Team Communication as a Collaborative Process

Felix Gervits[1]*, Kathleen Eberhard[2] and Matthias Scheutz[1]

[1] *Human-Robot Interaction Laboratory, Department of Computer Science, Tufts University, Medford, MA, USA,*
[2] *Department of Psychology, University of Notre Dame, Notre Dame, IN, USA*

Communication channels can reveal a great deal of information about the effectiveness of a team. This is particularly relevant for teams operating in performance settings, such as medical groups, military squads, and mixed human–robot teams. Currently, it is not known how various factors, including coordination strategy, speaker role, and time pressure, affect communication in collaborative tasks. The purpose of this paper is to systematically explore how these factors interact with team discourse in order to better understand effective communication patterns. In our analysis of a corpus of remote task-oriented dialog (cooperative remote search task corpus), we found that a variety of linguistic- and dialog-level features were influenced by time pressure, speaker role, and team effectiveness. We also found that effective teams had a higher speech rate and used specific grounding strategies to improve efficiency and coordination under time pressure. These results inform our understanding of the various factors that influence team communication and highlight ways in which effective teams overcome constraints on their communication channels.

Keywords: team communication, common ground, disfluency, workload, time pressure

INTRODUCTION

Grounding in Task-Oriented Dialog

The success of a team largely depends on how well teammates can coordinate their actions in an efficient manner. However, coordination can be difficult when teammates must dynamically adapt their decision-making, communication, and planning strategies (Serfaty et al., 1993). Sycara and Sukthankar (2006) have identified several capabilities that a team needs to plan and coordinate its actions effectively. These include an overall intention to execute a plan (*joint intention*), sharing of goals, plans and knowledge of the environment (*common ground*), and awareness of the roles and responsibilities, as well as the capabilities and limitations of one's teammates (*team awareness*). Of these elements, perhaps the most important for team success is establishing and maintaining common ground. Common ground facilitates efficient communication, particularly for teams working under stress, by serving as a mutual knowledgebase from which information about goals, plans, and perspectives may be shared (Clark, 1996).

In teams with open communication channels, common ground is often established through dialog interaction. The process by which this occurs is known as *grounding* and consists of two phases – the Presentation Phase and the Acceptance Phase (Clark and Schaefer, 1989). In the Presentation Phase, a speaker makes an utterance and seeks confirmation from the listener that the utterance was understood. This confirmation of understanding comprises the Acceptance Phase, wherein the listener provides evidence that they understood the message. Such evidence can take the form of an

overt acknowledgment (e.g., "Okay," "Got it," "Mhm," etc.), the relevant next contribution in the exchange (e.g., responding to a question), or simply through continued attention. Exchanges that contain a Presentation and Acceptance Phase are known as *contributions* and form the essential unit by which grounding occurs. "*Contributions, therefore are different from most standard linguistic units. They are not formulated autonomously by the speaker according to some prior plan, but emerge as the contributor and partner act collectively. Success depends on the coordinated actions by the two of them*" (Clark and Schaefer, 1989, pp. 292).

Grounding is especially important in task-oriented dialog, in which people need to use language to coordinate their actions and accomplish a joint task. Consider the following example from the cooperative remote search task (CReST) corpus (Eberhard et al., 2010) where a *director* (*D*) and a *searcher* (S) must communicate (remotely) to locate various colored boxes in a novel environment[1]:

> *Presentation Phase:*
>   D: Do you know where the s- the- where's the sixth green box?
> *Acceptance Phase:*
>   *Side sequence:*
>   S: Um [pause] u:m [pause] wait where number six is?
>   D: Yeah
>   S: Number six is in a room where there's like a pink box
>   D: Okay

Due to the highly disfluent Presentation Phase, the searcher is not able to understand what was said, and so must initiate a clarification request through the use of a *side sequence*. The Acceptance Phase thus consists of this side sequence, followed by the searcher's response to the initial question and the director's final acknowledgment ("Okay"). Though it took multiple turns, this exchange is considered a contribution since an utterance was both presented and accepted. At the end of this contribution, the team's common ground is successfully updated with mutual knowledge of the location of the sixth green box.

The extent to which common ground can be efficiently established and maintained depends largely on the constraints of the particular medium of interaction (Clark and Brennan, 1991). This is closely related to the *principle of least collaborative effort* (Clark and Wilkes-Gibbs, 1986), which holds that people seek to minimize the joint effort needed to ground a communicative exchange in a particular medium. For example, face-to-face interaction provides the most reliable basis for grounding due to factors, such as gaze direction and eye contact, which facilitate joint attention and permit the use of deictic (e.g., pointing) gestures to signal reference. Interaction in the email medium is further constrained due to lack of cotemporality and sequentiality, which requires people to adapt to longer turns and delayed feedback.

In mediums where verbal communication occurs simultaneously and remotely (e.g., CReST example above), people must manage additional constraints due to lack of co-presence and visual access. Gestures cannot be reliably used as communicative devices, so people would have to adapt by using additional words to describe what would ordinarily be signaled visually (Doherty-Sneddon et al., 1997). This presents a challenge to the interlocutors since verbal descriptions are not as reliable a basis for grounding as shared perception and joint attention. In sum, conversational exchanges in any medium involve a trade-off in effort, but people seek to minimize the overall *collaborative* effort needed to manage the constraints.

Constraints on Remote Communication

Teams interacting in settings where access to shared visual information is limited must dynamically adapt to various constraints on their communication channels. These constraints necessitate the use of additional techniques to manage the costs of grounding. To evaluate some of these grounding techniques, Clark and Wilkes-Gibbs (1986) asked participants to verbally describe ambiguous tangram shapes to a partner looking at the same shapes across an opaque screen. The results showed that in cases when the initial noun phrase description was not sufficient to resolve the referent, people employed a range of techniques to establish common ground, including (1) *self-repairs* – the director would modify his initial utterance to better describe the object, (2) *expansions* – the director would add an additional description at the end of the noun phrase to further elaborate what he meant, and (3) *replacement* – the matcher would reject the director's description and offer one from her own perspective that was also compatible; this replacement would then be accepted by the director.[2] The results also showed that directors did not initially refer to a figure from an egocentric perspective (e.g., "the ice skater"), but rather described the figure in more general terms on the first trial (e.g., " … looks like a person who's ice skating, except they're sticking two arms out in front"). This was done to establish a common perspective, which was then utilized to minimize collaborative effort in subsequent trials. In general, directors initially used more words and turns to describe a referent to a partner, but once the referent had been established in common ground, they tended to shorten their description. This fits in line with the principle of least collaborative effort and shows how people adapt to the constraints of the medium to efficiently establish common ground.

In another study examining the role of interaction and grounding in task performance, Clark and Krych (2004) carried out a collaborative building task in which a *director* was tasked with instructing a *builder* to construct various Lego models. In one condition, the builder's workspace was visible to the director, and in another condition it was not visible. In another non-interactive condition, the director was not even physically present. Instead,

---

[1]In the dialog examples, asterisks (*) indicate simultaneous speech, hyphens (-) indicate repaired segments, colons (:) indicate prolongations, commas indicate brief silent pauses, and longer pauses are indicated in brackets. For readability, the director will be referred to as male and the searcher as female, although the gender distribution varied between teams.

[2]An example of a *replacement* can be seen in the dialog example from the CReST corpus above. The searcher referred to the sixth green box as "number six" – and this new referent was then established in the team's common ground after the director's acknowledgment ("Yeah").

the builder listened to an audio recording of the instructions while constructing the model. The results showed that (not surprisingly) teams performed best when they had full visual access to their partners and could interact with them as needed. Teams were much less efficient when they could not monitor one another's workspaces, and they made eight times as many errors when there was no interaction between partners. This latter finding is significant because it suggests that when directors spoke without monitoring their listener (e.g., audio recording) their instructions were not easily followed. These results point to the important role that feedback and interaction play in the grounding process and show how these factors can affect team performance.

Another factor that can affect communication in remotely communicating teams is time pressure. Time pressure can increase cognitive workload by causing mental stress that disrupts a wide range of team-related factors, including planning and coordination (Entin and Serfaty, 1999). Language is particularly affected by cognitive load and can manifest through disfluencies, speech rate, and dialog-level properties (e.g., responsiveness, agreement, cohesion). In terms of disfluencies, Berthold and Jameson (1999) found that fragments, false-starts, repetitions, and pauses increased with higher workload, whereas speech and articulation rate were found to decrease. However, other studies have suggested that increased disfluency rate may be due to coordination processes rather than planning difficulties associated with workload (Swerts, 1998; Bortfeld et al., 2001; Clark and Krych, 2004). In terms of speech, Lively et al. (1993) found that utterance length decreases with workload, but more recent work by Khawaja et al. (2012) found *increases* in sentence length as well as in speech rate, words indicating disagreement, and the usage of plural personal pronouns (e.g., "we" and "us"). Finally, Urban et al. (1995) found that under high workload, effective teams asked fewer questions, made fewer requests, and made fewer responses to requests.

These findings should be interpreted with caution because the corresponding studies varied widely in task domain, team size, and team structure. Due to these conflating factors, the effects of workload on dialog and coordination in remotely communicating teams remains unclear. It is also unclear how factors such as team structure and speaker role interact with workload, and if effective teams can overcome the negative constraints on their communication channels by increasing collaboration in specific ways. Our study was designed to address these gaps in the empirical literature.

Present Study
Motivation
The main purpose of the study is to investigate effective communication in performance teams and determine to what extent various factors can influence this process. Of particular interest was the effect of workload, speaker role, interaction strategies, and grounding techniques. To explore these factors, we analyzed the annotated CReST corpus (Eberhard et al., 2010), which contains task-oriented dialog between two humans in a director/searcher hierarchical structure. Dialog was spontaneous and unscripted, with teammates communicating *via* remote headset in order to coordinate their actions and achieve a variety of objectives within a set time limit. Time pressure was introduced 5 min into the task, requiring the team to complete an additional objective with a timer counting down the remaining time.

Predictions
One specific question of interest is how time pressure and speaker role impact remote communication. Although some studies showed a decrease in speech rate, utterance length, and overt communication with increasing workload (Serfaty et al., 1993; Entin and Serfaty, 1999), these studies predominantly involved face-to-face interaction. However, there is evidence that remote communication and lack of visual monitoring impose additional constraints on communication. In one study, Krauss and Weinheimer (1966) found that people communicating by remote channels used more words when they received reduced or no concurrent feedback from the listener. Other findings suggest that people require additional verbal feedback in remote communication (Doherty-Sneddon et al., 1997). Lack of visual monitoring can also lead to decreased task efficiency, an increase in errors, and an increase in disfluencies (Clark and Krych, 2004). For these reasons, we expected speech rate and disfluency rate to increase with time pressure, particularly for directors because of their greater role in managing the task. Moreover, we expected directors to exhibit greater initiative under time pressure as reflected in an increase in directive utterances, with searchers exhibiting a corresponding increase in receptive utterances, such as replies and acknowledgments. These predictions are based on prior studies on two-person collaborative tasks, which show that directors typically take more initiative (Bortfeld et al., 2001; Clark and Krych, 2004).

It is important not to isolate the influence of speaker role and time pressure, as we also expect that these factors would strongly interact with grounding and coordination strategy. Currently, it is unknown how different coordination strategies may affect performance in teams with asymmetrical roles. We predict that effective coordination strategies will involve a sharing of the task responsibilities, as this sharing has been previously shown to distribute the effects of workload between teammates (Khawaja et al., 2012). For example, a strategy in which the director and searcher manage the task demands equally might be more effective than a strategy in which one role takes control. However, it is possible that strategies involving the searcher taking more initiative may also be effective, as this would allow the searcher to describe her location while simultaneously navigating the environment. Another possibility is that effective teams would overcome the deleterious effects of workload by switching to a more implicit mode of coordination (Serfaty et al., 1993; Entin and Serfaty, 1999). For example, Orasanu (1990) found that successful teams planned more during low workload periods, enabling them to use this built-up shared understanding (common ground) to adapt to increasing workload without the need for explicit communication. However, given the difficulty of the CReST task, it is predicted that task demands will increase the need for planning, and therefore, more *explicit* communication.

In general, we predict that teams that establish common ground with respect to objects and locations in the environment

should perform better than teams that fail to do so. Evidence for this may be found in the teams' distribution of conversational moves. One example of this would be speakers showing a greater responsiveness to their teammate by consistently seeking confirmation that a message was understood or that an action was successfully accomplished. Consequently, the receptive partner would make more dialog moves that signal understanding. Additional evidence of effective grounding can be obtained through analyzing disfluent speech. Though disfluencies may indicate production difficulty due to workload (Berthold and Jameson, 1999), they have also been shown to provide an interpersonal benefit (Bortfeld et al., 2001; Arnold et al., 2007). For example, Clark and Krych (2004) found that rates of self-repair disfluencies were high in a collaborative task, likely due to increased coordination and planning. Clark and Wilkes-Gibbs (1986) showed that in a referential communication task, people made many self-repair errors in the process of adjusting their verbal descriptions to a partner's perspective. Importantly, it has been demonstrated that the presence of disfluency in spontaneous speech does not negatively impact comprehension (Brennan and Schober, 2001). For these reasons, we expect effective teams to make more self-repair disfluencies. Successful grounding, either through specific conversational moves or self-repairs, would enable misunderstandings to be repaired before they can build up and affect performance (Levelt, 1983). We predict that effective teams would sustain this collaborative interaction even under time pressure.

MATERIALS AND METHODS

Approximately 8 min of annotated natural language data was extracted from each team of 10 teams in the CReST corpus (Eberhard et al., 2010), for a total of 20 dialogs, 2712 utterances, and 15194 words. The corpus was annotated for speech events, disfluencies, and dialog moves (see Disfluency and Speech Coding and Dialog Structure Annotation below), and teams were assigned to one of two performance groups (Effective vs. Ineffective) based on their objective score on the task (see Team Effectiveness).

Participants

The 10 teams (20 individuals) from the 2010 corpus were analyzed. The participants were all college-aged (19–25 years old) native speakers of English and were paid up to $10 for their participation. Five of the 10 teams were previously acquainted (either friends or roommates), while the other five teams were unacquainted. Six of the teams had a homogenous gender composition (two M/M; four F/F), while the other four had a mixed-gender composition. Five of the teams had a female director and the other five had a male director.

Cooperative Remote Search Task

In the task, pairs of individuals worked together to explore a physical environment and achieve a variety of objectives while under time pressure. One teammate was assigned the *Director* role while the other was designated as the *Searcher*. The director was not physically present in the environment, but rather was seated in front of a computer with an on-screen map of the environment and a headset with which to communicate with the searcher remotely. By contrast, the searcher was situated in the environment and had to interact with physical objects while communicating with her partner *via* headset. The environment consisted of a hallway and six connected office rooms which contained various colored boxes: eight empty green boxes, eight blue boxes (numbered 1–8) containing three colored blocks, eight empty pink boxes, and a cardboard box.

Several objectives needed to be completed within the time limit of 8 min. One objective was for the searcher to locate the cardboard box and place blue blocks from each of eight blue boxes into the target box. The director's map of the environment contained most of the locations of these blue boxes, but not all of them (see **Figure 1**). Another objective was for the director to mark on his map the location of eight green boxes. Since this information was only available to the searcher *via* exploration, a high degree of information exchange was required. To examine performance under time pressure, teams were given an additional set of instructions after 5 min. They were told that in the remaining 3 min they had to complete all previous objectives as well as a new one. This new objective required them to collect yellow blocks from the blue boxes and place them into the eight pink boxes. During these final 3 min, a timer was displayed on the director's screen counting down the remaining time.

Disfluency and Speech Coding

Disfluencies in the natural language annotations were coded based on the HCRC Disfluency Coding Manual (Lickley, 1998) and included *repetitions* (e.g., "in the b- box"), *substitutions* (e.g., "through the door-, er, window"), *insertions* (e.g., "walk to the door – the nearby door"), and *deletions* (e.g., "look toward the – just go back"). Pauses were common in the corpus but were not included in the analysis as they have been shown to reflect different kinds of processes (Swerts, 1998; Nicholson et al., 2010). Disfluency rates were reported for each participant as a proportion per every 100 words. *Speech rate* (words/minute, or w.p.m.) and *mean length of utterance* (MLU, or average number of words per turn at talk) were also calculated.

Dialog Structure Annotation

Each utterance in the transcribed annotation was coded as a type of conversational move using the scheme from Carletta et al. (1997). This scheme views dialog as a conversational game, wherein exchanges serve to fulfill some mutual purpose. Each utterance in this game is called a *conversational move* and can fit into one of three broad categories: *Initiation*, *Response*, and *Ready*. A conversational game is started with an *Initiation* move (see **Table 1**), which can be either a command (*Instruct*), a Wh- or Yes/No question (*Query*), or a statement (*Explain*). Queries can
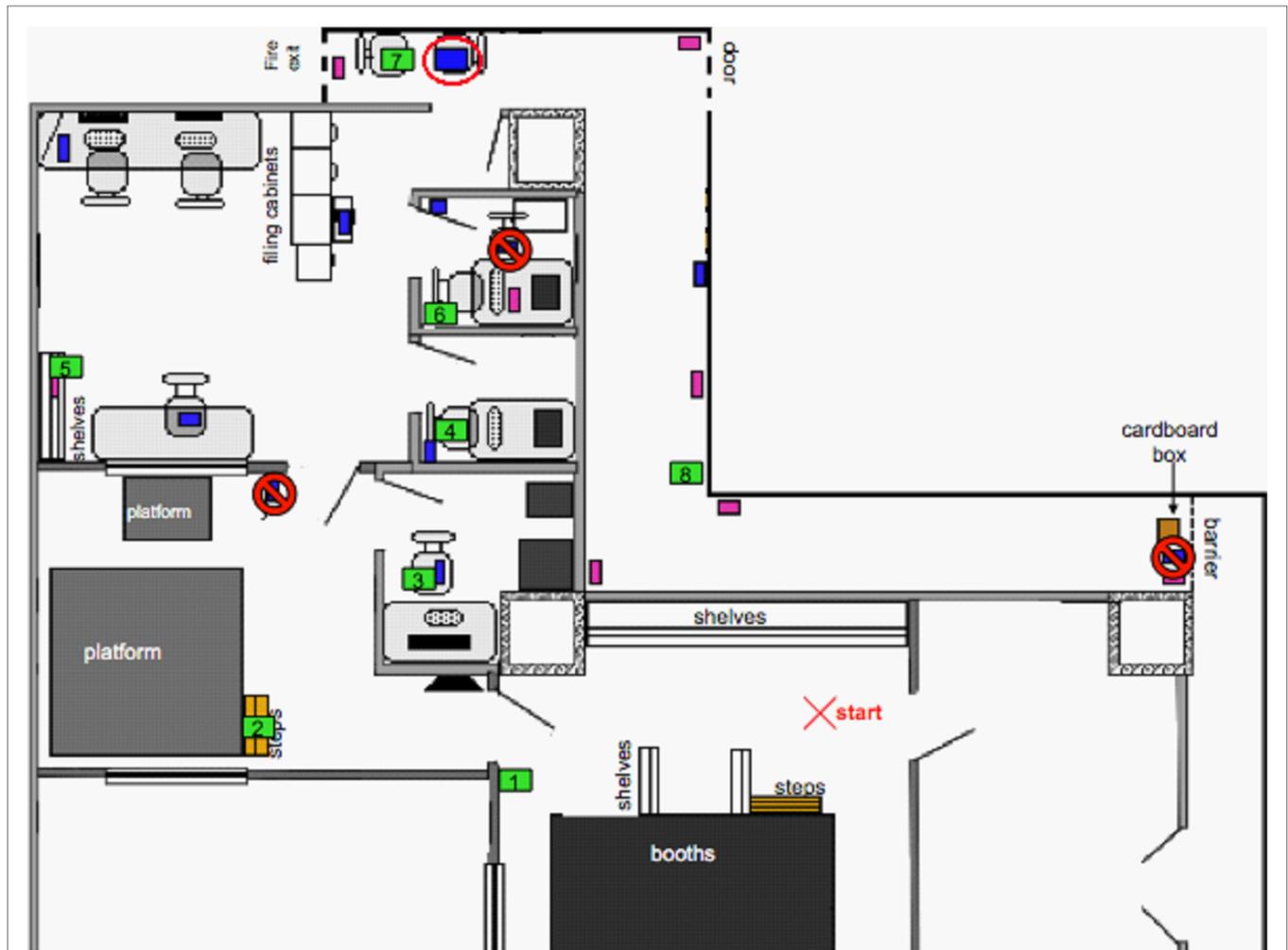
FIGURE 1 | Map of search environment showing the locations of the cardboard box, eight green boxes, eight blue boxes, and eight pink boxes, as well as the inaccurate locations of three blue boxes on the director's map (crossed in red). One additional blue box (circled in red) was not marked on the director's map but did exist in the environment. The color labels and locations of the green boxes are included for explanation but were not present for the director.

TABLE 1 | Initiation moves and examples.

Instruct	"Follow the corridor to the next room"
Explain	"I see a blue box in the room"
Query-Y/N	"Do you see a box there?"
Query-Wh	"Could you describe what's in the room?"
Check	"You said the box is in the corner?"
Align	"Does that make sense?"

be further broken down into *Checks* and *Aligns*. *Checks* involve paraphrasing the previous utterance to insure understanding of what was recently said, whereas *Aligns* are an explicit query to insure that the teammate has understood what was said before moving on.

Response moves (see **Table 2**) include *Acknowledgments* and *Replies* (Y/N and W-). *Yes/No replies* may be expressed in different ways, but always serve to indicate an affirmative or negative response. *Reply-W* moves are other replies that do not explicitly mean "yes" or "no."

Ready moves consist of an utterance-initial acknowledgment, such as "Okay" or "Alright," followed by an *Initiation* move. They are used to open a new segment of discourse and also may simultaneously close a preceding segment.

> *Ready:* "Alright, what's in the next room?"

The rates of producing each type of move were reported as a proportion based on the total number of utterances (e.g., Checks/Utterance).

Team Effectiveness

Performance was objectively scored with respect to the teams' successful completion of each of the three subtasks (i.e., the blue, green, and pink colored boxes). The maximum score for each subtask was 8, for a total maximum score of 24. The average score was 9.9 (range 1–19), and the median was 8. To examine performance differences, the 10 teams were assigned to one of

TABLE 2 | *Response* moves and examples.

Acknowledgment	"Okay," "Uh huh," "Mhm"
Reply-Y	"Yes," "Mhm," "Right"
Reply-N	"No," "Nope," "Nah"
Reply-W	D: "Should I take the box or just the blocks?"
	S: "**Just the blocks**"

two performance groups, *effective* or *ineffective*, based on whether they fell above or below the median score, respectively. The average score for the effective group was 14.8 (SD = 4.0), whereas for the ineffective group it was 5.0 (SD = 2.5).

The two groups differed slightly in demographic measures. For the effective group, four out of five teams were previously acquainted, and four of the five teams consisted of same-sex individuals (one M/M; three F/F); two directors were male and three were female. For the ineffective group, one out of five teams was previously acquainted, and two out of five teams consisted of same-sex individuals (one M/M; one F/F); three directors were male and two were female.

RESULTS

Quantitative Analysis

We performed a series of statistical analyses to test our hypotheses of the various factors influencing team communication in the CReST corpus.

Speech and Dialog Measures

One question of interest was what effect time pressure and speaker role would have on disfluency rate, speech rate, and MLU. The rates of these speech measures were calculated for the final 3 min (time pressure) and for the 5 min before that (no time pressure). A MANOVA was conducted with Speaker role (Director vs. Searcher) and Time Pressure (absent vs. present) as factors. There was a significant Speaker × Time Pressure interaction [$F(2,35) = 3.863$, $p = 0.030$], which was driven by speech rate [$F(1,36) = 4.424$, $p = 0.042$]. Under time pressure, directors spoke faster than the searchers [mean of 105.7 w.p.m. vs. 80.4 w.p.m.; $t(18) = 2.134$, $p = 0.047$; see **Figure 2**], but their MLU did not differ [5.0 vs. 4.5; $t(18) = 1.041$, $p = 0.312$]. However, the directors' MLU increased under time pressure [mean 4.1 vs. 5.0; $t(18) = -2.442$, $p = 0.025$], and there was a trend in faster speech rate [mean of 82.5 w.p.m vs. 105.7 w.p.m; $t(18) = -1.867$, $p = 0.078$].

The same MANOVA was conducted on rates of four types of disfluencies: repetitions, substitutions, insertions, and deletions. There was no significant difference between speaker roles – a finding which supports previous investigations of the same corpus (Nicholson et al., 2010). However, disfluency rates in the final 3 min (time pressure) did change significantly with Speaker role; directors had higher disfluency rates than searchers, $F(4,15) = 3.813$, $p = 0.025$. Univariate tests revealed that the directors' substitution rates were significantly higher, $F(1,18) = 7.564$, $p = 0.013$. Overall, our data support the prediction that disfluency rate and speech rate would increase with time pressure, especially for directors.
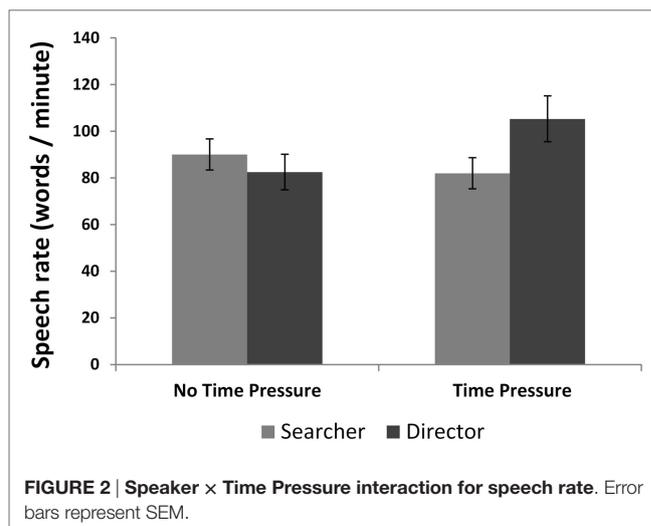

FIGURE 2 | Speaker × Time Pressure interaction for speech rate. Error bars represent SEM.

We also predicted that time pressure would impact the types of dialog exchanges that occur, but would affect each role differently. Specifically, directors would make more *Initiate* and *Ready* moves with time pressure and searchers would make more *Response* moves. Univariate Mixed 2 × 2 ANOVAs were conducted on the number of the different types of *Initiate* and *Response*, and *Ready* moves, with Speaker role as a between-subjects factor and Time Pressure as a within-subjects factor. For *Initiate* moves, there was a main effect of Speaker, with directors making more *Instruct* [$F(1,36) = 50.449$, $p < 0.001$], *Query* [$F(1,36) = 19.059$, $p < 0.001$], and *Ready* [$F(1,36) = 25.494$, $p < 0.001$] moves. There was a trend toward directors making more *Check* moves [$F(1,36) = 3.142$, $p = 0.085$]. However, searchers made more *Explain* moves [$F(1,36) = 32.980$, $p < 0.001$]. There was a Speaker × Time Pressure interaction for *Explain* [$F(1,36) = 6.136$, $p = 0.018$] and *Align* moves [$F(1,36) = 7.626$, $p = 0.009$], and a trend for *Query* moves [$F(1,36) = 3.284$, $p = 0.078$]. Directors produced more of these moves under time pressure, whereas searchers produced fewer. There was also a Speaker × Time Pressure interaction for *Check* moves [$F(1,36) = 50.449$, $p < 0.001$] due to searchers producing more under time pressure and directors producing less. Consistent with our predictions, directors produced more initiating moves than searchers, especially under time pressure. The one exception was that searchers produced more *Explain* moves, which upon further inspection, was due to their being responsible for describing the locations of the green boxes to the director.

In terms of *Response* moves, there was a Speaker × Time Pressure interaction for *Acknowledgments* [$F(1,36) = 10.732$, $p = 0.002$], resulting in directors making fewer acknowledgments and searchers making more under time pressure. For *Reply* moves, there was a main effect for Speaker [$F(1,36) = 15.856$, $p < 0.001$] and a significant Speaker × Time Pressure interaction [$F(1,36) = 4.669$, $p = 0.037$] due to searchers making fewer replies under time pressure. Despite this downward trend for replies under time pressure, searchers still had a numerically higher rate of *Response* moves than directors. For *Ready* moves, there was a main effect for Speaker [$F(1,36) = 25.494$, $p < 0.001$], resulting from directors producing more of these dialog moves. Overall,

these results support our predictions that directors would take more initiative in the task, whereas searchers would be more receptive (see **Figure 3**).

Performance Measures

Next, we examined the effects of performance and time pressure on the directors' and searchers' disfluency rates, speech rates, and dialog moves. We first tested for differences between the groups' disfluency and speech rates under no time pressure using unpaired t-tests. The results showed no differences between groups in the rates of the different types of disfluencies (all $ps > 0.05$).

Mixed 2×2 ANOVAs with Time Pressure (absent vs. present) as a within-subjects factor and Group (effective or ineffective) as a between-subjects factors were conducted on MLU and speech rate. There was no effect of Group on MLU ($F < 1$), which was contrary to the view that effective teams would exhibit an implicit coordination strategy (Orasanu, 1990). The effective directors' average speech rate was numerically greater than the searchers (101.4 w.p.m vs. 86.8 w.p.m, respectively), but the difference was not significant ($F < 1$).
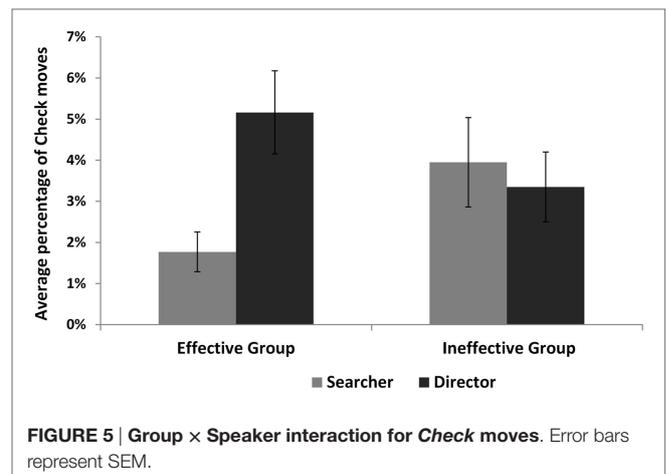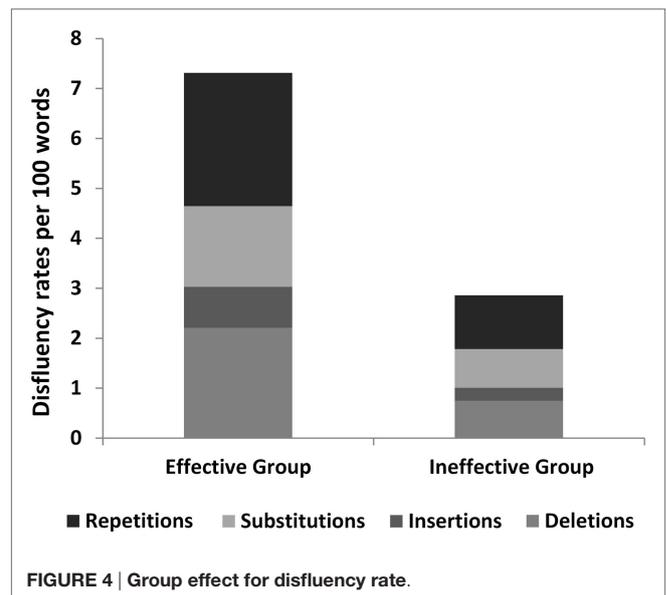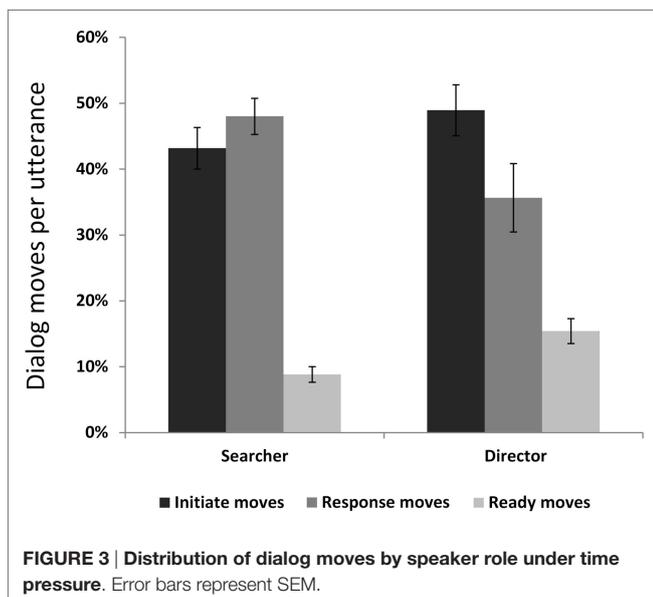
Next, we compared disfluency rate between performance groups. A MANOVA was conducted on the rates of the four types of disfluencies, with Group as a between-subjects factor and Time Pressure as a within-subjects factor. A significant Group effect was observed [$F(4,33) = 2.787$, $p = 0.042$] on rates of self-repair disfluencies (see **Figure 4**). The effective group displayed an increased rate of *Insertions* [$F(1,36) = 4.292$, $p = 0.046$], *Deletions* [$F(1,36) = 4.414$, $p = 0.043$], and a trending increase in *Substitutions* [$F(1,36) = 2.826$, $p = 0.101$]. The effective teams' increased self-repair rate is consistent with our prediction, which attributed it to additional planning and coordination.
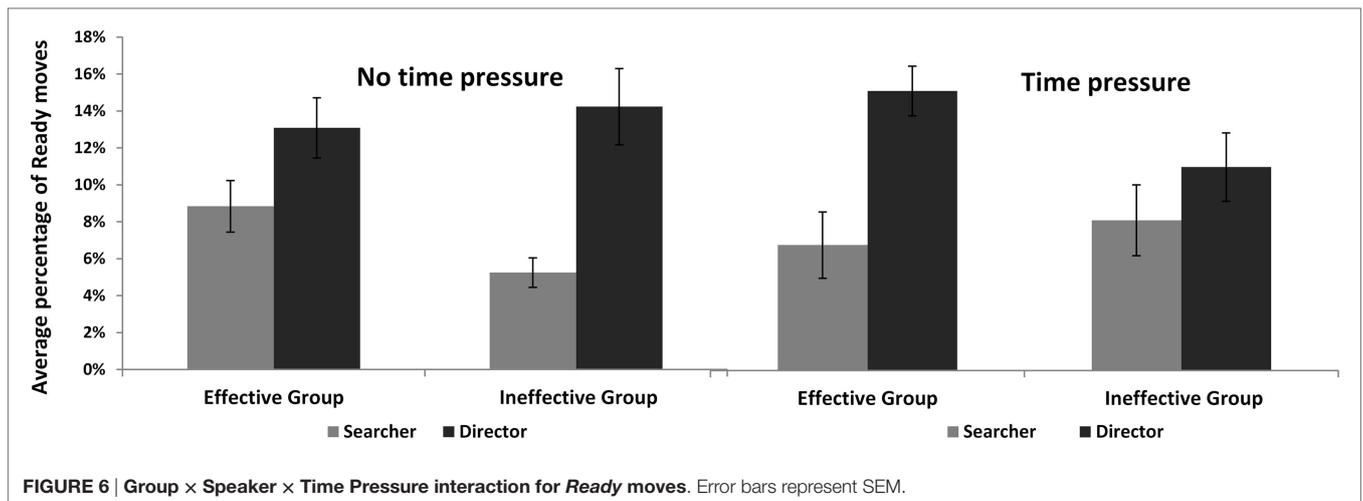
Next, we compared the distribution of conversational moves between both performance groups using $2 \times 2 \times 2$ mixed ANOVAs with Group and Speaker as between-subjects factors and Time Pressure as a within-subjects factor. We observed a Group × Speaker interaction for *Check* moves [$F(1,32) = 7.053$,

$p = 0.012$], with effective directors producing more Checks than the ineffective directors, and ineffective searchers producing more than the effective searchers (see **Figure 5**). In the analysis of *Ready* moves [$F(1,32) = 4.657$, $p = 0.039$], we observed a three-way interaction wherein effective directors' rate of *Ready* moves was higher than their teammates' with and without time pressure. However, ineffective directors only used more Ready moves than their teammate when there was no time pressure; under time pressure, the rate was not significantly different (see **Figure 6**). These results support our prediction that directors will take more initiative in the task, especially under time pressure. They also show that speaker role is a significant factor that interacts with dialog patterns and overall communicative effectiveness.

Qualitative Analysis

Though we found support for our hypothesis that effective teams would display more collaborative dialog, our speech and discourse measures may not quite tell the whole story. It is possible that

FIGURE 4 | Group effect for disfluency rate.

FIGURE 3 | Distribution of dialog moves by speaker role under time pressure. Error bars represent SEM.

FIGURE 5 | Group × Speaker interaction for *Check* moves. Error bars represent SEM.

FIGURE 6 | Group × Speaker × Time Pressure interaction for *Ready* **moves**. Error bars represent SEM.

the particular coordination strategies and grounding techniques that teams adopted may have influenced our results. Thus, it is important to evaluate examples from the teams' dialog exchanges to investigate how common ground was established in effective vs. ineffective teams. In the following analysis, we looked at several additional factors that may have influenced communicative effectiveness: coordination strategy, goal communication, and grounding techniques.

Coordination Strategy

Although each team was given the same set of instructions, there were significant differences in the approaches that teams used to perform the task. We have classified these *coordination strategies* into three major categories: *director leads*, *searcher leads*, and *shared leadership*. Each of these strategies differs with respect to the involvement of each role in managing the team's actions.

In a *director leads* strategy, the director has an increased speech rate and MLU, and produces more *Instruct* moves. By contrast, the searcher is the more receptive party, displaying a reduced speech rate and utterance length, but more *Response* moves (especially acknowledgments). The main role of the director in this strategy is to manage the searcher throughout the environment, while eliciting information about the green boxes scattered throughout. The main role of the searcher is to follow the director's instructions and provide information when it is requested. One would think that such a strategy would reduce interactive dialog due to the distinct separation between roles; however, it is still important to ground conversational exchanges. Since the two teammates are not colocated in the same physical environment, they still need to engage in perspective-taking when giving and responding to instructions. For example, the director needs to consider the searcher's movement in a 3D environment, and likewise, the searcher needs to tailor her descriptions to match the director's 2D floor plan. Of all the teams, one team from the effective group used this strategy throughout the task, and another effective team switched to it under time pressure. The same trend was found in the ineffective group as well, so this strategy on its own was not linked to improved performance (see **Table 3**).

As an example of the types of exchanges used by teams that employed this strategy, consider the following example from an ineffective team (#4) (dialog moves have been included for reference):

D: So are – you're walking into a big room (Instruct) am I right? (*Align*)
S: Uh yeh (*Reply-Y*), the room with the filing cabinets? (*Check*)
D: Okay (*Ready*), uh, no- no (*Reply-N*), leave that room make a right (*Instruct*)
S: Uh huh (*Acknowledge/Continue*)
D: Leave that room (*Instruct-repeat*), what do you see? (*Query-W*)
S: To the right there's a closed door (*Reply-W*).
D: No … {laugh} no, well can you open the door? (*Check*) No you can't open the door that's right. This is so odd (*Explain*).

Here, the director attempted to ground locational information by using more *Check*, *Align*, and *Query* moves. Despite the attempt, the team was unsuccessful because the teammates still had not identified one another's location. The director was not solely to blame either. The searcher did not volunteer information, but instead merely replied to the director's queries. As a result, the director had to (inefficiently) ask multiple questions to determine the searcher's location. This trend actually got to the point where, in the last utterance, the director asked and responded to his own question.

Now, compare this to an example from an effective team (#1) using this same strategy:

D: If you: [pause] turn around go out of that room (*Instruct*)
S: Okay (*Acknowledge*)
D: Straight in front of you there should be a chair (*Explain*)
S: Yes (*Acknowledge*)
D: At a table, there's a blue box there (*Explain*)
S: Yes (*Acknowledge*)
D: Okay (*Ready*). Get that (*Instruct*)

TABLE 3 | Team and participant composition.

Team	Performance group	Coordination strategy	Green score	Blue score	Pink score	Total score	Gender (director/searcher)	Relationship
1	Effective	Director	4	2	7	13	F/F	Acquainted
2	Effective	Searcher → shared	4	7	7	18	M/M	Acquainted
3	Effective	Searcher	5	5	5	15	F/F	Acquainted
5	Effective	Shared → director	5	3	1	9	M/F	Unacquainted
9	Effective	Searcher → shared	6	7	6	19	F/F	Acquainted
4	Ineffective	Director	1	0	0	1	F/F	Unacquainted
7	Ineffective	Searcher → director	4	1	1	6	M/F	Unacquainted
8	Ineffective	Searcher → shared	2	0	5	7	M/F	Unacquainted
10	Ineffective	Shared	2	0	2	4	F/M	Acquainted
11	Ineffective	Searcher	5	0	2	7	M/M	Unacquainted

Here, the director produced his utterances in installments, seeking confirmation that each step was understood before proceeding. Notice also how the director engaged in perspective-taking by referring to locations from the searcher's perspective (e.g., "Turn around" and "Straight in front of you"). Overall, the fact that these two teams used the same strategy, yet performed differently, suggests that this particular coordination strategy may not reflect performance. Factors, such as perspective-taking and grounding, may have more of an impact on team performance rather than the coordination strategy used.

A *searcher leads* strategy is the reverse of *director leads*. Here, the searcher is the one initiating more utterances, along with an increased speech rate and utterance length. The utterances mainly serve to update the director about her actions and movements, as well as to indicate the location of the green boxes. One common dialog pattern found in teams using this strategy was the searcher volunteering information unprompted. This allowed the searcher to give an update while navigating the environment, which helped to efficiently establish and maintain common ground.

Consider the following example of an effective team (#3) using this strategy:

> S: I'm going straight [pause] and then there's a doorway and at the very very left hand side, there's the first green box

In this example, the searcher took initiative by providing information as she navigated the environment. At first glance, the *searcher leads* strategy seems to be the most effective since it minimizes joint effort as the director is continually updated on the searcher's location and current plan. However, the team still needs to establish common ground with respect to the locations being described, which may be difficult if the searcher is operating from an egocentric perspective.

Consider the following example from an ineffective team (#11) using a *searcher leads* strategy:

> D: Okay where are those?
> S: Okay [pause] um [pause] the first one [pause] as I came into this room [pause] and I walked straight ahead *um*
> D: *Okay*
> S: [pause] just as I was about to turn right [pause] there's kind of this uh stage in front of me a:nd there's steps up to it and the green box is right in front of that on [pause] the- on the step
> D: Okay

Notice how the searcher used egocentric language, such as: "as I came into this room," "just as I was about to turn right," "this stage in front of me," etc. The searcher was not pausing to seek confirmation, but was rather stringing together lengthy descriptions that may or may not have been understood correctly. However, even though the director was responding with acknowledgments to all this, it was likely difficult for him to orient to the searcher's exact location on the map. This pattern of communication, more so than the coordination strategy, may have contributed to this team's poor performance.

Overall, this strategy requires the same degree of perspective-taking and grounding that a *director leads* strategy would require, and so it is not surprising that it is not well correlated with performance. In particular, one team in each performance group used this strategy, but two additional teams in each group changed to it during time pressure (see **Table 3**). Again, we find no evidence of coordination strategy alone being indicative of performance.

The final strategy we observed is the *shared leadership* strategy. Here, the director and searcher play a relatively equal role in managing the task. This strategy is characterized by an equal speech rate, as well as an equal rate of *Initiate* and *Response* moves; teammates alternate giving and receiving instructions depending on the situation. Such a strategy should require the most collaboration between teammates since they each equally contribute to the dialog. This strategy was slightly more common in the effective teams, with one team using it only when time pressure was absent, and two teams switching to it under time pressure. Interestingly, both of these teams that switched were the top performers. Only two of the ineffective teams used this strategy, with one of them using it for the duration of the task, and the other team switching to it under time pressure (see **Table 3**).

One pattern of behavior observed in effective teams using this strategy was the director keeping the searcher up to date about the remaining time, as in this example from Team #2:

> D: We only have like 30 seconds just to let you know
> S: Alright

This was a short, efficient utterance which served to establish common ground with respect to time remaining. Since the director had a timer on his map, he deemed it informative for the searcher to have this information as well. There is evidence that sharing such information may distribute the workload between members of a team (Khawaja et al., 2012).

Below is another example from the same team using this strategy:

S: Well [pause] see the two pink boxes?
D: Yes
S: On the right corner – the inside corner
D: Yes
S: There's a green box on that corner
D: Okay, alright
S: Okay

This exchange was initiated by the searcher, who explicitly sought confirmation at the end of each utterance to ensure that both teammates formed an accurate shared understanding of the location being described. This highlights another benefit of the *shared leadership* strategy, which is that it takes elements from the other role-focused strategies – in this case, the searcher volunteering information unprompted. However, this strategy is also susceptible to the weaknesses of those strategies, as can been in the following example from an ineffective team (#10):

D: So was the: – the: [pause] block that you received from the first blue box [pause] yellow?
S: Yeah [pause] ah yeah [pause] *it was*
D: *Okay* well that's wonderful [pause]
S: Mhm, I mean there were a couple in there and I just happened to grab the yellow one
D: Okay [pause] um [pause] have you taken a block out of the: – [pause] uh the- the second blue box?
S: Ye:s

Here, the team was spending valuable time establishing what the searcher had previously done, when this information could have been provided earlier in a more efficient manner. For example, the searcher could have stated which blocks she was taking out of which boxes as she was doing so. This shows how a *shared leadership* strategy is really a trade-off between the other role-focused strategies. In some cases, it could be beneficial, but the team still needs to ground their exchanges efficiently.

Overall, coordination strategies may have some effect on performance, but they do not tell the whole story. Thus, we found no evidence that coordination strategy alone would explain performance. We will now consider how teams communicated the new goal information and how they grounded their communicative exchanges.

Goal Communication

Another important factor determining success in the task was how well the directors communicated the new goal. Recall that directors received a new set of instructions after 5 min, which they had to convey to the searcher. Some directors were able to communicate this more effectively, giving them momentum going into the final 3 min. All of the effective directors accurately conveyed the new goal, but two of the five ineffective directors incorrectly explained the new goal; one failed to mention the yellow blocks, and another failed to mention the pink boxes (see example below). This could explain some of the poor performance

observed in the ineffective teams, especially with regards to their score on the new task. Below is an example of goal communication from an effective team (#2).

D: Alright, can you hear me?
S: Yeah
D: Okay. Alright, you- me- you- your new task is to place a yellow block into each of the pink boxes
S: Okay

This exchange took about 6 s and consisted of two contributions (in four turns). This statement also made use of the team's shared knowledge, since there was no explicit mention that the yellow blocks are located inside the blue boxes or what the pink boxes are. This was all established as common knowledge from prior interaction, so the team was able to minimize joint effort by recruiting this information. Since the searcher replied "Okay," this signaled that the message was understood and that there was no need for further explanation.

Next, we turn to an example of an ineffective team (#4) incorrectly communicating the new goal:

D: Alright, new task. Listen. In every blue box there's gonna be a yellow block. Okay?
S: Uh huh
D: I need you to grab the yellow block and put it into [pause] um a yellow box. Okay?
S: Into a yellow box?
D: Yeh, apparently there are more yellow boxes. Okay.
S: Okay
D: Okay so basically look we were just – okay we weren't oriented, we have three minutes by the way so we're gonna have to hurry up.

The new goal was to collect yellow blocks from the blue boxes and put them into each of the pink boxes. This was erroneously communicated by the director, likely due to cognitive load impacting his ability to retain the instructions. The searcher even questioned this instruction by initiating a side sequence ("into a yellow box?"), but the director confirmed the original erroneous instruction. As a result, neither teammate knew what the real goal was, and so (unsurprisingly) the team scored no points on this new task. This kind of miscommunication is especially problematic for a team using a *director leads* strategy, since this is the role entrusted to take initiative and make decisions.

Overall, we found important differences in goal communication between effective and ineffective teams. However, since three out of five ineffective teams communicated the goal correctly and still displayed poor performance, we must consider additional factors.

Grounding Techniques

Another factor that may have influenced performance is how teams grounded their exchanges. Grounding can involve a number of techniques, including seeking confirmation that a message has been understood (Clark and Brennan, 1991), engaging in perspective-taking (Brennan et al., 2010), using shared referents (Clark and Wilkes-Gibbs, 1986), monitoring

one's listener (Bavelas et al., 2000), breaking up longer utterances (Clark and Brennan, 1991), and using self-repairs to clarify the message (Bortfeld et al., 2001). We have already seen some of these in previous examples, but additional cases will be discussed below.

Consider the following example from an effective team (#3) that highlights many of these properties:

S: So now if you look straight like- okay if there like- do they show you how there's two mini rooms attached?
D: Yes
S: Okay, so the first mini room which is parallel with that desk, right?
D: Yes
S: There's a table with um a- one chair
D: Mhm
S: And there's a green box there
D: Okay

In this exchange, the searcher led the interaction and sought confirmation after each utterance. The director in turn provided an acknowledgment after each installment to ensure that common ground was maintained. Notice also the extensive perspective-taking by the searcher. Her first utterance "if you look straight" initially contained an egocentric statement (since the director does not have a 3D perspective of the room) but upon noticing this, she self-repaired it to a query in order to accommodate her partner's perspective. By engaging in many of the grounding techniques that characterize effective coordination, the searcher was efficiently able to share the location of the green box.

Monitoring of one's conversational partner is another important grounding technique used by effective teams. Consider the following example below from team #2:

S: So in every pink box I'm putting a:
D: Yellow cube
S: Yellow cube in, kay
D: Okay

In this example, the prolongation after "I'm putting a:" signaled difficulty in retrieving the next word, so the director made an accurate prediction and finished the sentence to help out his teammate and minimize joint effort. This suggests active listening and monitoring, both hallmarks of collaborative dialog.

Another important grounding technique utilized by effective teams is establishing shared referents (i.e., labeling) for objects in the environment. By using shorthand names for various rooms and locations, it becomes easier (and more efficient) for teams to refer to them without needing to give a full description.

Consider the following exchange from an effective team (#2) that shows this technique:

S: I'm going back into room one
D: Okay room one, like the very first starting room?
S: Yeah
D: Okay

This particular labeling strategy is known as *refashioning* (Clark and Wilkes-Gibbs, 1986) and involves one person proposing a label and the other person agreeing to it. Importantly, this exchange involved a side sequence to clarify that "room one" referred to the starting room. This completed the contribution and allowed the label to be added to the team's common ground. However, if the director had not sought clarification about "room one," then both parties would be operating from an egocentric perspective, and the intended referent would not be shared in common ground. This could actually hinder the team, as is seen in the following example from an ineffective team (#11):

S: And the:n [pause] as I came in this room that I called th- the server room *um*
D: *Mmm*
S: There's a green box there

In this example, the searcher referred to "the server room," but this was not previously established in common ground. As a result, the director could not resolve the referent and produced a hesitation ("Mmm"). In general, egocentric labeling makes it difficult for teams to take their partners' perspectives and discuss locations with one another efficiently.

Overall, teams that did not establish common ground with respect to the various referents in their environment were less efficient at coordinating their actions. This was particularly evident in the following exchange by Team #4, which took 29 s to communicate the location of a single green box:

S: By the way there's a green box here in this hallway.
D: Where which hallway?
S: The- the- okay so you know where the cardboard box is where I told you?
D: Yeah yeah uh huh
S: So right at the end of that wall um to the right like right at the edge is …
D: What number?
S: Box number 8.
D: 8? And it's on- to that- next to the cardboard box on the right?
S: Well okay so the cardboard box is at the very end of that hallway [pause] and at the other end of that little hallway.
D: Oh okay
S: On the right is the green box, number 8.
D: Okay. I gotcha.
TIME UP

This exchange took 98 words over the course of 12 conversational turns. Because the team was not successfully grounded, they had to spend excess time and resources to communicate a small piece of information. If these teammates had established shared knowledge of the floor plan and the various referents (e.g., the hallway, box locations, etc.), then this exchange of information could have been accomplished more efficiently.

DISCUSSION

Summary and Interpretation of Results

One result of our analysis was that factors, such as time pressure and speaker role, interact with speech and dialog measures. First, we found that time pressure was associated with an increase in speech rate, MLU, and disfluencies – particularly for directors. The effect for speech measures was not tied to a performance group, but rather seemed to result from increased workload demands imposed by the task (see **Figure 2**). Previous studies have not reached a consensus about the effects of time pressure on these speech measures, but here we provide evidence that they increase in a two-person cooperative, remote search task. The finding that disfluencies increased for directors supports Bortfeld et al. (2001). However, our results extend this finding by showing that time pressure and team performance interact with speaker role to affect rates of self-repairs.

We had several hypotheses about how effective teams might communicate in the task. In terms of dialog, we found that directors produced more initiating conversational moves (*Instruct, Query, Ready*) than searchers did, and their rates generally increased with time pressure (see **Figure 3**). However, Searchers produced more *Explain* moves, which was due in part to their being responsible for describing the locations of the green boxes to the directors. Searchers also produced more *Checks* and *Acknowledgments* under time pressure. These results suggest that directors generally took more initiative, whereas searchers were more receptive on the whole. This is consistent with previous findings (Bortfeld et al., 2001; Clark and Krych, 2004), but we extend this result by showing that coordination strategy may have interacted with these dialog patterns (see **Figure 7**). We found no support for the view that

effective teams would shift to a more implicit communication strategy (Entin and Serfaty, 1999). Given the nature and overall difficulty of the task, it was very unlikely that teams could succeed without explicit communication. Instead, we observed that more effective teams tended to produce more disfluencies and a trend toward more speech. Moreover, the observed increase in disfluency rate was not simply due to longer utterances in effective teams, as their MLU did not differ from ineffective teams.

Our results support the view that information provided by speech disfluency could be essential to the effective coordination of a team. Due to the fast-paced and difficult nature of the task, repairs were inevitable. However, if such disfluencies were solely indicative of production difficulty then we might have seen an increase for all teams, or only for the ineffective teams. However, since they increased for the *effective* teams, disfluencies may indicate something else besides increased workload (see **Figure 4**). Teams that used more self-repairs may have been able to minimize collaborative effort by clarifying and adjusting their utterances to their partner's perspective. Another benefit may have come from the improved recall associated with words preceded by disfluent speech (Corley et al., 2007). In this way, disfluencies can be viewed as a type of grounding mechanism, which may have been utilized by effective teams to (ironically) enhance the clarity and accuracy of their utterances.

In our analysis of conversational moves, we found an increase in the rate of *Check* moves for effective directors compared with ineffective directors (see **Figure 5**). Effective directors consistently used these dialog moves to establish common ground and manage the various task-induced constraints on their communication channels. In general, we expected directors to use more *Check* moves, since their floor map was a less reliable basis for grounding than the searchers' direct experience in the environment. We also found a similar effect for *Ready* moves, with effective directors displaying a higher rate than their teammate under time pressure, whereas ineffective directors' rates did not differ from their teammates under time pressure (see **Figure 6**). *Ready* utterances are often short confirmations (e.g., "OK" and "Right") used to end an existing dialog sequence and prepare the start of a new one. Effective directors' consistent use of *Ready* moves with and without time pressure supports our prediction that this role would take the greater initiative in managing and guiding the team.

With regards to coordination strategies, we found that teams generally coordinated their activities in one of three ways: *director leads*, *searcher leads*, and *shared leadership*. We did not find that any particular coordination strategy was more effective than others, as teams from both performance groups used a mixture of these strategies. This suggests that coordination strategy alone was not indicative of effective performance. However, these strategies may have interacted with speaker role, making it difficult to establish whether some of the effects we found were due to speaker role alone or to an interaction with coordination strategy. For example, though we found that directors had an overall increase in initiating utterances, this may not have held for all directors. A director on a team using a *searcher leads* strategy did not display the same increased speech rate and initiative that we found in directors from teams using a different strategy. Because

FIGURE 7 | Distribution of dialog moves by speaker role for each coordination strategy.
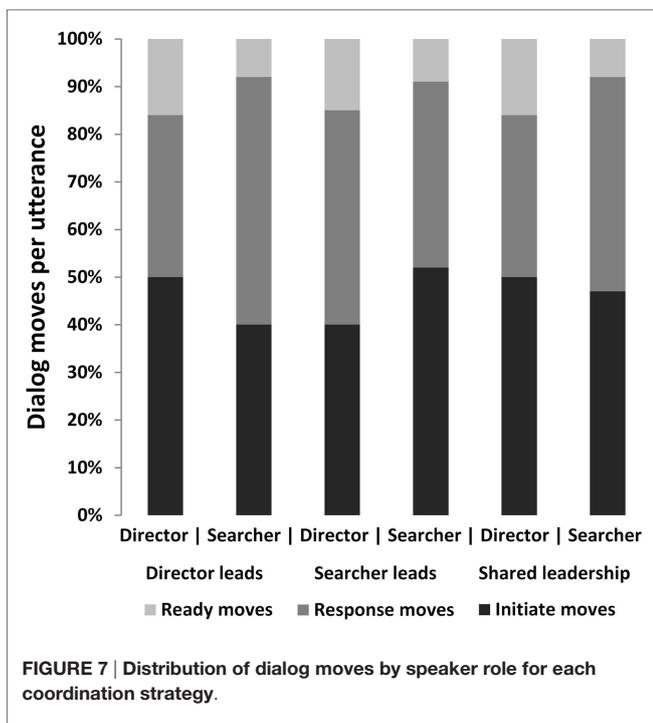
we only had five participants per group, we were limited in our ability to tease apart these differences (however, see **Figure 7** for some general trends). Future studies will need to incorporate larger samples to compare these different strategies and test how they affect team communication and performance.

Team Performance

While our results suggest that effective teams displayed improved performance due to increased collaboration and grounding, we must also consider how and why ineffective teams failed. One possibility is that ineffective teams failed to establish common ground, which increased collaborative effort and made coordination more costly and effortful. We found some evidence for this, as directors on the ineffective teams made fewer dialog moves indicating cohesion (e.g., *Check, Ready*) with their teammate. As a result, these teams had problems grounding their exchanges, which led to an inability to determine their partner's location and coordinate their actions with respect to objects and locations in the environment.

Another explanation for poor performance is that some teams failed due to reasons unrelated to coordinated dialog (e.g., workload, coordination strategy, miscommunication, or team composition). Regarding workload, we did not find evidence that the ineffective teams were differentially affected by the time pressure in our task. Our analysis showed that speech rate, average utterance length, and disfluency rates were not higher in ineffective teams. Since these measures have been associated with increased workload in previous studies (Berthold and Jameson, 1999; Khawaja et al., 2012), we cannot conclude that ineffective teams experienced greater workload. However, it has been shown that different operationalizations of workload (time pressure vs. task/resource demands) can lead to different effects on team performance (Urban et al., 1995). Our study did not disambiguate these factors since the workload condition involved both time pressure and increased task demands. Furthermore, workload was not measured apart from indirectly through our disfluency and dialog measures. This made it difficult to rule out if certain roles or teams were differentially affected by time pressure. One resulting possibility is that only the director was affected by time pressure because the timer was only visible on his map. Another possibility is that effective teams may not have actually been under any workload at all. These scenarios are unlikely because (1) our workload manipulation involved both time pressure and additional task demands, so it was unlikely that people were entirely unaffected, (2) disfluency rates, which have been associated with workload, increased for effective teams, and (3) the median score of all teams was an 8 (out of 24), and even the best team only scored a 19 – suggesting that the task was fairly challenging for all teams. Still, future studies will need to objectively measure workload to determine if the effects of workload are variable across performance groups.

One additional factor that may have affected performance is team composition. Due to the limited sample size, our two groups were not equally matched for gender or familiarity relationship. Though real-world teams are known to be heterogenous with respect to these factors, it is possible that teams in our study performed differently as a result of their gender composition or prior interactions (Myaskovsky et al., 2005; Bear and Woolley, 2011).

For example, Bortfeld et al. (2001) found that men produced more fillers and repetition utterances than women, especially when they were in the lead role of a task. This should not have affected our results since the gender composition between groups was fairly equal, and we did not examine fillers. There is also some evidence that team familiarity may affect coordination and performance (Jehn and Shah, 1997), but that it has no effect on disfluency rates (Bortfeld et al., 2001). Our performance groups did differ with respect to familiarity relationship (see **Table 3**), so this may have contributed to the performance disparity. Although it is possible that gender and familiarity may have influenced the performance of our groups, this does not change the finding that teams of varying effectiveness exhibited differing patterns of communication. It was these patterns, and their relation to team coordination, that our study sought to investigate.

Another possible explanation for poor performance was the miscommunication of the new goal in two of the teams. This did seem to result in a lower score for the new task, but did not significantly affect these teams' overall performance. The mean score for ineffective teams that *correctly* communicated the goal was 5.7 (SD = 1.5), whereas the mean score for the two teams that *incorrectly* communicated it was 4.0 (SD = 4.2). Though the means are numerically different, the two teams that incorrectly communicated the goal were the highest- and lowest-scoring teams in the ineffective group. As a result, we cannot claim that this factor alone contributed to their poor performance. Finally, we have already ruled out the likelihood that coordination strategy alone influenced performance, as teams from both performance groups used a mixture of these strategies.

Overall, our analysis suggests that patterns of communicative effectiveness were largely influenced by efficient grounding techniques that facilitated the establishment of common ground. Moreover, these patterns were maintained under time pressure, as effective teammates efficiently grounded their exchanges and engaged in perspective-taking of their partners. These results fit in line with the view that language is a collaborative activity, and speakers are necessarily and actively engaged in monitoring their listeners (Brennan et al., 2010). According to this view, language is not only for communicating information but also for making sure that the listener has understood the message. Our findings are consistent with this position and suggest that effective communication strategies are the basis of coordination in teams.

CONCLUSION

Team communication is a complex process with many factors influencing both coordination and performance. Our study examined time pressure, speaker role, grounding, and coordination strategy in order to better understand how these factors relate to communicative effectiveness, and ultimately, team performance. Of these factors, the most important one for determining effective performance was the ability to efficiently establish and maintain common ground with one's teammate through task-oriented dialog. We have shown that particular grounding techniques used by effective teams were more conducive to managing common ground under time pressure and changing task demands. This supports previous findings that collaboration

in joint tasks may be enhanced to offset costs associated with increasing task complexity (Kirschner et al., 2009; Khawaja et al., 2012). Future studies will include larger, more controlled samples, and will examine additional factors that might influence communication (e.g., task/team structure, coordination strategy, etc.). Objective measures of workload will need to be collected to examine whether the effective strategies seen in high-performing teams serve to actually reduce cognitive load. Another direction for future work is to investigate the extent to which the observed effects also apply to face-to-face interaction and other mediums within which people interact. By better understanding the factors that influence team communication and performance, we can improve training programs for organizational teams, ameliorate the effects of workload on group performance, and design artificial agents that interact closely with humans.

REFERENCES

Arnold, J. E., Kam, C. L., and Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: the on-line attribution of disfluency during reference comprehension. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 914–930. doi:10.1037/0278-7393.33.5.914

Bavelas, J. B., Coates, L., and Johnson, T. (2000). Listeners as co-narrators. *J. Pers. Soc. Psychol.* 79, 941–952. doi:10.1037/0022-3514.79.6.941

Bear, J. B., and Woolley, A. W. (2011). The role of gender in team collaboration and performance. *Interdiscip. Sci. Rev.* 36, 146–153. doi:10.1179/030801811X13013181961473

Berthold, A., and Jameson, A. (1999). "Interpreting symptoms of cognitive load in speech input," in *UM99, User Modeling: Proceedings of the Seventh International Conference*, ed. J.Kay (New York: Springer Wien), 235–244.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., and Brennan, S. E. (2001). Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Lang. Speech* 44, 123–147. doi:10.1177/00238309010440020101

Brennan, S. E., Galati, A., and Kuhlen, A. K. (2010). Two minds, one dialog: coordinating speaking and understanding. *Psychol. Learn. Motiv.* 53, 301–344. doi:10.1016/S0079-7421(10)53008-1

Brennan, S. E., and Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *J. Mem. Lang.* 44, 274–296. doi:10.1006/jmla.2000.2753

Carletta, J., Isard, S., Doherty-Sneddon, G., Isard, A., Kowtko, J. C., and Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Comput. Linguist.* 23, 13–31.

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

Clark, H. H., and Brennan, S. A. (1991). "Grounding in communication," in *Perspectives on Socially Shared Cognition*, eds L. B.Resnick, J. M.Levine, and S. D.Teasley (Washington, DC: APA Books), 127–149.

Clark, H. H., and Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *J. Mem. Lang.* 50, 62–81. doi:10.1016/j.jml.2003.08.004

Clark, H. H., and Schaefer, E. R. (1989). Contributing to discourse. *Cogn. Sci.* 13, 259–294. doi:10.1207/s15516709cog1302_7

Clark, H. H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition* 22, 1–39. doi:10.1016/0010-0277(86)90010-7

Corley, M., MacGregor, L. J., and Donaldson, D. I. (2007). It's the way that you, er, say it: hesitations in speech affect language comprehension. *Cognition* 105, 658–668. doi:10.1016/j.cognition.2006.10.010

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., and Bruce, V. (1997). Face-to-face and videomediated communication: a comparison of dialogue structure and task performance. *J. Exp. Psychol. Applied* 3, 105–125.

Eberhard, K. M., Nicholson, H., Kubler, S., Gundersen, S., and Scheutz, M. (2010). "The Indiana 'cooperative remote search task' (CReST) corpus," in *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010* (Malta), 17–23.

Entin, E. E., and Serfaty, D. (1999). Adaptive team coordination. *Hum. Factor.* 41, 312–325. doi:10.1518/001872099779591196

Jehn, K. A., and Shah, P. P. (1997). Interpersonal relationships and task performance: an examination of mediation processes in friendship and acquaintance groups. *J. Pers. Soc. Psychol.* 72, 775–790. doi:10.1037/0022-3514.72.4.775

Khawaja, M. A., Chen, F., and Marcus, N. (2012). Analysis of collaborative communication for linguistic cues of cognitive load. *Hum. Factor.* 54, 518–529. doi:10.1177/0018720811431258

Kirschner, F., Paas, F., and Kirschner, P. A. (2009). Cognitive load approach to collaborative learning: united brains for complex tasks. *Educ. Psychol. Rev.* 21, 31–42. doi:10.1007/s10648-008-9095-2

Krauss, R. M., and Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *J. Pers. Soc. Psychol.* 4, 343–346. doi:10.1037/h0023705

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition* 14, 41–104. doi:10.1016/0010-0277(83)90026-4

Lickley, R. (1998). *HCRC Disfluency Coding Manual*. Edinburgh: University of Edinburgh, Human Communication Research Centre.

Lively, S. E., Pisoni, D. B., Van Summers, W., and Bernacki, R. H. (1993). Effects of cognitive workload on speech production: acoustic analyses and perceptual consequences. *J. Acoust. Soc. Am.* 93, 2962–2973. doi:10.1121/1.405815

Myaskovsky, L., Unikel, E., and Dew, M. A. (2005). Effects of gender diversity on performance and interpersonal behavior in small work groups. *Sex Roles* 52, 645–657. doi:10.1007/s11199-005-3732-8

Nicholson, H., Eberhard, K. M., and Scheutz, M. (2010). "'um. I don't see any': the function of filled pauses and repairs," in *Proceedings of DiSS-LPSS Joint Workshop* (Tokyo), 89–92.

Orasanu, J. (1990). *Shared Mental Models and Crew Decision Making*. Technical Report No. 46. Princeton, NJ: Princeton University, Cognitive Science Laboratory.

Serfaty, D., Entin, E. E., and Volpe, C. (1993). Adaptation to stress in team decision-making and coordination. *Proc. Hum. Factor Ergonom. Soc. Ann. Meet.* 37, 1228–1232.

Swerts, M. (1998). Filled pauses as markers of discourse structure. *J. Pragmat.* 30, 485–496. doi:10.1016/S0378-2166(98)00014-9

Sycara, K., and Sukthankar, G. (2006). *Literature Review of Teamwork Models (Tech. Rep. No. CMU-RI-TR-06-50)*. Pittsburgh, PA: Robotics Institute.

Urban, J. M., Bowers, C. A., Monday, S. D., and Morgan, B. B. Jr. (1995). Workload, team structure, and communication in team performance. *Milit. Psychol.* 7, 123–139. doi:10.1207/s15327876mp0702_6

AUTHOR CONTRIBUTIONS

FG was responsible for acquiring the data from the corpus, data analysis, and drafting the paper. KE was responsible for developing and providing the corpus used for analysis as well as contributing to the data analysis and making major revisions to the paper. MS was responsible for the conception and design of the study as well as major input and revision in the writing process. All authors agree to be accountable for the integrity of the data.