and temporal data to predict group cohesiveness. Our investigation used data collected in a collaborative setting in which participants from different countries worked together to complete a software development project.

This study focused on the following aspects of the construct: first, it compared previous collaboration measures that had been used to analyze computer-based communications and determined which of these measures were most effective for predicting *Task Cohesion*; second, we developed a measure based on content features and determined if it could be used to improve *Task Cohesion* prediction; finally, we added temporal-based measures to the content measures in order to determine if it could improve the *Task Cohesion* prediction.

1.3. Objective

The main goal of this research was to investigate whether it is possible to improve the prediction of *Task Cohesion* within groups by using a combination of collaboration, content, and temporal features to analyze group communication activities. More specifically, we sought to answer the following questions:

- i. From existing collaboration measures, which approach is most effective at predicting *Task Cohesion* levels within a global collaborative setting?
- ii. Can we augment *Task Cohesion* prediction rates by adding content features to existing measures?

- iii. Which content features are more appropriate for describing a particular type of interaction?
- iv. Can we improve *Task Cohesion* prediction by adding temporal data to existing measures?

In order to seek answers to the research questions posed above, we performed the following activities.

- Exploring different collaboration measures that can be used to predict group cohesiveness. The previously used measures of *Linguistic Style Matching*, *Density*, and *Information Exchange Similarity* were augmented by newer measures such as *Reply Similarity*, *Reply Rate*, and *Information Exchange Rate*. Measures were then evaluated at both the individual and group level to determine their predictive power of *Task Cohesion*.
- Defining a set of content features for automatic classification of communication classes in global teams. Measures denoting the similarity among a group's messages pertaining to work, social, and planning words were created to classify communication content.
- Integrating content classes into previous collaboration measures to improve their prediction level. The above measures were then combined with similarity and rate measures to improve their predictability.
- Exploring the use of temporal data to improve *Task Cohesion* prediction of studied measures. New measures of *Pacing Similarity*, *Pacing rate*, and *Coherence Similarity* were created to determine if they could improve the predictability of *Task Cohesion*.

1.4. Contributions

This study sought to identify the team characteristics that could successfully predict an individual's perception of *Task Cohesion* in a global software development setting. This work contributes to the field of global software teams in the following ways. First, while group cohesiveness is a well-studied variable in teams (mostly in co-located teams), the effect of specific content variables (such as work, social, and planning) on group cohesion perception has not been thoroughly explored in other studies. Second, the development of an automatic text classification method and then adapting it to the content variables found in software development teams is relatively unique. Finally, the development of temporal variables that can characterize the rhythmic nature of global

software teams represents an important contribution to the knowledge of how distributed teams actually work. More specifically, we used a method based on time series analysis to represent the temporal synchrony between team members, which to our knowledge is fairly unique. All the measures, proposed in this work, provide new information on the study of software development groups as well as virtual teams. More importantly, they provide a guide to how particular variables might be able to predict *Task Cohesion* levels among groups, which may result in better team performance.

1.5. Dissertation Outline

This dissertation is organized as follows. Chapter two presents some background on the use and definition of the cohesion construct in previous studies, and an overview of related work on the measures used to predict cohesion. Chapter three describes the measures that were used in this study and how the measures were calculated. Chapter four reports on the results of the experiments, and Chapter five presents the final conclusions.

CHAPTER 2

BACKGROUND

The following chapter provides a description of the research literature related to the Task Cohesion construct that was used in this study. This section begins by presenting the definition of group cohesion along with a brief description of various dimensions often used to represent the cohesion construct. Various team variables and group characteristics that may be related to group cohesion are also discussed. In addition, we describe some of the techniques that have been used to measure group cohesion. Research related to the different measures used to predict Task cohesion are also discussed at length.

2.1. Cohesion Definition

Cohesion is an important emergent state that is usually defined as "a dynamic process that is reflected in the tendency for a group to stick together and remain united in the pursuit of its goals and objectives" [8]. According to the group dynamics literature, any group cohesion process can be observed from both a group and individual view [7], i.e. Individual cohesion and Group cohesion. Moreover, each of these two perspectives can be analyzed from different contexts i.e. Social and Task oriented. Therefore, according to [9], cohesion can be represented by four dimensions: 1) Group Interaction - Task (GIT), which represents an individual's perception about his or her group members' tendencies to remain in the group because of the task; 2) Group Interaction - Social (GIS), which describes an individual's perceptions about the groups' tendencies to remain in the group because of the social interaction; 3) Individual Attraction to the Group - Social (ATGS), which means the individual's intention to stay in the group because of the social interaction; and 4) Individual Attraction to the Group - Task (ATGT), which represents the individual's intention to stay in the group because of the task.

Another definition of cohesion is called *Perceived cohesiveness*, which describes an individual's feeling about his membership in a particular group [2]. This definition focuses on the degree of belonging or morale feelings that the individual has toward a group.

In this work we focus on one of the dimensions cited in [9], Group Interaction - Task; since the focus of this work is on small work teams. Thus, it seemed appropriate to look at an individual's perception about the group members' tendencies to remain in the group because of the task.

2.2. Cohesiveness and Other Group Variables

Members in a cohesive group are generally cooperative, supportive of one another, and have open communication [50] with their group members. Thus, the cohesion construct is generally seen as having many relationships to other group processes and outputs. For example, [30] found a relation between cohesion and the following teamwork variables: a) Transition processes, which are actions that teams perform between performance episodes such as goal analysis and planning; b) Action processes, which are activities that occur as the team works toward its goal; c) Interpersonal processes, which represent those team activities that are focused on interpersonal relationships. Also, group cohesiveness has been seen as a mediator between transformational leadership and performance [1]. Moreover, there have been a number of studies that have linked cohesion to group performance [49].

Unfortunately, strong group cohesion can sometimes have negative results for a team. For example, a study completed by [40] shows that high social cohesion within a group can sometimes lead to poor performance. The authors speculate that the reason for this poor performance is that the high social cohesion levels can often lead to higher levels of group conformity and a reluctance to criticize other's performances. If team members are reluctant to criticize one another, then the team's overall performance may eventually suffer. Since a healthy dose of criticism is a necessary component of every successful team, groups need to develop a cultural climate that allows members to feel free to criticize ideas.

2.3. Team Characteristics and Behavior

Teams come in various sizes and often come together to socialize and/or work. As a result, the cohesion construct has been studied in many different environments. As recent research shows, the strength of the relationship between group cohesion and performance seems to be affected by the group's task [29].

In comparison with co-located teams, virtual teams tend to be less cohesive [18], although the performance in both distributed and co-located 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 [42]. Other elements that affect group cohesiveness include team size, degree of democratic behavior within a group, participation, and satisfaction [40].

2.4. Work Related to Measuring Cohesion

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 research paper [9] uses the Group Environment Questionnaire (GEQ) to measure different levels of cohesion within groups. In this particular study, the authors measure cohesion levels across four different dimensions: Group Interaction - Task, Group Interaction - Social, Individual Attraction to the Group - Social, and Individual Attraction to the Group - Task. Other researchers have proposed similar dimensions and use surveys to measure their different group cohesion constructs [45] [41] [20] [6]. In addition to measuring group cohesiveness through surveys and self-reports, researchers have developed techniques for measuring the quantitative aspects of a team's interactions. More specifically, [39] 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 relative position of an agent (i.e., message sender) 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's 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 exchange among same 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 statistically significant relationships between cohesion and performance [32].

In this dissertation, we used survey data and a form of LSM to measure different aspects of group cohesiveness. The survey used in this study was developed by [5] and was based on the GEQ survey [9]. This particular survey measures cohesiveness among work groups along three different constructs: Group Interaction - Task (GIT), Group Interaction - Social (GIS), and Individual Attraction to the Group - Social (ATGS). In this research we used only 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.

2.5. Group Communication Types

Communication activities within a group have always been linked to cohesion among team members. Thus, the measurement and analysis of communication behaviors has been an ongoing focus of team research. The most common way to analyze communication data is to look at the low-level quantitative measures such as the number of communications or the length of a message [51]. However, these numbers and amounts do not generally capture the semantics of the communication activities. Thus, more content-oriented approaches have been suggested in order to discover underlying team processes. These content approaches usually involve the construction and application of a coding scheme that represents all the interesting categories of a particular type of communication such as the rules being displayed in the conversation [16], types of speech [13], or the actual meaning of the discussion [17]. The transcribed discourse is then sorted into different categories and tagged [16]. The tagged content can then be analyzed either as frequency counts of the categories themselves or as a series of events.

Researchers have developed many different types of coding schemes for analyzing group communication behavior. For example, [31] introduced a coding system that includes six categories that cover aspects of communication, joint information processing, coordination, relationship management, and motivation. [27] proposed a model that argues that the major types of communication behaviors within a collaborative setting are: giving and receiving help; exchanging resources and

information; explaining and elaborating information; sharing existing knowledge with others; giving and receiving feedback; challenging others' contributions; advocating increased effort and perseverance among peers; engaging in small group skills; monitoring each other's efforts and contributions. [14] used a similar approach to develop a coding system that was designed to characterize behaviors associated with positive social interdependence, as opposed to those behaviors linked to a more individualistic and competitive learning environment [14]. The coding scheme consists of five categories including planning, contributing, seeking input, social interaction, and reflection. [43] then used these same categories to analyze different communication behaviors within global software development projects to determine their effects on performance. The authors of this study found that successful (as compared to unsuccessful) global software development teams generally have some communication activities in all five categories, with the highest proportion of behaviors in the contributing and planning areas, in that order. Moreover, the least successful teams seemed to have more communication activities related to socializing and planning, as compared to contributing.

More recently, researchers have developed software that essentially automates the human process of assigning categories to text. One of the more frequently used software packages was developed by [36] and is designed to categorize content words into 74 different categories. An extension of this original work was the development of the LSM technique, which was a procedure used to measure the degree of similarity of words used among group members. As previously stated, the authors of this work found a statistically significant relationship between LSM and cohesion rates among team members, and a statistically significant, but weak, relationship between LSM and performance.

Both LSM and the Curtis & Lawson content categories [14] were used in this study to explore the relationship between message content and *Task Cohesion*. Since this study was interested in predicting cohesion rates in teams, it was felt that both these measures could provide clues into the type of communication that would promote *Task Cohesion* within a group. Thus, LSM and the Curtis & Lawson categories were applied to logs of students chats and discussion forums to characterize the interactive communication behavior patterns that existed among the student groups who participated in this study.

2.6. Temporal Communication Patterns

While the previous section established the need to look at the content of messages to predict cohesion, there is also evidence to support the need to look at temporal characteristics of group work. For example, [33] applied Time Series Analysis to data coded from student online tasks to characterize changes in students' emotions. [15] implemented time-line analysis in their research design to track the changes in student participation over time, and [28] introduced Lag-sequential analysis in conjunction with a coding scheme to show that temporal patterns were statistically significant related to variations in group performance. Unfortunately, the hand-coded temporal analysis of a teams behaviors that was used in most of these studies tends to be slow and time consuming.

Temporal factors have also been important to the success or failure of global software teams. Software development projects generally have milestones that occur throughout the project. These milestones, or phases, are accompanied by a spike in the exchange of information or code among team members. Each milestone for each phase represents a temporal instant that allows team members to determine if they have met their objectives, and that they are on pace and making progress toward the goal. Several researchers have tried to analyze the temporal aspects of the software development life cycle. For example, [25] found that frequent and timely responses to team members' posts resulted in more trust and cohesion within groups. Similarly, [4] found that faster response times among team members led to better performance, as measured by scores on a software development project. These findings seem to suggest that groups may be able to develop certain temporal patterns that engender team interactions and promote greater cohesion.

To describe these cycles, one must look to researchers outside the computer science area to find a coherence measure [22] that can be applied to student pairs within a team. To study different aspects of child-parent interactions, Gottman defined a coherence coefficient, similar to a correlation coefficient, that indicates whether two cycles peak at similar frequencies. Through such analysis, it is possible to determine if a pair is in sync. Gottman found the temporal patterns among pairs by performing cross-spectral analysis that occurred between each child-parent dyad, and then analyzed the resulting squared coherence plots. Squared coherence plots can be thought of as a measure of

correlation at each frequency. The more highly correlated a dyad is at both high and low frequencies, the more cohesive that dyad is. A low squared coherence for a dyad across all frequencies is an indicator of team fragmentation. By using such measures, it should be possible to examine the coherence rates among global software teams. Thus, this study used measures related to pacing and cross spectral analysis as predictors of *Task Cohesion* levels among software development teams.

2.7. Summary

This chapter presents an overview of the important literature that was used to guide this study. These ideas were then formulated into theoretical constructs that could be used to predict different *Task Cohesion* rates among software development teams. The research literature created a very powerful tool for defining the various factors that could be used to examine cohesion levels among team members. Thus, the research literature led to the initial development of the major factors that were used in the model presented in the next chapter.

CHAPTER 3

METHODOLOGY

In this chapter, we present the measures and methodology that were used to predict task cohesion within a global software development group. Three different types of measures were developed to help establish the relationship between different variables and a group's *Task Cohesion* scores. The first set of measures attempts to capture a team's collaboration-oriented interactions. The second set of measures address the use of different content classes that characterize the different types of communications that can be used to improve the *Task Cohesion* prediction. Finally, we describe the cohesion prediction measure through the use of temporal measures. All the different predictive measures were applied to data collected from the interactions among global software development student teams. A more detailed description of team composition and the assigned projects now follows.

3.1. Global Software Development Projects

The data set that was used in this dissertation was drawn from three student global software development projects that occurred between 2014 and 2015. Students who participated in these projects were enrolled in computer science or information technology courses at Universities in the US and Mexico. While each of the three projects contained some aspect related to the software development process, the actual assignments often varied in terms of team size, team composition, length of project, and specific task. For example, two of the projects asked students to redesign a non-profit website. The specific tasks consisted of redesigning three sections of the website (i.e., the home page, the events page, and the contribution page) and implementing a database that could support the various operations that were needed to maintain the three pages. A third project consisted of the creation of a website that would allow students to learn about an optimization algorithm. The various elements of the website included a section where users could read information about the algorithm, as well as a sections where users could test the alogorithm.

Project	Project A	Project B	Project C
Task	Web page development	Web page development	AI web page development
Duration	4 weeks	4 weeks	4 weeks
Surveys	Group Potency, Task Cohesion and Culture	Moral-Cohesion, Task Cohesion and Culture	Moral-Cohesion, Task Cohesion and Culture
Participants	15 teams, 4-5 members by team, 73 students	10 teams, 4-5 members by team, 47 students	10 teams, 5-8 members by team, 60 students

TABLE 3.1. Description of projects

Team composition among the three projects also varied slightly. Teams who participated in the 2014 projects had between 4-5 members, while the 2015 teams had between 5-8 members. A summary of the different project variables is presented in Table 3.1.

3.2. Software

Students who participated in the three projects were asked to communicate with one another using a project management web application. This application platform supported several collaborative tools including chat, forums, wikis, document sharing, etc. Moreover, the particular application software was enhanced so that it recorded and timestamped all interactions among group members and transferred this information to a centralized database. Additional refinements were made to the platform in order to provide feedback to some groups concerning their collaborations and participation levels.

3.3. Measures

The major objective of this research was to determine which factors are most predictive of a team's cohesiveness within a global software development context. This was achieved by focusing on specific factors that tend to affect team cohesion. These factors were derived from the research literature [19, 20, 46], and from smaller experiments performed by the author that looked at issues related to predicting the group constructs of *Task cohesion* [10, 11].

3.3.1. Cohesion Captured at the Individual vs Group Level

Cohesiveness is a construct that measures the degree to which team members are working together. Although some researchers have proposed capturing cohesiveness levels through a group-level survey (i.e., members complete a single survey while sitting together face-to-face), a much more common, and easier, approach is to have team members complete individual surveys. The individual survey approach is also more appropriate within a virtual team context [24], because members of the team are not co-located. Therefore, this research focused on the prediction of group cohesion through an individual survey intended to capture an individual's perception of their team as defined by questions on the *Task Cohesion* section of the Group Environment Questionnaire (GEQ) [5]. A group's cohesion measure was then obtained by averaging each group's individual responses to the survey. In this work, the cohesion values obtained from this survey were then aggregated and called *Task Cohesion*.

Having defined the measure of *Task Cohesion*, we then asked the question of which factors relate to the degree to which an individual perceives the team's ability to complete the task. Thus, the specific cohesion measure that we attempted to predict was *Individual Task Cohesion*.

3.3.2. Collaboration Measures

In order to determine which measures successfully predict *Task Cohesion*, we began by performing a set of experiments that were aimed at capturing different aspects of a team's collaboration activities. We hypothesized that there were six major collaboration measures: reply similarity, reply rate, information exchange similarity, information exchange rate, density, and Linguistic Similarity (LSM). Each of these measures was examined at both the individual and group levels. The reason we began with these particular measures was because these particular factors have been identified in previous research as having some affect on the cohesion construct. Thus, we wanted to see if these same variables had similar effects on our global software learning teams.

All of the collaboration measures were obtained by examining messages sent from and to other group members. In a virtual work environment, it is common to use communication tools such as chat (individual or group), forums, emails and file sharing to communicate with other members of the team. A common feature of these types of communication tools is that the user can specify

User	Message
member01	Hello. My name is Bryan and I am an student at University of North Texas...
member02	Hi, I am Carlos.
member02	I live in Panama.
member02	I study Computer Engineering.
member01	Nice two meet you
member01	I just read the project description and it seems that ...
member03	Hello I live in Texas, and I study Computer Science

TABLE 3.2. Example of communication among three team members.

which members receive the messages or information; e.g., when you write an email (or forum post) you usually define a recipient.

However, in a group chat all the posted messages are seen by all members of the team. So, the message recipient in a chat conversation could be all or one specific member. In the projects analyzed for this research, participants used a group chat that sent messages to all members of the team. Thus, we created a count that reflected when a message was delivered to one specific participant (i.e., the last participant in the communication).

Usually, messages extracted from chat conversations are much shorter than forum posts. In addition, participants often reply to other chat posts with a sequence of short messages. To handle counts for the latter case, we counted sequences of messages from the same member as a single message sent *reply*.

An example of a typical sequence of chat messages can be found in Table 3.2. In this example, we observe a conversation among three participants. Although all members of the team receive the same message, it is obvious that certain messages are intended for only one person. For example, there are three messages from *member02* to *member01*, two messages from *member01* to *member02* and one message from *member03* to *member01*. The total message counts for each participant for this example is shown in Table 3.3. Also, the number of replies by each person is shown in Table 3.4, i.e. *member02* replied once to *member01*, *member01* replied once to *member02* and *member03* replied once to *member01*.

Based on the example presented in the previous table, we proposed a new collaboration measure called *Reply similarity* as a way to predict individual *Task Cohesion*. For this measure, we assumed that each reply had one specific recipient. The *Reply similarity* for a person was then

Source/recipient	member01	member02	member03
member01	0	2	0
member02	3	0	0
member03	1	0	0

TABLE 3.3. Message interaction matrix.

Source/recipient	member01	member02	member03
member01	0	1	0
member02	1	0	0
member03	1	0	0

TABLE 3.4. Message reply matrix.

Member	Word count
member01	29
member02	12
member03	10

TABLE 3.5. Word count totals.

calculated by using the following formula,

$$(1) \quad \text{replysimilarity}_i = 1 - \frac{\text{abs}(\text{reply}_{ij} - \text{reply}_{ji})}{\text{reply}_{ij} + \text{reply}_{ji}}$$

Where reply_{ij} is the analyzed reply variable from person i to person j , and count_{ji} is the analyzed reply variable from person j to person i . This formula produces a value between 0 and 1. As stated above, this calculation was called *Reply similarity*. This measure was introduced with the assumption that people view their group's cohesiveness as being a combination of their own participation as well as others. Thus, the *Reply similarity* measure was hypothesized as a variable that can predict *Individual Task Cohesion*, as cited in [11].

We also calculated a measure called *Reply Rate*, which was defined as simply the number of replies sent by one person to any other person on the team. Similarly, we calculated group-level reply measures called *Group-level Reply Similarity* and *Group-level Reply Rate*, which consisted of summing up the *Reply similarity* and *Reply rate* scores for each team member of the team and then averaging these scores.

Another measure that was used to characterize team collaboration was the similarity of *Information exchanged* [46] among team participants. For this research, *Information exchange*

similarity was calculated by counting the number of words typed by each participant, with the assumption that these words were being transmitted to every other team member in the group (i.e. each word in each message was perceived as some type of participatory exchange). For example, the word count for each member who participated in the conversation shown in Table 3.2 is presented in Table 3.5. This calculation was computed as follows:

$$(2) \quad informationexchange_i = 1 - \frac{abs(wordcount_i - wordcount_j)}{wordcount_i + wordcount_j}$$

Where $wordcount_i$ and $wordcount_j$ are the number of words used by user i and j , respectively.

In addition, we calculated an *Information exchange rate*, which was the number of words sent by a person to any or every other person on the team. Similarly, we computed a *Group-level Information Exchange Similarity* and a *Group-level Information Exchange Rate* count by averaging the individual information exchange counts of each member on the team.

In social network analysis (SNA), a density measure is often used to describe a subset of nodes that are highly interconnected [48] within a large community. This calculation has also been used to determine the degree of connectivity (or collaboration) within small group settings [19]. The density measure consists of counting the number of existing links between any two team members in the group and then dividing this number by all possible links for that group. As observed, this is generally considered a group measure. However, since our analysis was developed to examine *Task cohesion* at the individual level, we modified the group density measure to calculate an *Individual-level density* score. This score was computed as follows:

$$(3) \quad individualleveldensity_i = \frac{links_i}{n - 1}$$

Where $links_i$ are the number of members that receive/sent messages from/to a user i during the project; and n is the number of participants in the team.

Again, this measure was calculated at the group-level using the normal group-level density measure as a way to predict *Task Cohesion*.

Finally, the sixth collaboration measure that was used to predict *Task Cohesion* was called *Linguistic Matching Similarity (LMS)*, as defined in [20]. While some previous studies found a statistically significant relationship between LMS and group cohesion in chat communications, others found no relationship between LMS and group cohesion in email communications [32]. Thus, it seemed appropriate to include this measure as part of this research to determine whether LMS has a relationship with *Task Cohesion* in this particular context.

The LMS measure is designed to show the degree to which two people share similar linguistic styles; assuming that people who have similar linguistic styles feel more connected to one another than those who do not have similar styles. LMS uses a special Linguistic Inquiry and Word Count (LIWC) software tool [23] that analyzes text on a word-by-word basis and calculates a percentage of words that occur within a pre-defined set of linguistic categories. The authors of the original LIWC work used these counts to develop similarity measures for nine function word categories: auxiliary verbs, articles, common adverbs, personal pronouns, indefinite pronouns, prepositions, negations, conjunctions, and quantifiers. Words in each of these categories are summed and averaged. For example, the personal pronouns (pp) similarity measure is calculated as follows.

$$(4) \quad LSM_{pp_i} = 1 - \frac{abs(pp_i - pp_j)}{pp_i + pp_j}$$

Where LSM_{pp_i} is the LSM for personal pronouns for person i ; besides pp_i and pp_j are the percentage of personal pronouns in total words typed by user i and j , respectively. Using the LIWC software tool, similarity scores were calculated for each of the nine function word categories. An individual similarity score was obtained by averaging individual function word similarity values. Once the individual scores were computed, they were averaged for the team and labeled as a group LSM. We used this same approach to obtain the linguistic similarity scores for each of the individuals and teams that participated in the projects. We then tried to determine their predictive capability for *Task Cohesion*.

As previously stated, this research began by looking at the relative strength of each of the six collaboration variables (as defined above) to predict *Task Cohesion*. This was done at both the

individual, and/or where appropriate, the group level. Once this was completed, we then selected the variables that seemed most able to predict *Task Cohesion* and tried to determine whether other variables could improve our prediction capabilities.

3.3.3. Content Measures

In addition to the nine linguistic categories discussed in [20], studies have discovered that other types of content word categories have had an effect on a person's perception of a group's *Task Cohesion*. For example, previous research [44] found that the word categories of Contribution, Seeking input, Reflection, Social and Planning can also affect team performance. Similarly, words associated with social content also seem to influence the level of trust among team members as well as create a more pleasant environment within a team. Although these word categories may not have a direct effect on *Task Cohesion*, it is natural to assume that they may have some indirect effect on *Task Cohesion*; for example, the absence of social behavior may result in a decrease in *Task Cohesion* among group members.

Therefore, we examined the relationship between *Task Cohesion* and the following content categories:

- Social: those messages which contain social interactions.
- Planning: those messages which contain verbs, nouns and dates related to organizing the project development.
- Work: those messages that contain information about the project. Borrowing from the original research [14], the work category included words related to Contribution and Seeking input labels.

The Reflection category, which was included in the original research, was not used in this study because, as other studies found, students' messages that contain words associated with reflection are extremely rare and, as a result, do not usually occur in sufficient numbers to be compared.

Before classifying the various messages into the appropriate content categories, we removed all sentences that contained *agreement* messages, so that these types of short messages were not

Message classification process
1. Remove agreement messages.
2. With training dataset
3. Extract unigrams.
4. Extract LIWC categories.
5. Train a Support Vector Machines classifier.
6. Train a Naive Bayes classifier.
7. Train a Random Forest classifier.
8. Select classifier with best performance.
9. Use classifier on the unlabeled set.

TABLE 3.6. Message classification.

included in any of our counts for the proposed categories. Agreement messages are those that confirm or deny some previous message such as "ok," "sure," "good," etc.

The algorithm that was used to classify the messages into the different work categories is described in Table 3.6.

Once the messages were placed into their appropriate content categories, we calculated the reply similarity and reply rate for each of those categories: *Social reply similarity*, *Social reply rate*, *Planning reply similarity*, *Planning reply rate*, *Work reply similarity* and *Work reply rate*. To calculate these scores, we used the same similarity approach as described in section 3.3.2.

3.3.4. Temporal Measures

Another way to examine the recorded data from software development groups is to analyze the time spent participating in paired communication synchrony. The two measures that capture this type of communication are *Pacing* and *Coherence*.

The *Pacing* measure was defined as the average number of seconds between messages from the same user. We then computed a *Pacing similarity* score for each member by comparing the pacing score of user i to every other team member's pacing score, one at a time. These paired scores were then averaged and defined as an individual's *Pacing similarity* score. A *Pacing rate* score was also computed for each subject by simply counting the number of seconds between messages from a user.

The second major temporal measure deals with the synchrony of messages sent between two individuals; where the messages from each user are defined as a time series, and then these

FIGURE 3.1. Frequency graph of 4 team members.

frequencies are then compared to one another, one at a time [23]. For instance, Figure 3.3.4 shows, as a data series, the number of message sent by four participants in a team.

Then we estimated the spectral density from each data series by creating a periodogram using a Fast Fourier transform. The Figure 3.3.4 shows the periodogram, a diagram of frequencies by amplitude.

Then we calculated the *Coherence* between two time series which is the correlation between pair members of each frequency. As a result we got a *Coherence* series for each combination of a pair in the team ($c(2, n)$), see Figure 3.3.4.

This resulted a vector of *Coherence* values (each value representing *Coherence* in a specific frequency) between two subjects. To provide a unique temporal *Coherence* value between two individuals, we took the highest coherence score. Finally, for each participant in a team, we calculated a *Coherence* with the rest of the team members and averaged those scores. This measure was called *Temporal coherence*.

FIGURE 3.2. Periodogram of 4 team members.

3.4. Models

Using the previous metrics, we proposed three sets of experiments: 1) An assessment of the relationships between the collaboration measures and *Task cohesion* (as described in Section 3.3.2), 2) An assessment of the performance of content-based similarity features (as described in Section 3.3.3) in the prediction of *Individual Task cohesion*, and 3) An assessment of the performance of temporal features (as described in Section 3.3.4) as predictors of *Task cohesion*.

In the first set of experiments, we examined the relationship between *Task Cohesion* and the three measures based on previous works: *Linguistic Style Matching*, *Individual density* and *Reply similarity* (published as *Individual cohesion*). We also examined the relationship between *Task Cohesion* and some new collaboration measures: *Information exchange similarity*, *Information exchange rate* and *Reply rate*. As can be seen in Table 3.7, all the models were controlled by *team size*.

FIGURE 3.3. Coherence graph of each member pair.

TABLE 3.7. Interaction models

Models
Team Size + Individual density = Task Cohesion
Team Size + Reply Similarity = Task Cohesion
Team Size + Linguistic Style Matching = Task Cohesion
Team Size + Reply Rate = Task Cohesion
Team Size + Information Exchange Similarity = Task Cohesion
Team Size + Information Exchange Rate = Task Cohesion

We then analyzed the same set of variables at the group level by using *Group LSM*, *Density*, and *Group Reply similarity*, and *Group Reply similarity* + *Group Information exchange*, as seen in Figure 3.8.

In the second set of experiments, we assumed that the *Reply similarity* model would be a statistically significant factor in predicting *Task Cohesion*. However, we believed that not all messages are relevant to cohesion prediction. So, we split the *Reply similarity* feature into three

TABLE 3.8. Group Interaction models

Models
Team Size + Density = Task Cohesion
Team Size + Group Reply Similarity = Task Cohesion
Team Size + Group Linguistic Style Matching = Task Cohesion
Team Size + Group Reply Rate = Task Cohesion
Team Size + Group Information Exchange Similarity = Task Cohesion
Team Size + Group Information Exchange Rate = Task Cohesion

FIGURE 3.4. Content model.

content measures: *Social reply similarity*, *Work reply similarity*, and *Planning reply similarity*. This particular model is presented in Figure 3.4.

In the third set of experiments, we analyzed a team's interaction tempo by adding temporal data into the model. Thus, we considered *Temporal coherence*, *Pacing Similarity*, and *Pacing Rate* measures in the next proposed model. We also chose between *Pacing Similarity* and *Pacing Rate* as the best pacing representation to create a model. The model is presented in Figure 3.5.

Results obtained from these experiments helped clarify which particular variables have strong relationships to *Task Cohesion* and how these various constructs interacted with one another.

FIGURE 3.5. Temporal model.

CHAPTER 4

RESULTS

Communication and survey data were collected from three different global software development student projects. This data was then used to determine which set of variables were most successful at predicting *Individual Task Cohesion*. More specifically, this study examined the relationships between a team's collaboration activities (as measured by *Reply similarity*, *Reply rate*, *Information exchange similarity*, *Information exchange rate*, *Density*, and *Linguistic Similarity (LSM)*), message content (as measured by *Social reply similarity/rate*, *Planning reply similarity/rate*, and *Work reply similarity/rate*), and temporal activities (as measured by *Pacing similarity*, *Pacing rate* and *Coherence*) and *Task Cohesion*.

4.1. Sample Characteristics

A total of 5583 messages were transmitted during the three projects. A total of 167, out of a possible 180, *Task Cohesion* surveys were collected. Since we had 23 missing surveys, we reduced the data set to include only those messages sent by students who had completed the questionnaire; thus, we had a total of 5446 messages in the final data set.

4.1.1. The Culture Effect

We received cohesion surveys from people who were born in eleven different countries. However, the majority of the completed surveys came from students born in India, the US, and Mexico (i.e. $n > 20$), while only a few surveys came from students born in other countries (i.e. $n < 4$). As a result, we reduced the data set even further and used data from only students born in India, the US, and Mexico; thus, ending up with a final count of 4849 messages sent by 153 individuals. We then compared the *Task Cohesion* mean values between countries and found that students from India tended to have higher *Task Cohesion* perceptions than either US or Mexican students (see Table 4.1). As a result, we used the *Culture* factor as a control variable in our analyses.

TABLE 4.1. Task cohesion values by culture. $*p < 0.05$.

Country	mean	India	US
India	8.01		
US	6.21	1.8057*	
Mexico	6.48	1.5310*	0.2746

TABLE 4.2. Partial correlations of collaboration measures and density to Task Cohesion, controlled by culture. $*p < 0.10$, $**p < 0.05$.

Measure	Task Cohesion
Density	0.037
Reply similarity	0.060
LSM	0.114*
IES	0.152**

4.1.2. Collaboration Measures

In order to determine the relationship between our collaboration measures and *Task Cohesion*, we began by computing a partial correlation between the major measures (i.e., *Reply similarity*, *Individual Density*, *Linguistic Style Matching* and *Information exchange similarity* and *Task Cohesion*), controlling for *Culture*. The results of these correlations are reported in Table 4.2.

Results suggest that the *Density* measure, at the individual level, is unrelated to *Task Cohesion* ($r=0.037$, $p=0.325$). Since weights assigned to different edges in the reply graphs tended to vary widely (varying from 1 to 70), we suspect that the large variance among members' exchanges may have affected the strength of the correlation between *Density* and *Task Cohesion*. Similarly, we found no correlation between *Reply similarity* and *Task Cohesion* ($r=0.060$, $p=0.230$). One reason for the lack of a correlation between these two variables may be explained by the lack of communication activity of one or two members in a group. For example, we noted, in most cases where one or more group members did not participate in the group's discussions, the *Reply similarity* scores were very low.

On the other hand, *Linguistic Style Matching* has a positive, although weak, statistically significant correlation with *Task Cohesion* ($r=0.114$, $p=0.081$). An analysis of the use of function words by culture for the first week of each project shows that the student teams from the countries

who participated in Project A had a statistically significant difference in their use of personal pronouns ($p=0.027$) and quantitative words ($p=0.054$). Also, student teams from the countries who participated in Project C show a statistically significant difference in their use of impersonal pronouns ($p=0.055$); however, Project B participants show no difference in function-word usage by country. It is important to note that the second project (Project B) consisted largely of students who were born in either Mexico or the US; while the other two projects consisted mainly of students born in either Mexico or India. Thus, it appears that US and Mexican students have more similar linguistic patterns than Mexican and Indian students.

Finally, *Information exchange similarity* has a positive and statistically significant correlation with *Task Cohesion* ($r=0.152$, $p=0.031$). Although *Information exchange similarity* and *Reply similarity* are somewhat related, in terms of what they are measuring, the simple word-based metric of *Information exchange similarity* shows a higher correlation with cohesion perception, possibly because words rather than replies (which tend to be sentences) produce more data. The amount of data that are used to analyze the relationship among variables may affect the degree of statistical significance that can be attained. Word counts may also be a better metric of different levels of interaction, since individuals who tend to be more engaged in the project will probably communicate more, which in turn may affect the perception of the group's cohesiveness.

4.1.3. Relation of Quantity-Based Measures

We next tried to determine if the intensity of the communications, as measured by *Reply rate* and *Information exchange rate*, might affect our cohesion measure. The reader might recall that both *Reply rate* and *Information exchange rate* are aggregate counts of their respective communication variables. For example, *Information exchange rate* is simply a count of the number of words that an individual generates. Results of this analysis are shown in Table 4.3. The first column repeats the correlation scores for both *Reply rate* and *Information exchange rate*, while the second column shows the correlation scores for the rate version of those same two variables.

Our results indicate that *Reply rate* has a positive, although weak, statistically significant correlation with *Task Cohesion* ($r=0.108$, $p=0.093$). *Information exchange rate* also has a statistically significant positive correlation with *Task Cohesion* ($r=0.175$, $p=0.016$). In comparison with the

TABLE 4.3. Partial correlations of quantity-based measures to Task Cohesion, controlled by culture. * $p < 0.10$., ** $p < 0.05$.

Measure	Similarity	Rate
Reply	0.060	0.108*
Information exchange	0.152**	0.175**

TABLE 4.4. Partial correlations of group measures to Task Cohesion, controlled by culture. * $p < 0.10$.

Measure	Task Cohesion
Group reply rate	0.113*
Group information exchange rate	0.112*
ROT reply rate	0.076
ROT information exchange rate	0.052

regular collaboration measures, the more quantitative-based measures have higher correlation values. This seems to suggest that quantity based measures that capture amounts, and perhaps degree of engagement, may be better predictors of group cohesion than measures that try to assess different collaboration similarity constructs in a team's exchanges.

4.1.4. Effect of Group Collaboration Measures

We then took the two variables that appeared to be statistically significant related to *Task Cohesion* and calculated group level scores for each of these two measures. Thus, we created a *Group reply rate* and *Group information exchange rate* variable and examined the relationship between these two factors and *Task Cohesion*. We also looked at the relation between *Task Cohesion* and *Rest-Of-the-Team (ROT) reply rate* and *ROT information exchange rate*, both variables are obtained by computing scores for the group, but minus the individual data. Table 4.4 shows the correlations for each of these four variables.

Group reply rate and *Group information exchange rate* show a positive, but weak, statistically significant correlation with *Task Cohesion*, i.e. $r=0.113$, $p=0.083$ and $r=0.112$, $p=0.085$, respectively. However, *ROT reply rate* and *ROT information exchange rate* show no correlation with *Task Cohesion* i.e. $r=0.076$, $p=0.175$ and $r=0.052$, $p=0.264$, respectively. Since the difference between a group variable and a ROT variable is the member's data, these results again suggest that an

individual's engagement in communication is a better predictor of *Task Cohesion* perception than a group measure.

4.2. Summary of Collaboration-based Measures

This research examined a number of different collaboration measures to determine whether certain types of factors were better predictors of *Task Cohesion* than others. Similarity measures, drawn from research and the author's own works, and quantitative measures were created for this study. These measures were then applied to data generated from three global software learning projects.

Individual similarity measures intended to predict *Task Cohesion* had mixed results. For example, individual *Density* was not correlated with *Task Cohesion*. At the same time, *Reply similarity* had no statistically significant relation to cohesion. The lack of statistically significant results for these two variables can be explained by looking at the number of exchanges between members in the co-located teams. Students seemed to have fewer exchanges between team members in their own country as opposed to team members in the other country. Possibly, co-located team members may have had offline conversations that were not measured. It may also be the case that the measure itself needs to be redefined, since scores for both *Density* and *Reply similarity* are highly affected by inactive members. On the other hand, *Linguistic Style Matching* had a small, but statistically significant relationship to group cohesiveness. We speculate that the statistical significance was low because of the large difference between the use of function words by Indian students as compared to Mexican students. *Information exchange similarity*, which is a word-based calculation, also achieved a statistically significant correlation with *Task Cohesion*. This seems to suggest that the simple measure of number of words provides a better representation of cohesion than either messages or replies, perhaps because such a measure generates more data for the analysis.

Since communication within a virtual learning team tends to vary, metrics based on interaction intensity were also proposed i.e. *Reply rate* and *Information Exchange Rate*. Both seemed to predict *Task Cohesion* much better than the similarity version of these two variables. It should be noted that the word-based factor of *Information exchange rate* was a better predictor of *Task Cohesion* than *Reply rate*. Although we computed group versions of *Reply rate* and *Information*

Exchange Rate we found only weak statistically-significant relationships between either one of the variables and *Task Cohesion*.

These results seem to indicate that previously cited similarity measures used to predict *Task Cohesion* may be limited to analyzing groups that are highly interactive and that work on short-term tasks. Data from this study suggests that similarity measures may be affected by less-active members and the presence of cross-cultural teams, both of which can impact message length and word usage. Since global software teams have both of these characteristics, similarity measures may have limited value in computing cohesiveness for distributed learning projects. However, this study also found that several quantitative measures that captured both reply and word rates were useful predictors of a group's cohesiveness within a global software development learning team. Thus, further research was done to determine whether a content analysis approach (which may detect how participants are talking about the project or when there exists a conflict); in combination with a rate approach (which may help to identify when good-performance teams start working), could produce better predictions of the perceived cohesiveness within a distributed team.

4.3. Content-based Features as Cohesiveness Predictors

Following the examination of the relationship between previously known collaboration measures and cohesion, we began looking at the actual content of students' messages to determine whether this could increase the predictive power of our model. As part of the procedure, we developed an automated text analysis program that could classify students' messages into different categories, as defined in [14]. Our test data set consisted of 1866 messages that had been manually-annotated for inclusion in a previous work [26]. The different word categories (as defined by [14]) and their distribution for the initial data set is listed in Table 4.5.

As previously mentioned, we did not try to classify messages related to the *Reflection/Monitoring* category, because these types of messages occur very rarely within a virtual team environment. Similarly, we removed those messages that contained small phrases that indicated agreement such as *ok, good, alright*. We eliminated these types of phrases by looking at the *Agreement* category value (as determined by the LIWC tool) for each message. Those messages that contained a value higher

TABLE 4.5. Initial dataset

Class	Instances
Social	340
Planing	166
Contributing	927
Seeking input	394
Reflection/Monitoring	39

TABLE 4.6. 3-classes dataset

Class	Instances
Social	305
Planning	166
Work	1279

or equal to 0.5 were removed. Following this procedure, we removed 77 messages (from *Social* and *Contributing* classes) from the final data set.

In addition, we collapsed the *Contributing* and *Seeking input* categories into a new class called *Work*. The rationale for combining these two classes was that any message that contained references to work (e.g., project, program, etc.) were related to the actual programming project and should, thus, be assigned to a work related category. The final list of categories and their distribution is presented in Table 4.6.

The features used to help seed the classification process were those found in research related to the LIWC software tool. Using this tool, we investigated 73 features. We also computed the unigrams of each message, obtaining a dictionary of more than 3000 entries.

We tested the ability to use both feature sets (i.e., Unigrams & LIWC) to predict the target class. Similarly, we tested the use of only LIWC features for predicting cohesion. For this experiment, we created a test dataset that had classified messages into one of the three stated categories.

The test data set was then use to compare the performance of three classifiers: Support Vector Machines, Random Forest, and Naive Bayes. The performance metric used for comparison was the F1-score, which is the harmonic mean of precision and recall.

TABLE 4.7. F-score values for classification algorithms.

Feature set	Support Vector Machines	Random Forest	Naive Bayes
LIWC+Unigrams	0.821	0.777	0.522
LIWC	0.752	0.776	0.495

TABLE 4.8. 3-classes distribution in the three projects.

Class	Instances
Social	936
Planning	4528
Work	392

As seen in Table 4.7, all classifiers obtained better results by using both the LIWC and Unigrams features as compared to using only LIWC features. In addition, the best performances were obtained by the Support Vector Machines classifier (F1-score=0.821).

As a result of this study, we used the Support Vector Machine classifier to label all the messages in our three experimental projects. The result of applying the classifier to the messages in the experimental dataset can be found in Table 4.8.

Some of the features that provided more information for classifying each message into a category are shown in Table 4.9. The Work label was found to be related to unigrams that are sometimes associated with performing activities such as give, talk, data, instructor, right. However, the Work label was also found to be related to three LIWC categories: 1) *Function words* (FUNCTION), a category that includes words such as prepositions, articles, and common adverbs, which are often related to the idea of formal thinking [37]; 2) *Work*, a category that includes words such as accomplish, work, and success, which are often related to the idea of well-performed teams [11]; 3) *Interrogatives* (INTERROG), a category that includes words such as how, when, what, which are often used to represent the students' seeking input process.

Similarly, words related to the Social category included words such as fun, lets, thanks, hello, later, nice - all of which seem to suggest characteristics related to social interaction. The Social category also included more concrete relations with specific LIWC categories: 1) *Informal*, a category that consists of words such as Netspeak (lol, btw, thx), Swear words (fuck, damn, shit),

TABLE 4.9. Relevant words by category, according to classifier.

Work	Social	Planning
Right	LIWC.INFORMAL	meeting
LIWC.FUNCTION	fun	night
LIWC.WORK	Lets	tonight
LIWC.INTERROG	Thanks	Thursday
data	Oh	whenever
talk	LIWC.FOCUSPRESENT	must
instructor	later	yet
could	LIWC.SOCIAL	schema
modification	Nice	yours
give	hello	Monday

Nonfluencies (err, mmm); 2) *Present tense* (FOCUS PRESENT), a category that includes words such as today, is, now, which are words related to more personal information sharing; 3) *Social*, a category that consists of words such as mate and they. Again, LIWC provided the research with relevant key categories related to specific conversation types.

On the other hand, the main word features that comprised the Planning label did not include any specific LIWC category. Instead, the classifier produced unigrams related to project management activities such as night, whenever, days of week (Thursday, Friday, Monday). We also saw patterns that included words related to specific management activities such as meeting and schema (which was probably because of the type of projects that were assigned, e.g., database schema). The Planning label also included words related to future events, which are often used in planning tasks. Given this particular list of words, it should be possible to create a LIWC category that would automatically identify these types of communication.

These results shows the importance of the LIWC tool and the use of unigrams to obtain the appropriate label for the messages interchanged by team members.

After placing the messages into their various content categories, we computed the correlations between *Social similarity*, *Planning Similarity*, and *Work similarity* and *Task Cohesion*. In addition, we computed correlations between *Social Rate*, *Planning rate*, and *Work rate* and the same target variable.

TABLE 4.10. Correlations of content-based variables with Task Cohesion.* $p < 0.1$, ** $p < 0.05$

Content type	Similarity	Rate
Social	0.052	0.136**
Planning	0.001	0.060
Work	0.101*	0.063

Table 4.10 shows the correlations of these variables when controlled by team size and culture. The *Work similarity* variable shows a nearly statistically significant correlation with *Task Cohesion* ($r=0.101$). However, neither *Social similarity* or *Planning similarity* are correlated with the cohesion construct.

On the other hand, there exists a statistically significant correlation ($r=0.136$) between the *Social rate* variable and *Task Cohesion*. But, again *Work rate* and *Planning rate* do not show a statistically significant correlation.

The statistically significant *Work similarity* correlation can be explained by looking at previous results that show the perception of Group cohesiveness is often affected by a group's perception that all members are doing their fair share of the work. So, if evidence shows that all members are participating in the communication, then an individual's perception of the group's cohesiveness should be higher. Moreover, the similarity between members' conversations about work (i.e., *Work similarity*) seems to be more important than the rate at which these exchanges occur, i.e.

On the other hand, *Social rate* shows a stronger correlation with *Task Cohesion* than *Work rate*. This result may demonstrate the importance of having some social communications among group members within a virtual environment. Despite the possibility that participants' social interactions may sometimes affect the accurate assessment of *Task Cohesion*, as discussed in [40], we believe that this did not occur in the context studied in this research, as evidenced by the large number of work-related communications as compared to social messages that were transmitted among group members, as shown in Table 4.11.

In contrast, there was no correlations between *Planning similarity* and *Task Cohesion*, nor between *Planning rate* and *Task Cohesion*. This lack of statistical significance between *Task*

TABLE 4.11. Average and sum of similarity and rate content measures

Content type	Similarity (average)	Rate (sum)
Social	0.216	559
Planning	0.081	203
Work	0.408	2046

TABLE 4.12. Comparison of similarity measures and metric measures models.

* $p < 0.05$, ** $p < 0.01$

	Measures
Similarity	0.298*
Rate	0.314**
Similarity+Rate	0.351*

Cohesion and the planning construct may be due to the relatively small number of messages that were generated in this category.

Once the initial analysis was complete, we then aggregated the content data and constructed three linear models(i.e., 1) *Similarity measures*, 2) *Rate measures*, and 3) *Similarity+Rate measures*), to assess their predictive capabilities. Table 4.12 shows that, in comparison, rate metrics show a higher correlation to *Task Cohesion* than similarity metrics. Moreover, *Similarity+Rate measures* together produced an even stronger correlation with Group Task cohesiveness. Based on these results, it seems reasonable to conclude that students who are in teams that have both work and social communications tend to perceive their group as being more cohesive.

We also evaluated the performance of these same metrics at a group-level i.e. we calculated *Group Social Similarity*, *Group Work Similarity*, *Group Planning Similarity*, *Group Social Rate*, *Group Work Rate*, *Group Planning Rate*. We first removed those teams that had one or more participants who were not identified with one of the major countries i.e. United States, India and Mexico. Also, we computed the *Group Task cohesion* measure by averaging the *Task Cohesion* perception of the team members. When a survey was not completed by a student, *Group Task Cohesion* was estimated using a missing data technique based on systematic non-response [34]. A total of 18 teams (out of 35) were used in the group-level analysis.

TABLE 4.13. Correlations of content-based variables at group-level with Task Cohesion.

Content type	Similarity	Rate
Social	-0.235	0.088
Planning	-0.124	0.192
Work	-0.062	-0.038

TABLE 4.14. Temporal metric statistics

Project	Instances used	Average Pacing Similarity	Average Pacing Rate
Project A	94.52%	0.56	4d 02h 52m
Project B	82.97%	0.54	4d 01h 09m
Project C	88.33%	0.53	6d 03h 38m

Results shown in Table 4.13 indicate that none of the group metrics had a statistically significant correlation with *Task Cohesion*. This result may have occurred because of the small sample size for the comparison. Thus, more data may be required to determine whether the individual-level metrics, which we found to be related to *Task cohesion* (i.e., whether (*Work similarity* and *Social rate*)), have similar effects when aggregated at the group level.

4.4. Temporal Features for Predicting Task Cohesion

In addition to the described collaboration and content-based approaches for predicting *Task Cohesion*, we developed temporal measures and analysed their effect on Group cohesiveness. As discussed in Chapter 3, we calculated *Pacing similarity*, *Pacing rate*, and *Coherence Similarity* for all individuals and groups in each project.

As indicated in Table 4.14, in the column labeled "Instances used," some of the participants were removed from this set of data. Students who posted only one message were not included in the data for this category because the pacing similarity measure can not be calculated for a single instance. Thus, students who transmitted only one communication were eliminated from the computations for the pacing variable.

Data analyzed for the remaining subjects appears to indicate that *Pacing similarity* has a similar value across all three projects; however, a comparison of the *Pacing Rate* shows that the communication rhythms within the three projects was dissimilar. More specifically, the rate

TABLE 4.15. Pacing correlation with Task Cohesion. $*p < 0.1$

Project	Correlation
Pacing Similarity	0.049
Pacing Rate	-0.129*

TABLE 4.16. Effect of coherence similarity to Task Cohesion. $**p < 0.05$

Measure	Correlation
Coherence Similarity	0.170**

of student replies was much slower in the last project. Using this data, we then calculated the correlation between *Pacing Similarity* and *Task Cohesion* and between *Pacing Rate* and *Task Cohesion* at the Individual-level.

Table 4.15 shows that *Pacing Similarity* has no effect on *Task Cohesion*. On the other hand, *Pacing Rate* has a weak, but statistically significant effect on *Task Cohesion*. This correlation, albeit weak, suggests that frequent communication tends to lead to an increase in an individual's perception of *Task Cohesion*, possibly because the more that team members discuss information about either social or work activities, the more team members trust one another and have positive feelings towards one another. The lack of a relation between *Pacing similarity* and *Task Cohesion* suggests that individuals prefer frequent, although erratic, communication (*Pacing Rate*) with team members, over a more uniform, but less frequent, rate of communication.

In addition to our pacing measures, we calculated the temporal variable of *Coherence similarity* for team members. This measure represents the average of the highest synchronies of an individual with the rest of the team. The results of this analysis, as presented in Table 4.16, suggests that *Coherence similarity* is a very good predictor of *Task Cohesion*. Therefore, this metric seems to capture the synchrony of collaborations among team members, regardless of the frequency of when those communications occur (e.g. communications every 12 hours, versus communications every 24 hours, etc).

A model using the *Pacing Rate* and *Coherence Similarity* was created to predict *Task Cohesion*. Table 4.17 shows that the temporal model representing (Pacing Rate+Coherence Similarity) statistically-significant predicts the variability of the *Task Cohesion* variable ($r=0.358$). Thus, our

TABLE 4.17. Linear temporal models to predict Task Cohesion. *** $p < 0.01$

Model	Correlation
Pacing Rate + Coherence Similarity	0.358**
Group Pacing Rate + Group Coherence Similarity	0.733**

model indicates that temporal-based measures at the individual-level are helpful for understanding the perceptions of Task cohesiveness.

As with other metrics, we tested the prediction power of these measures at the Group-level. Results in the last row of Table 4.17 shows Group-level temporal measures (Group Pacing Rate + Group Coherence Similarity) are also good predictors of *Task Cohesion*. In this particular case, the variability of the temporal measures may be reduced by aggregating them by group.

4.5. Summary of Results

Our final results showed mixed results for our three proposed models (as presented in Chapter 3). For example, the experiments that looked at the relationships between *Task Cohesion* and the collaboration variables of *Linguistic Style Matching*, *Individual density*, *Reply similarity*, *Information Exchange* found that only *Information Exchange* and *Linguistic Style Matching* were statistically significant related to *Task Cohesion*. We were somewhat more successful with experiments concerning the relationship between the two collaboration rate variables of *Information exchange rate* and *Reply rate*. Both of these variables were positively related to *Task Cohesion*. At the group-level, both *Group reply rate* and *Information exchange rate* were positively related to *Task Cohesion*.

We also had mixed results with the content-related model. The content variables of *Work similarity* and *Social rate* were the only factors found to be related to *Task Cohesion*. *Social* and *Planning similarity* were unrelated to *Task Cohesion*, nor were *Work* and *Planning rates*. None of the group-level content-based variables had a significant relationship with *Task Cohesion*.

Our final model, depicting the relationship between the temporal variables and *Task Cohesion*, showed statistically significant results between the *Pacing* and *Coherence* variables and *Task Cohesion*, although *Pacing similarity* was not statistically-significant related to *Task Cohesion*. Both *Pacing rate* and *Coherence similarity* had strong and positive correlations with *Task Cohesion*. This

was also true at the group-level of our comparisons. This positive relationships between pacing and *Task Cohesion* was even stronger when *Pacing rate* and *Coherence* were combined.

Although some of our measures do not seem to be good predictors of *Task Cohesion*, others show great promise. For example, the quantitative measures that determined different rate patterns, such as *Reply rate*, *Information rate*, *Social rate*, and *Pacing rate*, were generally better predictors than our similarity measures. Moreover, our Temporal variables seem to be very good predictors of *Task Cohesion*, even at the group level. Unfortunately, only a few of the content-related measures appear to be good predictors of *Task Cohesion*.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The major goal of this study was to determine which feature sets could predict *Task Cohesion* for individuals and teams who were engaged in global software development projects. The motivation for this work was the expectation that information about the cohesion levels among individuals and teams in global software development projects could lead to better interactions among team members, and ultimately better group performance. Although many researchers have explored relationships between certain collaboration variables [45, 4, 35], few have examined the multiple factors of collaboration, content, and temporal variables. Previous research indicated that linguistic similarity, information exchange, and message content could help determine cohesion levels within groups [20, 47, 3]. Using these initial variables as a starting point for this research, we developed several additional measures that related to a team's communication activities related to collaboration factors (*Reply Similarity*, *Reply Rate*), content factors (*Work Similarity*, *Work Rate*, *Social Similarity*, *Social Rate*, *Planning Similarity*, *Planning Rate*), and temporal factors (*Pacing Similarity*, *Pacing Rate*, *Coherence similarity*). Each of these factors was examined at both the individual and group level.

To test the adequacy of our proposed features, new and previously tested measures were used to determine the correlations between the various factors and *Task Cohesion*. The data used in the study was collected from three global software development student projects. The data included student responses to a *Task Cohesion* survey [5] and a profile questionnaire and the communication activities (I.e., postings, chats, forums, etc.) between team members in each group and project. Each of these communication activities was time-stamped and stored for the purpose of analysis. The data was then organized by individual and team and used to measure each of the proposed constructs. For example, the similarity and rate measures were computed for each person and then aggregated at the team levels. Correlations were then obtained using *IBM SPSS Statistics*.

During the development of the analysis, we found that *Culture* (represented as native country) is an important variable which determine the perception of *Task Cohesion*. Specifically, in our

experiments, we found that people from United States and Mexico have lower group cohesion's perception than participants from India. Therefore, we controlled our proposed models by *Culture* and *Team Size*.

5.1. Findings

To test the adequacy of our proposed features, new and previously tested measures were used to determine the correlations between the various factors and *Task Cohesion*. During the development of the analysis, we found that *Culture* (represented as a student's native country) is an important variable that can determine an individual's perception of *Task Cohesion*. Specifically, in our experiments, we found that people from United States and Mexico have lower perceptions of their groups cohesion's than participants from India. We also found that the size of the team could sometimes affect an individuals *Task Cohesion* scores. Therefore, we the proposed that our models should be controlled by both *Culture* and *Team Size*.

Results of the first analysis, which concentrated on the relation between the collaboration-based features and *Task Cohesion*, showed that the variables of *Individual density* and *Reply similarity* had no correlation with *Task Cohesion*. It seems that both measures are influenced by the number of participants and so require a minimum participation from most (if not all) team members. These metrics could still be useful in settings where the size of the teams are large and everyone is an active participant such as in [11]. On the other hand, we found that *Linguistic Style Matching* and *Information exchange* had a statistically significant correlation with *Task Cohesion*. Perhaps the *Linguistic Style Matching* measure was able to capture the linguistic mimicry that sometimes occurs when the people within a group begin to like one another. This particular phenomenon was discussed in [20], which found that LSM had an effect on cohesion related to the task [40]. Similarly, the *Information exchange* measure may have represented the participation intensity and similarity of communication within a team, since it was based on word counts (information). We also found that the different collaboration rate metrics, such as *Information Exchange rate* and *Reply rate*, have a statistically significant and positive correlation to Task Cohesion. We believe that both calculations describe participants' engagement in the software development project development. This relation was also found at the group level by examining the variables of *Group reply rate* and

Group information exchange. All these results suggest that metrics based on *rate* can be very useful for predicting Task Cohesion in groups that may lack access to face-to-face communication such as what may occur in a remote virtual setting.

A model based on the content within the communications (social, work and planning) was also developed. First, we generated a message classifier that performed accurately, particularly, when classifying social and work messages. Using output from the classifier, we were able to analyze their relationship with *Task Cohesion*. This model showed mixed results. Only *Work similarity* and *Social rate* were found correlated to *Task Cohesion*. The statistically significant relation between *Work similarity* and *Task Cohesion* can be explained by looking at previous results that show that the perception of cohesiveness is often affected by a group's perception that all members are doing their fair share of the work. So, if evidence shows that all members are participating in the communication, then an individual's perception of the group's cohesiveness should also be higher. On the other hand, the strong correlation between *Social rate* and *Task Cohesion* shows that it is important to have at least some social communications during a software development project, since these types of interactions can lead to an increase in trust within the group, which may result in an increase in group cohesiveness[32]. We then combined the content measures into three categories (i.e., similarity metrics, rate metrics, and similarity+rate metrics) to determine whether we could increase the strength between our measures and *Task Cohesion*. All models were found correlated to *Task Cohesion*, but the best model was *Similarity+Rate*, which suggests that models that include social and work communication categories (as defined by similarity and rate) are good predictors of team cohesiveness.

The final set of variables attempted to describe the temporal features found in global software development communications. These temporal factors were defined as *Pacing Similarity*, *Pacing Rate*, and *Coherence Similarity*. Results showed a statistically significant negative correlation between the *Pacing Rate* and *Task Cohesion*, which suggests that frequent and perhaps sporadic rhythmic communications about different social and work themes increases the cohesion among team members. On the other hand, *Pacing similarity* was not found to be related to *Task Cohesion*, which suggests that a minimum (although not equal) participation is necessary for individuals to

perceive that their group is cohesive. Thus, *Pacing similarity* is not relevant to *Task Cohesion*. We also found a positive correlation between *Coherence Similarity* and *Task Cohesion*, which suggests the importance of establishing a rhythm within a team. Finally, the temporal models constructed at individual and group-levels were found to be good predictors of *Task Cohesion*, which indicates the existence of a strong effect of frequent and rhythmic communications on cohesion related to the task.

Most of models found a correlation between their respective variables and *Task Cohesion*, in particular those models that used rates and temporal factors. These findings reflect the importance of looking at other, more fine-grained types of communications (i.e., content variables) and interaction metrics (i.e., temporal variables) as a way to characterize team interactions. Also, it is important to note that these models performed reasonable well under sparse-data conditions. Even though the teams were small (4-5 people), and the projects were short (4-6 weeks), we were able to find statistically-significant relationships between many of our factors and team cohesion. We would expect these models to improve their predictability as more communication data becomes available.

5.2. Contributions

Group cohesion is an important construct to study, given its relation to performance. Traditional methods for estimating values for the group cohesion construct have used both group and individual surveys. Most of the previous research has tested the various measures using groups located in the same setting and have not examined teams in remote settings working on medium-sized projects.

This work provided an analysis of the effectiveness of previous collaboration measures as well as the development of new metrics. Using both the new and previous measures, we found that some collaboration measures were unable to predict *Task Cohesion*, probably due to the small number of messages in the sample. However, other collaboration measures performed well, even with the small sample size. More specifically, we found that the collaboration rate metrics such as *Reply rate* and *Information exchange rate* had a positive relation to *Task Cohesion*. We believe that these collaboration measures tend to capture the engagement characteristics of the communication within a global software development team.

This work also presented the results of our content-based measures and showed that there were mixed results between *Task Cohesion* and content. The two variables that had the most statistically significant relation to *Task Cohesion* were *Social rate* and *Work similarity*. Again, a larger communication data set may produce better results. Nevertheless, the study does show the importance between social and work communication and group cohesiveness.

Our last contribution was the development and analysis of temporal metrics for analyzing group cohesion. Both *Pacing Rate* and *Coherence Similarity* were found to be statistically-significant related to *Task Cohesion*. Knowledge obtained from this study should provide insight into current empirical research on global virtual teams by defining the different temporal patterns that occur in these projects. It can also inform software development studies by suggesting how real-world teams have similar work rhythms. More detailed information about the temporal rhythms of distributed teams should contribute to the improvement of our overall understanding of how people use distributed technology to work together.

5.3. Future Work

Although the results presented in this dissertation have demonstrated that collaboration, content and temporal metrics can help in the estimate of group cohesiveness, the applied metrics can be further improved in a number of ways.

Content-based measures were derived from super-categories presented by [14]. However, this same work suggests that there are more fine-grained categories that might be helpful for identifying more specific work-related subcategories that might be more related to group cohesion.

The metric *Coherence Similarity* extracts similarity of rhythm communication patterns by using a frequency measure. In this work, we only used the maximum coherence value in any frequency. Therefore, it could be possible that other frequencies with higher values might provide more information. So, it might be possible to extract more information from the spectral analysis if we can create more or a unified metric that represents collaboration synchrony.

Finally, we believe that all of these metrics could be used in a different setting (i.e. co-located teams) and with possibly larger groups. The measures could then be used to help validate an overall model that characterizes the cohesion perceptions of an online distributed teams. The new found

model could then be implemented in a web-based product that could monitor a teams real-time activities and provide clues for assessing a teams cohesion.

BIBLIOGRAPHY

- [1] Bernard M Bass, Bruce J Avolio, Dong I Jung, and Yair Berson. Predicting unit performance by assessing transformational and transactional leadership. *Journal of applied psychology*, 88(2):207, 2003.
- [2] Kenneth A. Bollen and Rick H. Hoyle. Perceived Cohesion: A Conceptual and Empirical Examination. *Social Forces*, 69(2):pp. 479–504, 1990.
- [3] Ian Brooks and Kathleen Swigger. Using sentiment analysis to measure the effects of leaders in global software development. In *Collaboration Technologies and Systems (CTS), 2012 International Conference on*, pages 517–524, 2012.
- [4] Ian Brooks and Kathy Swigger. Leadership’s effect on overall temporal patterns of global virtual teams. In *Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on*, pages 371–379. IEEE, 2013.
- [5] Sally A. Carless and Caroline De Paola. The Measurement of Cohesion in Work Teams. *Small Group Research*, 31(1):71–88, February 2000.
- [6] Peter J Carrington, John Scott, and Stanley Wasserman. *Models and methods in social network analysis*, volume 28. Cambridge university press, 2005.
- [7] A. V. Carron and L. R. Brawley. Cohesion: Conceptual and Measurement Issues. *Small Group Research*, 31(1):89–106, February 2000.
- [8] Albert V Carron. Cohesiveness in sport groups: Interpretations and considerations. *Journal of Sport psychology*, 1982.
- [9] Albert V Carron, W’Neil Widmeyer, and Lawrence R Brawley. The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of sport psychology*, 1985.
- [10] Alberto Castro-Hernandez, Kathleen Swigger, Fatma Cemile Serce, and Victor Lopez. Classification of Group Potency Levels of Software Development Student Teams. *Polibits*, pages 55 – 62, 2015.

- [11] Alberto Castro-Hernandez, Kathleen Swigger, and Mirna P Ponce-Flores. Effects of cohesion-based feedback on the collaborations in global software development teams. In *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2014 International Conference on*, pages 74–83. IEEE, 2014.
- [12] Francois Chiocchio. Project team performance: a study of electronic task and coordination communication. *Project Management Quarterly*, 38(1):97, 2007.
- [13] Noshir S Contractor and S Grant. The emergence of shared interpretations in organizations: A self-organizing systems perspective. *Cycles and dynamic processes in communication processes*, pages 216–230, 1996.
- [14] David D Curtis and Michael J Lawson. Exploring collaborative online learning. *Journal of Asynchronous learning networks*, 5(1):21–34, 2001.
- [15] Thanasis Daradoumis, Alejandra Martínez-Monés, and Fatos Xhafa. An integrated approach for analysing and assessing the performance of virtual learning groups. In *International Conference on Collaboration and Technology*, pages 289–304. Springer, 2004.
- [16] Philip Emmert and Larry Lee Barker. *Measurement of communication behavior*. Longman Publishing Group, 1989.
- [17] Blaz Fortuna, Eduarda Mendes Rodrigues, and Natasa Milic-Frayling. Improving the classification of newsgroup messages through social network analysis. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pages 877–880. ACM, 2007.
- [18] K. Furumo and J.M. 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*, volume 1, pages 26c–26c, January 2006.
- [19] Luciano Gamberini, Francesco Martino, Anna Spagnolli, Roberto Ba, and Michela Ferron. Your Team Cohesion is Low: A Systematic Study of the Effects of Social Network Feedback on Mediated Activity. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan,

- Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, A. Ant Ozok, and Panayiotis Zaphiris, editors, *Online Communities and Social Computing*, volume 6778, pages 172–181. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [20] Amy L. Gonzales, Jeffrey T. Hancock, and James W. Pennebaker. Language Style Matching as a Predictor of Social Dynamics in Small Groups. *Communication Research*, 37(1):3–19, February 2010.
- [21] Paul Goodman, Elizabeth Ravlin, and Marshall Schminke. Understanding Groups in Organizations. *Tepper School of Business*, January 1987.
- [22] John M Gottman. Detecting cyclicity in social interaction. *Psychological Bulletin*, 86(2):338, 1979.
- [23] John Mordechai Gottman. *Time-series analysis: A comprehensive introduction for social scientists*, volume 400. Cambridge University Press Cambridge, 1981.
- [24] Andrew M. Hardin, Mark A. Fuller, and Joseph S. Valacich. Measuring Group Efficacy in Virtual Teams New Questions in an Old Debate. *Small Group Research*, 37(1):65–85, February 2006.
- [25] James Howison, Keisuke Inoue, and Kevin Crowston. Social dynamics of free and open source team communications. In *IFIP International Conference on Open Source Systems*, pages 319–330. Springer, 2006.
- [26] Matthew Ray Hoyt. *Automatic Tagging of Communication Data*. PhD thesis, University of North Texas, Denton Texas, August 2012.
- [27] David W Johnson and Roger T Johnson. Cooperation and the use of technology. *Handbook of research for educational communications and technology*, 1:1017–1044, 1996.
- [28] Manu Kapur. Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning*, 6(1):39–56, 2011.
- [29] Steve W.J. Kozlowski and Daniel R. Ilgen. Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest*, 7(3):77–124, December 2006.

- [30] Jeffery A LePine, Ronald F Piccolo, Christine L Jackson, John E Mathieu, and Jessica R Saul. A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, 61(2):273–307, 2008.
- [31] A Meier, E Voyiatzaki, G Kahrmanis, N Rummel, H Spada, and N Avouris. Teaching students how to improve their collaboration: Assessing collaboration quality and providing adaptive feedback in a cscl setting. In *Submitted as part of the symposium New Challenges in CSCL: Towards adaptive script support (Nikol Rummel and Armin Weinberger)*, ICLS, 2008.
- [32] Sean A Munson, Karina Kervin, and Lionel P Robert Jr. Monitoring Email to Indicate Project Team Performance and Mutual Attraction. 2014.
- [33] Hanni Muukkonen, Kai Hakkarainen, Kari Kosonen, Satu Jalonen, Annamari Heikkilä, Kirsti Lonka, Katariina Salmela-Aro, Juha Linnanen, and Kari Salo. Process-and context-sensitive research on academic knowledge practices: Developing cass-tools and methods. In *Proceedings of the 8th international conference on Computer supported collaborative learning*, pages 545–547. International Society of the Learning Sciences, 2007.
- [34] Daniel A. Newman and Hock-Peng Sin. How Do Missing Data Bias Estimates of Within-Group Agreement? Sensitivity of SD WG, CVWG, rWG(J), rWG(J) * , and ICC to Systematic Nonresponse. *Organizational Research Methods*, 12(1):113–147, January 2009.
- [35] Kate G Niederhoffer and James W Pennebaker. Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4):337–360, 2002.
- [36] James W Pennebaker, Martha E Francis, and Roger J Booth. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, 2001.
- [37] J.W. Pennebaker. *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury USA, 2013.
- [38] Anne Powell, Gabriele Piccoli, and Blake Ives. Virtual teams: a review of current literature and directions for future research. *ACM Sigmis Database*, 35(1):6–36, 2004.
- [39] C. Reffay and T. Chanier. How Social Network Analysis can help to Measure Cohesion in Collaborative Distance-Learning. In Barbara Wasson, Sten Ludvigsen, and Ulrich Hoppe,

- editors, *Designing for Change in Networked Learning Environments*, number 2 in Computer-Supported Collaborative Learning, pages 343–352. Springer Netherlands, January 2003.
- [40] Esa Rovio, Jari Eskola, Stephen A. Kozub, Joan L. Duda, and Taru Lintunen. Can High Group Cohesion Be Harmful? A Case Study of a Junior Ice-Hockey Team. *Small Group Research*, 40(4):421–435, August 2009.
 - [41] Wm. David Salisbury, Traci A. Carte, and Laku Chidambaram. Cohesion in Virtual Teams: Validating the Perceived Cohesion Scale in a Distributed Setting. *SIGMIS Database*, 37(2-3):147–155, September 2006.
 - [42] Victoria L. Schwanda, Kyle Barron, Jennifer Lien, Gretchen Schroeder, Ashley Vernon, and Jeffrey T. Hancock. Temporal Patterns of Cohesiveness in Virtual Groups. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW '11*, pages 709–712, New York, NY, USA, 2011. ACM.
 - [43] Kathleen Swigger, Robert Brazile, Brian Harrington, Shawn Peng, and Ferda Alpaslan. A case study of student software teams using computer-supported software. In *Proceedings of the 2005 International Symposium on Collaborative Technologies and Systems, 2005.*, pages 167–173. IEEE, 2005.
 - [44] Kathleen Swigger, Matthew Hoyt, Fatma Cemile Sere, Victor Lopez, and Ferda Nur Alpaslan. The temporal communication behaviors of global software development student teams. *Computers in Human Behavior*, 28(2):384–392, March 2012.
 - [45] Kathleen Swigger, Ferda Nur Aplaslan, Victor Lopez, Robert Brazile, Gerooge Dafoulas, and Fatma Cemile 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, SIGMIS CPR '09*, pages 187–196, New York, NY, USA, 2009. ACM.
 - [46] Yla R Tausczik and James W Pennebaker. Improving Teamwork Using Real-Time Language Feedback. 2013.
 - [47] Yla Rebecca Tausczik. Changing group dynamics through computerized language feedback. 2012.

- [48] Stanley Wasserman. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.
- [49] Sheila Simsarian Webber and Lisa M Donahue. Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of management*, 27(2):141–162, 2001.
- [50] Barbara A Wech, Kevin W Mossholder, Robert P Steel, and Nathan Bennett. Does work group cohesiveness affect individuals’ performance and organizational commitment? A cross-level examination. *Small Group Research*, 29(4):472–494, 1998.
- [51] Youngjin Yoo and Maryam Alavi. Emergent leadership in virtual teams: what do emergent leaders do? *Information and Organization*, 14(1):27–58, January 2004.