Feedback on Collaborative Skills in Remote Studio Design

Gina Turner, Michael F. Schober
New School for Social Research
TurnG183@newschool.edu, schober@newschool.edu

Abstract

The study reported here examines how feedback affects remote collaborative brainstorming and product design in a studio setting. 10 teams of 4 students at the Parsons School of Design met remotely via a chatroom interface for three 30-minute sessions to design a T-shirt. Team members who proposed ideas were rated as better collaborators, and team members who critiqued ideas without offering alternatives were rated as poor collaborators. This was also reflected in team members' word use, as measured by Pennebaker & Francis' Linguistic Inquiry and Word Count too; for example, team members who used more affect words were rated as poorer collaborators. A case study of one team suggests that the ideas proposed by team members rated as the best collaborators may be most likely to survive. Peer feedback can affect remote studio design discussion in ways comparable to what is found in other settings, despite task differences.

1. Introduction

Studio artists increasingly design in collaborative virtual groups, generating and critiquing ideas entirely via email, chat and shared graphics files. They may, in fact, never meet each other face to face, and they may live in different cities or on different continents. Skill at being a collaborator is thus a more and more integral part of the creative design process, and understanding what constitutes collaborative skill is an increasingly important aspect of developing more sophisticated computer-supported work spaces. The extent to which a remote environment supports feedback about how one's team members are reacting to what one is doing—from the kinds of facial and bodily cues available in face-to-face settings to more explicit ratings—is likely to affect the nature of the creative process and what is created.

Design educators are increasingly recognizing that an important component of training designers is training them to be good collaborators. But to what extent is collaborative skill trainable? And what constitutes collaborative skill in remote virtual groups as opposed to in face to face groups?

2. Assessing collaborative skill

Defining successful collaboration, even in face to face teams, is not straightforward [see for review 8, 25, 27]. While generating (good) ideas and keeping the team on track are clearly good things, a brief well-timed acknowledgment of a good idea can be just as important, and levity at the right moment can restore equilibrium. In addition, the right move may vary for different tasks and at different times.

The remote chatroom setting brings additional complexities. Instead of relying on facial cues to assess their team members' reactions and emotional tone, remote teams must rely on textual and timing cues. Instead of relying on vocal back channels for ongoing confirmation about whether they have been understood, team members must adapt to the different nature and timing of grounding cues in remote chat [3].

In the study described here, we examined how feedback about collaborative success affected team members' subsequent collaborative discourse in a chatroom-based studio art project. We were particularly interested in two major questions:

(1) *Do the effects of feedback on group discussion demonstrated in other domains (face to face and remote) extend to the studio art setting?* There are a number of reasons to expect, from prior findings in other domains, that

1530-1605/07 $20.00 ©2007 IEEE

feedback on one's skills and ideas will affect the nature of one's contributions to the discussion and ultimately the final products. Does this hold in artistic settings, or does artistic collaboration require an entirely non-evaluative tone?

(2) *How do better and worse collaborative moves manifest textually?* What happens in a chat room is by definition linguistic. Does the language that is used by participants rated as better collaborators differ from the language used by poorer collaborators, and if so how? If there are reliable textual indicators of better and worse collaboration, one could imagine developing new ways of intervening to train people to be better collaborators.

2.1 Peer assessment and feedback effects

Prior work in several domains lends plausibility to the idea that feedback on collaborative skills might be helpful in the studio art domain.

2.1.1 Peer assessment of collaborators

Face-to-face peer assessment has been shown to be valid and reliable in both field and laboratory studies in non-artistic domains, producing higher reliability than either supervisor ratings or psychological tests [see 15]. These days, computer-assisted peer assessment for face to face team members is increasingly common [25]. For example, a peer assessment tool created by Goldfinch and Raeside [11] allows for two things: assessing each team member individually, and allowing the team members to assess each other on specific group functions (e.g., helping the group to function well as a team, generating good ideas, etc.). A web-based version of this peer assessment tool [9] is based on feedback from students who felt that an overall team assessment was unfair if there had been unequal contributions from the members of the team. The fact that it is web-based gives the ability to provide immediate scores to the recipients, rather than having to tally those scores manually.

So what does peer assessment do? In face to face settings, peer appraisal of others' performance has been shown to improve relationships and task focus. One study [7] of 80 self-managing work groups of organizational behavior students examined how ongoing structured peer feedback affected students' perception of group effectiveness. In the study, groups of 5 or 6 students completed a group research project and made a formal presentation of the results. The finding was that peer appraisal had a positive impact on the students' perceptions (as self-reported on questionnaires) of open communication, task motivation, social loafing, group viability, cohesion, and satisfaction. Other studies have also shown that peer appraisal can reduce social loafing. In one study, getting feedback on how much or little effort a teammate has put into a task the first time it is performed affected how much effort the participant was willing to invest the second time the task was performed [13].

In addition, taking the time for peer appraisal in the production process can provide the opportunity for a "temporal shift": a pause in the working routine that can allow for meta-analysis of the work process [27], interrupting the work flow in a constructive way, and allowing for reflection and evaluation of the work being done. These sorts of temporal shifts can lead to positive group outcomes. Smith, Cooper, and Lancaster [25], in fact, found that participating in peer evaluation actually decreases students' resistance to the idea of receiving this feedback.

2.1.2 Brainstorming studies

Another line of research [e.g., 5, 6, 24] has focused on brainstorming tasks, in which collaborators are charged with generating as many ideas as possible, ideally with a non-evaluative tone. In remote brainstorming group systems, the number and quality of ideas generated can increase when participants can provide feedback on idea quality, although there seems to be no benefit of feedback in face to face brainstorming groups. Note that unlike in the peer assessment studies, where collaborators rate their peers as collaborators, here the focus is on rating the *ideas*. Thus the feedback is anonymous in a quite different sense than it can be in the peer assessment studies, where even if identities are hidden the raters still consider each idea-generator's qualities overall, rather than one idea at a time.

A sample piece of evidence from a brainstorming study using a remote support system [18] is that groups that received feedback in a shared table (that could be read by all group members) displaying the quantity of group activity on a variety of measures got higher marks on productivity, creativity and had more activity in general than those who did not receive this feedback. Intriguingly, negative evaluative tone in the feedback is not necessarily harmful. In a study on brainstorming to solve a complex university parking problem [5], negative evaluative tone (as injected by confederates) was shown increase the number of creative contributions and comments. This was despite the real possibility that fear of being evaluated might cause members of a group to withhold suggestions due to the possibility of negative evaluation.

Another finding in the brainstorming literature is that while the anonymity that remote systems allow can reduce bad feelings, it can also lead to social loafing. But this can be counteracted by adding in (anonymous) social comparison: on-line feedback about where one falls relative to the group on productivity. For example, in one study [24], participants in groups generating ideas for solving the School Business Task produced reliably more ideas when they got feedback about how their production compared to others'.

2.2 Feedback in studio art groups

Do these sorts of findings extend to artistic settings? How exactly should feedback affect remote studio design? Very little is known about the creative process in group discussion by studio artists, let alone in remote discussion. The very few studies [2, 15] that have examined peer assessment in artistic settings have been in the arena of music performance, and so it is unclear how they apply. Findings from the peer assessment studies and remote brainstorming studies have been in non-artistic topics and settings, and the nature of the brainstorming and decision-making have been different from what artists do: the studio art design setting involves visual design that includes esthetic judgment. It involves brainstorming, but evaluation of the brainstormed ideas cannot help but be tied to the identities of the originators.

One might imagine that a focus on evaluative feedback in an artistic setting can actually dampen creativity and hinder the open expression of risky ideas. But if the findings from studies in other settings extend to remote design by artists, then peer evaluation of collaborative performance during the design process could lead to improved collaboration, task outcome, and satisfaction. This should be above and beyond improvements in collaborative performance that come simply from collaborating over time (that is, from practice at collaborating) [e.g., 14, 17, 18].

It is particularly unclear whether prior findings on negative group affect will apply to the studio art setting. Beyond increasing brainstorming productivity in a non-design task [5], negative affect has been shown to enhance design creativity [10] in an engineering setting (designing helicopters). But in freer artistic expression, it seems self-evident that overly negative and harsh feedback can't always be a good thing for individual and thus group creativity. During a creative discussion, having to focus on controlling one's emotional responses quite plausibly could lead to less task focus, less willingness to take risks, and less satisfaction with the group. Focusing on others' collaborative performance (as opposed to the task at hand) could distract from the design task. Similarly, a focus on *being* evaluated could also distract from the design task. Perhaps the collaborative artistic process requires a non-evaluative tone for creativity to be maximally stimulated.

2.3 Textual correlates of collaborative skill

In virtual chat room settings, it is through the text itself that the work gets done and the emotional tone is established. This suggests that as researchers we may be able to exploit this fact to discover the textual correlates of collaborative skill in a particular task setting.

Our strategy in the current study is to examine potential textual correlates of collaborative skill using the Linguistic Inquiry and Word Count (LIWC) tool [21]. This text analysis tool is a dictionary-based word counting tool that counts words in different content categories and thus highlights patterns of content. LIWC assumes

that psychological information can be gleaned from the words used by a speaker or writer, independently of the words' individual and contextual meanings. LIWC contains a dictionary of over 2300 words and word stems, which have been categorized to reflect more than 79 linguistic dimensions.

Word-level analysis of this sort can be illuminating not only about textual content but also about people's cognitive styles, emotional states, and personalities. For example, Pennebaker and King [20] demonstrated that an individual's style of language use can be indicative of personality dimensions. Other analyses of transcripts have revealed diagnostic power for psychiatric disorders [12] and the ability to predict the process of bereavement [22]. The use of relational pronouns (e.g., we, us, our) can reflect degree of involvement with one's conversational partner and one's disposition [4].

Because chat room communication is textually based, it therefore provides us with an excellent opportunity to see how peer ratings of collaborative skill are reflected in particular textual choices. This may open the door into new insights into what counts as collaborative skill.

3. Method

In our study, 10 teams of 4 students at the Parsons School of Design met remotely for 3 30-minute sessions. Each team was composed of students who had not collaborated with each other in previous studio settings. The average age was 24, and 33% were women. Each team was assigned the same task: over the course of three 30-minute meetings, the team was to design a T-shirt representing the Parsons Design & Technology program. The task is concrete enough to allow comparison across teams but also open enough to allow for quite different solutions.

The team members were to type anything they needed to in order to accomplish the task, but only during chat sessions. Team members could exchange visuals and any related web sites by sending links in the chat room. They were told that the design the teams produced was to be evaluated by a jury composed of experts from four fields: graphic design, marketing, fashion, and illustration. The jury would be evaluating the results using several criteria: the extent to which the design communicated the character of the program, originality, and visual appeal. Teams were told that the design judged to be the most effective was to be produced.

All teams collaborated remotely via a standard chat interface. All teams filled out surveys at the end of each session. Six teams anonymously evaluated each member of their team at the end of each session based on the approach embodied in Freeman and McKenzie's [9] system. Under this approach, each team member rates each other team member on a 5-point rating scale on "Helping the group to function well as a team," "Understanding what is required", "Suggesting ideas," "Level of enthusiasm and participation," "Performing tasks efficiently," "Organizing the team and ensuring things get done," and "Providing constructive feedback." The remaining four teams filled out an unrelated online questionnaire of the same length and format at the end of each session: rating New York tourist attractions.

Within 24 hours after the ratings were carried out, members of the evaluation teams were emailed averaged ratings on each scale for all the members of their team. In addition to the emailed feedback, members of the evaluation teams also received the same feedback at the start of next session. The feedback consisted of graphs of each team member's average rating on each of the 7 evaluation questions. Non-evaluation teams received no feedback.

Note that the evaluations were averages and represented the consensus of the other team members. This makes the individual ratings entirely anonymous, and diminishes the impact of any one extreme outlier rating.

At the end of the third 30 minute session, all teams submitted their designs to a central repository.

3.1 Coding design chat moves

We created a coding scheme to identify the design moves in the chat sessions specific to the current context. Although the scheme was inspired by other schemes that examine design interaction, chat in the current context was sufficiently different so as to require a new scheme. For example, Anderson et al.'s [1] scheme codes each speaking turn based on three

dimensions: whether they were entirely social, entirely task-oriented, or a meta-commentary on completing the task; we needed more detail. Medland's [16] scheme looked specifically at communication activities in industrial and engineering design: delegation, reporting, awareness, and problem handling; the design behaviors in the current setting included a broader range of brainstorming and other activities that the Medland scheme doesn't capture. Another coding scheme, Eckert & Stacey's [8] taxonomy of types of discussion, is more focused on group processes across multiple settings than the content of the communication. This scheme, again, doesn't examine specific content at quite the right level for our purposes.

In our scheme, design moves included design proposals (e.g., "I'm thinking pirates"); defending one's proposal (e.g., "well to me simple is fancy"); rejecting other's proposals without providing an alternative proposal (e.g., "no, that's creepy"); and off-task commentary (e.g., "the internet is for geeks. I only read books by james joyce"). Table 1 shows the complete list of categories in our scheme.

Table 1. Codes and definitions

Code	Definition
Design proposal	Specific, concrete suggestions to be included in design – even silly suggestions
Agreeing with a proposal	Clearly expressing a positive reaction without further comment
Critiquing or considering a proposal	Providing constructive criticism of another's proposal
Defending your own proposal	Backing up your own design proposal against the critique of others
Idea-generating image	Sending a link to an image as a prompt or fodder for additional design discussion
Image proposal	Link to image sent that is actual design proposal
Off-task commentary	Off-task conversation that is playful, silly, joking, etc.
Soliciting/Procedural proposal	How to go about the design process; what exactly is required
Requesting information about a proposal made	Asking for clarification about someone else's design proposal
Rejecting a proposal	Clearly expressing a negative reaction without further comment
Statement about Technology	Specifically about the software they're using, or technological problems
Withdrawing a proposal	Person taking back their own proposed idea
Other	Chat about who's signed on, is it time to leave, hellos, goodbyes

The interrater reliability, calculated on 100% of the chat corpus, was good, with an overall Cohen's kappa of .79. Individual reliabilities for the pertinent categories were .85 for design proposals, .78 for defending a proposal, .85 for image proposals, .84 for statements about technology, and .89 for off task commentary.

4. Results

The chat sessions, which ranged from 49 to 263 entries ("moves"), covered wide-ranging and divergent topics; they had a freewheeling character that was rather different from what one sees in engineering design and management contexts. (Although the comparison isn't entirely fair, given different tasks and coding schemes, the percentage of jokes and off-task commentary in the current study was notably higher than in [5]. As is always the case in chat, the conversational moves weren't necessarily linearly sequential; people could respond to moves several turns earlier. All led to a final design.

With access to intermediate and final designs, we could also see which design proposals survived. For example:

user17_1 well why doesn't it have graphics on it like pizza and coffee stains and solder
user17_4 like a dirty t-shirt from a design student staying up all night
user17_4 there we go…
user17_3 now we're talking

The idea of "solder" from this chat made it (in misspelled form) onto a graphical proposal (Figure 1):

Figure 1. Proposed design with element from chat.

Similarly, the following snippet of chat shows its impact in the final design, which was indeed red and black:

user17_2	what color
user17_2	rainbow may be good
user17_4	no rainbow
user17_4	just red and black
user17_3	rainbow is gay color..

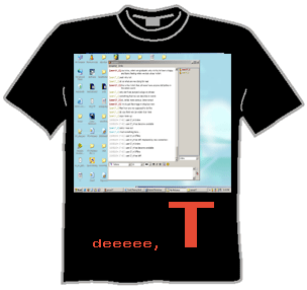

Figure 2. Final design with elements from chat.

4.1 Ratings and design chat moves

In general, people's ratings of team members' collaborative skill reflected the design behaviors we coded. People who proposed any designs were rated marginally higher than those who didn't on "Understanding what is required," $F(1,22) = 3.35$, $p = .08$, and reliably higher on "Level of enthusiasm and participation," $F(1,22) = 4.61$, $p < .05$). Design proposals—specific, concrete suggestions for what should be included in the design—included proposals like "Maybe the spaceship can look like sushi."

People who defended their own design proposals were rated higher on four questions: "Understanding what is required" ($F(1,22) = 5.44$, $p<.05$); "Suggesting ideas" ($F(1,22) = 5.66$, $p<.05$); "Level of enthusiasm and participation" $F(1,22) = 5.29$, $p<.05$); and "Performing tasks efficiently" ($F(1,22) = 4.55$, $p<.05$. Defending a design proposal –backing up one's own design proposal against the critique of others—included backing up a design statement like "…I also thought about working with 'render, solder, create, bleed" by additionally stating, in a separate entry: "it doesn't have as much to do with parsons, as it is a fun phrase…"

People who proposed images—providing a link to an image file or a link to a website with an image that was an actual design proposal—were also rated higher on the questions "Organizing the team and ensuring things get done" ($F(1,22) = 5.6$, $p<.05$). Examples of images proposed can be seen in Figure 3.

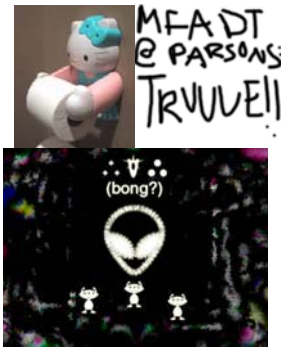

Figure 3. Proposed images.

People who made statements about the technology were rated higher on "Understanding what is required" ($F(1,22) = 10.4$, $p<.01$); "Level of enthusiasm and participation" ($F(1,22) = 9.66$, $p<.01$); "Performing tasks efficiently" ($F(1,22) = 6.06$, $p<.05$); and "Organizing the team and ensuring things get done" ($F(1,22) = 9.08$, $p<.01$. Statements about technology ranged from things like "well maybe we don't need to upload (an image) since we share the same drop box" to "ok is there (the design program) Illustrator on these PCs?" In other words, technology statements were specifically about the software they were using, or technological problems they were having with it.

4.2 Ratings and word choice

People's ratings of team members' collaborative skill also reflected the team members' word choice in the chat rooms, in particular in the LIWC categories of affect words, anger words, cognitive mechanism words, and "we" words. The affect category consists of all emotionally charged words in LIWC, positive or negative. Some of the words laden with negative affect (e.g., suck, crap, stupid) are also evaluative, and likely to indicate a negative opinion about a design proposal in the current setting. Anger words are a subset of the affect words. The cognitive mechanism words reflect cognitive processes (e.g., know, understand, wish). "We" words indicate group membership (e.g., we, our, let's).

People who used more affect words were reliably rated lower on all but one of the seven ratings ("Helping the group to function well as a team": r = .-.62, p <.005; "Understanding what is required": r = -.59, p<.005; "Level of enthusiasm and participation": r = -.53, p<.01; "Performing tasks efficiently": r = -.50, p<.05; "Organizing the team and ensuring things get done": r = -.49, p<.05; and "Providing constructive feedback": r = -.54, p <.01). Affect words were used especially often in entries such as "this sucks" or "the mfa shirts are so ugly and dumb" (in reference to the department's t-shirt design created from a contest the previous year). In addition to this, we saw some correlation between various word choices and design behaviors. Those who more frequently rejected others' proposals used a higher rate of anger words, r = .39, p<.05, as in "that's stupid" or just simply "ugly." And those who used more words in the cognitive mechanism category produced fewer off task contributions to the discourse, r = -.43, p<.01. This makes sense; the cognitive mechanism word "justify" used in the statement "try to justify your color choice" seems to indicate active engagement with the task at hand.

4.3 Effects of peer evaluation

How did evaluation affect subsequent chat? In the second session, team members who had evaluated each other at the end of the Week 1 session chatted differently than team members who had not. In particular, after evaluating each other and seeing the evaluations, members of evaluation teams used a greater percentage of words related to the self—"I" and "we" words—while members of the non-evaluation groups decreased their percentage of self-related words, F (1,8) = 6.94, p<.05 (see Figure 4).

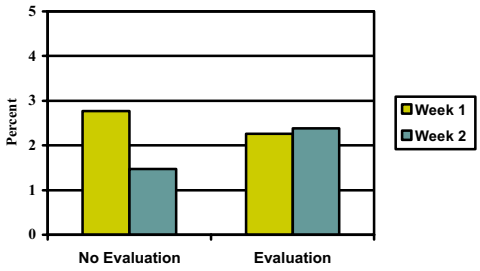

Figure 4. Word choice: Use of "we" words.

These word findings reflect the broader tendency for members of the evaluation groups to present more task-related proposals like "Let's go", "Let's make a shirt that…" "We should…" or "We need…" and seem to reflect greater task focus and increased activity.

After evaluating each other and seeing the evaluations, members of the evaluation teams also reduced their use of affect words, while members of the non-evaluation groups increased their use of affect words, F (1,8) = 10.23, p< .05 (see Figure 5).

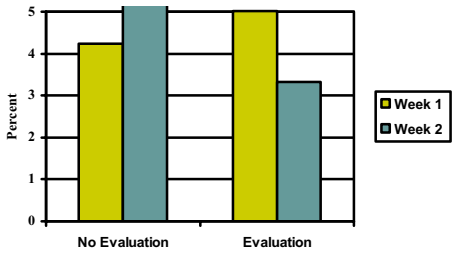

Figure 5. Word choice: Use of affect words.

This makes sense in light of the finding that those who used more affect words rejected more proposals, and those who rejected more proposals were rated lower. This seems to reflect a broader tendency for members of the evaluation groups to present fewer broad-based

evaluative statements and to focus more on details of the T-shirt design. For example, in the non-evaluation group, we see negative and non-constructive statements like "it's like annoying to read" and "no, that's creepy." In the evaluation group we see more constructive statements like "I think it'll look good on white t-shirts" and "what theme did people like before?", which provide specific critiques and solicit specific types of input.

A case study of one team, the "chocolate bar team," suggests the intriguing possibility that final products were more likely to come from contributors who were rated as having greater collaborative skill. We examined 3 weeks of discussion leading to this team's final shirt design and coded design proposals for elements that ultimately ended up in the final design (see Figure 6).

Figure 6. Final Design: Chocolate Bar Team.

Some examples of these surviving proposals are: "It should be a plain white shirt"; "I kind of drew this chocolate bar? Thingy with its wrapper half open"; "and the chocolate itself is like a keyboard.."; "the wrapper has a DT logo on it"; and "yeah I'd put the logo on the wrapper definitely". Team members' ratings of each others' collaborative skill reflected surviving design proposals; the rank of the team members' ratings correlated exactly with the rank of the number of each person's surviving design proposals (see Table 2).

Table 2. Design Proposals that Survived in the Final Design – Chocolate Bar Team

	Average rating overall	Surviving proposals
User1	2.88	14
User2	2.7	8
User3	2.57	6
User4	2.24	0

Of course, we can't tell whether designers who have better ideas are rated as better collaborators, or whether designers who have greater collaborative skill are more effectively able to make their design proposals stick, or both. But this case study does suggest that design schools' interest in fostering collaborative skill is a reasonable aim.

5. Discussion

Obviously these findings are based on a relatively small sample of studio design teams in a very particular context, and much remains to be known. We don't know if the particular kinds of word choice that reflected peer evaluation ratings would extend to different work settings and contexts. We don't know if more or less constrained tasks would lead to different kinds of discussions and different effects of evaluative feedback. We didn't collect satisfaction ratings from team members, and so we don't have evidence about whether satisfaction with the interaction was affected by evaluating each other. Nonetheless, the findings allow us to draw some preliminary conclusions on the questions we raised at the start.

Regarding Question 1, some of the effects of feedback on remote group collaboration demonstrated in different domains can indeed extend to a studio art setting. Our teams of remote studio designers were reliably able to rate each other's collaborative performance; there was general agreement on which team members were better and poorer collaborators. This extends findings about the reliability of such ratings in face-to-face and non-artistic tasks to the remote studio setting.

Second, as in other settings, a studio art designer's performance as a remote collaborator can be improved (as measured by peer ratings)

by feedback about which aspects of their performance are lacking. The mere fact of evaluating each other increases behaviors that get higher ratings and decreases behaviors that get lower ratings. This corresponds with the intuitions of design educators about the effectiveness of evaluative feedback on collaborative skill, and it suggests that this kind of feedback effectively can translate to remote settings. Note that this is despite the fact that the task discussion had a freewheeling "artistic" character that isn't what one tends to see in studies of business, management or engineering group discussion.

An important next step would be to examine the actual task outcome—juried evaluations of the final designs—on a larger scale with a larger number of teams. Our data do not show one way or another whether peer evaluations of collaborative skill correspond with the quality of final group products, as they sometimes do in other domains. Further testing is needed in artistic design settings before recommendations can be made about best training practices. But, again, the anecdotal evidence from design educators is lent some preliminary support.

Regarding Question 2, on the textual manifestation of collaborative skill: Our findings demonstrate that the language used in remote chat settings can reflect ratings of collaborative success, and evaluative feedback about collaborative skill can affect word choice in subsequent chat. Of course, word choice is only one linguistic marker; there are many other potential levels and layers of linguistic analysis (syntactic form, discourse markers, politeness cues, etc.) that could be carried out, and it remains to be seen how telling word choice turns out to be as a correlate of deeper matters. But we propose that the fact that there are any textual correlates of rated collaborative skill at all raises the possibility that feedback about language use (here, word choice) could stand in for more complex judgments about collaborative moves, and could complement peer evaluation of team members. It may well be that a new inroad into sensitizing remote team members to their collaborative behaviors is to provide feedback on and to focus on the text they produce.

6. References

[1] Anderson, A.H., Mullin, J., Katsavaras, E., McEwan, R., Grattan, E., & Brundell, P. (1999). Understanding multiparty multimedia interactions. In S. E. Brennan, A. Giboin & D. Traum (Eds.), *Psychological Models of Communication in Collaborative Systems (AAAI Fall Symposium)* (pp. 9-16). Menlo Park, CA: AAAI.

[2] Bergee, M. J. (1993). A comparison of faculty, peer, and self-evaluation of applied brass jury performances. *Journal of Research in Music Education, 41*, 19 – 27.

[3] Clark, H. H., and Brennan, S. A. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.). *Perspectives on socially shared cognition*. Washington: APA Books.

[4] Cegala, D. J. (1989). A study of selected linguistic components of involvement in interaction. *Western Journal of Speech Communication, 53*, 311-326.

[5] Connolly, T., Jessup, L. M., & Valacich , J. S. (1990). Effects of anonymity and evaluative tone on idea generation in computer-mediated groups. *Source Management Science, 36*(6), 689 – 703.

[6] Dennis, A. R. & Valacich, J. S. (1993). Computer brainstorms: More heads are better than one. *Journal of Applied Psychology, 78*, 531-537.

[7] Druskat, V.U., & Wolff, S.B. (1999). Effects of timing of developmental peer appraisals in self-managing work groups. *Journal of Applied Psychology, 84*, 58-74.

[8] Eckert, C., & Stacey, M. (2001). Dimensions of communication in design. In *13th International Conference on Engineering Design (ICED 01),* Glasgow, Scotland, UK, Design Management - Process and Information Issues, 473-480.

[9] Freeman, M., & McKenzie, J. (2002). SPARK, a confidential web-based template for self and peer assessment of student teamworks: Benefits of evaluating across different subjects. *British Journal of Educational Technology, 33*, 551-569.

[10] George, J. M., & Zhou, J. (2002). Understanding when bad moods foster creativity and good ones don't: The role of context and clarity of feelings. *Journal of Applied Psychology, 87*(4), 687 – 697.

[11] Goldfinch, J. & Raeside, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group

project. *Assessment & Evaluation in Higher Education, 15*(3), 210-225.

[12] Gottschalk, L. A., Stein, M. K., & Shapiro, D. H. (1997). The application of computerized content analysis of speech to the diagnostic process in a psychiatric outpatient clinic. *Journal of Clinical Psychology, 53*(5), 427 – 441.

[13] Harkins, S.G., & Jackson, J.J. (1985). The role of evaluation in eliminating social loafing. *Personality and Social Psychology Bulletin, 11*, 457 - 465.

[14] Hollingshead, A. B., McGrath, J. E., & O'Connor, K. M. (1993). Group task performance and communication technology: A longitudinal study of computer-mediate versus face-to-face work groups. *Small Group Research, 24*(3), 307 – 333.

[15] Hunter, D., & Russ, M. (1995). Peer assessment in performance studies. In S. Griffiths, K. Houston, & A. Lazenblatt (Eds.), *Enhancing student learning through peer tutoring in higher education: Section 3 – Implementing (Vol. 1, pp. 63 – 65)*. Colerain, Northern Ireland: University of Ulster.

[16] Medland, A.J. (1992). Forms of communication observed during the study of design activities in industry. *Journal of Engineering Design, 5*, 243 - 253.

[17] McGrath, J. (1991). Time, interaction, and performance (TIP): A theory of groups. *Small Group Research, 22*(2), 147 - 174.

[18] McGrath, J., Arrow, H., Gruenfeld, D. H., Hollingshead, A. B., O'Connor, K. M. (1993). Groups, tasks, and technology: The effects of experience and change. *Small Group Research, 24*(3), 406 - 420.

[19] Michinov, N., & Primois, C. (2005). Improving productivity and creativity in online groups through social comparison process: New evidence for asynchronous electronic brainstorming. *Computers in Human Behavior, 21*, 11 – 28.

[20] Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296 – 1312.

[21] Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic Inquiry and Word Count LIWC2001*. Mahwah, NJ: Erlbaum.

[22] Pennebaker, J.W., Mayne, T. J., & Francis, M.E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*(4), 863 - 871.

[23] Saavedra, R., & Kwun, S.K. (1993). Peer evaluation in self-managing work groups. *Journal of Applied Psychology, 78*, 450 - 462.

[24] Shepherd, M. M., Briggs, R. O. , Reinig, B. A. , Yen, J. & Nunamaker, J. (1996). Invoking social comparison to improve electronic brainstorming: Beyond anonymity. *Journal of Management Information Systems, 12*(3), 155 – 170.

[25] Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International, 39*, 71 - 81.

[26] Stahl, G. (2002). *The complexity of a collaborative interaction*. Paper presented at the International Conference of the Learning Sciences (ICLS 2002), Seattle, WA.

[27] Staudenmayer, N. Tyre, M., & Perlow, L. (2002).Time to change: Temporal shifts as enablers of organizational change. *Organization Science: A Journal of the Institute of Management Sciences, 13*(5), 583 - 597.

[28] Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249 – 276.