Running head: WRITTEN VERSUS SPOKEN FEEDBACK

To cite, please refer to: Fox Tree, J. E. & Clark, N. B. (2013). Communicative effectiveness of written versus spoken feedback. *Discourse Processes*, *50*(5), 339-359.

Communicative Effectiveness of Written versus Spoken Feedback

Jean E. Fox Tree and Nathaniel B. Clark

University of California Santa Cruz

Author Note

This research was supported by faculty research funds granted by the University of California Santa Cruz to Jean E. Fox Tree and by a graduate research fellowship funded by the Perlino Foundation, Santa Cruz, CA, to Nate Clark. We thank our many research assistants who aided in data collection and coding. We also thank Jeffrey Hancock and three anonymous reviewers for comments on this manuscript. Correspondence can be addressed to Jean E. Fox Tree or Nathaniel B. Clark, Psychology Department, Social Sciences II room 277, University of California, Santa Cruz, CA, 95064, {foxtree,nbclark}@ucsc.edu.

Abstract

How does written and spoken feedback influence communicators' effectiveness in a shared task? Groups of two to four participants engaged in a referential communication task. The director described an array of shapes to the rest of the group via streaming video chat. In each group, one to three matchers attempted to arrange cards depicting those shapes into the director's arrangement. In one condition, matchers could speak to each other and the director through the video chat. In the other condition, matchers could only type contributions into a shared text chat. Spoken feedback dyads successfully arranged more cards than written feedback dyads. But in groups with more than one matcher, feedback modalities were equally effective. Across group sizes and feedback types, when any matcher contributed more to the discourse, all matchers benefited. Spoken and written feedback groups differed in the amount and types of contributions matchers provided. Spoken feedback groups had more total discourse than written groups. Matchers in written groups made more bare requests for clarification, while matchers in spoken groups made more contributions of new information. These differences are explained in terms of constraints on grounding across the modalities.

Keywords: grounding, spontaneous communication, instant messaging, video chat, texting

Communicative Effectiveness of Written versus Spoken Feedback

The internet offers many communication options for geographically disparate interlocutors, and these options vary in the modalities and temporal nature of the communication. We focus on communication by audiovisual video chat and text-based instant messaging, both of which allow for simultaneous or near-simultaneous exchange. The ability to provide synchronous feedback to a speaker is associated with greater understanding of speakers' utterances (Schober and Clark, 1989). Because people attempt to ground communication differently depending on the medium in which they are communicating, different modalities of providing feedback have strengths and weaknesses (Clark and Brennan, 1991; Fox Tree, Mayer, & Betts, 2011). Different modality types are commonly found in computer-supported collaborative work. Communicators may talk on the phone and type messages at the same time, perhaps also communicating with others via audiovisual chat such as Skype, or with communicators co-present in the room.

We assessed the utility of written versus spoken feedback, and the role of participants' varying amount of contributions, to both individual and group success in a referential communication task. In this task, participants attempted to get various numbers of fellow participants to order abstract shapes into a specified order. In one condition, participants could see and hear each other, producing spoken dialogue such as the following from a group of four participants, with quotations indicating spoken contributions:

(1)
Director:	"Okay. The third one looks like a guy sitting with his knees kind of like this."
Matcher 1:	"Against a wall?"
Director:	"Yeah kind of like against a wall."
Matcher 2:	"Is he facing the left side?"
Director:	"Yeah. He's facing left."
Matcher 2:	"Okay cool."
Matcher 1:	"I think I got that one."

Director: "Yeah?"
Matcher 1: "Yeah."
Matcher 3: "Uh, I don't know."
Matcher 1: "He kind of has his legs bent?"
Director: "Yeah his legs are like bent into like a triangle shape."
M1: "Okay."
Director: "Kind of like, like that, sorta."
Matcher 2: "Alright."
Director: "You guys found it?"
Matchers 1, 2, and 3 simultaneously: "I think so./Yeah./Yeah."

In the other condition, participants could still see each other, but had different verbal communication channels. The director, who knew the order of the shapes, could speak to the other participants. The matchers, who were reconstructing the order, could write (but not speak) to the whole group. This design produced a combination of spoken and written communication such as the following from a group of three participants, with quotations indicating spoken contributions and lines without quotations indicating written contributions (orthography of the written contributions is exactly as typed):

(2) Director: "The next one looks like the top half of a person with their head and their arms up towards the air, but without the bottom half - they have no legs. Umm, they're really like skinny body and like I said the arms are both up in the air pointing up towards the sky."
Matcher 1: okay, i think i got that one
Matcher 2: is the person facing left or right?
Director: "And the person is facing towards you, looking at you. It's not facing right or left. The head looks like you're looking at the person straight on."
Matcher 2: ok

This method of communication models a popular format adopted by distance learning educational programs where instructors can be heard by all their students, but students write to their instructors via instant messaging text.

Grounding Across Media

In the *collaborative theory of language use*, communicators take into account each other's *common ground* in preparing and interpreting communication (Clark, 1996). Common ground includes things like prior history, shared environments, and information from the discourse at hand. As they take turns making contributions to the discourse, communicators check each others' understanding by *grounding* their contributions – making sure that what's contributed is understood well enough for the current purposes of the communication (Clark & Wilkes-Gibbs, 1986).

Grounding across written and spoken feedback conditions is constrained in a number of ways that impacts communicative effectiveness (Clark & Brennan, 1991). Feedback provided via a spoken channel has the advantage of a rich signal that can provide information beyond the words uttered. But spoken channels can also be disadvantageous, if communicators must discuss something they could otherwise avoid discussing, such as noise generated by others in the vicinity of the computer.

In a *referential communication task*, *directors* instruct *matchers* about the order in which to place abstract shapes in an array. Spoken responses allow directors both the costs and the benefits of hearing matchers. Written responses from matchers remove some of the costs (such as discussing irrelevant things) and some of the benefits (such as being able to hear a confused tone in a matcher's voice). It is unknown whether the crystallization of information in the form of written feedback to a director is an overall gain or loss.

Spoken feedback has other potential advantages in its *cotemporality* (production and comprehension take place in synchrony) and *simultaneity* (as something is produced, it is understood, and as understanding occurs, production is also possible; Clark & Brennan, 1991).

Matchers who communicate with speech receive cues that directors are listening while they are speaking. They also know that a quizzical *huh?* can cause a speaker to adjust what's being said mid-stream (cf. Clark & Krych, 2004). But with written responses, understanding starts when production finishes. The director cannot read the message until the matcher has sent it, and the matcher does not know when the message will be read and responded to.

Written feedback has its own suite of advantages. One is its persistent record (Gergle, Millen, Kraut, & Fussell, 2004; McKinlay, Arnott, Proctor, Masting, & Woodburn, 1993), which may allow directors to notice recurring areas of confusion. Instant messaging may also have a reduced social cost for help-seeking, because it is seen as a less formal, more playful means of communication (Honeycutt, 2001), and is in some cases preferred over email when discussing tasks between peers (Honeycutt, 2001). Another advantage of written feedback is that multiple matchers can supply feedback simultaneously (McKinlay et al., 1993; Whittaker, Brennan, & Clark, 1991). With spoken feedback, speaking simultaneously would result in incomprehensibility. Finally, writing may require fewer of the social aspects important in face-to-face talk (conversational openings and closings, politeness) and therefore be a more efficient form of communication (see discussion of *lean media* in Rau, Gau, & Wu, 2008). Table 1 summarizes of the pros and cons of spoken and written feedback.

Table 1. Summary of Pros and Cons of Spoken and Written Feedback.

	Pros	Cons
Spoken Feedback	Rich signal beyond words uttered. Can supply cotemporal and simultaneous feedback about comprehension. Speaker can expect timely response from addressee.	Need to account for irrelevant auditory events. Multiple addressees cannot supply feedback simultaneously. No readily reviewable record of conversation Increased cost for help-seeking.

		Potentially less efficient because it requires social niceties.
Written Feedback	No irrelevant auditory events to accommodate. Readily reviewable record of miscommunication or confusion. Reduced social cost for help-seeking. Multiple matchers can supply feedback simultaneously. Potentially more efficient because it can avoid some social niceties.	No supplementary acoustic information co-occurring with words. Message cannot be read until sent. Writer does not know when addressee will respond to message.

Although grounding constraints provide a framework for thinking about how communication proceeds with different modes of feedback, the effects of these constraints are largely unknown. Directors may respond to the different expectations for feedback by looking or listening for understanding with spoken communication, but explicitly requesting written evidence of understanding with written communication. The explicit providing of confirmation may have consequences for accuracy. Matchers providing explicit responses about their level of comprehension may be more likely to get the help they need. If explicit responses are more common with writing, matchers using written communication should outperform matchers using spoken communication. On the other hand, matchers may be able to more effectively hide their problems with typing than with speaking, and therefore avoid getting the help they need.

The effects of grounding constraints may also vary among conversational groups of different sizes. For example, when only two people are participating in the task, the advantage of having multiple simultaneous messages in a text-chat interface is minimal, but when four are participating, this advantage may be far more important. In one study of writing-only

communication (typing and drawing) in threesomes, 41% of participant contributions were produced in parallel (5% of spoken dialogue overlaps; Whittaker et al., 1991). Similarly, in a one-on-one video-chat in the spoken communication condition, it may be easier for a director to attend to the non-lexical signals of the single matcher than it would be for the director to attend to all three matchers' non-lexical expressions in a four-participant video chat. Thus, the richer signal advantage of spoken communication may be more important in smaller groups, while the advantages of simultaneous text-based feedback may be more important when there are multiple addressees.

One the other hand, group size may have no effect. In one study on the ability to reach consensus, there was no relationship between group sizes of three and six depending on whether participants communicated via face-to-face or text-to-text (McKinlay et al., 1993), although the researchers suggested that one group may have masked a result in the predicted direction of the larger group's performing worse than the smaller (McKinlay et al., 1993).

Active versus Passive Participation

According to the collaborative theory of language use, communication is maximally effective when communicators can provide feedback about their understanding (Clark, 1996). In an interaction, one can participate more or less actively. For example, in a classroom environment, students who ask the instructor several questions should outperform students who passively listen in on lessons but rarely participate (similar to learning by observing or *vicarious learning*, Sutton, 2001). Direct participants (those who actively engage with their interlocutors) can negotiate until they have achieved mutual understanding, while participants in passive roles must do their best with the information they can glean.

In accordance with this prediction, direct conversational participants outperform overhearers in a variety of communicative tasks. Schober and Clark (1989) demonstrated this advantage in a referential card task similar to the one used in this study. Similarly, in a tutorial design featuring an electronic tutor, direct tutorial participants outperformed vicarious participants who were over-hearing and over-viewing the interaction (Craig, Driscoll, & Gholson, 2004). Direct participants also outperformed nonparticipants on a movie retelling task (Kraut, Lewis, & Swezey, 1982) and in a model building task (Clark & Krych, 2004).

Vicarious participants do better when observing behavior produced under active learning conditions, where students answer questions posed by an instructor, engage in problem-solving activities, or otherwise become actively involved (e.g, Mayer, Dow, & Mayer, 2003; Moreno, Mayer, Spires, & Lester, 2001). For example, participants performed better when overhearing more active communication (dialogue) versus more passive communication (monologue). People who listened in on dialogues between two other people doing a referential card task outperformed those who listened in on monologues created by a single instructor (Fox Tree, 1999). People who overheard virtual tutors and tutees displayed greater learning when overhearing information presented in a dialogue format than in a monologue-like format (Driscoll et al., 2003).

While active versus passive has been treated as a dichotomous variable in most research studies, we consider active and passive participation as two poles of a continuum. This has some precedent in the literature – Kraut et al. (1982) had participants listen to movie summaries that were created with full feedback from an addressee, partial feedback (restricted to backchannels and one-word answers), or no feedback. Participants who heard partial or full feedback gave higher quality and more complete retellings of the summaries than those listening in on

summaries created with no feedback. But more importantly, both direct addressees and overhearers provided better summaries when speakers were given more feedback, though the listener providing feedback benefitted more than the overhearer.

Further, while it is a step forward to conceptualize active and passive participation as varying continuously, it is important to remember that the amount of participation is only part of the story. In a classroom, a student who asks "Can you please repeat that?" every four minutes is providing a large amount of participation, but this participation may not be as meaningful or helpful to the class as another student's single contribution that illustrates the connections between the current lecture and a topic from the previous week. In this study, we address this issue in two ways: by counting the number of words produced by each participant in a session, and by categorizing matchers' turns based on the amount of novel material they introduce into the group's ongoing discourse.

While actively participating in communication improves comprehension, the collaborative theory of language use does not make clear predictions about what modality of feedback is most effective for successful instruction. The instructional setting creates a situation where one communicative participant, the instructor, has priority both in terms of the content of the communication (for example, the material to be covered in a lesson plan) and in the priority of the communication (the order in which participant contributions are addressed). Should feedback be in the same mode as instruction, as is usually the case (the instructor and the learners are face-to-face in a traditional classroom)? Or is it better for feedback to be in a mode different from instruction (learners can send their questions and comments as text while the instructors are presenting)?

Current Study

While existing research has described the *media-richness* of various forms of communication (for example, face-to-face communication is more rich than email), few have tested learning across richness rates (see Rau et al., 2008, for review). Furthermore, as outlined above, it is not necessarily the case that a media-rich form of communication will be the most effective in a task-focused environment. Communicative forms have both costs and benefits; for example, grounding a matcher's face-to-face contribution may take more time than grounding a text message contribution.

In the current study, two types of substantive matcher feedback were compared, spoken responses and written responses in a visual, but not auditory, setting. In a group decision-making task, participants reached more consensus face-to-face than text-to-text (McKinlay et al, 1993). Holding visual setting constant allows focusing on spoken versus written feedback. It also models classroom communication where instructors can see a room full of people but not communicate with each of them simultaneously. Communicative effectiveness was assessed by matchers' accuracy in ordering abstract shapes. In addition, the amount and type of addressee feedback was assessed in order to test the role of matchers' participation.

There is reason to expect no difference across conditions. Earlier studies suggested that the primary task (here, identifying abstract shapes) will be prioritized over secondary tasks (such as interpersonal management) in such a way as to ensure success of the primary task regardless of the communicative medium, with differences in media observable through how the communication proceeds rather than communicative success (McCarthy & Monk, 1994). In addition, as people become more accustomed to communicative technologies, they become more

facile at accommodating to grounding constraints (Fox Tree et al., 2011), further suggesting little difference in performance across modalities.

However much work in mediated communicative settings has been done with homo-modal communication; for example, face-to-face or text-to-text, but not face-to-text (cf. Condon & Čech, 2001; Gergle et al., 2004; Hancock, 2004; Hancock & Dunham, 2001a, 2001b; McCarthy & Monk, 1994; McKinlay et al., 1993; Walther, Loh, & Granka, 2005). There have been few tests of hetero-modal mediated communicative settings. In one study of question-asking and answering in homo-modal and hetero-modal settings (e.g., the question was spoken but the answer was written), researchers found an advantage for hetero-modal communication, with more correct answers produced and more errors caught (Chevalier & Fox Tree, 2012). In another study of voice communication versus pen-and-voice communication, having both written and spoken information was more efficient than having written alone or spoken alone (Daly-Jones, Monk, Frohlich, Geelhoed, & Loughran, 1997).

Although these results suggest that hetero-modal mediated communicative settings may be more effective than homo-modal mediated communicative settings, the way communication proceeds may also be an important factor. In a study of communication using shared visual information, participants preferred to negotiate with speech (talking over telephone headsets with the shared whiteboard) rather than by visual input alone (writing and drawing on the whiteboard), even though other activities, such as creating a permanent display of the content of the communication, were similar across the with- and without-speech conditions (Whittaker et al., 1991).

People also may provide different kinds of information depending on whether they are communicating with speech or writing. A study of two- to three-person groups using a text-based

tool to communicate documented that participants produced more explanations (37% of messages) than clarifications (questions and answers, 19%) and elaborations ("relating new information to prior knowledge," 5%; Münzer & Xiao, 2005, p. 386). But dyads carrying out a voice-to-voice referential card task had more similar rates of *requests for clarification*, such as requests for repetitions of what was said (46.5%), and *requests for potential elaboration*, such as asking "Backward?" when provided with the information "It looks like someone falling" (53.5%; Hupet, Chantraine, & Nef, 1993, p. 343). In the current study, we evaluated the rates of requests for more information, elaborations, and alternative descriptions across spoken and written feedback conditions.

It is unknown to what extent richer signals outweigh simultaneity and social advantages. Spoken feedback may be more effective because it is more media-rich. But the potential efficiency of written feedback may make it more effective. There are reasons to believe both that matchers will have an easier time providing feedback when they can speak to directors and that they will have an easier time providing feedback when they can write. Furthermore, the effectiveness of different kinds of feedback may vary depending on the number of matchers. The more feedback provided to directors, regardless of whether it comes from one person or many, the more directors will be engaged in negotiating perspectives, which should benefit all matchers (cf. Fox Tree & Mayer, 2008).

Method

In this experiment, one participant (director) provided instructions for arranging tangram cards via a streaming video chat to either one, two, or three other participants (matchers). In the spoken feedback condition, the matchers provided feedback using the same video chat as the

director; in the written feedback condition, they provided feedback in a group text chat. The participants had five minutes to complete the task.

Time constraints (such as those typical in distance learning settings with an instructor moving through a spoken lecture and students responding with text) may highlight differences that are not observable when communication proceeds until all participants are satisfied with their understanding. Like others testing how media affects communication (Daly-Jones et al., 1997; McCarthy & Monk, 1994), we limited communication time.

Participants

In total, 89 University of California Santa Cruz undergraduates participated in this experiment. Participants received course credit for their time. A breakdown of participants by experimental condition can be found in Table 2. Directors and Matchers were strangers.

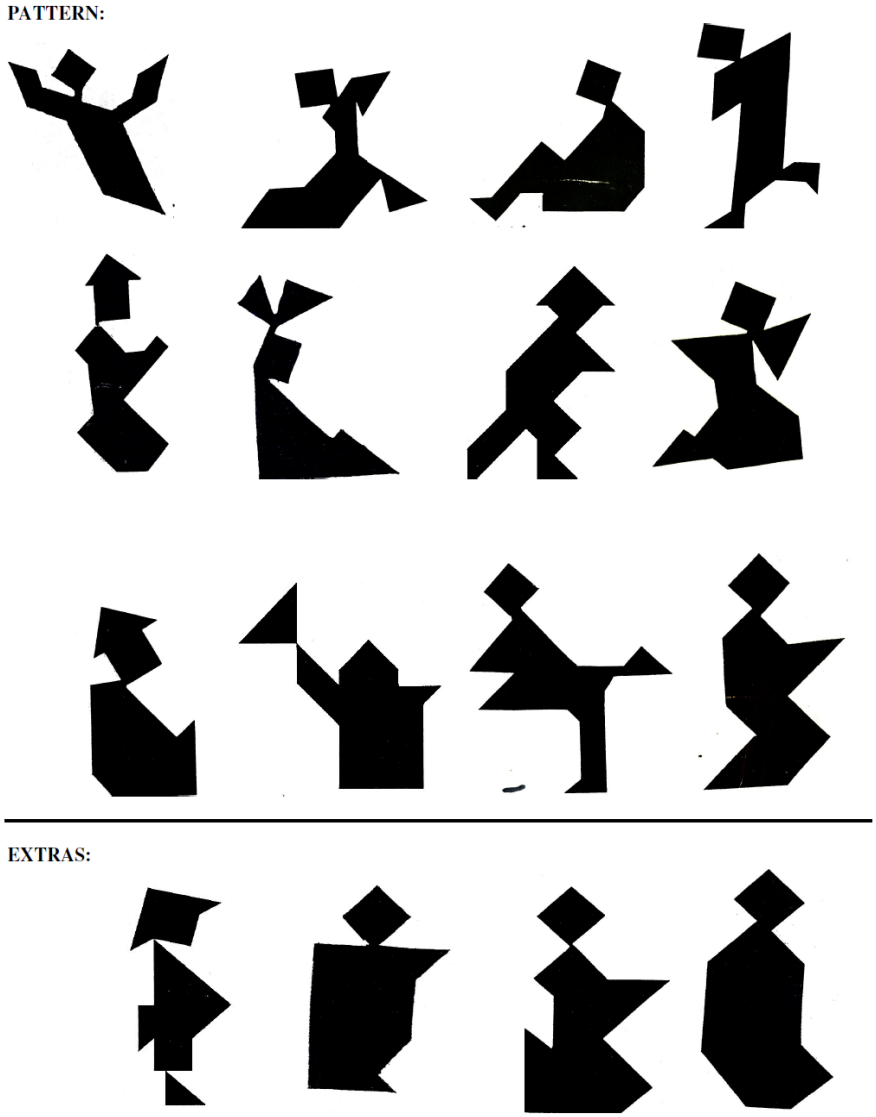
Table 2. Number of groups of each size and total number of participants by feedback type.

	Groups of 2 (1 Director, 1 Matcher)	Groups of 3 (1 Director, 2 Matchers)	Groups of 4 (1 Director, 3 Matchers)	Total Number of Participants
Written	5	5	5	45
Spoken	5	6	4	44

Materials

Directors received a master sheet indicating a 3 x 4 arrangement of shapes with 4 additional distractor shapes. Matchers received 16 cards depicting the 12 target and 4 distractor shapes (see Figure 1).

Figure 1. Stimuli used in experiment.



Procedure

Participants were scheduled in groups of up to 4. Upon arrival at the lab, one participant was randomly assigned to be the director, and the rest, matchers. Participants sat in separate booths that were connected by video chat, and, in the written feedback condition, also connected by a group text chat. Thus, participants could not physically see each other, but saw each other on computer screens. All participants could see all written feedback in the group text chat, and all participants were mutually visible in both feedback conditions.

The video chat window was maximized so that each participant could see the other participants as large as possible. In the written feedback condition, the text window was placed at the bottom of the screen, obscuring only each participant's small panel showing an image of themselves. In the written feedback trials, the matchers' microphones and directors' speakers were turned off so that no audio information could be transmitted. While the participants were visible to one another, the matchers had to type to communicate, which placed constraints on their ability to gesture to each other or the director. That is, in the audio feedback group, participants communicated freely using video chat, but could not use written chat. In the written feedback group, participants could see each other, but only the director's voice could be heard. Matchers were instructed to communicate by written chat. In both conditions, participants were not able to see each other's cards.

After experimenters ensured that all connections among booths were working properly, they gave the sheet with the target card arrangement to the director, and told the director to begin. The experimenter entered the director's booth and stopped the session after five minutes.

Screen-capture software SnapzProX (Ambrosia Software Inc., 2008; http://www.ambrosiasw.com/utilities/snapzprox/) was used to record the sessions. This software takes a video of the screen of the computer from which it is recording, and can add both the audio input and audio output of that computer as the audio track for the video. This way, recording a session from a single computer could capture all of the text and audio information the matchers provided, as well as the director's descriptions.

Data Scoring and Coding

Following the session, each matcher was scored for the number of cards attempted, correctly placed, and incorrectly placed. Recordings were transcribed to calculate a *participation*

index, defined as the number of words produced by a participant divided by the total number of words produced by all members of the group. For the spoken feedback group, the participation index was the number of words spoken by the matcher divided by the total number of words spoken. For the written feedback group, the participation index was the number of words written by the matcher divided by the total number of words produced (number of matchers' written words plus number of director's spoken words).

In addition to coding the amount of participation from each matcher and the director, we also coded five categories of matcher feedback: (a) acknowledging understanding, for example, nodding yes or saying, "Got it;" (b) signaling lack of understanding, for example, a quizzical eyebrow movement or saying, "huh?"; (c) requesting more information, for example, asking, "What do the arms look like?"; (d) elaborating ongoing description, for example, asking, "Is one leg thicker than the other?"; or (e) suggesting new/alternative description, "Does the waving guy kind of look like a watering can?" The distinction between an alternative description and an elaboration hinged on whether the contribution reconceptualized the item and described it from a different perspective, or whether the contribution suggested additional details that were consistent with the current perspective. Fox Tree and Mayer (2008) found that participants who overheard descriptions with multiple as opposed to single perspectives performed better at identifying abstract shapes similar to those used in our task.

Results

Table 3 displays means and standard deviations of participants' performance in the referential card task. To avoid violations of independence, all inferential statistics reported below treat each group as one datapoint, whether that group consists of 2, 3, or 4 individuals.

Table 3. Mean (SD) of correct card placements and errors by group size and feedback type.

	2 (1 Director, 1 Matcher)		3 (1 Director, 2 Matchers)		4 (1 Director, 3 Matchers)		Summed across group sizes	
	Correct	Errors	Correct	Errors	Correct	Errors	Correct	Errors
Written	4.0 (2.2)	3.4 (2.1)	4.6 (2.3)	1.7 (1.4)	5.8 (2.9)	2.1 (1.5)	4.8 (2.4)	2.4 (1.7)
Spoken	9.2 (2.3)	1.4 (1.1)	5.2 (1.4)	2.1 (1.4)	5.5 (1.0)	2.1 (2.9)	6.6 (2.5)	1.9 (1.7)

Two-way between-groups ANOVAs were conducted on each group's number of cards correct and their errors, using group size (2, 3, or 4 participants), feedback modality (text or video), and their interaction as independent variables. The ANOVA on success rate reached overall significance, $F(5,24) = 3.70$, $p = .013$. There was a main effect of feedback modality, $F(1,24) = 5.43$, $p = .029$, and an interaction between group size and feedback modality, $F(2,24) = 4.73$, $p = .019$. Follow-up t-tests revealed that, for groups of one director and one matcher, groups in the spoken feedback condition ($M = 9.2$, s = 2.3) outperformed those in the written feedback condition ($M = 4.0$, $s = 2.2$), $t(8) = 3.64$, $p = .007$. The differences in performance between groups using spoken feedback and those using written feedback at larger group sizes were not significant, and thus it appears that the substantial advantage for spoken feedback in a single-addressee setting drove the significant main effect for feedback modality.

The ANOVA conducted on errors by feedback modality and group size did not reach overall significance, $F(5,24) = 0.76$, $p = .588$, indicating that errors were equiprobably distributed across groups of different sizes, using different feedback modalities.

There were also substantial differences in the strategies participants used to negotiate card identities and locations, even among groups with similar outcomes. Table 4 displays the means and standard deviations of the total words produced in each session, and of the directors' participation index (i.e., the portion of the total words produced by the director). Larger

participation index values represent the directors' producing a larger share of the total words produced during the session.

Table 4. Mean (SD) word count and director participation index by group size and feedback type.

	2 (1 Director, 1 Matcher)		3 (1 Director, 2 Matchers)		4 (1 Director, 3 Matchers)		Summed across group sizes	
	Words	Dir PI	Words	Dir PI	Words	Dir PI	Words	Dir PI
Written	417 (69)	.90 (.06)	453 (97)	.86 (.06)	554 (74)	.90 (.04)	475 (96)	.89 (.05)
Spoken	782 (50)	.73 (.07)	688 (96)	.75 (.08)	644 (83)	.74 (.20)	708 (94)	.74 (.11)

Two two-way between-groups ANOVAs were conducted, on the total words produced in each session, and on the director's participation index from each session, using group size, feedback modality, and their interaction as independent variables. The ANOVA on total words produced reached overall significance, $F(5, 24) = 15.51$, $p < .001$. There was a main effect for feedback modality, $F(1, 24) = 60.37$, $p < .001$, and an interaction between feedback modality and group size, $F(2, 24) = 6.08$, $p = .005$. More words were produced in sessions using spoken feedback than in those using written feedback. Across spoken-feedback groups, larger group sizes had fewer total words. In contrast, across written-feedback groups, larger group sizes had more total words. The most words were produced in one-on-one, spoken feedback sessions, while the least words were produced in one-on-one, written feedback sessions.

In terms of the participation index data, the ANOVA reached overall significance, $F(5, 24) = 3.96$, $p = .009$. The only significant effect was the main effect of feedback modality, $F(1, 24) = 18.79$, $p < .001$. Directors produced a greater share of the words in written feedback groups than spoken ones. Conversely, matchers produced a greater share of the words in spoken feedback groups as compared to written ones.

Further, the participation index of a session's director was negatively correlated with the group's success at the task (Pearson's $r = -.40$, $df = 29$, $p = .029$), indicating that the more matchers contributed to the group's discourse, the better the group did as a whole. Given this correlation and the significant effect of feedback modality on the director's participation index, one might expect spoken-feedback matchers to outperform their written-feedback counterparts; however, as discussed above, this was only the case in the single-addressee sessions.

To help resolve this apparent paradox, one must remember that participating actively means more than just producing many words. Accordingly, we also examined the different types of contributions provided by matchers across the two feedback conditions. As previously discussed, we coded three specific types of behavior used by matchers to negotiate grounding: turns that requested more information without contributing new content, turns that elaborated on current descriptions, and turns that suggested alternative descriptions. Table 5 displays the average number of times each behavior appeared in a session, across feedback modality and group size.

Table 5. Mean (SD) of sessions' grounding behavior by group size and feedback type.

		Group size			
Feedback type	Behavior type	2 (1 Dir., 1 Matchers)	3 (1 Dir., 2 Matchers)	4 (1 Dir., 3 Matchers)	Summed across group sizes
Written	Request	1.8 (0.8)	1.4 (1.1)	1.8 (1.5)	1.7 (1.1)
	Elaboration	1.8 (0.8)	2.4 (2.0)	3.4 (1.1)	2.5 (1.5)
	Alt. description	0.0 (0.0)	0.2 (0.5)	0.0 (0.0)	0.1 (0.3)
Spoken	Request	1.8 (1.5)	2.3 (1.9)	2.5 (2.1)	2.2 (1.7)
	Elaboration	5.2 (2.4)	5.8 (2.7)	6.3 (4.9)	5.7 (3.1)
	Alt. description	2.8 (3.0)	2.2 (2.3)	1.3 (1.0)	2.1 (2.2)

Note: This table displays the average number of times a grounding behavior was used per session, not the average number of times an individual matcher used these behaviors.

Because these behaviors are discreet events being categorized and counted, rather than proportions of correct and incorrect responses, ANOVAs are not appropriate for examining their distribution across different group sizes and feedback modalities. Instead, we used hierarchical loglinear analysis. This procedure for dealing with multidimensional categorical data is analogous to Pearson's χ^2, but extended to more than two categorical dimensions (Tabachnick and Fidell, 2007, p. 858). For this analysis, we constructed a three-dimensional contingency array, comprising the total counts of each behavior type, in each combination of group size and feedback modality. This 3x3x2 contingency array serves as the observed values for all loglinear models, as well as the expected values for the fully-saturated model, which includes all three main effects of behavior type, group size, and feedback modality, all three two-way interactions, and the three-way interaction.

Following typical procedure for hierarchical loglinear analysis (Tabachnick and Fidell, 2007, p. 873), we started with the fully saturated model and iteratively removed a single term with the highest order among remaining terms (i.e., first removing the three way interaction, then each two way interaction, then each main effect). When the likelihood ratio G^2 statistic (analogous to Pearson's χ^2) indicated that the current model provided significant departure from an adequate fit for the data, we returned the most recently removed term and considered the resulting model to be the most parsimonious model still providing adequate data fit.

By this procedure, we converged on a model containing the three main effects, the two-way interaction of feedback modality and behavior type, and the two-way interaction of feedback modality and group size.[1] This model adequately fit the data, Likelihood-ratio G^2 (8) = 7.13, p =

[1] Like Pearson's χ^2, loglinear analysis also requires that all expected values be greater than 1, and that 80% of expected values be greater than 5. Given that there were no alternative descriptions observed in written feedback groups of 2 or 4 people, this assumption has been violated, which may lead to a loss of power. Accordingly, the same analysis was repeated after collapsing elaborations and alternative descriptions into a single category. Because

.523. The two-way interaction of feedback modality and group size indicates that, collapsed across our three behavior types, the amount of behavior being counted in groups of a given size depended on whether those groups were using written or spoken feedback. This makes sense given the interaction effect of group size and feedback modality in the ANOVA on words produced in each session – more total words results in more contributions of any type. Because this interaction effect is not directly relevant to our questions about the *types* of contributions matchers made, it is not further discussed.

The other interaction effect, of feedback modality and behavior type, indicates that, collapsed across different group sizes, the distributions of our three behaviors of interest varied across groups using written and spoken feedback. In particular, the odds of a bare request for information were 2.29 times larger for the written feedback condition than the spoken feedback condition, while the odds of an alternative description were only 0.06 times as large for written than spoken feedback (or conversely, the odds of an alternative description were 16.95 times larger for spoken than written feedback). For one-on-one groups, the odds of a bare request for information were 4.44 times larger in the written feedback condition than for spoken feedback. These distributions suggest that simply requesting more information may not be as helpful as providing additional content in a turn. The fact that only one turn across all group sizes in the written feedback condition contained an alternative description makes a statistical analysis of the role of alternative descriptions in task success impossible; however, it is worth noting that the largest count of alternative descriptions was for the one-on-one spoken feedback groups, which also had the most cards correctly placed.

---

this 2x3x2 hierarchical loglinear analysis converged on the same most parsimonious model (including the three main effects, the two-way interaction of feeback modality and group size, and the two-way interaction of feedback modality and behavior type), we report on the 3x3x2 model.

Discussion

This study demonstrates that, for task-oriented, one-on-one conversations, interlocutors communicated more effectively when given a channel for spoken feedback, but for task-oriented conversations including more people, spoken and written feedback were equally effective. This is consistent with the collaborative theory of language use, and can easily be explained in terms of the constraints on grounding imposed by the communicative modality. In one-on-one conversations, the rich signal advantage for spoken communication is present, but there is no corresponding benefit from the possibility of simultaneous feedback in a text setting, because there is only ever one person who could be giving feedback at a time.

Grounding constraints also easily explain the lack of differences observed for conversations including more people across written and spoken feedback conditions. For conversations of three or four people, the richer signal of spoken feedback balanced out the persistent record of communication and simultaneous response capability of written feedback.

Overall, this is good news for distance learning settings where spoken feedback from many learners is not possible, but written feedback is. When online courses include synchronous interaction, these conversations usually include many participants. There was no clear cost to communicative effectiveness when restricting learners' communication to written feedback in a setting with time pressure and more than one addressee.

When spoken feedback was available, however, there was a communicative advantage for group feedback. Matchers in the spoken feedback condition participated in communicative exchanges about three times as frequently as matchers in the written-feedback condition, and these matchers' turns also contributed novel information to the discourse more frequently than

did written matchers' turns. These results align with findings that people prefer to negotiate with speech (Whittaker et al., 1991).

While differences in the interactivity of groups larger than two were unrelated to task performance with the current referential card task, in other tasks that rely more on generating creative ideas, the interactivity advantage might lead to improved task performance. In the whiteboard study, participants only doodled on the whiteboard when they had a speech channel available (Whittaker et al., 1991); the researchers suggested that this was because without speech, doodling may have been taken as a meaningful contribution to the communication, or taken up whiteboard space needed for meaningful contributions. Doodling may be helpful for creativity (Schott, 2011), and thus a task that leads to more doodling, or in our case more interactivity, may also lead to more creativity. It's also possible that without the visual component, the written feedback condition might have resulted in worse communicative outcomes than we observed here, allowing for an even greater advantage for spoken feedback conditions over written-but-not-visual feedback conditions. Written-but-not-visual is the format typically found in distance learning educational programs.

In conclusion, results demonstrate that written communication is not necessarily an impoverished communicative medium, and that it can be just as effective as spoken communication for conversations with more than two people. At the same time, if spoken feedback is possible, greater overall levels of interactivity are possible.

References

Ambrosia Softworks, Inc. (2008). *SnapzPro X*, v. 2.0.3, http://www.ambrosiasw.com/utilities/snapzprox/.

Chevalier, J. W. & Fox Tree, J. E. (2012). Using heteromodal communication to optimize knowledge and awareness. *American Journal of Psychology, 125*(3), 361-368.

Clark, H. H. (1996). *Using Language*. New York: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 127-149). Washington, D. C.: American Psychological Association.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*, 62-81.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Condon, S., & !ech, C. (2001). Profiling turns in interaction: Discourse structure and function. In *Proceedings of the 34th Hawai'i International Conference on Systems Sciences, 4*. Washington, D.C.: IEEE Computer Society Press.

Craig, S., Driscoll, D. & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia, 13*(2), 163-183.

Daly-Jones, O., Monk, A., Frohlich, D., Geelhoed, E., Loughran, S. (1997). Multimodal messages: The pen and voice opportunity. *Interacting with Computers, 9*, 1-25.

Driscoll, D. M., Craig, S. D., Gholson, B., Ventura, M., Hu, X., & Graesser, A. C. (2003). Vicarious learning: effects of overhearing dialog and monologue-like discourse in a virtual tutoring session. *Journal of Educational Computing Research, 29*(4), 431-450.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes, 27*, 35-53.

Fox Tree, J. E., & Mayer, S. A. (2008). Overhearing single and multiple perspectives. *Discourse Processes, 45*(160-179).

Fox Tree, J. E., Mayer, S. A., & Betts, T. E. (2011). Grounding in instant messaging. *Journal of Educational Computing Research, 45*(4), 455-475.

Gergle, D., Millen, D., Kraut, R. E., & Fussell, S. R. (2004). Persistence matters: Making the most of chat in tightly-coupled work. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 431-438). New York: ACM Press.

Hancock, J. T. (2004). Verbal irony use in face-to-face and computer-mediated conversations. *Journal of Language and Social Psychology, 23*(4), 447-463.

Hancock, J. T. & Dunham, P. J. (2001a). Language use in computer-mediated communication: The role of coordination devices. *Discourse Processess, 31*(1), 91-110.

Hancock, J. T. & Dunham, P. J. (2001b). Impression formation in computer-mediated communication revisited: An analysis of the breadth and intensity of impressions. *Communication Research, 28*(3): 325-347.

Honeycutt, L. (2001). Comparing e-mail and synchronous conferencing in online peer response. *Written Communication, 18*(1), 26-60,

Hupet, M., Chantraine, Y., & Nef, F. (1993). References in conversation between young and old normal adults. Psychology and Aging, *8*(3), 339-346.

Kraut, R. E., Lewis, S. H., & Swezey, L. W. (1982). Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology, 43*(4), 718-731.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: what works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*, 806-813.

McCarthy, J. C. & Monk, A. F. (1994). Measuring the quality of computer-mediated communication. *Behaviour & Information Technology, 13*(5), 311-319.

McKinlay, A., Arnott, J., Proctor, R., Masting, O., & Woodburn, R. (1993). A study of turn-taking in a computer-supported group task. *People and Computers HCI93 Conference.* Cambridge: Cambridge University Press.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*, 177-213.

Münzer, S. & Xiao, B. (2005). Small groups learning synchronously online at the workplace: The interaction of factors determining outcome and acceptance. *Journal of Universal Computer Science, 11*(3), 378-393.

Rau, P.-L. P., Gao, Q., & Wu, L.-M. (2008). Using mobile communication technology in high school education: Motivation, pressure, and learning performance. *Computers & Education, 50*, 1-22.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*(2), 211-232.

Schott, G. D. (2011). Doodling and the default network of the brain. *The Lancet, 378*(9797), 1133 - 1134.

Sutton, L.A. (2001). The principle of vicarious interaction in computer-medicated communications. *International Journal of Educational Telecommunications, 7,* 223-242.

Tabachnick, B.G., and Fidell, L.S. (2007). Multiway frequency analysis. In B.G. Tachnick and L.S. Fidell, *Using Multivariate Statistics, 5th Edition* (pp. 858-912). San Francisco: Pearson.

Walther, J. B., Loh, T., & Granka, L. (2005). Let me count the ways: The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology, 24*(1), 36-65.

Whittaker, S. J., Brennan, S. E., & Clark, H. H. (1991). Coordinating activity: An analysis of interaction in computer-supported cooperative work. Proceedings, CHI '91, Human Factors in Computing Systems , pp. 361-367. New Orleans, LA: Addison-Wesley.