How Social Cues Shape Task Coordination and Communication

Allison Sauppé, Bilge Mutlu
Department of Computer Sciences, University of Wisconsin–Madison
1210 West Dayton Street, Madison, WI 53706 USA
{asauppe, bilge}@cs.wisc.edu

ABSTRACT
To design computer-supported collaborative work (CSCW) systems that effectively support remote collaboration, designers need a better understanding of how people collaborate face-to-face and the mechanisms that they use to coordinate their actions. While research in CSCW has studied how specific social cues might facilitate collaboration in specific tasks, such as the role of gestures in video instruction, less is known about how a range of communicative cues might facilitate activities across many collaborative settings. In this paper, we model the predictive relationships between facial, gestural, and vocal cues and collaborative outcomes in three different tasks, drawing conclusions on how each cue might contribute to these outcomes in a given task and how such relationships generalize across tasks. The resulting models provide a quantitative understanding of the relative importance of each type of social cue in predicting collaborative outcomes, as well as a more thorough understanding of how the role of each social cue changes across tasks. Additionally, our results provide confirmation and illumination of prior findings in face-to-face and computer-mediated communication research.

Author Keywords
Computer-supported collaborative systems (CSCW); computer-mediated communication (CMC); face-to-face communication; social cues; design

ACM Classification Keywords
H.5.3. **Information Interfaces and Presentation**: Group and Organization Interfaces—*collaborative computing, computer-supported cooperative work, computer-mediated communication*

INTRODUCTION
Successful collaboration requires effective communication and coordination of actions and intentions. Individuals engaging in collaboration make use of shared knowledge, such as visual or verbal common ground, to predict how their partners in collaboration will behave and to modify their own behaviors in response to that prediction [43]. This ability to predict others' behaviors during the collaboration requires paying attention to their *social cues*—implicit behaviors that express thoughts or intentions. For example, in the absence of utterances, physical positioning and gaze cues facilitate communication by directing the partners' attention to referents [19].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CSCW'14, February 15 - 19 2014, Baltimore, MD, USA
Copyright 2014 ACM 978-1-4503-2540-0/14/02 15.00.
http://dx.doi.org/10.1145/2531602.2531610

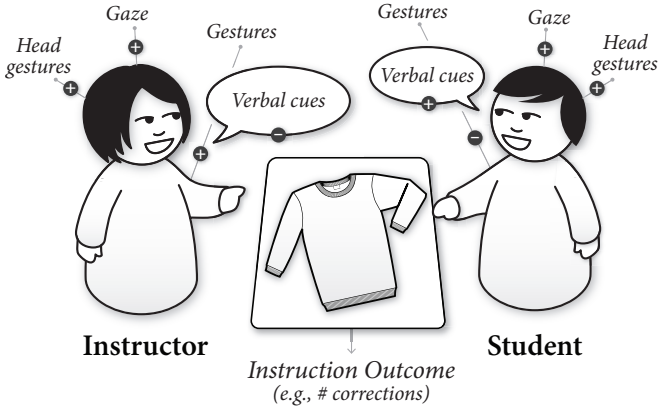

Figure 1. An illustration of the instructional scenario used in this work, highlighting the four categories of social cues we studied: head gestures, gaze, arm and hand gestures, and verbal cues.

In face-to-face collaborations, partners utilize each other's social cues to establish common ground. Individuals who are unable to meet face-to-face, on the other hand, rely on communication and collaboration technologies establish such common ground and collaborate at a distance. While these technologies enable participants to collaborate when the opportunity would not otherwise exist, they lack the level of social presence afforded in face-to-face interactions due to the limited number of *channels*—mediums that convey a subset of cues—that they convey (e.g., a telephone only mediates communication via the audio channel) [49, 51].

Previous research in CSCW has explored the role of specific social cues on collaboration in specific tasks as supported by specific communication channels, such as how gestures might support collaboration over the video channel in physical tasks [18]. However, a more comprehensive understanding of the role of social cues in collaboration, particularly how the broader range of social cues shape objective and subjective outcomes in collaboration across different roles and tasks, might inform the design of collaboration tools that provide support for a wider range of social cues and that adapt to the changing priorities of communication across different tasks.

To provide recommendations for designers of CSCW systems, we chose to study dyadic collaboration in its natural context, focusing on face-to-face dyadic communication. This approach allowed us to study the range of cues participants produced and was thereby not restricted by the channels available in a given CSCW system. While CSCW research has long been informed by studies of face-to-face communication, our work presents a more comprehensive perspective by examining how a wide range of social cues presented by all collaborators shapes interactions and how these interactions vary across contexts.

Based on the face-to-face interactions studied, we present a set of models that provides a holistic understanding of how numerous social cues impact communication in a variety of contexts. More specifically, these models describe the relationship between social cues, such as gaze, gestures, and verbal cues, and collaborative outcomes, such as task performance, perceived task success, and social outcomes, across three common collaborative activities: cooperation, instruction, and negotiation (Figure 1). These models are derived from a regression-based analysis of video data collected from eight human-human dyads in a laboratory setting. From these models, we draw the following contributions to CSCW:

- A quantitative understanding of the relative importance of each type of social cue in predicting collaboration outcomes.
- A better understanding of the variability in the role that each type of social cue plays in predicting these outcomes across different scenarios of collaboration.
- Confirmation and illustration of prior findings in face-to-face and mediated communication.

The remainder of this paper provides a review of related work, describes the methods of data collection, analysis, and results of the analysis, discusses the theoretical and design implications of these findings, and acknowledges some of the limitations of this work.

BACKGROUND

While remote collaboration draws on many of the same communicative cues and mechanisms as face-to-face communication [26], technologies that mediate such collaboration do not always support the channels necessary to perceive them, as audio and video communication restrict social cues that are commonly used in face-to-face interactions, such as gaze and nodding behaviors and shared view of gestures, respectively. These behaviors are crucial to the effectiveness of communication and collaboration; for example, gaze plays a particularly important role as it adds information to the dialogue [42] and indicates where attention is directed [19].

The mismatch between the range of behaviors people use to communicate and coordinate and the channels that CSCW systems support has been acknowledged by various theories of computer-mediated communication (CMC) [49, 51]. *Bandwidth theory*, initially implemented in more traditional media of the 1970s, discussed the effectiveness of a system in terms of the number of cues a system could convey (i.e., the system's *bandwidth*) and how it helped establish the perception of co-location between participants [44]. Previous work has shown that when the bandwidth of the system was increased, for example, by adding video to an audio-only system, participants' feelings of co-location would also increase [44]. An extension of bandwidth theory for collaboration systems is the *cues-filtered-out model*, which proposed a one-to-one attribution of cues to communicative functions [8]. According to this model, any decrease in the number of cues present in a system would directly decrease communicative function, since less information was communicated and social presence was lowered [8]. A third theory of CMC, *media richness theory*, aimed to explain the degree of communicativeness of the system as a function of the number, i.e., the richness, of cues communicated relative to the system task [9]. Based on this theory, systems that communicated a larger number of cues were better suited to more complicated tasks, such as explaining how to perform a surgical procedure, while systems that communicated a smaller number of cues were better suited to simpler tasks, such as setting up an appointment.

One challenge for CSCW system designers is to prioritize the cues most necessary for a given task, similar to the perspective presented by media richness theory that a more complex task will require that more cues are communicated [29]. While prior work has highlighted the superfluousness of video in many contexts [39], including simple tasks that require no spatial information [34], elimination of the visual channel in complex tasks or tasks that require spatial information can lead to degradation in task outcomes in remote collaboration. For example, instructional compliance during remote assistance for surgical procedures slows when the video screen is not visible to collaborators [32]. Many channels that enable the communication of social cues in face-to-face communication have been explored in the context of CSCW systems in an effort to increase the efficiency of future designs. The paragraphs below review previous work on how visual and audio channels support the communication of social cues.

Visual Channel

Gaze, one of the most important social cues in face-to-face interactions, has similar repercussions in remote collaboration that are enabled by CSCW systems. One of the most important functions of gaze is grounding, which improves communication, particularly in the execution of complex or spatial tasks [19]. Synchronized gaze can have additional positive effects on task performance and can also reduce overlap in conversational turn-taking [48]. Communicators also use head movements to direct attention and pass the floor in conversations [33]. Head movements also include nodding behavior, which communicators can use to signal that they are engaged in the conversation [27]. Nods can also take the place of verbal feedback, similarly indicating agreement with a speaker [46].

Although the region of the head is essential in conveying many of the important social cues, hands and arms also produce gestures that augment communication. Gestures can be used to replace speech turns [17] or achieve quicker grounding, reducing the number of speech overlaps [30]. Prior work shows that providing visual access to the gestures of remote collaborators in spatial tasks improves task outcomes as compared to

audio-only communication by enabling participants to more precisely communicate spatial goals and thus make better decisions [14]. Past research also includes a number of systems that have been developed to augment visual channels to better capture and communicate the various gestures that occur in collaboration [7, 35].

Audio Channel

The use of the audio channel in the design of CSCW systems enables collaborators to use language in order to compensate for social cues that cannot be perceived due to design or technical limitations. Prior work that compared the use of audio-only and audio and video communication in spatial tasks has shown that the absence of a video feed cause participants to query their partner more for status updates [34], while the addition of a visual channel lead to the modification of dialogue structure to incorporate visual feedback [17]. Even in tasks that do not require spatial reasoning, the presence of video results in more fluent negotiations, enabling participants to revisit information more due to the reduced overhead of conversational turn-taking. The improvements in dialogue flow result in more mutually beneficial negotiations [10].

While these findings highlight the importance of social cues in specific collaborative activities and offer some guidelines for design, less is known about how the broader range of social cues contribute to these activities and how these relationships might generalize. Previous work provides only a limited understanding of this design space due to their focus on the role of social cues in the context of specific collaborative activities that are supported by specific technologies, limiting our ability to make comparisons across the findings of different studies. Our work seeks to investigate the knowledge gap in understanding the relationships between social cues and the outcomes of collaboration activities by building predictive models across a wider range of social behaviors and collaborative activities. A better understanding of these relationships will inform future theoretical explorations and designs of collaboration technologies.

METHOD

In this work, we collected data from dyads engaged in face-to-face collaboration and built regression-based models of the relationships between their social cues and the outcomes of their collaboration to inform future research and design in CSCW. To better understand how these relationships changed across different types of collaborative tasks, we studied three different tasks. For each task, we considered an objective measure of the effectiveness of the collaboration and a measure of the perceived success of the collaboration by the participants. At the conclusion of the study, we also collected a set of measures for social outcomes, such as the participants' perceptions of how much rapport was formed between themselves and their partner. The model for each outcome offers insight into which cues are significant predictors of the outcome and therefore should be supported by a CSCW system designed to improve task outcomes.

This section discusses the selection of the collaborative activities that are considered in this work, provides information on the specific instantiations of these activities for data collection, and describes the data corpus, measures, and the statistical modeling approach.

Approach

Our choice of studying face-to-face communication instead of remote interactions mediated by a CSCW system was motivated by our desire to explore the dynamics of the wider range of social cues across multiple collaborative contexts. The design of CSCW systems is informed either by studies of face-to-face communication or studies of communication using a specific collaboration technology. However, both types of studies are typically situated in a specific collaborative context, such as instruction, cooperation, or negotiation, and focused on specific communicative cues, such as gaze or gestures, making it difficult to identify how a wider range of social cues jointly affect collaboration outcomes and compare findings across contexts. Additionally, studies that use a specific collaboration technology to provide further design recommendations necessarily limits the number of social cues available to the parties in the interaction. Prior work has shown that users adapt their use of communicative cues to the channels supported by the technology used [20, 42]. Thus, it is difficult to discern what social cues the users would have naturally employed, what adaptations they would have made when these cues are not supported by the specific system used, and how future designs might support or deter such adaptations.

In contrast, studying face-to-face interactions provides the opportunity to gain a holistic understanding of how social cues shape collaboration without restricting the cues available for study. Furthermore, studying interactions in multiple contexts allows us to understand how these relationships vary across different forms of collaboration. From these studies, we can draw conclusions about which social cues are most important, the relative contributions of cues to collaboration outcomes, the tradeoffs between objective and subjective outcomes, and how participant behavior might change across tasks. These conclusions can inform the design of future CSCW systems as to which channels should be prioritized and enable predictions about how the elimination of particular channels might affect task and interaction outcomes.

Collaborative Tasks

Collaborative work encompasses a wide range of tasks. For instance, employees of a company might receive training and instruction on a regular basis, members of work teams might frequently cooperate on creative or physical tasks, and company representatives might engage in negotiations with clients or partners. As these tasks become increasingly distributed, it is increasingly important that the technologies facilitating such tasks are as effective as possible. In this study, we consider three common collaborative tasks: *instruction*, *cooperation*, and *negotiation* [3, 52]. The paragraphs below provide a brief review of prior work in CSCW that has explored collaboration in these tasks.

Instruction

Previous work has explored remote instruction in training and instruction contexts and has studied how specific cues and

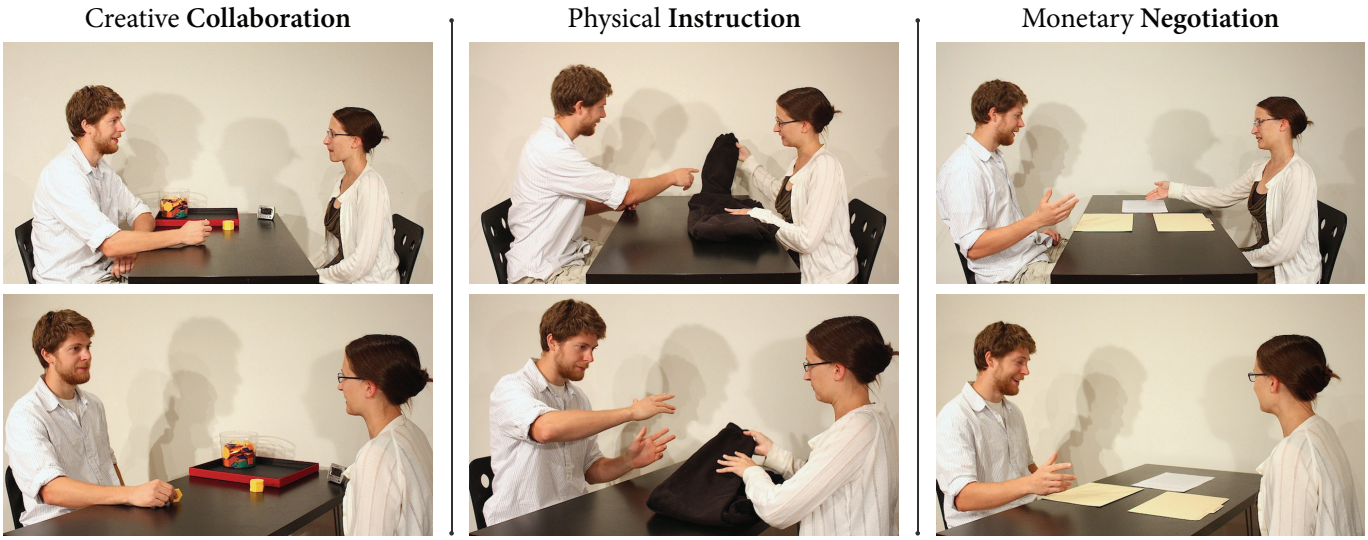

Figure 2. The three collaborative tasks used in our study: brainstorming a packing list in the creative *cooperation* task; learning to fold a pair of pants in the physical *instruction* task; and negotiating prices for two travel packages in the monetary *negotiation* task.

affordances affect instruction [16, 18, 21]. For instance, studies of video communication for surgical training in a hospital setting highlighted the importance of and increased need for a shared visual space among participants [32].

Cooperation

Cooperation has been explored in formal work contexts such as managing the London Underground [23] and air traffic control [5], as well as in informal cooperations such as brainstorming [36] and play dates [28]. These CSCW systems tend to display the person via video, the context via video, or both. Previous work has developed guidelines to help future development of cooperative systems [50]. This work emphasizes the need to carefully select a medium for communication based on the goals of the system. Developing a stronger understanding of the importance of significant behavioral cues in cooperative systems will better inform future developers as they choose a medium on which to focus for a particular application.

Negotiation

Negotiation has been well-studied in CSCW research, ranging from implicit negotiations that arise in face-to-face interactions [1] to explicit negotiation tasks [4]. Preliminary research into interaction patterns undertaken during negotiations that use collaboration systems shows that these patterns change when using text, audio, or video to perform the negotiation. For example, participants negotiating over video trade smaller pieces of information during a turn than those negotiating over text or audio [10]. A preliminary model of negotiation addresses the high-level structure of how computer-mediated negotiation differs from face-to-face negotiation [37].

Participants and Data Corpus

A total of 48 participants from a university campus participated in this study. The participants studied a diverse range of fields and were aged between 19 and 60 (M = 23.02, SD = 6.70). All participants were native English speakers. Participants were assigned into dyads of two participants who were unfamiliar with one another to jointly perform the experimental tasks. We assigned dyads, conversational roles, and task orderings so that each was fully stratified by gender, creating an equal number of female-female, male-male, and female-male dyads.

The experimental design stratified participants by gender to reflect earlier findings on gender effects in collaborative work and other social situations. For example, gender has been shown to affect participant behaviors in face-to-face negotiation [47], self-esteem [31], and approach to leadership [12].

In our experimental setup, the members of the dyads sat across from each other at a table, as illustrated in Figure 2. The data collection equipment included three high-definition cameras at 1080p resolution and 30p frame rate. Two of the cameras were positioned across from each participant, capturing their upper torsos as visible from across the table. The third camera captured the speech and nonverbal behaviors of both participants from the side using a wide-angle lens. The final data corpus consisted of 72 recorded tasks, with three tasks for each dyad and three angles for each task. These 72 tasks amounted to 4 hours and 17 minutes of audio and video data, and across tasks participants interacted for an average of 14 minutes and 11 seconds. The average video length for the instruction, cooperation, and negotiation tasks were 3 minutes and 4 seconds, 5 minutes, and 6 minutes and 7 seconds, respectively.

Procedure

Following informed consent and an overview of experimental procedures, both participants were seated across from one another at a table. Since the participants were strangers, an acclimation task was performed before starting the experimental tasks. During the acclimation task, participants were given three minutes to discuss their school or work life with one another. Participants then performed in each of the three experimental tasks—instruction, cooperation, and negotiation—in a stratified order. Following each experimental task, participants were asked to complete a questionnaire regarding their perceptions of themselves, the other participant, and their joint collaboration in completing the task. Upon completion of

all three tasks, participants completed a longer questionnaire that evaluated group characteristics and dynamics, as well as demographic information. The experimenter then debriefed the participants. All three tasks and questionnaires, as well as the post-experiment questionnaire, took between 45 minutes to an hour to complete. This variation arose from individual differences in completion times for tasks, particularly in the negotiation task. Participants were paid $10 for their time.

Experimental Task

We created the three collaborative activities considered in this study as part of a higher-level collaboration concerned with planning a trip abroad. The following sections describe the design of each of these experimental tasks.

Instruction

In the *instruction* task, one participant taught their partner how to fold a pair of pants, reflecting the theme of preparing for a trip abroad. One participant took on the role of the *instructor*, while the second participant took on the role of the *student*. To simulate the knowledge disparity that is often found in instructional tasks, the instructor learned how to fold the pants using a non-traditional set of steps prior to teaching it. To reflect the context disparity found in remote collaboration, the instructor was not allowed to touch the pants while teaching the student.

To introduce the folding technique to the instructor, the student was asked to leave the room while the instructor watched a video that explained the pant-folding procedure. The instructor was allowed to fold the pants along with the video and was given as much time as necessary to review the video. When the instructor felt that they had memorized the necessary steps to folding the pants, the student was brought back into the room. The instructor then taught the student how to recreate the steps necessary to reproduce the folds shown in the video.

Cooperation

In the *cooperation* task, participants worked together to identify items to pack for their trip abroad through brainstorming, following the structure of a common "creativity task" to build a cooperative atmosphere by creating a temporary problem-solving group [22]. The development of this task built on the definition of cooperation as "a style for direct interaction between at least two co-equal parties voluntarily engaged in shared decision-making as they work toward a common goal" [6]. During the cooperation task, participants were on equal ground as they worked to complete their packing list and built off of one another's ideas and experiences.

For this task, participants were told that the items must be able to fit in luggage that they can take on a plane. They were instructed not to count quantities (e.g., "three pairs of pants" would be reduced to "pants") to encourage them to rely on their creativity when identifying items. Participants were asked to think of as many items as possible in five minutes. One participant was given a bucket of counter chips. Each time the participants agreed upon a new item to pack, the assigned participant was instructed to move one of the counter chips into a second bucket, indicating that they wished to add this item to their packing list.

Negotiation

Participants in the *negotiation* task were given opposing roles in a business transaction. We strove to create a task that encouraged participants to follow a traditional argumentation-based negotiation model in which participants' goals conflict and cooperation cannot be assumed, forcing participants to use a series of arguments to persuade their partner to pursue some action [45]. Additionally, similar to many negotiation tasks, participants were not completely aware of their partner's goal. We chose money as the basis for our negotiation task so that we could quantify an objective measure of task success often not available in negotiation tasks. Participants were offered the chance at a small prize (a candy bar) in addition to their monetary compensation to help offset a lack of motivation to negotiate in the laboratory.

Participants were assigned to a role as either a *travel agent* or a *traveler*. The experimenter explained that each participant would be working within their respective role to negotiate prices for two different travel packages. Each participant was given a folder containing their price limits for each travel package. For the travel agent, this price was the *lowest* price the travel agent was willing to sell the travel package for. For the traveler, this price was the *highest* price the traveler was willing to pay for the travel package. The prices were chosen such that the travel agent's limit was substantially lower than the traveler's limit. The goal for the travel agent was to maximize how much they could convince the traveler to spend on the package over the travel agent's limit, while the goal for the traveler was to maximize how much they could save on the package by convincing the travel agent to sell the package for less than the traveler's limit.

Participants were given their instructions individually by the experimenter while their partner was outside of the room, and they were told not to share the information in their folder with the other participant. To further motivate an authentic negotiation, each participant was told that if they were able to minimize the amount of money they lost beneath a certain threshold, at the end of the experiment they would be given a small prize in addition to their monetary compensation. The "winning" participant received a candy bar at the end of the experiment. Participants would barter on one package at a time and were required to come to an agreement on both packages. When the participants agreed on a price, the travel agent would write the price on a sticky note and affix it to the back of that travel package.

Measurements

Our dependent variables reflected both objective and subjective outcomes. For each task, one objective measure described the effectiveness of the two participants at achieving the task goal. Participant responses to questionnaires on seven-point rating scales measured perceived task effectiveness for the participant, their partner, and their ability to work as a team (cooperation: 19 items, Cronbach's α = .837; instruction: 9 items, Cronbach's α = .793; negotiation: 16 items, Cronbach's α = .855). Additionally, participants answered a larger set of questions at the conclusion of all three tasks. These questions measured several group outcomes including rapport (24

items, Cronbach's α = .920), teamwork (7 items, Cronbach's α = .811), collaborativeness (4 items, Cronbach's α = .701), empathic concern (7 items, Cronbach's α = 764), perspective-taking (7 items, Cronbach's α = .820), interpersonal solidarity (14 items, Cronbach's α = .774), and homophily (9 items, Cronbach's α = .812).

Our objective measures varied based on task. In the instruction task, we measured the number of times the instructor offered a correction to the student ("No, unfold that part. You need to angle it more to the right when you fold it."). The objective measure for the cooperation task was the total number of unique items that were generated by participants in brainstorming. In the negotiation task, we measured the total discrepancy between the actual selling price and the median of the high point (how much the traveler was allowed to spend on a package) and the low point (the lowest amount for which the travel agent was allowed to sell a package).

In order to compare the relative importance of behaviors across tasks, we transformed our data so that a positive trend in the outcome measure indicated a more effective performance of the task. For example, in the cooperation task, the higher the objective measure (i.e., number of items brainstormed) was, the better the participants did at completing the task. The opposite is true for the instruction and negotiation tasks, however. The lower the number of corrections in the negotiation task and the lower the price discrepancy in the instruction task, the better participants did at completing the task, according to the transformed data. As we did with the measures for the cooperative task, we transformed our objective measures for the instruction and negotiation tasks. The negotiation measure was transformed so that the outcome that was most beneficial to both participants had the highest number, and the outcome which had the least beneficial outcome was given the lowest transformed number. Likewise, the measure for the instruction task was transformed so that the highest number reflected the fewest number of corrections, while a lower number reflected more corrections. Under the transformations of objective measures across tasks, the number associated with the most beneficial outcome became the highest number, and the number associated with the least beneficial outcome became the lowest number. For example, in a data set where $p_0 = 5$, $p_1 = 7$, and $p_2 = 11$, the transformation would result in $p_0 = 11$, $p_1 = 9$, and $p_2 = 5$.

Based on findings from prior work and a preliminary analysis of our data set, we identified 10 types of behavioral cues to serve as dependent variables requiring further analysis: gazing at one's partner, gazing at the shared context, iconic gestures, metaphoric gestures, deictic gestures, beat gestures, head nods, smiling, verbal filler, and laughter. All variables were coded in a binary manner—0 for *not occurring* and 1 for *occurring* events—every 33.3 milliseconds. The data was coded by two coders such that there was a 10% overlap with which to evaluate inter-coder reliability. One coder was responsible for creating a detailed description of each code, such as examples of the code, counter-examples of the code, and minimum requirements for the code (e.g., a participant must gaze away for at least 100 ms to count as a shift in gaze). This coder then trained the second coder with the given description, providing examples. The inter-rater reliability showed substantial agreement for all attributes; participant gazing at their partner (84% agreement, Cohen's κ = .69), participant gazing at the shared context (82% agreement, Cohen's κ = .65), head nods (89% agreement, Cohen's κ = .74), smiles (83% agreement, Cohen's κ = .67), filler speech (76% agreement, Cohen's κ = .65), laughter (75% agreement, Cohen's κ = .63), iconic gestures (82% agreement, Cohen's κ = .66), metaphoric gestures (78% agreement, Cohen's κ = .64), deictic gestures (78% agreement, Cohen's κ = .63), and beat gestures (76% agreement, Cohen's κ = .62).

Our resulting data set consisted of 15 data tables. Each row in a data table took the following form:

$$< \boldsymbol{y x_1 x_2 \ldots x_n} >$$

where $\boldsymbol{y}$ represents the *response* variable, and $\boldsymbol{x_i}$ represents a single *predictor* variable—one (of possibly many) variables that can be used to predict $\boldsymbol{y}$. We discuss the data tables for each dependent variable below.

Objective Measures

For each objective measure, the analysis considered the predictors x_i as the aggregate occurrences of each behavior for each *dyad* during that task. We chose to initially consider the individual roles of a dyad separately, compiling one data table for each participant, resulting in two data tables for each task, each data table with 24 tuples of 10 values each. After constructing models for each role, we then considered the significant predictors of each role in a single model to measure the relative importance of *both* participants' cues on the overall outcome of the task. This combined model considered 24 tuples for each task, with the number of variables varying with the task. In this model, it is possible for each value to appear twice, once for each participant in the dyad. Our final data corpus contained three data tables for objective measures, with one data table for each task.

Task Perception Measures

For the task perception measures, the analysis considered the predictors as the aggregate occurrences of each behavior for each *participant* during that task. We chose to consider the participants of a dyad individually since each participant had their own perception of the task outcome. This resulted in tuples which contained 10 values; one for each of the 10 behaviors of a single participant.

The dependent variable for the task perception measures was dependent on the task under analysis. For the cooperative task, participants had equal roles since both had an equal responsibility to brainstorm ideas. Thus, we chose to consider all participants in a single data table, resulting in 48 tuples for the cooperative task.

However, both the instruction and negotiation tasks split participants into different roles: instructor and student for the instruction task, and travel agent and traveler for the negotiation task. Because participants in one role may have had different perceptions of task outcome than participants in the other role, we chose to consider roles in both the instruction

and negotiation task separately. This resulted in four data tables (two for instruction and two for negotiation), each of which had 24 tuples.

Subjective Measures

For the subjective measures, the analysis considered the predictors as the aggregate occurrences of each behavior for each *participant* across all three tasks. We chose to consider the participants of the dyad individually, as each participant had their own perception of task success. This consideration resulted in tuples that contained 10 values, one for each of the 10 behaviors of a single participant. The outcome variable for the subjective measures was the average of the scale items for that particular participant.

The tuples for all three of the outcomes discussed above were normalized across each tuple by dividing the values of the tuple by the length of the video for that task or, in the case of the final subjective measures, the aggregate time of all three tasks. Finally, each variable (i.e., each column) was normalized to vary between 0 (least frequent) to 1 (most frequent).

Analysis

Our analysis took a two-step approach. First, we analyzed our data using *stepwise multiple linear regression* to model the relationships among the independent and dependent variables. We then discovered significant themes of findings by combining an examination of these models and observations of the video data. This approach has been utilized in past HCI research to better understand how a range of behaviors shape task outcomes [2, 24].

Each analysis in Stepwise Multiple Linear Regression followed the formulation below:

$$\boldsymbol{y} = (\beta_0 + \beta_1 \boldsymbol{x_1} + \beta_2 \boldsymbol{x_2} + \ldots + \beta_n \boldsymbol{x_n}) + e$$

Where β_0 is a constant that represents the y-intercept, whereas $\beta_1 \ldots \beta_n$ are coefficients for each of the n predictors. The values of each predictor for a given situation are represented by $x_1 \ldots x_n$. The error for the model is represented by e.

In its simplest form, Linear Regression models the relationship between one predictor variable x and one outcome variable y with some error e. Given a set of x, y data points, a line is fitted to the data such that the sum of squared residuals is minimized. The resulting model provides a β value (measured in standard deviations) for x, which describes the change in y for a one unit change in x. A positive β indicates that the corresponding predictor has a similar trend with the outcome variable (e.g., an increase in the occurrence of the predictor x increases the outcome variable y), whereas a negative β indicates that the corresponding predictor has an inverse relationship with the outcome variable. The larger $|\beta|$ is, the more important x is in predicting y.

To expand regression modeling to more than one predictor variable requires Multiple Linear Regression. While there are numerous approaches to constructing Multiple Linear Regression models, we discuss our particular approach below, called Forward Stepwise Multiple Linear Regression. Here, the final model is constructed by gradually including predictors in the model according to their contribution to the model [13]. For the purposes of our study, at each step we included any predictor with a p-value below .250. The final model is then comprised of predictors which are statistically significant ($p < .100$). Because the β coefficients are measure in standard deviations, each β_k provides the relative contribution of each independent variable x_k in predicting the dependent variable y.

With the final models, we returned to the original video data collected from the study to offer explanations for the results in our models. Videos were reviewed in conjunction with the models to discover themes in the data. These themes focused on a number of different results, ranging from the importance of one cue across tasks, to the importance of multiple cues within a single task to produce a particular effect.

RESULTS

Our analysis resulted in a total of 13 models that characterized the relationship between observed cues and task outcomes. Figure 3 details the models and statistical tests of the significance of these relationships. Drawing on the quantitative findings from these models and qualitative observations from the video data, we constructed four themes under which these findings fell. Below, we provide examples from the videos and results from our models and, where appropriate, discuss the findings in the context of prior work for each theme.

Theme 1: Rapport in Negotiation

Prior work suggests that rapport can be an important component in mutually beneficial negotiation outcomes [11, 38]. Our data from the negotiation task showed that dyads that successfully negotiated a fair price also displayed behaviors that are associated with rapport. For example, in the exchange below, the traveler highlights his or her monetary situation as a student in an attempt to establish more rapport and increase the travel agent's empathy for him or her. When the travel agent rejects that attempt, the traveler laughs, acknowledging that the negotiation strategy would likely prove futile. Figure 4, Left illustrates both the traveler and the travel agent laughing during negotiation.

> **Traveler:** You know, as a student, I don't have much money. Can you give me a lower price?
> **Travel Agent:** Unfortunately, I really can't go any lower than that, with cutbacks and everything...
> **Traveler:** [laughter from traveler] OK, OK. It was worth trying.

Exhibited behaviors that are associated with an increase in rapport included a greater use of gazing toward the partner for both the traveler, $\beta = .116$, $t(23) = 2.66$, $p = .057$, and the travel agent, $\beta = .473$, $t(23) = 7.83$, $p = .016$, and—in the case of the traveler—laughing, $\beta = .323$, $t(23) = 5.73$, $p = .039$, as shown in Model 3 in Figure 3. The use of cues associated with rapport, particularly gazing toward the partner, was also a significant predictor of empathy between the participants, as shown in Model 8, $\beta = .761$, $t(47) = 4.10$, $p = .036$.

These behaviors help the participants in the negotiation to develop a relationship with each other, which results in more

Outcome (y)	Function ($\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n + \varepsilon$)	R_A^2	Significance
Objective Collaborative Task Outcomes			
1 Cooperation *# creative solutions*	$-.472 + .567\boldsymbol{E_p}^l - .322\boldsymbol{E_t}^l + .842\boldsymbol{G_d}^l - .307\boldsymbol{G_b}^l + .578\boldsymbol{H_n}^l + .575\boldsymbol{S_f}^l + .552\boldsymbol{G_b}^r + .441\boldsymbol{H_n}^r - .231\boldsymbol{S_f}^r + .758\boldsymbol{S_l}^r + .036$	.737	E_p^l $p < .001^{***}$ S_l^l $p = .001^{**}$ E_l^p $p = .005^{**}$ G_b^r $p = .012^{*}$ G_d^l $p = .017^{*}$ H_n^r $p = .013^{*}$ G_d^l $p < .001^{***}$ S_f^r $p = .011^{*}$ H_n^l $p = .004^{**}$ S_l^r $p = .091^{\dagger}$
2 Instruction *# corrections*	$.019 + .268\boldsymbol{G_d}^s + .210\boldsymbol{H_n}^s - .427\boldsymbol{S_l}^s + .201\boldsymbol{E_p}^i + .079\boldsymbol{G_i}^i + .314\boldsymbol{G_d}^i + .326\boldsymbol{H_n}^i + .048$	.826	G_d^s $p = .061^{\dagger}$ G_i^i $p = .047^{*}$ H_n^s $p = .015^{*}$ G_d^i $p = .014^{*}$ S_l^s $p = .019^{*}$ H_n^i $p = .023^{*}$ E_p^i $p = .017^{*}$
3 Negotiation *Monetary agreement*	$.109 + .473\boldsymbol{E_p}^a - .465\boldsymbol{G_d}^a + .323\boldsymbol{S_l}^a + .116\boldsymbol{E_p}^t - .247\boldsymbol{G_d}^t + .493\boldsymbol{H_s}^t + .512$	.364	E_p^a $p = .057^{\dagger}$ E_p^t $p = .016^{*}$ G_d^a $p = .026^{*}$ G_d^t $p = .045^{*}$ H_l^a $p = .005^{**}$ H_s^t $p = .039^{*}$
Perceived Collaborative Task Outcomes			
4 Cooperation *Perceived success*	$.700 - .611\boldsymbol{G_i}^l + .708$	.315	G_i $p = .016^{*}$
5 Instruction *Perceived success*	Instructor — *No significant model* Student — *No significant model*	*N/A* *N/A*	*N/A* *N/A*
6 Negotiation *Perceived success*	Travel Agent — *No significant model* $.567 - .392\boldsymbol{G_i}^t - .532\boldsymbol{H_n}^t - .647\boldsymbol{S_l}^t + .901$	*N/A* .431	*N/A* G_i^t $p = .038^{*}$ S_l^t $p = .008^{**}$ H_n^t $p = .041^{*}$
Social Outcomes			
7 Collaborativeness	*No significant model*	*N/A*	*N/A*
8 Empathy	$.638 + .761\boldsymbol{E_p} - .683\boldsymbol{E_t} + .844$	.064	E_p $p = .036^{*}$ E_t $p = .030^{*}$
9 Homophily	$.411 - .402\boldsymbol{G_d} + .740$	.129	G_d $p = .028^{*}$
10 Interpersonal Solidarity	*No significant model*	*N/A*	*N/A*
11 Perspective-taking	*No significant model*	*N/A*	*N/A*
12 Rapport	$.601 - .341\boldsymbol{E_p} - .407\boldsymbol{H_n} + .281\boldsymbol{S_f} + .634$	.322	E_p $p = .025^{*}$ S_f $p = .048^{*}$ H_n $p = .010^{*}$
13 Teamwork	$.679 - .266\boldsymbol{E_t} + .318\boldsymbol{G_i} - .446\boldsymbol{H_n} + .604$	.332	E_t $p = .036^{*}$ H_n $p = .003^{**}$ G_i $p = .018^{*}$

Figure 3. Models for the objective collaborative outcomes of each task, the perceived collaborative outcomes of each task, and the social outcomes overall. The variables represented are gazing toward the participant (E_p), gazing toward the shared context (E_t), iconic gesture (G_i), metaphoric gesture (G_m), beat gesture (G_b), deictic gesture (G_d), head nodding (H_n), smiling (H_s), verbal filler (S_f), and laughter (S_l). Superscripts l and r, i and s, and a and t denote the left and right participants in the collaboration scenario, instructor and student in the instruction scenario, and travel agent and traveler in the negotiation scenario, respectively.

empathy between the participants toward each other's position and working towards a more mutually beneficial agreement.

Theme 2: Establishing Procedural Grounding Using Gaze

Previous research has highlighted the importance of gaze in monitoring attention and understanding toward achieving procedural grounding in collaboration [15, 41]. The excerpt below illustrates how collaborators might use gaze to assess understanding, where the instructor monitors the student after issuing an instruction:

> **Instructor:** Put it to your left at a 90 degree angle. [watches student's face]
> **Student:** [performs the instruction correctly]

Our models showed visual monitoring of a partner to predict task success across tasks. The roles that demanded gazing toward one's partner included the instructor, $\beta = .201$, $t(23) = 7.25$, $p = .017$, in the instruction task (see Model 2), the traveler, $\beta = .116$, $t(23) = 2.66$, $p = .057$, and the travel agent, $\beta = .473$, $t(23) = 7.83$, $p = .016$ in the negotiation task (see Model 3), and the left participant in the cooperation task (see Model 1), $\beta = .567$, $t(23) = 11.82$, $p < .001$. Findings from the instruction task are contrary to prior work, which found little importance of monitoring in instruction [16]. On the other hand, gaze was shown to be more important for both participants in the negotiation task, which is consistent with prior work that has shown the addition of video to provide important benefits in negotiation and conflict resolution [10].

In the instruction task, monitoring seemed to play a particularly important role; instructors monitoring the actions and behaviors of their students when giving instructions and students offering feedback on their understanding of the instruction were both strong predictors of success in task outcomes as shown in Model 2. Predictors of task success included students who offered nonverbal feedback such as head nods $\beta = .210$, $t(23) = 8.12$, $p = .015$. While laughter can serve as a form of feedback, our models showed that laughter by students had a significant negative effect on the overall outcome of the task, $\beta = -.427$, $t(23) = -6.90$, $p = .019$, which might indicate students that are not taking the task seriously or not attending to the instructor's commands.

In tasks such as instruction, the introduction of a shared visual context requires participants to split their attention between their partner and the shared context. Visual monitoring enables participants to maintain the necessary level of awareness of

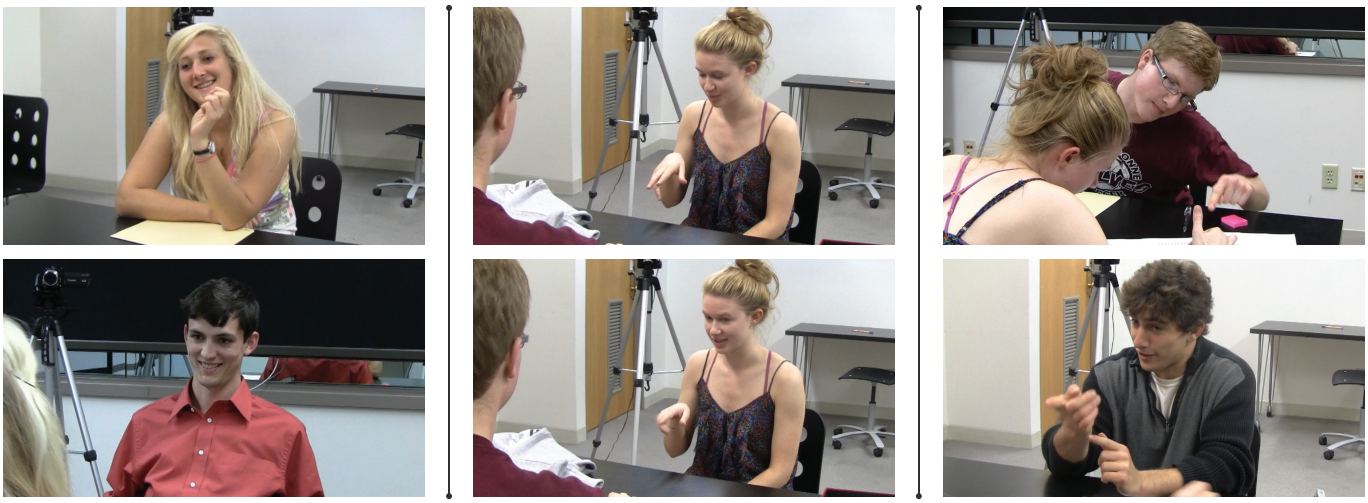

Figure 4. *Left*: An example of the agent rejecting a compromise proposed by the traveler. The traveler then laughs about the attempt, building rapport with the agent despite the failed negotiation. *Middle*: An example use of gaze for establishing procedural common ground. The instructor first uses her gaze combined with a deictic gesture aimed at the workspace to indicate the next step and then gazes at the student to monitor the student's face for indications of understanding. *Right*: Examples of the differential use of deictic gestures in the cooperative and negotiation tasks. Pointing was found to be a positive predictor of task success during cooperation, but to negatively predict task success in negotiation.

the shared context and their partner. Figure 4, Middle illustrates the instructor shifting her gaze to direct attention toward and/or monitor the task space and then to monitor her student's understanding while providing instruction. This process is further facilitated by nonverbal feedback from partners whose understanding and actions are being monitored.

Theme 3: Deictic Gestures in Shared Contexts

Speakers frequently use deictic gestures either in place of or in conjunction with verbal utterances [17] in order to facilitate their communication of spatial goals in tasks that involve a shared visual context [14]. The excerpt below from the instruction task illustrates the instructor using a deictic gesture to provide the student with the spatial information necessary to complete the next step of the task:

> **Instructor:** Now bring that piece in your hand to here [points].

In both the instruction and cooperation tasks, increased use of deictic gestures resulted in better task performance. In the instruction task, deictic gestures were used to disambiguate locations and objects by both the instructor, $\beta = .314$, $t(23) = 8.47$, $p = .014$, and the student, $\beta = .268$, $t(23) = 3.01$, $p = .061$ as shown in Model 2. In the cooperative task, the participant who controlled the counting chips pointed to emphasize a particular concept (see Model 1), $\beta = .842$, $t(23) = 12.75$, $p < .001$, as their hands often held the counting chips and were therefore unable to perform other gestures. Deictic gestures were also frequently used in the negotiation task, for instance, to highlight components of the written package description. However, we found that an increased use of deictic gestures lowered task outcomes (see Model 3), regardless of whether the gesture was performed by the travel agent, $\beta = -.247$, $t(23) = -2.94$, $p = .057$, or traveler, $\beta = -.465$, $t(23) = -3.20$, $p = .045$.

Across all tasks in our study, we observed deictic gestures to be regularly used to replace or augment verbal utterances. Particularly in tasks that demanded communication of spatial information, deictic gestures played a key role in conveying complex information or instructions that may be difficult or impossible to verbally describe. Communicators might also use deictic gestures to save the energy that they would put into providing complex verbal descriptions or to provide sufficient redundancy in their language to prevent misunderstandings. In negotiation, however, gestures pointed toward partners might be perceived as a sign of an aggressive strategy for negotiation. Figure 4, Right illustrates deictic gestures directed toward the task space in the cooperation task and toward the partner in the negotiation task.

Theme 4: Nonverbal Feedback to Avoid Interruption

Throughout all three tasks, nonverbal feedback positively predicted task performance. Such feedback can be used to provide input without interrupting the speaker [27, 46]. The example below from the cooperation task illustrates this behavior:

> **Participant 1:** We could bring a passport, cash, ...
> **Participant 2:** [nods]
> **Participant 1:** ...traveler's checks.

Here, Participant 1 is listing off items to pack for the trip, and Participant 2—not wanting to interrupt—nods to indicate their agreement with the items that Participant 1 listed thus far.

Head nods were a common form of nonverbal feedback observed in both the instructor, $\beta = .326$, $t(23) = 6.02$, $p = .023$, and the student (see Model 2), $\beta = .210$, $t(23) = 2.32$, $p = .061$, as well as both participants in the cooperation task (see Model 1), $\beta = .578$, $t(23) = 8.97$, $p = .004$ and $\beta = .441$, $t(23) = 7.28$, $p = .013$. Additionally, laughter was a significant predictor of task success for the left participant in the cooperative task (see Model 1), $\beta = .758$, $t(23) = 2.56$, $p = .091$, and the traveler in the negotiation task (see Model 3), $\beta = .323$, $t(23) = 4.73$, $p = .039$.

We observed participants to use nonverbal feedback to express their understanding or agreement with their partner's utterances without causing disruption. These cues enabled participants to provide feedback while maintaining the fluidity of the interaction.

DISCUSSION

Our findings make three main contributions to research and design in CSCW: (1) they provide a set of models that quantitatively characterize the relative importance of a set of social cues in predicting specific outcomes for a number of collaboration scenarios; (2) they enable comparisons of these predictive relationships across collaboration outcomes and scenarios; and (3) they confirm and further detail prior findings from face-to-face and mediated communication. The paragraphs below discuss the research and design implications of each type of contribution. Additionally, we discuss some of the key limitations of our study.

Relative Importance of Social Cues

The models presented in the previous section inform the design of technologies to support particular collaboration scenarios. For example, in the instruction task, deictic gestures were weighted more heavily than gaze, suggesting that if designers must choose a single channel to display, such as choosing between displaying either a partner or the workspace, choosing the workspace would provide greater benefits to task outcomes, as suggested by findings reported in Theme 3. However, these weights might vary across different outcomes within the same task, which are tradeoffs designers must consider. For instance, although our data did not show student gaze to affect task outcomes in instruction, it showed gazing toward one's partner to have a positive impact on empathy and gazing toward the shared workspace to decrease empathy.

While technologies would ideally make all essential channels simultaneously available to the user, this goal can be difficult to achieve due to technical, cost, or design constraints, such as limited screen space or the challenges of arranging multiple sources of channels in a useable layout. Designers might alleviate these limitations by providing users with control over how video feeds are displayed, allowing overlaying of channels, such as the ClearBoard design that involved a shared workspace for sketching that overlaid the partner's video [25], or mechanisms for quickly switching between channels to view the channel that best supports the current needs of the communication.

Change of Relative Importance of Cues Across Scenarios

Our models also inform designers about how the role each type of social cue plays in collaboration changes across different forms of collaboration. While the tasks considered in this study were each carried out and analyzed independently of the other tasks, real-world use of collaboration technologies might involve more fluid transitions in tasks. For example, a sales manager might use a collaboration tool to create a sales plan with another manager at a satellite branch and then later use the same system to negotiate product pricing with a distributor. In the cooperation task, the combination of a large number of verbal, facial, and gesture cues are equally important, while the negotiation task relies heavily on deictic gestures and gaze (see Theme 1). While a face-only video feed combined with a view of the shared visual context may be sufficient to support the negotiation task, a successful cooperative task might require a broader view of each participant that includes their hands and arms to support the communication of gestures.

The potential use of CSCW systems across different forms of collaboration requires these systems to support several modalities of communication, such as a view of the partner and their workspace and audio of the primary speech channel and of verbal feedback. All channels may not be necessary at all times, either because the social cues communicated in that channel do not affect task outcomes or because a channel is not currently available or useful. Future systems should consider dynamically prioritizing their support for channels based on the ongoing task. This goal might be achieved by providing users with a set of task-based profiles, such as "negotiation" or "instruction," to enable them to easily adapt their use of channels to the demands of the specific task at hand. Future systems might also detect the type of task users are engaged with in order to automatically adapt their support for different channels to specific task requirements.

Confirmation of Prior Results

Our work confirms a number of findings from previous research, demonstrating that our tasks are representative of collaboration scenarios explored in prior work. We corroborate findings from both face-to-face and computer-mediated communication studies, suggesting that our work might have implications for future design of CSCW systems.

Gaze has often been identified as an important cue in communication. In face-to-face interaction, gaze can be used to disambiguate or reinforce references to objects in instruction tasks [40]. Similar results have been found in mediated communication, where gaze synchronization between partners results in an increase in task performance [48]. As seen in Theme 2, across all three tasks, gaze was repeatedly one of the most significant predictors of objective outcomes. Participants' gaze was also a significant predictor of outcomes such as increased empathy, leading to an improvement in communication effectiveness in cooperative tasks. However, our findings contradict work that found that instructors do not use gaze towards their partner as an integral part of their communication during teaching tasks [16]. We found deictic gestures to be important predictors of objective outcomes when a shared workspace was in use, confirming prior work concerning the role of deictic gestures in augmenting communication, leading to a decrease in task time and an increase in task comprehension in instructional tasks [18]. Additionally, feedback behaviors such as nodding and verbal backchannels were found to be significant predictors of task outcomes in the instruction and cooperation tasks, as shown in Theme 4. This result is consistent with prior work concerning the role of backchannels in communicating acknowledgement and understanding of a partner's utterances [17], as well as work indicating that backchannels can improve task performance and communicate task engagement [27].

Limitations

Our investigation offers many insights into the role of social cues in collaboration, although the study and the research approach presented here have a number of limitations. First, our regression-based approach used aggregate counts of social cues to characterize their use in the entire duration of the task, which does not offer insight into the temporal characteristics of these behaviors and their implications for collaboration. Therefore, we are able to draw only high-level conclusions about the importance of each of these behaviors.

Second, while we expect our findings to inform the design of CSCW systems, we also anticipate limitations in how much findings from face-to-face interactions can predict outcomes of mediated interactions, as previous work has shown that communicators often adapt their behaviors to the limitations and opportunities introduced by remote collaboration technologies in order to achieve effective communication [20, 42].

Finally, as mentioned previously, additional factors such as the dialogue content or the environment in which the collaboration is carried out may serve as additional predictors of the objective and subjective outcomes for the tasks considered in this study. Our work is not an exhaustive exploration of additional factors, as it focuses only on visible or audible behaviors in a distraction-free laboratory environment. Future work should include a more nuanced exploration of factors such as discourse markers and thematic shifts in dialogue in noisy environments, which may offer a richer understanding of how social cues, features of dialogue, and environmental factors work together to shape collaboration.

CONCLUSION

With the increasing use of technologies that support remote collaboration, future technologies must be designed to facilitate a wide range of collaborative activities and support transitions between these activities. To help guide the development of such future systems, we conducted a study of face-to-face collaborations to better understand what communicative behaviors significantly predict task and social outcomes and thus might serve as important features that designers must seek to support in remote collaboration. Using a regression-based approach, we modeled the predictive relationships between facial, gestural, and vocal cues and collaborative outcomes in cooperation, instruction, and negotiation tasks. Our models provide a quantitative understanding of how each type of social cue contributes to shaping collaborative outcomes in each task and insights into how such relationships generalize across tasks. Our results also provide confirmation and illumination of prior findings from face-to-face and computer-mediated communication research. These findings will inform the design of future CSCW technologies by providing designers with an understanding of the role of social cues in collaboration.

ACKNOWLEDGMENTS

We would like to thank Brandi Hefty and Jilana Boston for their assistance in data collection and analysis, Catherine Steffel for her help with the editing of the paper, and our anonymous reviewers for their comments and suggestions. National Science Foundation award 1149970 supported this work.

REFERENCES

1. Birnholtz, J., Grossman, T., Mak, C., and Balakrishnan, R. An exploratory study of input configuration and group process in a negotiation task using a large display. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 91–100.
2. Brendgen, M., Bowen, F., Rondeau, N., and Vitaro, F. Effects of friends. characteristics on children.s social cognitions. *Social Development 8*, 1 (1999), 41–51.
3. Carnevale, A. Workplace basics: The skills employers want. *Community, Technical, and Junior College Journal 59*, 4 (1989), 28–33.
4. Carnevale, P., and Isen, A. The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. *Organizational behavior and human decision Processes 37*, 1 (1986), 1–13.
5. Conversy, S., Gaspard-Boulinc, H., Chatty, S., Valès, S., Dupré, C., and Ollagnon, C. Supporting air traffic control collaboration with a tabletop system. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, ACM (2011), 425–434.
6. Cook, L., and Friend, M. Principles for the practice of collaboration in schools. *Preventing School Failure: Alternative Education for Children and Youth 35*, 4 (1991), 6–9.
7. Cornelius, C. J., Nguyen, M. A., Hayes, C. C., and Makena, R. Supporting virtual collaboration in spatial design tasks: Are surrogate or natural gestures more effective? *Human-Machine Systems, IEEE Transactions on 43*, 1 (2013), 92–101.
8. Culnan, M., and Markus, M. *Information technologies.* Sage Publications, Inc, 1987.
9. Daft, R., and Lengel, R. Information richness. a new approach to managerial behavior and organization design. Tech. rep., DTIC Document, 1983.
10. Dong, W., and Fu, W.-T. One piece at a time: why video-based communication is better for negotiation and conflict resolution. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 167–176.
11. Drolet, A. L., and Morris, M. W. Rapport in conflict resolution: Accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. *Journal of Experimental Social Psychology 36*, 1 (2000), 26–50.
12. Eagly, A., and Johnson, B. Gender and leadership style: A meta-analysis. *Psychological bulletin 108*, 2 (1990), 233.
13. Efroymson, M. Multiple regression analysis. *Mathematical methods for digital computers 1* (1960), 191–203.
14. Farmer, S. M., and Hyatt, C. W. Effects of task language demands and task complexity on computer-mediated work groups. *Small Group Research 25*, 3 (1994), 331–366.
15. Fernald, A., Zangl, R., Portillo, A. L., and Marchman, V. A. Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. *Language Acquisition and Language Disorders 44* (2008), 97.
16. Fussell, S., Setlock, L., and Parker, E. Where do helpers look?: gaze targets during collaborative physical tasks. In *CHI'03 extended abstracts on Human factors in computing systems*, ACM (2003), 768–769.
17. Fussell, S. R., Kraut, R. E., and Siegel, J. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM (2000), 21–30.
18. Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E., and Kramer, A. D. Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction 19*, 3 (2004), 273–309.

19. Gergle, D., and Clark, A. T. See what i'm saying?: using dyadic mobile eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, ACM (2011), 435–444.
20. Gergle, D., Kraut, R. E., and Fussell, S. R. Language efficiency and visual technology minimizing collaborative effort with visual information. *Journal of Language and Social Psychology 23*, 4 (2004), 491–517.
21. Gergle, D., Rosé, C., and Kraut, R. Modeling the impact of shared visual information on collaborative reference. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 1543–1552.
22. Grawitch, M., Munz, D., Elliott, E., and Mathis, A. Promoting creativity in temporary problem-solving groups: The effects of positive mood and autonomy in problem definition on idea-generating performance. *Group dynamics: Theory, research, and practice 7*, 3 (2003), 200.
23. Heath, C., and Luff, P. Collaborative activity and technological design: Task coordination in london underground control rooms. In *Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work*, Kluwer Academic Publishers (1991), 65–80.
24. Huang, C.-M., and Mutlu, B. Evaluating narrative gestures for humanlike robots. In *Robotics: Science and Systems 2013* (2013).
25. Ishii, H., and Kobayashi, M. Clearboard: a seamless medium for shared drawing and conversation with eye contact. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (1992), 525–532.
26. Jonassen, D., and Kwon, H. Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development 49*, 1 (2001), 35–51.
27. Jung, M. F., Lee, J. J., DePalma, N., Adalgeirsson, S. O., Hinds, P. J., and Breazeal, C. Engaging robots: easing complex human-robot teamwork using backchanneling. In *Proceedings of the 2013 conference on Computer supported cooperative work*, ACM (2013), 1555–1566.
28. Junuzovic, S., Inkpen, K., Blank, T., and Gupta, A. Illumishare: sharing any surface. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 1919–1928.
29. Karsenty, L. Cooperative work and shared visual context: An empirical study of comprehension problems in side-by-side and remote help dialogues. *Human-Computer Interaction 14*, 3 (1999), 283–315.
30. Kirk, D., Rodden, T., and Fraser, D. S. Turn it this way: grounding collaborative action with remote gestures. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 1039–1048.
31. Kling, K., Hyde, J., Showers, C., and Buswell, B. Gender differences in self-esteem: a meta-analysis. *Psychological bulletin 125*, 4 (1999), 470.
32. Koschmann, T., LeBaron, C., Goodwin, C., and Feltovich, P. The mystery of the missing referent: Objects, procedures, and the problem of the instruction follower. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, ACM (2006), 373–382.
33. Kraut, R. E., Fussell, S. R., and Siegel, J. Visual information as a conversational resource in collaborative physical tasks. *Human-computer interaction 18*, 1 (2003), 13–49.
34. Kraut, R. E., Gergle, D., and Fussell, S. R. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, ACM (2002), 31–40.
35. Kuzuoka, H., Kosuge, T., and Tanaka, M. Gesturecam: a video communication system for sympathetic remote collaboration. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, ACM (1994), 35–43.
36. Laufer, L., Halacsy, P., and Somlai-Fischer, A. Prezi meeting: collaboration in a zoomable canvas based environment. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, ACM (2011), 749–752.
37. Munier, M., Baïna, K., and Benali, K. A negotiation model for cscw. In *Cooperative Information Systems*, Springer (2000), 224–235.
38. Nadler, J. Rapport in negotiation and conflict resolution. *Marq. L. Rev. 87* (2003), 875.
39. Ochsman, R. B., and Chapanis, A. The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. *International Journal of Man-Machine Studies 6*, 5 (1974), 579–619.
40. Ou, J., Oh, L. M., Yang, J., and Fussell, S. R. Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2005), 231–240.
41. Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M., and Poggi, I. A model of attention and interest using gaze behavior. In *Intelligent Virtual Agents*, Springer (2005), 229–240.
42. Schober, M. Spatial perspective-taking in conversation. *Cognition 47*, 1 (1993), 1–24.
43. Sebanz, N., Bekkering, H., and Knoblich, G. Joint action: bodies and minds moving together. *Trends in cognitive sciences 10*, 2 (2006), 70–76.
44. Short, J., Williams, E., and Christie, B. *The social psychology of telecommunications*. John Wiley and Sons Ltd, 1976.
45. Sierra, C., Jennings, N., Noriega, P., and Parsons, S. A framework for argumentation-based negotiation. *Intelligent Agents IV Agent Theories, Architectures, and Languages* (1998), 177–192.
46. Stivers, T. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction 41*, 1 (2008), 31–57.
47. Stuhlmacher, A., and Walters, A. Gender differences in negotiation outcome: A meta-analysis. *Personnel Psychology 52*, 3 (1999), 653–677.
48. Vertegaal, R., and Ding, Y. Explaining effects of eye gaze on mediated group conversations:: amount or synchronization? In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, ACM (2002), 41–48.
49. Walther, J., and Parks, M. Cues filtered out, cues filtered in: Computer-mediated communication and relationships. *Handbook of interpersonal communication 3* (2002), 529–563.
50. Weiseth, P., Munkvold, B., Tvedte, B., and Larsen, S. The wheel of collaboration tools: a typology for analysis within a holistic framework. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, ACM (2006), 239–248.
51. Whittaker, S. Theories and methods in mediated communication. *The handbook of discourse processes* (2003), 243–286.
52. Zimmermann, P. Negotiation in the workplace. *Emergency nurse: the journal of the RCN Accident and Emergency Nursing Association 11*, 6 (2003), 12.