

# Exploring Socio-Cognitive Effects of Conversational Strategy Congruence in Peer Tutoring

Tanmay Sinha  
Articulab  
School of Computer Science  
Carnegie Mellon University  
tanmays@cs.cmu.edu

Ran Zhao  
Articulab  
School of Computer Science  
Carnegie Mellon University  
rzhao1@cs.cmu.edu

Justine Cassell  
Articulab  
School of Computer Science  
Carnegie Mellon University  
justine@cs.cmu.edu

## ABSTRACT

Conversational interaction is a dynamic process in which information is conveyed and relationships are negotiated via the use and timing of specific conversational strategies. In this work we examine similarity in use and timing of the relationship-oriented communicative strategies self-disclosure, reference to shared experience and praise, during a reciprocal peer tutoring interaction. We computationally model two kinds of similarity that quantify whether and how students are similar or different in their use of the strategies over time, and differentiate the effects by gender, relationship status and session. In order to assess their impact, we leverage learning and self-reported rapport as outcome variables. Our results show significant effects in cumulative use as well as in the pattern of timings of conversational strategy usage by partners in a dyad, along with interesting relationships to socio-cognitive processes.

## Categories and Subject Descriptors

H.1.2 [Information Systems]: Models and Principles—*User/Machine Systems, Human Factors*

## Keywords

Conversational Strategy; Rapport; Learning

## 1. INTRODUCTION

Social interaction has long been recognized as contributing to cognitive development [35]. In that vein, peer tutoring provides opportunities for developing social and communication skills, positive attitudes towards social relationships and the process of learning, as well as domain-specific skills. It has been emphasized that these relationships are likely to affect task-mastery and creativity, especially in social contexts [13]. The kinds of relationships that support well-being and learning in this way depend crucially on the development of interpersonal closeness over time. For that reason, the study of social relationships and their contribution to learning and other cognitive tasks must also focus on time-bound dyadic processes that occur in the growth of social bonds.

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 the author(s) 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](mailto:Permissions@acm.org).

INTERPERSONAL '15, November 09-13 2015, Seattle, WA, USA

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-3986-5/15/11 ...\$15.00

DOI: <http://dx.doi.org/10.1145/2823513.2823519>.

Partners regulate their interaction through the use of conversational strategies (particular ways of talking) that contribute to - or have an effect on - building, maintaining (or, sometimes, destroying) a budding relationship. As computers take on more and more teaching tasks, it is important to understand what conversational strategies people use, whether they use some more than others, and whether particular ones are more useful than others in particular learning or social contexts (perhaps some work better with boys than girls, or some with old friends rather than new, or some might work better in social situations while others in classroom work etc). In addition, it is important to understand the time-bound patterns of reciprocity, synchrony, or the lack thereof in the use of these conversational strategies to see if not just the type of strategy but also the mutuality of its use, and the dynamics of its use over time plays a role in the impact of its use.

Therefore, in this work, we leverage structure from the patterns of conversational strategy usage (both cumulative and temporal) by students engaged in reciprocal peer tutoring, and examining their impact on learning and rapport. Specifically, we looked at the conversational strategies of self-disclosure, reference to shared experience and praise and operationalized two measures to quantify the similarity in their usage by two partners : a)first, absolute difference in the number of strategies used, b)second, a dynamic time warping based distance to capture alignment in the timings of strategy usage. Furthermore, since researchers have long posited that friends learn better together than do strangers, we investigated how friends and strangers differ in their conversational strategy exchange patterns. In addition, as prior work has found differences between boys and girls in the use of relationship-building talk, we looked at gender differences. Finally, we examined how interlocutors differ while working in task vs. social conversations, as well as in their interaction with each other over a period of weeks.

In what follows, we describe relevant related work that motivates our current research. The study context and methodology are described in section 3 and 4 respectively. We then present detailed analyses and results for each of our questions in sections 5 and 6, the discussion and conclusion in section 7, finally ending with future work in section 8.

## 2. RELATED WORK

For clarity, we divide related work into a first sub-section that provides theoretical background on identification of conversational strategies in situations of interpersonal closeness, as well as their relationship to rapport and learning. The

second sub-section describes related work that motivates our analysis of conversational strategy synchrony.

## 2.1 Conversational Strategies, Rapport and Learning

Rapport or the feeling of harmony or connection with another, is an important aspect of human interaction. This phenomenon has been shown to have powerful effects in the domain of education [5]. It has been proposed that rapport should be examined at the dyadic level instead of focusing on the individual [4]. Our prior work [36] has in fact developed a dyadic computational model that explains how dyads manage rapport through the use of specific conversational strategies, which in turn function to fulfill social goals that make up rapport - *face management*, *mutual attention*, and *coordination*.

Foundational work by [33] conceptualizes the interpersonal nature of *face* as a desire to be recognized for one's social value and individual positive traits. Face-boosting strategies such as *praise* serve to create increased self-esteem in the individual and increased interpersonal cohesiveness or rapport in the dyad. *Mutual attention* facilitates learning about the other person by leading dyads to provide information about themselves through the strategy of *self-disclosure*. As the relationship proceeds, self-disclosures become more intimate in nature. In addition, over time interlocutors increasingly *coordinate* less to sociocultural norms set by the outside world, and more to interpersonally determined norms. *Referring to shared experience* allows interlocutors to index commonality and differentiating in-group and out-group norms.

Prior work provides theoretical [5] and empirical evidence to support the positive impact of rapport on learning - rapport that could be facilitated, paradoxically, via rudeness or insults, which index the fact that interlocutors have reached the interpersonal norms stage of a relationship [24] or behavioral convergence along low-level linguistic features [31, 32]. However, the literature lacks an examination of how usage of discourse-level conversational strategies affects rapport and learning over time. What leads us to ask the question this way? As we have discussed above, usage of self-disclosure, reference to shared experience and praise are ways of building interpersonal connectedness between interlocutors. It has been suggested that greater similarity in verbal behaviors such as these is an index of *increased* connectedness and interpersonal rapport which, in turn, leads to greater willingness to examine misconceptions, and hence to improved learning [6]. In addition, it has been argued that rapport allows disagreement in such a way as to lead to reconciliation of views, and hence change in beliefs [20]. The conceptual framework of [22] distinguishes between the notion of *moves* (actions bearing on the course of interaction between participants) and *tactics* (moves selected in a deliberate, principled way, aimed at achieving a particular short-term goal), and *strategies* (algorithm for selecting from a set of tactics based on current circumstances, and aimed at achieving a long-term goal). We thus understand ways in which conversational moves can affect goals, such as tutoring another successfully.

## 2.2 Conversational Strategy Congruence

The social exchange theory, which defines social behaviors as an exchange [16, 15] motivates our analysis. During the

development of relationships, social exchange is regulated by a series of obligations [11] - how we feel entitled to respond based on the behaviors we expect from others (for e.g. desire to be approved of). Some prior work has proposed that reciprocity is a very important social norm in the early stages of a relationship [1].

However, other researchers have described a number of factors that attenuate the reciprocity of interpersonal exchange with the development of a relationship - for e.g. established trust and closeness do not require dyads to strictly follow the norms of reciprocal exchange. On the other hand, [23] suggests that increased closeness and trust do not reduce the obligation of reciprocating, but rather serve to extend the time frame of interpersonal exchange. Using the strategy of self-disclosure during peer interactions facilitates the development of a relationship [34]. In a similar vein, [21] points out that peers are thought to acquire similar behavioral repertoires during shared experiences, and therefore they are mutually attracted and seek each other out for further interaction. Overall, these findings suggest conversational strategy exchange patterns to be both a function of time as well as the relationship status of individuals.

However, the link between social interaction, peer familiarity and these conversational strategies is still unclear. For instance, [21] suggests that with growing familiarity, children learn similar response repertoires - a kind of social reciprocity - that promotes the maintenance of further social interactions. Children make overtures more frequently, as well as spend more time in interactions with familiar peers. Likewise, arousal reinforcement theory [3] suggests that being in an unfamiliar context arouses the child to a non-optimal level for interaction or exploration, while presence of a familiar peer reduces the child's arousal to a more optimal level. This leads us to look at interaction time-based effects in dyads of friends and strangers.

Conversational strategies can be perceived as social bids. The degree to which interlocutors are successful determines how these exchange patterns will vary over time. For instance, there is always risk and benefit involved during self-disclosures [25]. The costs of disclosing are increased vulnerability and less privacy. The benefits are increased trust, rapport and reciprocation, which could outweigh the costs [17]. Therefore, we are interested in investigating how peer-directed social bids, facilitated by the conversational strategies of self-disclosure, reference to shared experience and praise, affect learning and rapport outcomes. The literature on entrainment [9] emphasizes that one of the central insights behind the concept of communicative accommodation is that speakers are motivated to reduce communicative differences between themselves and the interlocutor for a myriad of reasons ranging from desire for approval to desire for more efficient communication. Thus, in our work, in addition to looking at the total count of conversational strategies, we investigate the similarity in cumulative as well as in pattern of timings of self-disclosure, reference to shared experience and praise.

## 3. STUDY CONTEXT

Reciprocal peer tutoring data was collected from 12 American English-speaking dyads (6 dyads were friends and 6 strangers; 6 were boys and 6 girls), with a mean age of 13 years, who interacted for 5 hourly sessions over as many weeks (a total of 60 sessions, and 5400 minutes of data), tu-

toring one another on procedural and conceptual aspects of an algebra topic. Prior work demonstrates that peer tutoring is an effective paradigm that results in student learning [30], making this an effective context to study dyadic interaction and learning. Each session began with social chit-chat, after which the first tutoring period started, followed by another small social interlude, a second tutoring period with role reversal between the tutor and tutee, and finally the final social time. In the following sections, we will use *social* period to refer to all the 3 social periods in a session, and *task* period to refer to both the tutoring periods in the session. Note that having 5 interaction sessions over many weeks allows us to investigate longitudinal differences in conversational strategy usage.

### 3.1 Learning Outcome

For every reciprocal peer tutoring session, the tutee was provided with a working sheet comprising of  $\approx 10$  questions on linear equations, which were to be solved and briefly explained step-by-step. In addition, the tutor was given a correctly solved version of the working sheet that he/she used to guide the tutee in the tutoring period.

To assess learning outcomes during and after the process of tutoring, we computed the following two measures reflecting problem correctness: a) **L1 attempted**: Total percentage of problems correctly solved by the tutee in each of the peer-tutoring sub-sessions in a session, out of total problems attempted in the working sheet. A question in the working sheet was marked as attempted, if at least one step was partially or fully solved, b) **L2 solved**: Total percentage of problems correctly solved by the tutee in each of the 2 peer-tutoring sub-sessions in a session, out of total problems present in the working sheet.

### 3.2 Rapport Outcome

After each session, both participants in the dyad completed 7 point likert scale (1 = Disagree Strongly; 7 = Agree Strongly) questionnaires, reflecting the dimensions of **Attentiveness** (3-item scale indexing interest, attention and respectfulness of the partner towards the speaker, Cronbach  $\alpha = 0.42$ ), **Positivity** (2-item scale indexing friendliness and warmth towards the partner,  $\alpha = 0.72$ ), **Coordination** (3-item scale indexing whether partners felt in sync, could say everything that they wanted to say and that the interaction was not frustrating,  $\alpha = 0.64$ ), and **Long Term Rapport** (3-item scale indexing whether the partners felt that they knew each other, were more comfortable and had greater liking compared to the previous interaction session,  $\alpha = 0.78$ ). In addition, the questionnaire asked about **Self Efficacy** (7-item scale indexing whether the partners thought they were good tutors, learned a lot from tutoring and were concerned about tutoring quality, motivation and impact on the tutee,  $\alpha = 0.5$ ).

In order to compute a dyadic measure from individual questionnaire ratings, we computed the following two measures: a) **R1 total**: Total score for each questionnaire dimension, calculated by addition of individual questionnaire scores, b) **R2 mean&sd**: Mean of the score for each questionnaire dimension for the dyad, subtracted by the standard deviation. Intuitively, this metric will be higher if average questionnaire scores are higher for the dyad, as well as individual variability from the mean is lower, and vice versa.

## 4. METHODOLOGY

### 4.1 Coding of Conversational Strategies

In order to construct a reliably annotated corpus, we employed 3-5 human annotators to code conversational strategies that prior work has shown to contribute to rapport. In this paper we use the annotations finished to-date, specifically the conversational strategies of self-disclosure, reference to shared experience and praise. Annotators were provided with explicit definitions and examples to use in making their judgment. Inter-rater reliability of conversational strategy annotations, computed via Krippendorff's alpha was 0.753 for self-disclosure, 0.798 for reference to shared experience and 1.0 for praise. After achieving high enough inter-rater reliability, most of the sessions were coded by independently by the annotators. Below, we briefly describe the rationale behind our coding manuals, along with example utterances from our dataset depicting these three conversational strategies.

#### 4.1.1 Self-Disclosure

Self-disclosure refers to the verbal expressions used by people to reveal aspects of themselves to others. It occurs when people reveal personal private information about themselves – meaning information that otherwise would not be seen or known by the person being disclosed to (or would be difficult to see or know). A lot of psychological literature talks about the ways people reveal facts about themselves as ways of building relationships, but we are the first to look at the role of self-disclosure during social and task interactions by the same dyad, particularly for adolescents engaged in reciprocal peer tutoring. While there are several categories of personal and private self-disclosures, we were interested in two types: a) self-disclosures that reveal the long-term aspects of oneself that one may feel are deep and true, and therefore important to reveal in the context of a relationship (e.g, “I love playing hockey”), b) self-disclosures about one's transgressive (forbidden or socially-unacceptable) actions, which may be a way of making the other person feel better by disclosing that one is not perfect (e.g, “I did badly on the pre-test”). We were interested in these two because they are excellent examples of personal private information that would not be known if not revealed, and because there are abundant examples of each in our adolescent data set, which makes us think there is something important there to examine. In addition to coding for the presence of self-disclosure, we also coded its dimensions of valence, intimacy and indexing commonality.

#### 4.1.2 Referring to Shared Experience

Referring to shared experience is an important way of showing that the two interlocutors have known each other and interacted outside of the context of the current peer tutoring interaction. To be clear, we did not code aspects of one's experience that were similar (e.g, both partners liking video games). These would fall into the category of “common (similar) interests”. We only coded for experiences that the two participants had or would have together (e.g, going to the mall together last week). Reference to shared experiences is a way of indexing a special bond or relationship that other people don't have, which can in turn build rapport. In addition to coding for the presence of shared experience, we also coded its two dimensions: ROE (reference to

outside current experience, meaning shared activities that are outside the experiment; e.g, “Did you see my Facebook post last night?”) and RIE (reference to inside current experience, meaning peer tutoring related experience, such as referring to the pre-test or post-test; e.g, “I remember you helped me with a problem like this before”).

### 4.1.3 Praise

Praise is the expression of a favorable judgment of an attribute, behavior or product of the other person. We were interested in two types of praise: labeled and unlabeled. Labeled praise is an expression of a positive evaluation of a specific attribute, behavior or product of the other person (e.g, “great job with those negative numbers”), while unlabeled praise is a generic expression of positive evaluation, without a specific target (e.g, “Perfect”).

## 4.2 Operationalizing Similarity Constructs

We utilized dynamic time warping (DTW) to obtain a global distance that can characterize how conversational strategy usage for each partner in the dyad is aligned in time. Originally presented by [19] for speech recognition purposes, this technique allows two time series that are similar but locally out of phase to align in a nonlinear manner. In our case, each element in the two time series refers to time from the start of a peer tutoring session (in seconds) at which each individual in the dyad used certain conversational strategy. Concretely, given these two time series, say  $A = [a_1, \dots, a_n] \in \mathbb{R}^{1 \times n}$  and  $B = [b_1, \dots, b_n] \in \mathbb{R}^{1 \times n}$ , DTW is a technique to align  $A$  and  $B$  such that the sum of the Euclidean distances between the aligned samples is minimized. In order to perform this alignment, DTW can distort (or, as [19] call it, “warp”) the time axis - compressing it at some places and expanding it at others.

Thus, by viewing the time axis as a stretchable one, DTW is able to match (via construction of a warping path) a point of time series  $A$  even with surrounding points of time series  $B$ . Minimum global dissimilarity, or DTW distance can be assumed as the stretch-insensitive measure of the inherent difference between two given time series. Furthermore, the shape of the warping curve itself provides information about which point matches which, i.e., the pair wise correspondences of time points can be easily inspected.

This warping path is a central part of comparing the two time series, since it determines which points match and are to be used for calculating the distance between the two time series. One simplistic way, for instance, is linear matching, that aligns the  $i^{th}$  point of the first curve with the  $i^{th}$  point of the second curve. Because this matching is very sensitive to small distortions in the time axis, a more computationally expensive way is to perform a complete matching. This technique calculates the distance between every point of the first curve and every point of the second curve. For every point, the smallest distance to the other curve is decided. These distances are summed and divided by the number of points. Each point can match with no more than one point of the other curve.

In contrast, a non-linear (elastic) alignment provided by DTW produces an intuitive similarity measure, allowing similar shapes to match even if they are locally out of phase on the time axis, by allowing elastic shifting in order to detect similar shapes with different phases. [26] provide a more comprehensive technical description of the DTW algorithm.

To investigate similarity in the pattern (timing) of conversational strategy usage, we employed the *dtw* package in *R* [12], and selected the following parameters: a)First, we utilized the *symmetric2* step-pattern [29] that lists transitions allowed while searching for the minimum distance path between the two time series, with the constraint that one diagonal step costs as much as the two equivalent steps along the sides. Intuitively, step-patterns limit the maximum amount of time stretch and compression allowed at any point of the alignment, b)Second, we performed an open-ended alignment, meaning that we freed the endpoint of time series  $B$  in order to allow for a partial match. Intuitively, relaxing the end-point constraint results in computing the alignment which best matches all of time series  $B$  with a leading part of time series  $A$ , c)Third, we normalized the DTW distance by the length of the two input time series, in order to accommodate time-series of varying lengths (certain dyads, for instance, do lot more self-disclosure relative to other dyads).

Finally, since conversational strategy usage by the partners in the dyad not only varies in terms of expressive frequency, but also in terms of time progression, we computed two measures reflective of similarity between partners in the dyad: a)**Diff**: This metric measures the absolute difference between the number of a specific conversational strategy usage by both students in a dyad. Thus, a lower difference would imply that the count of a specific conversational strategy is very similar for the dyad, b)**Diff+Time**: Since **Diff** metric cannot take into account the temporal distribution of conversational strategies, we utilized the normalized DTW distance obtained from the time warping algorithm as a second measure reflective of the similarity in the pattern of timings of conversational strategy usage.

## 5. EXPLANATORY ANALYSIS

We conducted a four-way  $2 \times 2 \times 2 \times 5$  repeated measures ANOVA to investigate the effect of gender (male, female), relationship status (friends, strangers), period (social, task) and session (1, 2, 3, 4, 5) on the a)total conversational strategy usage by partners in a dyad (section 5.1), b)absolute & DTW difference in the number of usages by partners in a dyad (section 5.2). Period and session were used as within-subject repeated measures.

For all statistically significant effects ( $p < 0.05$ ) reported, we also looked at effect size (Cohen’s  $d$ ) in order to assess the practical significance of these results. In essence, effect size is the difference between two means (e.g., friends minus strangers) divided by the pooled standard deviation (adjusted with weights for the sample sizes) of the two conditions [10].

### 5.1 Conversational Strategy Usage (Total)

#### 5.1.1 Self-Disclosure

With total number of self-disclosures as the dependent variable, the interaction between period and relationship status was significant ( $F(1, 96) = 6.579, p = 0.01$ ), with the mean value of total number of self-disclosures being higher for friends in the task period compared to their social period, while being lower for strangers in task period as compared to their social period (figure 1.a).

While the nature of the task focuses partners more on the math problems instead of on personal information, the social periods are less restrictive. However, strangers are

more likely to follow “stereotypical behaviors” [5], such as adherence to the conventions of the task, in turn leading them to use fewer self-disclosures during the task. On the other hand, friends are more likely to ignore experimental rules and follow interpersonal norms regardless of whether it is a social or task period. They are also more likely to focus on the relationship than the task.

In addition to this interaction effect, we found a significant main effect for relationship status ( $F(1, 96) = 19.921, p < 0.001$ ; Cohen’s  $d=0.742$ ), with friends ( $M=16.15, SD=21.22$ ) disclosing more than strangers ( $M=4.482, SD=5.43$ ). In sync with prior work [27], a more intimate relationship leads to self-disclosing to maintain the relationship [27], clarify who one is and solidify the relationship [25]. Conversely, in less intimate relationships (i.e., strangers), the reasons for disclosure mainly involves impression management, and acquiring knowledge on which to build a relationship.

Finally, we also found a significant main effect for gender ( $F(1, 96) = 6.993, p = 0.009$ ; Cohen’s  $d=0.478$ ), with the number of self-disclosures being higher for females ( $M=14.55, SD=22.72$ ) compared to males ( $M=6.75, SD=5.81$ ). As [25] points out, this classic gender difference in self-disclosure results from variations in how men and women are socialized and gender-role expectations [28].

### 5.1.2 Reference to Shared Experience

With total number of reference to shared experience as the dependent variable, only the main effect for relationship status was significant ( $F(1, 80) = 16.882, p < 0.001$ ; Cohen’s  $d=0.871$ ), with friends ( $M=3.37, SD=3.63$ ) having higher number of reference to shared experience compared to strangers ( $M=0.88, SD=0.82$ ).

On investigating the two dimensions of reference to shared experience, we found that a) friends ( $M=4.53, SD=4.73$ ) had significantly ( $F(1, 58) = 26.65, p < 0.0001$ ) higher ROE than strangers ( $M=0.07, SD=0.25$ ), b) friends also had significantly ( $F(1, 58) = 7.14, p = 0.009$ ) higher RIE ( $M=2.2, SD=1.86$ ) than strangers ( $M=1.1, SD=1.27$ ).

These results seem to suggest that friends use a lot of references to shared experience in the peer tutoring interaction, while strangers, who even though share some experience inside their current interaction (RIE) don’t refer much to it. As far as ROE is concerned, friends use ROE to distinguish in-group from out-group individuals, while strangers lack a shared base [8] outside their current interaction, leading them to have significantly lower ROE than friends.

### 5.1.3 Praise

With the total number of praise utterances as the dependent variable, only the main effect for period was significant ( $F(1, 80) = 4.620, p = 0.03$ ; Cohen’s  $d=0.91$ ), with task periods ( $M=3.52, SD=4.69$ ) having more praise utterances than social periods ( $M=0.46, SD=0.79$ ). Intuitively, this result makes sense since praise serves as an important motivational strategy for explicit encouragement or feedback in the task periods and is less likely to be used very frequently during a social chit-chat.

Moreover, rapport management requires face management, which in turn relies on behavioral expectations and interactional goals [33]. Since these interactional goals are more explicit in the task period compared to the social period, greater face management is required, which leads interlocutors to use more praise.

## 5.2 Conversational Strategy Usage (Similarity)

### 5.2.1 Self-Disclosure

By employing absolute difference in the number of self-disclosures (**Diff**) as the dependent variable, we found significant interaction effects between period and relationship status ( $F(1, 96) = 5.273, p = 0.02$ ), with friends having higher similarity in the mean number of self-disclosures in social sessions as compared to their task sessions, and strangers having higher similarity in number of self-disclosures in task sessions as compared to their social sessions (figure 1.b).

In addition to this interaction effect, main effects for gender ( $F(1, 96) = 6.26, p = 0.014$ ; Cohen’s  $d=0.48$ ) and relationship status ( $F(1, 96) = 10.928, p = 0.001$ ; Cohen’s  $d=0.59$ ) were significant, with male dyads ( $M=2.02, SD=2.29$ ) and stranger dyads ( $M=1.66, SD=1.69$ ) being more similar in the number of self-disclosures done by the 2 partners, compared to female ( $M=4.125, SD=5.97$ ) and friend ( $M=4.32, SD=5.87$ ) dyads respectively. Females were more unbalanced in the number of self-disclosures done by both the partners in the dyad. We did not find any significant differences across gender, relationship status, period and session for the DTW distance (**Diff+Time**), reflective of similar timings of self-disclosure by the partners.

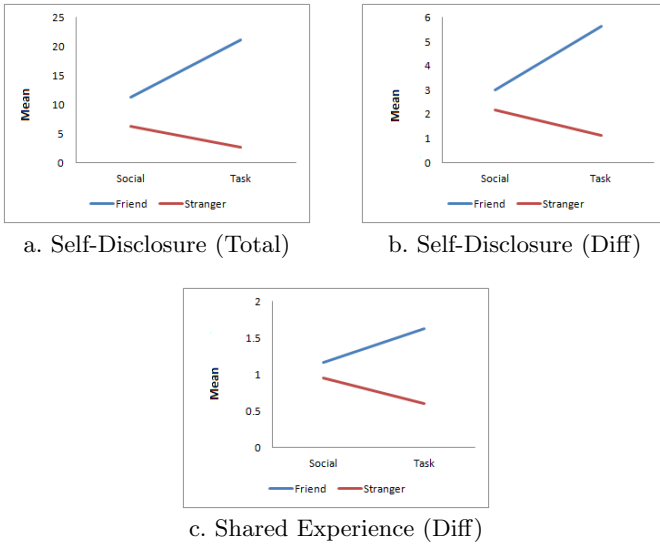
### 5.2.2 Reference to Shared Experience

By employing absolute difference in the number of reference to shared experience (**Diff**) as the dependent variable, we found that these results followed the same trend as the results for absolute difference in self-disclosures. There was a significant effect for interaction (figure 1.c) between period and relationship status ( $F(1, 80) = 4.33, p = 0.04$ ). Furthermore, the main effect for relationship status was significant ( $F(1, 80) = 8.584, p = 0.004$ ; Cohen’s  $d=0.56$ ), with stranger dyads ( $M=0.775, SD=0.8$ ) being more similar in the number of reference to shared experience done by both partners in the dyad, compared to friends ( $M=1.4, SD=1.26$ ). We did not find any significant differences across gender, relationship status, period and session for the DTW distance (**Diff+Time**), reflective of similar timings of shared experience by the partners.

### 5.2.3 Praise

With the absolute difference in number of praise utterances (**Diff**) as the dependent variable, only the main effect for period was significant ( $F(1, 80) = 5.837, p = 0.01$ ; Cohen’s  $d=0.796$ ), with social periods ( $M=0.38, SD=0.6$ ) more similar in number of praise utterances used by both individuals, compared to task periods ( $M=2.96, SD=4.54$ ). In order to understand this result, one can turn to the following mechanism - in a conventional tutoring session (where the tutor and tutee do not reverse roles in the interaction), a one-way mechanism such as praise is more likely to be prevalent in the task sessions, since it is a direct means of bolstering the tutee’s confidence.

From the interactional perspective too, the tutor’s commands to the tutee can potentially threaten tutee’s negative face and therefore, the tutor is likely to use more praise in order to hedge such face threatening acts. Therefore, an imbalance in number of praise strategies during a task is more natural. However, in our case, since our peer tutoring session was reciprocal (both students in the dyad got to play the roles of tutor and tutee), a higher imbalance in praise



**Figure 1: Plots depicting significant interaction effects from the repeated measures ANOVA**

utterances seems to suggest that one of the students used less praise than his partner. We did not find any significant differences across gender, relationship status, period and session for the DTW distance (**Diff+Time**), reflective of similar timings of praise by the partners.

## 6. IMPACT OF SIMILARITY

Next, we computed correlations (Pearson  $r$ , Spearman Rank  $\rho$ ) to find relationships among the outcome variables of learning (section 3.1), rapport (section 3.2) and operationalized measures of conversational strategy similarity (section 4.2). Significance of the correlation was assessed via two tailed t-test. Since the total conversational strategy usage hides information about the exchange patterns of the individuals involved in peer tutoring, we were interested in testing the impact of only conversational strategy similarity on socio-cognitive outcomes.<sup>1</sup>

### 6.1 Self-Disclosure

Overall, for the 12 dyads, we found significant correlations with learning. Greater similarity in the number of self-disclosures by the partners (**Diff**) was positively associated with learning ( $r = -0.47, p = 0.0008$ ). However, the less similar a dyad was in their pattern of timings of self-disclosure (**Diff+Time**), the more they learned ( $\rho = 0.31, p = 0.031$ ). In other words, the less the tendency for reciprocity in the conversational exchange pattern over time, the more the dyad learned during that session.

For friend dyads greater similarity in pattern of timings of self-disclosure was positively associated with higher ratings (**R2 mean&sd**) for attentiveness ( $\rho = -0.37, p = 0.04$ ) and coordination ( $\rho = -0.37, p = 0.04$ ) on the questionnaire dimension of self-reported rapport. These results suggest a positive impact of similarity in timings of conversational strategy usage by partners on rapport, while a negative impact on learning. We found similar results for female dyads.

<sup>1</sup>Since results mostly follow a similar trend with both the learning outcomes, we just report correlations with **L1 attempted**. Wherever appropriate, correlations with **L2 solved** is reported.

Looking at gender and relationship status together, we found that lower difference (greater similarity) in the number of self-disclosures done by the individuals in a dyad (**Diff**) was associated with a) higher learning ( $r = -0.53, p = 0.05$ ) in female-friend dyads, b) lower learning (**L2 solved**) in female-stranger dyads ( $\rho = 0.76, p = 0.04$ ). For male-friend and male-stranger dyads, we didn't find any significant effects for learning. However, interestingly, the more these dyads differed in the number of self-disclosures done by the partners (**Diff**), higher were the total (**R1 total**) coordination ( $\rho = 0.51, p = 0.05$ ) and self-efficacy ( $r = 0.58, p = 0.05$ ) ratings provided.

### 6.2 Reference to Shared Experience

Overall, again, we found significant correlations with learning. Greater similarity in number of references to shared experience by the partners (**Diff**) was positively associated with learning ( $r = -0.50, p = 0.007$ ). For female dyads, in addition, higher differences in number of reference to shared experience was associated with higher positivity (**R2 mean&sd**) ratings ( $r = 0.54, p = 0.05$ ) in the questionnaire. However, the more similar a female dyad was in their pattern of timings of shared experience usage (**Diff+Time**), higher were the total positivity (**R1 total**) ratings ( $r = -0.62, p = 0.02$ ).

Looking at gender and relationship status together, we found interesting differences between male-friend dyads and female-friend dyads. Lower difference (greater similarity) in the number of reference to shared experience usage by individuals in the dyad (**Diff**) was associated with a) higher learning (**L2 solved**) in female-friend dyads ( $r = -0.63, p = 0.03$ ), b) lower learning ( $r = 0.54, p = 0.05$ ) in male-friend dyads.

### 6.3 Praise

Overall, for our entire study sample (12 dyads), we found significant correlations with self-efficacy, attentiveness and coordination. Greater similarity in number of praise usage by individuals in a dyad (**Diff**) was positively associated with self-efficacy (**R2 mean&sd**) ratings ( $\rho = -0.57, p = 0.02$ ). Praise - a form of "verbal persuasion", has been identified by [2] as one way of increasing self-efficacy. We also found that the more similar a dyad was in their pattern of timings of praise (**Diff+Time**), the higher were the total (**R1 total**) attentiveness ( $r = -0.58, p = 0.02$ ) and coordination ( $r = -0.52, p = 0.04$ ) ratings. Surprisingly, there was no significant correlation between similarity in usage of the strategy of praise and learning.

Next, for strangers, we found that greater similarity in number of praise usage by individuals in a dyad (**Diff**) was positively associated with self-efficacy (**R2 mean&sd**) ratings ( $\rho = -0.81, p = 0.04$ ), while greater similarity in the pattern of timings of praise (**Diff+Time**) was associated with higher ratings (**R2 mean&sd**) for positivity ( $\rho = -0.85, p = 0.03$ ). Since praise also serves to boost face of the interlocutor, the relationship to positivity (that indexes friendliness and warmth) makes sense. There were no significant correlations for friends with respect to difference in praise usage.

Finally, to close the loop, we looked at gender-based differences. While male dyads had just greater similarity in number of praise usage by individuals in a dyad (**Diff**) positively associated with ratings (**R1 total**) for positivity ( $\rho =$

-0.88,  $p = 0.05$ ) and self-efficacy ( $\rho = -0.89, p = 0.04$ ), female dyads had the exchange patterns of praise over time (**Diff+Time**) positively associated with ratings (**R1 total**) for attentiveness ( $r = -0.73, p = 0.01$ ), coordination ( $r = -0.68, p = 0.03$ ) and ratings (**R2 mean&sd**) for self-efficacy ( $r = -0.75, p = 0.01$ ).

## 7. DISCUSSION AND CONCLUSION

To summarize, in this work, we explored how the similarity in aggregated count of conversational strategies by partners, as well its temporal variations, relate to rapport and learning. We discovered significant effects that explain importance of gender and relationship status in deciphering conversational strategy exchange patterns, as well as a different set of rules governing social exchanges within these settings. Taken with our work on interpersonal influence and convergence of low-level linguistic features [32], the results in this paper signal that the phenomenon of behavioral congruence operates as well at the higher-level of conversational strategies and is a positive predictor of social and cognitive processes in peer tutoring.

In essence, we learned the following about the phenomenon of conversational strategy congruence - unexpectedly, friends differ from one another more than do strangers with respect to the number of self-disclosures and references to shared experience; that is, interpersonal synchrony in usage of these two kinds of conversational strategies is higher in strangers. Much as we found in [7], for friends coordination does not always mean similarity; perhaps because of their familiarity, they tend to break the social rules of conversation. This ties in to our finding that task periods (where the dyad is working on math) are particularly demonstrative of the difference between friends and strangers. Friends seem to care little for the rules; they spend more time self-disclosing than engaging in tutoring!

By demonstrating associations between exchange patterns of conversational strategies and social (rapport) and cognitive (learning) phenomena, our work has provided stronger support for our dynamic time warping approach capturing an actual representation of the social dynamics of the conversation, and not just random fluctuation. In essence, this methodology has allowed us to learn that for friends, similarity in the timing of conversational strategy usage (self-disclosure and reference to shared experience) is positively correlated with rapport, but negatively correlated with learning.

In order to understand this counter-intuitive result, we must first acknowledge that every peer tutoring session has a relational goal (manage-relationship) as well as a transactional or instrumental goal (tutor-math). Interactional goals such as these form one of the bases of rapport [33]. However, we believe that, given limited time, it is difficult for interlocutors to simultaneously achieve both goals. In addition, extreme attention paid to relational goals (via many exchanges of self-disclosure) can lead to reduced performance in the task at hand, in turn affecting future of rapport. Note that while similarity in timings of self-disclosure is negatively associated with learning overall, the similarity in number of self-disclosures is positively correlated with learning - although only for female-friend dyads, who are the ones who engage in the most self-disclosure (as prior literature has also described). In addition, more similarity in the number of reference to shared experience done by partners is associ-

ated with higher learning in female-friend dyads and lower learning in male-friend dyads.

Finally, more similarity in praise usage is related to more self-efficacy, while it is not significantly related to learning. Receiving praise has been identified by [2] as one way of increasing self-efficacy, and therefore we might expect that greater similarity in praise done by partners in the interaction might be associated with similar levels of mutual motivation, enhanced engagement and therefore learning. However, as [14] points out, the effect of a human tutor's praise on the tutee is actually quite complex. For instance, general praise may be received as insincere. And, in fact, the majority of praise utterances in our dataset were not very specific or directed at the tutee's performance or effort (for the most part tutors said "good job"). Overall, this result also echoes results from other studies such as [18], which emphasize that general feedback may contribute to motivation and associated states, but specific feedback matters more for learning. On the other hand, we did find that, not surprisingly, higher similarity in patterns of praise usage by the partners was positively associated with rapport.

## 8. FUTURE WORK

In future work, we will investigate finer-grained patterns of conversational strategies (sub-dimensions of self-disclosure that we coded for) between friends and strangers in both task and social periods. As these social exchanges include both verbal and nonverbal expressions, we have begun extending our framework to examine nonverbal immediacy as well. In addition, while in this work we have only looked at conversational strategy exchanges in isolation, in future an interesting research direction is to investigate exchange pattern across the different conversational strategies in order to represent a more complete picture of conversational strategy congruence. To further infer student conversational strategies that lead to increased rapport and higher learning outcomes in our reciprocal peer tutoring scenario, we will examine these sequences of social and task dialog moves over time, making note of the conditions under which they get used, how they tend to cluster etc.

Ultimately, we intend to dive into content level of the utterance and leverage quantitative methods to automatically analyze topics and opinion exchanges across different conversational strategies. Since dyads participating in reciprocal peer tutoring require a pattern of cognitive similarity that enables the tutor and tutee to anticipate one another's needs and actions and synchronize their work in a way that is synergistic toward meeting the dyad's ultimate goals, we have also started looking at the conversational text as a way of extracting, representing and analyzing such shared cognition via shared mental models. We will investigate the extent to which the mechanism of sharedness of mental models influences interpersonal processes as well as helps differentiate low and high performing dyads in tutoring. However, already the work presented here represents first steps toward a computational model that we can integrate into a virtual peer capable of using both social and task conversational strategies to increase learning gains.

## 9. ACKNOWLEDGMENTS

This work was supported by a graduate fellowship from the R.K. Mellon Foundation and a Focused Award from Google on improving adaptive learning technologies.

## 10. REFERENCES

- [1] I. Altman. Reciprocity of interpersonal exchange. *Journal for the Theory of Social Behaviour*, 3(2):249–261, 1973.
- [2] A. Bandura. *Self-efficacy: The exercise of control*. Macmillan, 1997.
- [3] D. E. Berlyne. Arousal and reinforcement. In *Nebraska symposium on motivation*. University of Nebraska Press, 1967.
- [4] F. J. Bernieri and J. S. Gillis. Judging rapport: Employing brunswik’s lens model to study interpersonal sensitivity. *Interpersonal sensitivity: Theory and measurement*, pages 67–88, 2001.
- [5] F. J. Bernieri and R. Rosenthal. Interpersonal coordination: Behavior matching and interactional synchrony. *Fundamentals of nonverbal behavior*, 401, 1991.
- [6] W. Cade, N. Dowell, A. Graesser, Y. Tausczik, and J. Pennebaker. Modeling student socioaffective responses to group interactions in a collaborative online chat environment. In *EDM*, 2014.
- [7] J. Cassell, A. J. Gill, and P. A. Tepper. Coordination in conversation and rapport. In *Proceedings of the workshop on Embodied Language Processing*, pages 41–50. ACL, 2007.
- [8] H. H. Clark. *Using language*. Cambridge university press, 1996.
- [9] J. Coupland, N. Coupland, and H. Giles. Accommodation theory. communication, context and consequences. *Contexts of Accommodation. Cambridge & Paris: Cambridge University Press & Éditions de la maison des sciences de lâŽhomme*, pages 1–68, 1991.
- [10] P. D. Ellis. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press, 2010.
- [11] R. M. Emerson. Social exchange theory. *Annual Review of Sociology*, 2:pp. 335–362, 1976.
- [12] T. Giorgino. Computing and visualizing dynamic time warping alignments in r: the dtw package. *Journal of statistical Software*, 31(7):1–24, 2009.
- [13] W. W. Hartup. Relationships and their significance in cognitive development. *Social relationships and cognitive development*, pages 66–82, 1985.
- [14] J. Henderlong and M. R. Lepper. The effects of praise on children’s intrinsic motivation: a review and synthesis. *Psychological bulletin*, 128(5):774, 2002.
- [15] C. T. Hill and D. E. Stull. Disclosure reciprocity: Conceptual and measurement issues. *Social Psychology Quarterly*, 45(4):pp. 238–244, 1982.
- [16] G. C. Homans. Social behavior as exchange. *American Journal of Sociology*, 63(6):pp. 597–606, 1958.
- [17] A. N. Ioinson and C. B. Paine. Self-disclosure, privacy and the internet. *The Oxford handbook of Internet psychology*, page 2374252, 2007.
- [18] G. T. Jackson and A. C. Graesser. Content matters: An investigation of feedback categories within an its. *Frontiers in Artificial Intelligence and Applications*, 158:127, 2007.
- [19] J. B. Kruskal and M. Liberman. The symmetric time-warping problem: from continuous to discrete. *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*, pages 125–161, 1983.
- [20] D. Laurillard. *Rethinking university teaching: A conversational framework for the effective use of learning technologies*. Routledge, 2013.
- [21] M. Lewis, G. Young, J. Brooks, and L. Michalson. The beginning of friendship. *Friendship and peer relations*. New York: Wiley, 197(5), 1975.
- [22] D. M. Morrison and V. Rus. Moves, tactics, strategies, and metastrategies: Defining the nature of human pedagogical interaction. *Design Recommendations for Intelligent Tutoring Systems*.
- [23] T. L. Morton. Intimacy and reciprocity of exchange: A comparison of spouses and strangers. *Journal of Personality and Social Psychology*, 36(1):72, 1978.
- [24] A. Ogan, S. Finkelstein, E. Walker, R. Carlson, and J. Cassell. Rudeness and rapport: Insults and learning gains in peer tutoring. In *Intelligent Tutoring Systems*, pages 11–21. Springer, 2012.
- [25] S. Petronio. *Boundaries of privacy: Dialectics of disclosure*. Suny Press, 2012.
- [26] L. Rabiner and B.-H. Juang. *Fundamentals of speech recognition*. 1993.
- [27] L. B. Rosenfeld and W. L. Kendrick. Choosing to be open: An empirical investigation of subjective reasons for self-disclosing. *Western Journal of Speech Communication*, 48(4):326–343, 1984.
- [28] Z. Rubin and S. Shenker. Friendship, proximity, and self-disclosure. *Journal of Personality*, 46(1):1–22, 1978.
- [29] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43–49, 1978.
- [30] A. M. Sharpley, J. W. Irvine, and C. F. Sharpley. An examination of the effectiveness of a cross-age tutoring program in mathematics for elementary school children. *American Educational Research Journal*, 20(1):103–111, 1983.
- [31] T. Sinha and J. Cassell. Fine-grained analyses of interpersonal processes and their effect on learning. In *Artificial Intelligence in Education*, pages 781–785. Springer, 2015.
- [32] T. Sinha and J. Cassell. We click, we align, we learn: Impact of influence and convergence processes on student learning and rapport building. In *Proceedings of the 2015 Workshop on Modeling Interpersonal Synchrony, 17th ACM International Conference on Multimodal Interaction*. ACM, 2015.
- [33] H. Spencer-Oatey. Face,(im) politeness and rapport. 2008.
- [34] D. A. Taylor and I. Altman. *Communication in interpersonal relationships: Social penetration processes*. 1987.
- [35] L. Vygotsky, E. Hanfmann, and G. Vakar. *Thought and language*. MIT press, 2012.
- [36] R. Zhao, A. Papangelis, and J. Cassell. Towards a dyadic computational model of rapport management for human-virtual agent interaction. In *Intelligent Virtual Agents*, pages 514–527. Springer, 2014.