

# Improving Teamwork Using Real-Time Language Feedback

Yla R. Tausczik

Human-Computer Interaction Institute  
Carnegie Mellon University  
ylataus@cs.cmu.edu

James W. Pennebaker

Department of Psychology  
The University of Texas at Austin  
pennebaker@mail.utexas.edu

## ABSTRACT

We develop and evaluate a real-time language feedback system that monitors the communication patterns among students in a discussion group and provides real-time instructions to shape the way the group works together. As an initial step, we determine which group processes are related to better outcomes. We then experimentally test the efficacy of providing real-time instructions which target two of these group processes. The feedback system was successfully able to shape the way groups worked together. However, only appropriate feedback given to groups that were not working well together from the start was able to improve group performance.

## Author Keywords

Feedback; linguistic analysis; CSCW; teamwork; CMC.

## ACM Classification Keywords

H.5.3 Information Interfaces and Presentation: Group and Organizational Interfaces: CSCW

## INTRODUCTION

Using social visualizations to improve group dynamics is a powerful new approach to supporting teamwork. Social visualizations monitor either non-verbal or verbal communication and make subtle aspects of the communication explicit by displaying them in real-time (e.g. [5, 19, 24]). The guiding principle of many of these implementations of social visualizations is that by drawing attention to aspects of the group dynamics, group members will become more aware, and as result will change the way they are communicating in beneficial ways. For example, DiMicco and colleagues [5] designed Second Messenger, a tool that displayed information about the inequalities in participation among group members. In response to their social visualizations, because of greater awareness, over-participants talked less.

Social visualizations provide a good instrument for supporting teamwork in a few ways. By tracking communication patterns, they are able to measure important aspects of group dynamics. Good communication is key to effective teamwork [14]. By providing real-time information on the group

dynamics, they provide a mechanism to shape teamwork. Through awareness of the current group dynamics individuals can make continuous adjustments to their behavior in the group.

Leshed and colleagues' GroupMeter, is an especially good example of a social visualization tool [18, 19]. GroupMeter tracks the words people write to each other over chat during the group task. It displays in real-time word use, such as level of agreement, that is relevant to the group's dynamics. Focusing on the words people use to communicate is an especially good source of information on group dynamics. Groups that work well together typically exchange more knowledge and establish good social relationships, which is reflected in the way that they use words [21, 28]. GroupMeter is a special type of social visualization which we will refer to as a **real-time language feedback system**, because it provides group dynamic information based on the words people use during a group interaction.

Despite the tremendous potential of social visualizations to change how groups work together, their advancement has been impeded in a few ways. First, these implementations of social visualizations strive to make members more aware of their group dynamics, yet these systems have focused on very limited dimensions of the group dynamics. Further, the few dimensions that have been displayed, such as balanced participation, are not necessarily the most important group processes. DiMicco and colleagues [5] successfully increased the balance of participation in their groups through social visualizations, however unequal participation may actually be better for producing high quality work [15, 16].

Second, these implementations of social visualizations have relied on individuals to figure out on their own how to change their communication; this is a daunting task. Individuals are already cognitively overloaded during teamwork. Further, individuals may not respond to greater awareness by changing their communication in beneficial ways. For example, Leshed and colleagues [19] found that in response to displaying agreement information group members agreed with each other more. However, more agreement led to lower quality work because group members passively agreed with each other instead of making substantial contributions to the group.

In this paper we tackle these two challenges to advance the design of real-time language feedback systems. Shaping group dynamics relies on understanding the basic question: *Why do some groups of people work well together while others do not?* Despite the substantial amount of research on groups, there is surprising little consistent evidence for which

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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

group processes promote good group outcomes [21, 20]. Intelligent Tutoring Systems, which serve a similar purpose for student learning by providing real-time feedback to students to guide their learning, have benefited from sophisticated psychological models of student learning [1, 8]. Without good models of learning they would not have been able to realize substantial gains over traditional classroom teaching. Better models of group dynamics will provide a basis for giving useful real-time group dynamic information. In this paper, we evaluate the importance of four group processes at promoting better group outcomes.

Group interventions have a long history. One technique in experimental psychology has been to use specifically crafted instructions to change group dynamics. For example, Nemeth and colleagues [22] proposed that conflict and dissension in groups was important for promoting group creativity. To test the importance of dissension as a group process they included a direction in the standard instructions for group members to disagree with each other during the group task (i.e. “most studies suggest that you *should* debate and even criticize each other’s ideas”). Specific instructions like this one are effective in changing the way that groups work together in a targeted way. In the design of our feedback system, we make use of specific instructions to change group dynamics.

In this paper we describe the development and evaluation of a real-time language feedback system that improves upon Leshed and colleagues’ GroupMeter [18, 19]. Like GroupMeter our feedback system monitors the words people write to each other over chat during the group task. However, instead of displaying the group dynamic information visually the feedback system gives instructions that target specifically selected group processes. Our feedback system addresses the problem of monitoring irrelevant group dynamics by targeting two group processes that we show to be related to better group outcomes. It also addresses the problem of relying of individuals to change their communication in beneficial ways by giving groups specific instructions that target these relevant group processes.

### “Good” Group Dynamics

The scientific study of groups is messy. There are no standard rules governing the behavior of groups and a web of theories, often with little empirical evidence, have been published that try to explain different aspects of group dynamics [20]. Very broadly, historical small group research has been categorized into three schools of thought [21]. Named after the universities of the main researchers in the 1950s-1980s, each approach adopted a different perspective on the most important aspect of groups to study. This research is informed most strongly by the approaches coming out of two of these schools.

The Harvard school focused on group interactions and processes [2, 3]. Bales and colleagues developed the Interaction Process Analysis to describe the ways in which group members were interacting during group discussions. They used this method to characterize groups along three dimensions: relative dominance, friendliness, and task focus. While

these group processes characterize the patterns of interactions in groups well, researchers have been unable to consistently relate these group processes to group outcomes [21]. The Illinois school studied small laboratory groups and focused exclusively on the conditions that increased task performance (e.g. [16]). They found that task performance was dependent on a number of conditions, including group composition, group structure, task type, and communication platform. However, the Illinois school exclusively studied the input conditions (e.g. group composition) under which groups performed well, while ignoring how groups were actually interacting in order to perform well [21]. We combine these two approaches by trying to identify *group processes*, which we define as patterns of interactions among group members, that produce better group outcomes, such as higher group satisfaction and performance.

The most compelling contemporary theory of group processes focuses on distinguishing between social regulation processes, which describe the social interactions between group members, and information processing processes, which describe how groups gather, organize, and process information as a group [21]. We describe a few of the most commonly discussed group processes as they are relevant to this theory.

Social regulation group processes operate by motivating, regulating, and coordinating groups of individuals through social processes. Groups are hypothesized to perform better, especially in the long run, if they are more cohesive and have deeper social relationships. Creating groups that have stronger social relationships is also an end in itself. Effective teams are ones in which individuals enjoy being a part of the team and want to work together again.

There are a few specific group processes that are relevant to social regulation. *Positivity*, the degree to which group members are encouraging, can enhance interpersonal relationships and motivate individuals to work harder [17]. However, positivity can also detract from task effort when it leads to off-topic conversations. Positivity is reflected in way group members communicate to each other; groups that are positive use more positive emotion words and make more statements agreeing with each other. Empirically when positivity is measured by analyzing communication patterns the outcomes of positivity are mixed. In one study, groups engaged in a flight simulation that used more positive emotion words and agreed with each other more during the simulation performed better at the task [6]. In a different study, successful coalitions of business students agreed with each other more during the negotiation task [12]. However, Leshed and colleagues [18] found that agreement was related to being passive, and led to lower group satisfaction and performance.

*Engagement*, the degree to which group members are paying attention and connecting with each other, can enhance group cohesion. If individuals are engaged with each other they are more likely to stay motivated and enjoy the task. Engagement is reflected in the degree to which group members tend to converge in the way they talk [13]. In general, groups tend to converge in the way that they talk, this is called Language Style Matching (LSM) [23]. When groups are more engaged

their language style converges more [13]. In one study of groups solving puzzles, groups that were more engaged felt greater group cohesion, although they did not perform better on the task [7]. However, in a naturalistic study of real hostage negotiations, higher LSM between hostage takers and police negotiators, which is suggestive of more engagement, was associated with more successful outcomes [29]. Empirical evidence suggests that engagement may be associated with greater group satisfaction and performance, although the results are inconclusive.

One of the most studied aspects of group dynamics, as well as commonly represented in social visualizations, is the balance of participation among group members. *Equal participation*, the degree to which all group members are equally involved, can enable everyone in the group to be active and enjoy the experience [10, 16]. However, equal participation in groups has been associated with both higher and lower quality work [16, 15, 9].

Information processing group processes focus on the cognitive aspects rather than the social aspects of group interactions. Groups can be thought of as information processing systems; through discussion, groups gather, organize, and process information [11, 21]. Limitations and advances in the group processes through which group members share and process information lead to changes in group performance. For example, the degree of *information exchange*, that is the degree to which more information and more relevant information is shared can effect group performance [27]. Despite the fact that all the relevant information to make a decision may be known in a group, group members may not share this information [4]. The degree of information exchange is reflected in group communication in part by the quantity of on task communication. Empirical evidence suggests, information exchange, as measured by quantity of communication and on task communication is related to greater performance. Groups engaged in a flight simulation that communicated more performed better on the task [26]. Groups of students solving problems together that made more task focused statements performed better [14].

We focus on these three social regulation group processes—positivity, engagement, equal participation—and this one information processing group process—information exchange. In theory all four group processes could substantially benefit groups, however there is insufficient empirical evidence examining the impact of these group processes on group outcomes.

## THE CURRENT STUDIES

A real-time language feedback system was designed to improve teamwork by targeting specific group processes. This paper had two objectives. The first was to establish which group processes should be targeted. The second was to experimentally evaluate whether our real-time language feedback system could improve teamwork. To address these two goals three studies were conducted: Pilot Study, Study 1, and Study 2. In each study, students took part in small group discussions over instant message to try to better understand a

few psychology theories. Students were tested before and after the group discussions to gauge learning. Group work was assessed by using a group discussion task because it is naturalistic, groups are often formed to learn from each other, and it allowed access to a large sample of groups.

### ***Research Objective 1: To identify group processes related to higher quality teamwork***

There are many theories about what constitutes good group dynamics. The focus was on four commonly discussed group processes: positivity, engagement, equal participation, and information exchange. These group processes can be measured linguistically and there is some theoretical and empirical evidence to suggest that they encourage better teamwork. However, the relationship between any group processes and effective teamwork is not well established. To address Research Objective 1 we determined which if any of these four group processes, as measured by language use, was correlated with better group satisfaction and performance across the three studies. The Study 1 was specifically designed to test Research Objective 1 (see Table 1), however we analyzed patterns across all three studies to determine the most robust effects.

### ***Research Objective 2: To capitalize on observed differences in group processes to improve teamwork using a real-time language feedback system***

By addressing Research Objective 1, it was possible to establish which if any of the four group processes was related to better teamwork. Study 2 was designed to experimentally evaluate whether we could improve teamwork by providing real-time language feedback that targeted these group processes by providing individualized instructions. Two of the group processes, engagement and information exchange, were selected because they were the most strongly related to better teamwork. In Study 2 a controlled experiment with 4 conditions (no feedback, engagement feedback only, information feedback only, both feedback types) was conducted to evaluate the effectiveness of the feedback.

## METHOD

A Pilot Study, Study 1, and Study 2 were conducted to develop and evaluate the real-time language feedback system.

### **Pilot Study**

*Participants* – In the Pilot Study, 494 students from two introductory psychology classes taught in Fall of 2010 participated in the group exercise as part of the class requirements. Of these participants 59 percent were female, they were aged 17 to 34 ( $M = 19$ ). Participants were randomly assigned to groups of 2-5 students ( $M = 2.9$ ) based on when they logged in to complete the assignment.

*Procedure and Measures* – Students logged into an education platform, Texas Online World of Educational Research (TOWER), at specified times to complete the group interaction task [25]. The education platform was an online course center where students took surveys, took quizzes, completed writing assignments, and participated in group chat. Prior to logging into the system students were instructed that in order

	Design	Feedback System	# Participants	# Groups	Satisfaction	Pre-quiz	Post-quiz
Pilot Study	Correlational	Yes	494	173	3.11 (0.98)	76.4% (19.5%)	83.9% (16.9%)
Study 1	Correlational	No	851	183	2.56 (0.80)	35.9% (14.7%)	46.0% (17.5%)
Study 2	Experimental	Yes	816 (621)	210 (160)	3.09 (1.00)	70.2% (16.1%)	72.1% (14.7%)

**Table 1. Overview of the three group discussion studies. Means (standard deviations) of group satisfaction as well as performance on quizzes before and after group discussion are given. For Study 2, due to technical errors, only 621 individuals in 160 groups were included in analyses.**

to complete the assignment they would need to read supplementary material on a few psychological theories (e.g. 10 pages of the textbook).

Once students logged into the educational platform they were directed to the first quiz. The quiz was 10 multiple-choice questions; it tested students' knowledge of the reading material. After completing the quiz they were randomly matched with other students currently waiting. When there were at least 2 students and no more than 5 students, individuals were directed to an instant messaging platform that was built into the educational platform. The group chat began as soon as someone typed the first message and lasted for 20 minutes.

During the chat groups received feedback messages this will be explained in more detail below. After 20 minutes, the chat window closed automatically. Students took a second 10 multiple-choice question quiz and rated their group members using the Interaction Rating Questionnaire (IRQ) modified to be appropriate for a group [23].

**Group Process Variables** – Four group processes were predicted to influence group satisfaction and performance based on the previous literature. Language markers were constructed for each of the four variables based on face valid measures and the previous literature. Language was analyzed using Language Inquiry and Word Count (LIWC) at the level of the group and individual as appropriate. Analyses control for group size, which can create statistical artifacts in the language markers.

**Positivity** Positive emotion words (*nice, good, great*) and asents (*yes, agree*) were converted to z-scores and summed as a measure of positivity [28].

**Engagement** The degree of language convergence, a proxy of engagement, was measured by calculating Language Style Matching (LSM) among pairs in the group (see [13]) and taking the average among all pairs in the group (see [7]). Pairwise LSM is the average difference of nine function word categories measured by LIWC.

**Equal participation** The Gini coefficient of individual level word count in the group was taken as a measure of equal participation. It was reverse scored, so that higher numbers show more equal participation.

**Information exchange** Word count and first-person singular pronouns were converted to z-scores, Word count minus first person singular was taken as a measure of information exchange. Word count is a commonly used crude measure of the total information exchanged (e.g. [26]). First-person singular pronouns suggest self-focus rather than topic focus [28], thus we refined the crude measure by penalized for self-focus rather than topic focus.

Hand coding confirmed that the linguistic markers were adequate proxies for the group processes. Thirty Mechanical Turkers rated a random sample of 30, 5-minute chat samples taken from the middle of the group discussions. Each sample was rated by 5 coders for the group process dimensions. Engagement, equal participant, and information exchange were coded with good inter-rater agreement (ICC = 0.68-0.72) and there was high agreement between the group process variables and their corresponding linguistic markers ( $r(28) = 0.49-0.68, p < 0.01$ ). Positivity was the most difficult to code; there was only slight agreement between the coders (ICC = 0.49) and rated positivity was not related to the linguistic markers of positivity ( $r(28) = 0.06, p = 0.74$ ). Coders may not have agreed on what being positive meant. The proposed linguistic markers of positivity were used nonetheless as a face valid measure of the number of times that group members agreed with each other and expressed positive emotion.

### A Real-time Language Feedback System

A feedback system was designed to shape group dynamics by giving real-time language feedback. Groups were given feedback messages instructing them how to improve their group dynamics. Instructions were created to target a specific group processes. For example, engagement feedback instructed participants to pay more attention to what the other group members were saying (e.g. *"Your group is working OK but could be improved. Be sure and pay attention to what others are saying."*) while information exchange feedback instructed participants to stay on topic and share more material from the readings (e.g. *"You are discussing the relevant material OK, you can improve. Be sure and share information from the readings."*).

Feedback messages provided individualized instructions based on the current group dynamics. Group dynamics were assessed every few minutes; each time the targeted group process (e.g. information exchange) was rated on a 4-point scale from "Poor" to "Good" based on the language markers of that group process. Groups received a different instruction based on their current level of the group process. For example, if a group scored "Poor" on a group process they would be told that they were doing poorly and needed to improve (e.g. *"You are not discussing the relevant material. Focus on the material and share more information from the readings"*). If a group scored "Good" on a group process then they would receive a message telling them they were doing well (e.g. *"You are doing a good job discussing the relevant material. Everyone seems to be paying attention to the topic."*). There were four instructions corresponding to the four ratings for each group process. For the first deployment of the feedback we chose to evaluate a group as a whole, so that group

members would have consistent instructions. The instructions were presented to the group using javascript pop-up messages above the chat box and logged for later analysis.

Current group dynamics were assessed by measuring language markers in nearly real time. Group chat messages from the last time interval (e.g. last 2.5 minutes) were collected and processed through LIWC running on the server. A group process for that time interval was rated based on benchmarks determined from previous class group discussions. For example, a group was rated as having “Poor” information exchange if their language markers of information exchange ranked in the lower quartile for language markers of information exchange in previous class discussions. This system for rating group processes worked for some group processes and not others, which we will return to later.

In the Pilot Study all groups were given engagement and equal participation feedback.

*Group Outcome Variables* – Group satisfaction and performance are two outcomes of teamwork, which are only slightly correlated ( $r(166) = 0.18$ ).

**Group satisfaction** Group satisfaction was operationalized as the average of the group members’ rated satisfaction with the group. Each group member rated how well they enjoyed the interaction, how well they ‘clicked’ with their group, and whether they would work with their group again on 5-point Likert scales [23]. The items were averaged to create an individual level measure of group satisfaction ( $\alpha = 0.86$ ). Individuals rated the group alone, however group members’ ratings of the group were related to each other justifying grouping individual level satisfaction scores at the group level ( $F(167,312) = 1.43, p = 0.02$ ).

**Group performance** Group performance was operationalized as the average group members’ improvement on the quiz after the group discussion compared to before the discussion. On average students scored better on the quiz after the group discussion than before. Individual level improvement was calculated by taking the individual’s score on the post-interaction quiz controlling for the score on the pre-interaction quiz. Individuals took the quizzes alone, however individuals improvement was related to which group they had been apart of justifying grouping individual level improvement scores at the group level ( $F(167,312) = 1.84, p < 0.001$ ).

### Study 1

Study 1 was conducted to identify which group processes were related to the group outcomes while no group processes were being manipulated. Thus, Study 1 was nearly identical to the Pilot Study, with the exception that there was no real-time language feedback. 851 students from two introductory psychology classes taught in Fall of 2011 participated in the Study 1 as part of the class requirements. Of these participants 62 percent were female, they were aged 17 to 48 ( $M = 19$ ). For Study 1, due to a procedural error group satisfaction was assessed a few days later rather than immediately after the group task.

### Study 2

Study 2 was conducted to experimentally evaluate the real-time language feedback system. Thus, Study 2 was also nearly identical to the Pilot Study, except students were randomly assigned to 1 of 4 feedback conditions: no feedback (control), engagement feedback, information exchange feedback, and both feedback types. 816 students from the same classes as Study 1 participated as part of the class requirements. Most of the students were the same in Study 1 and Study 2, however group assignments were different. By Study 2 students had familiarity with group chat, but had no experience with the feedback system.

Due to technical errors with the feedback system in Study 2, some individuals did not receive the appropriate number or type of feedback messages. 13 groups had to be removed because the group received conflicting feedback messages and/or messages at the wrong times. 37 groups were excluded because at least one member of the group received fewer than 50% of the feedback messages they were supposed to have received. This left 621 individuals in 160 groups in Study 2.

### IDENTIFICATION OF GROUP PROCESSES RELATED TO HIGHER QUALITY TEAMWORK

All four of the group processes were related to either group satisfaction or performance. There was variability in the relationship between the group processes and better teamwork across the studies, therefore to test for robustness the correlations for all three studies are presented.

#### Group Satisfaction

All four group processes were significantly correlated with group satisfaction (see Table 2). However, engagement was the most consistent predictor of group satisfaction. There was some variability in which group processes were related to group satisfaction in particular, Study 1 was different from the Pilot Study and Study 2. One explanation is that Study 1 had more difficult quizzes making the task harder than the other two studies.

#### Group Performance

The relationship between the four group processes and group performance was more complex. Information exchange was the most consistent predictor of group performance (see Table 3). Information exchange was related to better group performance in the Pilot Study and Study 1, however it did not predict group performance in Study 2. The other group processes were occasionally related to group performance. Engagement and equal participation were related to better performance in the Pilot Study and positivity was related to performance in Study 2.

#### Conclusion

The results show that there are a few identifiable ways in which groups work together that can predict how well the groups are working. In particular, the results suggest that engagement and information exchange are two important group processes that would be good targets of our real-time language feedback system. The results also show that there is

	Pilot Study df=165	Study 1 df=179	Study 2 df=157	Overall df=167
Positivity	0.32***	0.05	0.21**	0.19*
Engagement	0.23**	0.27***	0.25**	0.25**
Equal Participation	0.26***	0.08	0.30***	0.21**
Information Exchange	0.06	0.27***	0.16*	0.16*

**Table 2. Linguistic correlates of group satisfaction** (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

	Pilot Study df=165	Study 1 df=179	Study 2 df=157	Overall df=167
Positivity	0.01	-0.05	0.18*	0.04
Engagement	0.20*	0.12	0.07	0.13†
Equal Participation	0.13†	-0.02	0.13	0.08
Information Exchange	0.14†	0.24***	0.03	0.14†

**Table 3. Linguistic correlates of group performance** (†  $p < 0.10$  \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

variability in whether the group processes affect group outcomes. Variability creates a challenge to a designing a universal system to shape group work. Future work should examine the conditions under which specific group processes are important. Even small differences between the three studies may have influenced which group processes were important. The selected group process variables were better predictors of group satisfaction than group performance. Future work, should focus on other information processing dimensions of group processes rather than social regulation dimensions of group processes in identify stronger predictors of group performance.

### CAPITALIZING ON OBSERVED DIFFERENCES IN GROUP PROCESSES TO IMPROVE TEAMWORK

In Study 2 the feedback system was evaluated experimentally. Feedback was expected to change the way groups worked together by changing the group dynamics. It was predicted that engagement feedback would increase engagement and information exchange feedback would increase information exchange.

Feedback was also expected to improve teamwork by creating better group dynamics. Engagement feedback was predicted to increase group satisfaction, because of positive correlations between the two in the Pilot Study and Study 1. While information exchange feedback was predicted to increase group performance, because of positive correlations between the two in the Pilot Study and Study 1. We also speculated that engagement feedback might increase group performance and information exchange feedback might increase group satisfaction because they were sometimes correlated. Finally, we speculated that engagement and information exchange feedback might have synergistic effects by getting students to engage with each other and the material at the same time.

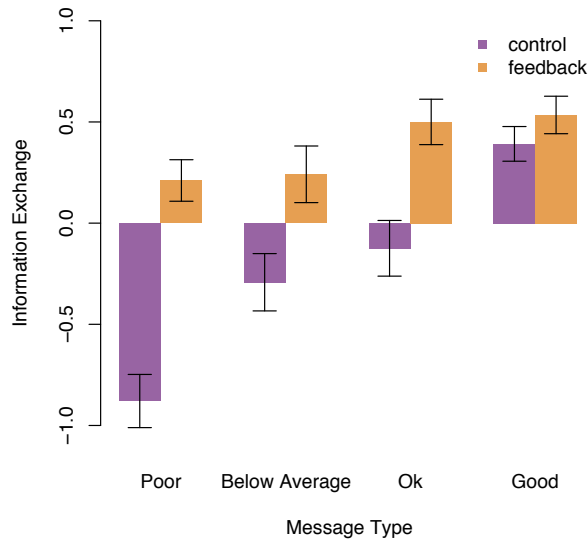
### Real-time Language Feedback System

Providing real time language feedback to shape group interactions is a simple concept, yet it requires a large number of technical specifications to work seamlessly. The first full deployment of the feedback system in Study 2 was a good initial trial. Students rated the usefulness of the feedback messages and provided comments on the system. On average they found the feedback to be slightly helpful ( $M = 3.13$ ,  $SD = 1.10$ ) along a 5-point scale where 5 corresponded to “very helpful” and 1 corresponded to “it made things worse”. There were significant differences based on the feedback type ( $F(2, 455) = 10.22$ ,  $p < 0.001$ ). Those in the information exchange condition found it to be the most helpful ( $M = 3.37$ ,  $SE = 0.08$ ); followed by those in the engagement feedback condition ( $M = 3.16$ ,  $SE = 0.09$ ); and those in the both feedback types condition found it to be slightly unhelpful ( $M = 2.80$ ,  $SE = 0.09$ ).

Those students who liked the feedback made comments about its usefulness such as “*Our group got back on task when the reminder popped up*”. Those that disliked the system provided valid critiques that will need to be addressed in future feedback systems. Many students commented that the messages were distracting: “*I think the feedback that we were getting actually distracted us a little*”. One student objected because they did not believe the feedback they were being given was accurate. He said “*We could have worked better if we actually had more time to finish our last questions and if the pop up window didn’t keep claiming we were off subject when our conversation was clearly relevant to the material*”.

There were a few technical issues that reduced the efficacy of the feedback system. The high computational demands of processing language use in real-time and providing feedback for hundreds of groups simultaneously proved to be a challenge and not all feedback messages were given. Missing feedback messages were not associated with feedback condition or feedback level. Also, benchmarks for the feedback messages in Study 2 were established based on the language use in Study 1. These benchmarks worked well for information exchange feedback. However, they did not work well for engagement feedback. Almost all the engagement feedback said the they were doing poorly. This had two consequences. First, groups received very negative messages for engagement feedback; feedback that was basically too harsh. Second, because the benchmarks were too harsh it was difficult to do better over time and receive messages in the “Ok” or “Good” range. This may have been discouraging for groups trying to use the feedback to improve.

Those students that received engagement feedback were more likely to comment about the feedback during the chat. Many of the comments were negative (e.g. “*yo this popup sucks*”, “*death to the green box*”). By calculating the percent of words exclusively making reference to the feedback system (e.g. ‘popup’) the degree to which groups discussed the feedback can be assessed. Discussion of the feedback system varied by condition ( $\chi(2) = 8.86$ ,  $p = 0.01$ ). 67% of groups that received engagement feedback, alone or with information exchange feedback, commented on the feedback system,



**Figure 1. Impact of information exchange feedback messages on information exchange.** X-axis displays the level of information exchange in the group in the previous 2.5 minute interval. Y-axis displays the level of information exchange in the group in the current 2.5 minute interval. Those in the feedback group received a message encouraging information exchange while those in the control did not.

whereas only 39% of groups that received information exchange feedback alone commented on the feedback system during the group discussion.

Overall, the Study 2 provided a good initial test of real-time language feedback system as a group intervention. Once some of the technical challenges from Study 2 are addressed, the feedback system is likely to be even more effective.

### Impact on Group Dynamics

Both feedback types had a significant impact on language use during the group discussion. Changes in language use suggest that the groups changed the way they were working together in response to the recommendations made by the feedback system. Engagement and information exchange feedback both increased the linguistic marker of engagement, LSM ( $F(3, 156) = 3.99, p = 0.009$ ). Groups that received engagement feedback alone ( $M = 0.76, SE = 0.02$ ), information exchange feedback alone ( $M = 0.78, SE = 0.01$ ), or both feedback types ( $M = 0.80, SE = 0.01$ ) all had higher LSM scores compared to groups that received no feedback ( $M = 0.72, SE = 0.02$ ). Pairwise t-tests showed that compared to the control, the feedback conditions had at least marginally significantly higher scores ( $t_{Eng}(156) = 1.6, p = 0.10$ ;  $t_{Inf}(156) = 2.8, p = 0.006$ ;  $t_{Bth}(156) = 3.1, p = 0.002$ ).

Information exchange feedback had a significant impact on the linguistic markers of information exchange ( $F(3, 156) = 6.14, p < 0.001$ ). Groups that received information exchange feedback alone ( $M = 0.46, SE = 0.18$ ) and those groups that received both feedback types ( $M = 0.34, SE = 0.18$ ) had higher values for the linguistic measure of information exchange than groups that received no feedback ( $M = -0.64, SE = 0.23$ ) or groups that received engagement related feedback alone ( $M = -0.22, SE = 0.23$ ). Pairwise t-tests

showed that these mean differences between the information exchange feedback conditions and the control were significant ( $t_{Inf}(156) = 3.8, p < 0.001$ ;  $t_{Bth}(156) = 3.3, p < 0.001$ ).

A multi-level model analyzing group dynamics in 2.5 minute time intervals was used to test whether some types of groups benefited from information exchange instructions more. The model tested whether the benefit of receiving feedback depended on the groups' previous information exchange level (e.g. "Poor" vs "Good"). There was a significant interaction between previous information exchange level and whether a group received a feedback message ( $t(956) = 2.27, p = 0.02$ ; see Figure 1). Those groups that received feedback messages indicating that their previous level of information exchange was "Poor", "Below average", or "Ok" were able to raise the amount of information exchanged, whereas those not given feedback maintained the same low levels of information exchange. There was not a large impact of giving feedback for those who were categorized as having "Good" levels of information exchange in the previous time interval.

The results show that both feedback types had an impact on the language used. Information exchange and engagement feedback impacted linguistic markers of engagement. Only information exchange feedback impacted linguistic markers of information exchange. Information exchange feedback was most useful for groups with low levels of information exchange.

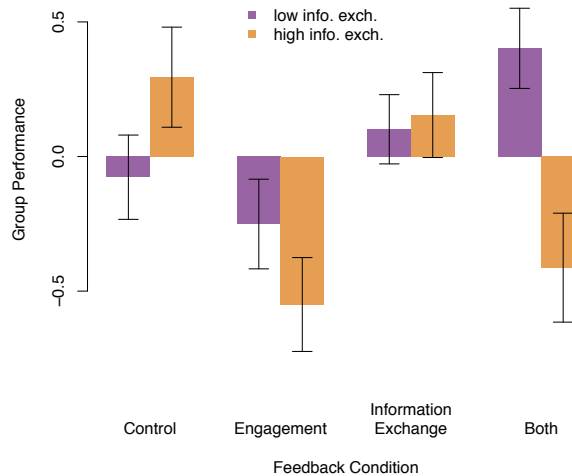
### Impact on Group Satisfaction

Receiving feedback had a significant impact on group satisfaction ( $F(3, 156) = 6.24, p < 0.001$ ). Groups that received information exchange feedback alone were marginally more satisfied ( $M = 3.34, SE = 0.09$ ) and those that received both feedback types were significantly less satisfied ( $M = 2.76, SE = 0.08$ ) compared to those that received no feedback ( $M = 3.14, SE = 0.10$ ) or engagement feedback only ( $M = 3.10, SE = 0.10$ ;  $t_{Eng}(156) = 0.31, p = 0.76$ ;  $t_{Inf}(156) = 1.6, p = 0.12$ ;  $t_{Bth}(156) = 2.7, p = 0.007$ ). This pattern of effects is consistent with ratings of the feedback systems overall, those in the information exchange condition preferred it to those in the engagement condition, and those that received both feedback types enjoyed the experience the least.

### Impact on Group Performance

Receiving feedback had a significant effect on group performance ( $F(3, 156) = 1.69, p = 0.046$ ). Those groups that received engagement feedback only performed significantly worse ( $M = -0.31, SE = 0.14$ ) than those that received no feedback ( $M = 0.05, SE = 0.12$ ), information exchange feedback only ( $M = 0.11, SE = 0.10$ ), or both feedback types ( $M = 0.13, SE = 0.13$ ). Those that received information exchange feedback, either alone or with engagement feedback, performed slightly better than the control, however the mean differences were not significant ( $t_{Eng}(156) = 2.1, p = 0.04$ ;  $t_{Inf}(156) = 0.4, p = 0.71$ ;  $t_{Bth}(156) = 0.4, p = 0.66$ ).

However, feedback had a different effect depending on groups' initial dynamics. Language analyses showed that feedback messages had a stronger effect on groups that had



**Figure 2. Group performance for groups in different feedback conditions. Groups are divided based on their initial levels of information exchange.**

lower levels of information exchange. To better understand how feedback might affect groups with different initial group dynamics, groups were divided based on their level of information exchange in the first 2.5 minutes. Those groups with “Good” levels of information exchange, were considered on task. The groups with “Poor”, “Below average”, or “Ok” levels were considered to be low in information exchange and to have room for improvement. The feedback condition had a significant effect on group performance depending on whether the groups began the discussion on task or if there was room for improvement. The interaction between feedback condition and levels of information exchange in the first 2.5 minutes was significant ( $F(3, 151) = 3.68, p = 0.01$ ; see Figure 2).

Those groups that had high levels of information exchange initially did not benefit from the information exchange feedback and did worse as a result of the engagement feedback. Pairwise t-tests show that the groups with high initial levels of information exchange performed worse when engagement feedback was given alone or with information exchange feedback compared to groups in the control, while those who received information exchange feedback performed the same ( $t_{Eng}(40) = 3.1, p = 0.004$ ;  $t_{Inf}(40) = 0.4, p = 0.59$ ;  $t_{Both}(40) = 2.9, p = 0.006$ ). Getting engagement feedback disrupted groups that were on task and information exchange feedback had no effect.

Those groups that were initially low in sharing information benefited from information exchange and engagement feedback when combined. Pairwise t-tests show that groups given engagement feedback alone or information exchange feedback alone performed no differently than control groups ( $t_{Eng}(112) = 0.8, p = 0.42$ ;  $t_{Inf}(112) = 0.8, p = 0.40$ ;  $t_{Both}(112) = 2.1, p = 0.04$ ). However, those groups given information exchange feedback and engagement feedback together performed better than control groups.

	Prediction	Result
Engagement Feedback	Increase engagement	Marginal
	Increase group satisfaction	No
	Increase (?) group performance	No
Information Exchange Feedback	Increase info. exchange	Yes
	Increase group performance	Some
	increase (?) group satisfaction	Marginal

**Table 4. Summary of Study 2 results.**

## Conclusion

Real-time language feedback shows great potential for shaping group dynamics. Feedback messages targeted at specific group processes were able to change the group dynamics (see Table 4). Engagement feedback led to a marginal increase in engagement. Information exchange feedback led to an increase in information exchange and engagement. Although not all of the changes to the group dynamics were predicted, many of the changes were in the right direction.

Feedback changed the group dynamics, but the changes in group dynamics did not always result in the predicted and desired outcomes. Engagement feedback had no effect of group satisfaction and actually decreased group performance. Engagement feedback failed most likely because it was too harsh. If given correctly engagement feedback might have been able to increase teamwork. However, we did learn a valuable lesson that harsh, negative feedback can be damaging.

Information exchange feedback led to a marginal increase in group satisfaction and had no visible effect on group performance when all groups were examined together. However, when groups were divided based on their initial levels of information exchange, groups that started out with low levels of information exchange benefited from information exchange feedback when combined with engagement feedback, while those with high levels of information exchange did not. This suggests that information exchange feedback may be useful, but only for groups that are not working well together from the start.

We speculated that combining feedback types might have synergistic effects; the results were mixed effects. In general, receiving both feedback types together was likely too cognitively taxing. Students that received both feedback types, liked the feedback the least and were more dissatisfied with their group experience. However, receiving both feedback types provided the only observable improvement in group performance, most likely because of synergistic effects.

## GENERAL DISCUSSION

### Toward a Model of Good Group Dynamics

The use of social visualizations to shape teamwork, has been significantly impeded by a poor understanding of the optimal group dynamics. Previous work in psychology, which social visualizations have relied upon, developed complicated theories of important group processes without being able to relate them to important outcomes [2, 21]. We focused on a



contemporary theory of group processes, which distinguishes between social regulation and information processing processes [21]. We found that social regulation processes, including positivity, engagement, and equal participation, were more consistently related to group satisfaction, and only occasionally related to group performance. Whereas, the degree of information exchange, an information processing process, was the group process most consistently related to group performance. However, even information exchange was only weakly related to group performance. A broader exploration of group processes is needed to identify good predictors of group performance.

The results suggest two fruitful paths toward developing a better model of good group dynamics. The selected group processes were more strongly related to group satisfaction than group performance. In part, because the majority of the selected group processes focused on social regulation. Social regulation is important in establishing smooth interactions, but in a short 20 minute group discussion it may not be as important for group performance. Ultimately, group performance is a result of collecting, organizing, and processing information. Thus, the focus should be on group processes involved in information processing. We selected the degree of information exchange as a simple group process involved in information processing. However, staying on task and sharing information was not always enough to promote understanding in the group discussions. It is clear that group processes that produce more complex information processing need to be identified and studied. One group process worth exploring is the development of roles based on expertise. For example, groups which develop roles in which those with less expertise ask questions and those with more expertise answer the questions may process the information that is shared more deeply. Identification of experts within a group was one technique that allowed decision groups to share more information [4].

Social regulation processes may still play an important role when they interact with information processing processes. We found that the feedback system was able to improve group performance for groups that started out exchanging little information when they received feedback that instructed them to stay engaged and exchange more information. In this case, engagement alone may not improve teamwork, however if the group is focused on the task and sharing information engagement may be able to enhance the value of sharing information. Another example of an interaction between social regulation and information processing may be the role of conflict and dissension in a group. Nemeth and colleagues found some evidence that conflict and dissension could improve creativity in a group [22]. Conflict in a group is likely to improve creativity, not because it solidifies relationships or enhances the emotional experience of group members, but because it coordinates individuals, by pitting them against each other, in a way that gets them to think more deeply.

Shaping group dynamics with real-time language feedback, not only has the benefit of improving group dynamics, it provides a strong experimental paradigm for establishing causal relationships between group processes and outcomes. Histor-

ically many psychologists have focused on inputs to groups (e.g. group composition) rather than group processes in part because it is difficult to experimentally manipulate group dynamics by changing group processes. This work relied on determining if there was a connection between group processes and outcomes using correlations, which is only a first step. Once the kinks in the feedback system are resolved, it will be possible to systematically test for causal relationships between group processes and outcomes.

### **Toward the Design of Real-time Language Feedback**

Real-time language feedback systems show great potential as group support systems to shape group dynamics. We advanced the design of these feedback systems in two ways. First, we used feedback to target two group processes that we thought could bring about better group satisfaction and performance. Second, our feedback system gave individualized instructions targeting these group processes. Individualized instructions help to explicitly direct groups to change their behavior in beneficial ways. The experimental evaluation of the feedback system provided both promising results as well as many failures.

The results of the evaluation provide strong lessons for the future design of real-time language feedback systems. First, a focus should be on using real-time language feedback to shape group dynamics that are relevant to group performance and not group satisfaction. We found that improving group satisfaction is especially challenging using feedback. Satisfaction or dissatisfaction with the feedback system overshadowed ratings of group satisfaction. Thus, changing group dynamics to improve group performance is a better goal. Many social visualizations have focused on aspects of social regulation, such as the amount agreement and equal participation (e.g. [19, 5]). This is misguided since we found that social regulation processes are more closely related to group satisfaction than group performance. Second, feedback should only be given to groups that are not working well together from the start. Good group dynamics are hard to establish. Feedback has the potential to disrupt groups that are doing well from the start, either because it is distracting or slightly inaccurate. Only those groups that are doing badly from the start may benefit from feedback. Third, providing accurate individualized feedback is important, yet challenging. Engagement feedback because it was too negative may have made groups perform worse. Fourth, providing more than one feedback message at a time may be too cognitively overloading. Once these challenges are overcome the feedback system shows great promise for being able to shape group dynamics and improve teamwork.

### **ACKNOWLEDGEMENTS**

Preparation of this manuscript was aided by funding from the Army Research Institute (W91WAW-07-C-0029) and National Science Foundation (NSCC-0904913).

### **REFERENCES**

1. Anderson, J. R., Corbette, A. T., Koedinger, K. R., and Pelletier, R. Cognitive Tutors: Lessons Learned. *Journal of Learning Sciences* 4 (1995), 167–207.

2. Bales, R. F. *Interaction Process Analysis: A method for the study of small groups*. The University of Chicago Press, Chicago, 1950.
3. Bales, R. F., and Strodtbeck, F. L. Phases in group problem-solving. *The Journal of Abnormal and Social Psychology* 46 (1951), 485–495.
4. Diehl, M., and Stroebe, W. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology* 53 (1987), 497–509.
5. DiMicco, J. M., Hollenbach, K. J., Pandolfo, A., and Bender, W. The Impact of Increased Awareness While Face-to-Face. *Human-Computer Interaction* 22 (2007), 47–96.
6. Fischer, U., McDonnell, L., and Orasanu, J. Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation Space and Environmental Medicine* 78 (2007), 86–95.
7. Gonzales, A. L., Hancock, J. T., and Pennebaker, J. W. Language style matching as a predictor of social dynamics in small groups. *Communication Research* 31 (2010), 3–19.
8. Graesser, A. C., Conley, M. W., and Olney, A. Intelligent tutoring systems. In *APA Handbook of Educational Psychology*, S. Graham and K. Harris, Eds., vol. 2579. American Psychological Association, Washington, DC, 2012, 451–473.
9. Hackman, J. R., and Kaplan, R. E. Interventions into group process: An approach to improving the effectiveness of groups. *Decision Sciences* 5 (1974), 459–480.
10. Hare, A. P. *Handbook of Small Group Research*. Free Press, New York, 1962.
11. Hinsz, V. B., Tindale, R. S., and Vollrath, D. A. The emerging conceptualization of groups as information processors. *Psychological Bulletin* 121 (1997), 43–64.
12. Huffaker, D., Swaab, R., and Diermeier, D. The language of coalition formation in multiparty online negotiations. In *In Proceedings of the International Communication Association Conference* (Montreal, Canada, 2008), 22–26.
13. Ireland, M. E., and Pennebaker, J. W. Language style matching in writing: Synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology* 99 (2010), 549–571.
14. Jonassen, D. H., and Kwon, H. Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development* 49 (2001), 35–51.
15. Kittur, A., and Kraut, R. E. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proc. CSCW 2008*, ACM Press (2008), 37–46.
16. Leavitt, H. J. Some effects of certain communication patterns on group performance. *Journal of Abnormal Psychology* 46 (1951), 38–50.
17. Lebie, L., Rhoades, J. A., and McGrath, J. E. Interaction process in computer-mediated and face-to-face groups. *Computer Supported Cooperative Work* 4 (1996), 127–152.
18. Leshed, G., Hancock, J. T., Cosley, D., McLeod, P. L., and Gay, G. Feedback for guiding reflection on teamwork practices. In *Proc. GROUP 2007*, ACM Press (2007), 217–220.
19. Leshed, G., Perez, D., Hancock, J. T., Cosley, D., Birnholtz, J., Lee, S., McLeod, P. L., and Gay, G. Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. *Proc. CHI 2009* (2009), 537–546.
20. Levine, J. M., and Moreland, R. L. Progress in small group research. *Annual Review of Psychology* 41 (1990), 585–634.
21. McGrath, J. E. Small group research, that once and future field: An interpretation of the past with an eye to the future. *Group Dynamics: Theory, Research, and Practice* 1 (1997), 7–27.
22. Nemeth, C. J., Personnaz, B., Personnaz, M., and Goncalo, J. a. The liberating role of conflict in group creativity: A study in two countries. *European Journal of Social Psychology* 34 (2004), 365–374.
23. Niederhoffer, K. G., and Pennebaker, J. W. Linguistic style matching in social interaction. *Journal of Language and Social Psychology* 21 (2002), 337–360.
24. Nowak, M., Kim, J., Kim, N. W., and Nass, C. Social Visualization and Negotiation: Effects of Feedback Configuration and Status. In *Proc. CSCW 2012*, ACM Press (2012), 1081–1090.
25. Pennebaker, J. W., Gosling, S. D., and Ferrell, J. D. Digital in-class teaching using the TOWER system: A pilot project of two 500-student introductory psychology classes. Tech. rep., Technical Report, Department of Psychology, University of Texas at Austin, 2012.
26. Sexton, J. B., and Helmreich, R. L. Analyzing cockpit communications: The links between language, performance, and workload. *Human Performance in Extreme Environments* 5 (2000), 63–68.
27. Stasser, G., Taylor, L. A., and Hanna, C. Information sampling in structured and unstructured discussions of three- and six-person groups. *Journal of Personality and Social Psychology* 57 (1989), 67–78.
28. Tausczik, Y. R., and Pennebaker, J. W. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29 (2010), 24–54.
29. Taylor, P. J., and Thomas, S. Linguistic style matching and negotiation outcome. *Negotiation and Conflict Management Research* 1 (2008), 263–281.