

# MEASURING SPEECH FLOW OF CO-LOCATED AND DISTRIBUTED COMMAND AND CONTROL TEAMS DURING A COMMUNICATION CHANNEL GLITCH

Preston A. Kiekel    Jamie C. Gorman    Nancy J. Cooke<sup>1,2</sup>  
Cognitive Engineering Research Institute<sup>1</sup>    Arizona State University East<sup>2</sup>  
Mesa, AZ    Mesa, AZ

Team cognition can be observed in the flow of communications among team members. This is shown in the context of a simulated unmanned aerial vehicle ground control station. Automatic measures of low-level team communication flow were used to assess high-level constructs of team cognition. Measures show support for the expected results of manipulations in this task. Co-location and channel degradation effects were successfully predicted by CHUMS, ProNet, and a cross-correlation function-based Dominance measure. Results grant concurrent validity to the measures, and highlight substantive effects of the manipulations. In particular, in geographically distributed teams, communication patterns are less stable, and the route planner exerts less communicative influence. Some co-location effects drop with task experience. During a mission containing a five-minute one-way communication channel cut, all teams communicate more like distributed teams, and team members do create alternate pathways to retain information flow.

## INTRODUCTION

Designers for team systems need a way to exploit team cognitive processes, in the same way that designers for individuals exploit individual cognitive processes. The most important advantage of cognitive data, over outcome data, is that cognitive data can be used to diagnose, and predict the team's future actions, even in the absence of an outcome measure. Fortunately, team communication is an inherent byproduct of team tasks, and so teams inherently yield the equivalent of a think-aloud protocol. However, the complexity of communication data makes it difficult to analyze. To address this need, we have developed automatic and semi-automatic measures of team communication, in order to extract team cognition (Kiekel, Cooke, Foltz, & Shope, 2001; Kiekel, Cooke, Foltz, Gorman, & Martin, 2002; Gorman, Foltz, Kiekel, Martin, & Cooke, 2003).

In this paper, we use our communication flow measures to identify manipulations of co-location and communication channel outages in a communication rich, interdependent team uninhabited air vehicle (UAV) task. We wanted to know how UAV teams would react to the isolation of geographic dispersion, especially when exacerbated by cutting a communication channel. These manipulations impact how the team can interact, but do not inherently impact the difficulty of the task. Consequently, we have found a much weaker impact of these manipulations on our performance measure than on

our process and cognitive measures. This is juxtaposed against a workload manipulation that was specifically intended to make the task harder, and did have a large impact on performance.

Communication data represent both team process and team cognition, and so we suspected that teams adapt to the process manipulations through communication patterns. If our measures possess adequate concurrent validity, then they should be able to discriminate between levels of these manipulations. Moreover, these measures allow us to explore the manipulation effects themselves, in a richer way than with a performance measure alone.

Concurrent validity is especially well tested by the communication outage or "glitch," because communication measures with adequate concurrent validity should certainly be able to detect and describe a 5-minute cut in a communication channel. With this manipulation, we have enough experimental control to make specific *a priori* predictions, to be confirmed or denied. We would expect teams to exhibit more communication patterns in the mission that includes the glitch. This is because during the glitch, the team member whose channel is cut would have to create an alternate path, through the remaining team member. This should also weaken that team member's ability to influence the discourse. Let us define "dominance" among team members to mean that the dominant member's behavior is more predictive of other member's behavior, than vice versa. Then discourse

dominance should generally be weaker in distributed teams.

## METHODS

Twenty teams of three members flew a simulated UAV, taking photographs of targets, over seven missions, each up to 40 minutes. Each team consisted of three randomly assigned interdependent roles: DEMPC (Data Exploitation, Mission Planning, and Control) plans the route, AVO (Air Vehicle Operator) flies the plane, and PLO (Payload Operator) takes pictures. Team members communicated over headsets. The presence or absence of speech by each team member was automatically recorded at each second, into a communication log (ComLog). The ComLog is a record of who spoke to whom, when, and for how long.

The experimental manipulations were 1) co-location, 2) workload and 3) a communication “glitch.” Each team was randomly chosen to be co-located in the same room, or distributed by portable walls and separate locations. After the fourth mission, the scenario was made more difficult to increase cognitive workload. Due to space limitations, we do not examine this effect in the present paper. During the sixth mission, a “glitch” was introduced. The communication channel from DEMPC to AVO was disabled for five minutes. However, AVO was still able to speak to DEMPC. We expected to see teams develop new communication patterns, such that the communication from DEMPC to AVO would be indirectly re-routed to go through the PLO.

We employed a 7 (missions, WS) x 2 (co-location, BS) model. Intra-mission analysis of the glitch effect is unavailable, because the precise timing of the glitch depended on the team's route, and hence could not be known in advance. Accordingly, we analyzed glitch effects with inter-mission repeated measures contrasts. We scrutinized changes in team communication from Mission 5 (pre-glitch) to Mission 6 (glitch), and from 6 to 7 (post-glitch). We examined differences in the Co-location effect between these missions, using WS interaction contrasts.

Communication measures are described next.

### ProNet

ProNet (Procedural Networks; Cooke, Neville & Rowe, 1996) is a form of lag sequential analysis, based on the Pathfinder algorithm. Pathfinder is used to identify linkages between nodes, using the set of

pairwise proximities among nodes. Nodes can be defined as events in a potential sequence, and proximity can be defined as a transition probability among these events. ProNet is the application of Pathfinder to such a sequence of nodes.

In this case, we use ProNet to identify representative sequences of discrete communication flow events in our ComLog. We define six nodes, one for each team member beginning and ending a speech act (i.e., Abeg, Pbeg, Dbeg, Aend, Pend, Dend). This is the most basic unit of communication, outside of content analysis.

With these six nodes, many possible patterns can be identified. We used these nodes to define three types of sequences. First, an *Xloop* means that person X begins and ends an uninterrupted utterance, then begins and ends again. This may indicate repetition. Second, an *XYcycle* means that person X speaks a complete utterance, then person Y does. Third, *XiY* means that person X interrupts person Y. We chose these sequence patterns because other chains were either too specific to be generally applicable (e.g. specific types of interruption), or were too short to be very meaningful (e.g. X begins then ends a speech event). We measure both the ProNet detection of these chains, and their frequency of occurrence within the ComLog.

### CHUMS

In the context of this discussion, we developed CHUMS (Clustering Hypothesized Underlying Models in Sequence; Kiekel, Cooke, Foltz, Gorman, & Martin, 2002) for measuring stability of communication patterns, in terms of relative amount of speech among team members. We separate the ComLog into one-minute intervals, and take the relative proportion of speech by each team member, within each minute. Minutes were chosen as our time unit, because it was long enough to capture behavior patterns, but short enough to reveal changes over time. Then we perform all pairwise  $\chi^2$  tests, among the proportion models at each minute. Proportions of speech quantities make up a multinomial model. Pearson  $\chi^2$  is an approximation test that is frequently used on multinomial data. We perform agglomerative cluster analysis on the minutes, using p-values as a distance metric, and alpha as a stopping rule. This approach allows us to see how many distinct patterns were present in the team's communication quantity behavior, at the one-minute level.

We measure communication pattern stability using the number of distinct models identified by CHUMS. Having more patterns means that teams used more communication strategies during the mission, so their communication is less stable over time.

### **Dominance**

We defined a measure representing the influence a team member's communication flow exhibits over the communication flow of other team members, in terms of amount of speech. Building on work by Budescu (1984), we defined a dominance measure based on the cross-correlations of speech quantity in the ComLog, among all pairs of speakers. Correlations among sequences of speech quantity were computed for each lag, among each pair of speakers. This constitutes the set of cross-correlation functions. Then, because we were not concerned with directionality of influence, we squared each correlation. Since we were working with  $R^2$ , Fisher's  $R$  to  $Z'$  transformation was not required, before performing arithmetic computations. Next, we took the weighted average of each squared cross-correlation function, with weighting being the inverse of the lag. This gives greater emphasis to influence revealed at early lags (i.e. speech events that are closer in time).

This process leaves us with a squared correlation for each pair of speakers, representing the capacity of each speaker's speech quantity behavior to predict each other speaker's. To obtain a relative dominance value for each speaker pair, we computed ratios of squared cross-correlations. So, for speakers  $X$  and  $Y$ , we took the correlation of  $X$  predicting  $Y$  ( $R^2_{xy}$ ), divided by  $Y$  predicting  $X$  ( $R^2_{yx}$ ), to obtain a dominance ratio. These ratios indicate dominance of each team member over each other.

In the above ratio, the dominance value ranges from 0 to 1 if  $Y$  is dominant (i.e. if  $R^2_{yx}$  is larger than  $R^2_{xy}$ ), but from 1 to positive infinity if  $X$  is dominant. To convert this to a symmetrical scale, we took the natural log of the dominance ratio. This yields a pairwise dominance measure ranging from negative infinity to 0 when  $Y$  is dominant, and from 0 to positive infinity when  $X$  is dominant. However, due to the variance stabilizing aspect of log transformations, the distribution of the pairwise dominance measure is approximately normal, with a mean of 0. Under these conditions, it is appropriate to perform arithmetic operations. To get a total

dominance score for each team member, we averaged the log of the ratios. The three dominance scores are  $Adom$ ,  $Pdom$ , and  $Ddom$ .

### **Predictions**

One would expect teams to detect the glitch, and form new communication patterns as a work around. For the mission that includes the glitch (Mission 6), teams should show more communication patterns, and hence more CHUMS models. Teams compensating for the glitch should show that the passage of information normally sent directly from DEMPC to AVO would be diverted through PLO. We would expect increases from Mission 5 to 6, for DPcycles, PDcycles, and PACycles. DACycles should drop.

The largest source of information flow is the DEMPC, who plans the route, so DEMPC should tend to have a high Dominance score. When DEMPC is cut out, either by physical absence or glitch, his or her influence should drop. Therefore, DEMPC dominance should be lower for distributed teams, and should decrease during the glitch.

## **RESULTS**

### **CHUMS Results**

Distributed teams had more CHUMS models than co-located teams for Missions 2, 4, and 5 (respectively  $t(89) = 2.08, p = .04$ ;  $t(89) = 3.00, p = .003$ ;  $t(89) = 2.95, p = .004$ ). The number of models increased from Mission 5 to the glitch (mission 6) ( $t(89) = 2.03, p = .045$ ), particularly for co-located teams (see Figure 1 for means that have not been corrected for repeated measures. Standard errors were not included because these means are not the adjusted means used in the hypothesis tests). Co-location's impact drops and reverses from Missions 5 to the glitch (mission 6) ( $t(89) = -2.61, p = .011$ ), and the reversal becomes more extreme from 6 (the glitch mission) to 7 ( $t(89) = -2.62, p = .010$ ).

### **Dominance Results**

AVO is reactive in Co-located teams, but dominant in distributed teams ( $F(1, 16.78) = 16.41, p = .001$ , see Figure 2a). DEMPC is moderately dominant for Co-located, but reactive for Distributed,  $F(1, 16.51) = 12.84, p = .002$ . During the glitch, Co-located teams become AVO-dominated ( $t(89) = 2.08, p = .040$ ), and PLO-reactive ( $t(89) = -1.88, p = .063$ , Figure 2b). Also, DEMPC is dominant in Mission 6

( $t(89) = 2.05, p = .043$ ), but reactive in all other mission.

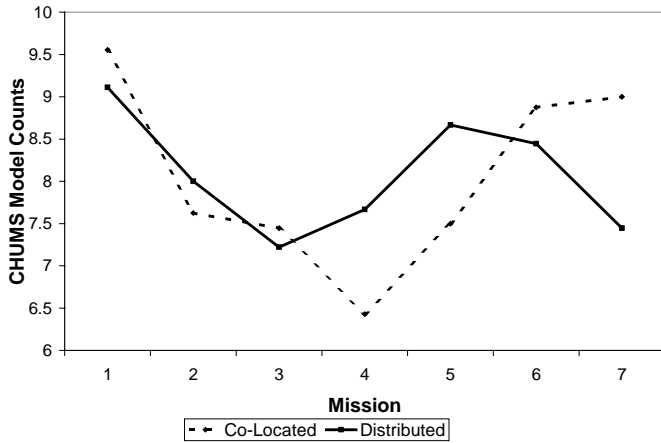


Figure 1. Uncorrected Co-location\*Mission means for CHUMS Models. (Mission 6 contains the glitch).

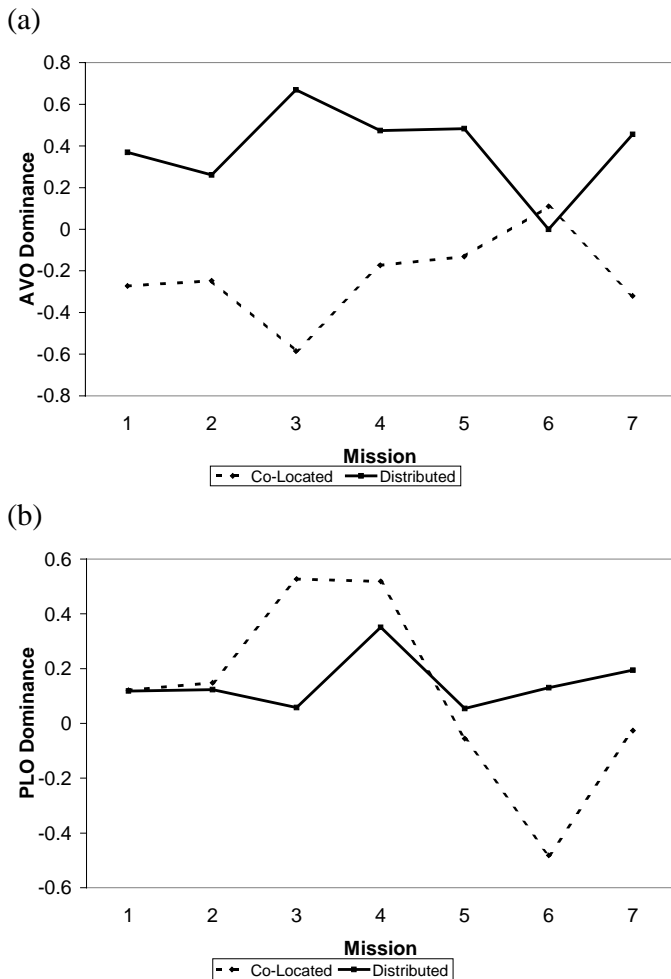


Figure 2. Uncorrected Co-location\*Mission means for Adom (a, top) and Pdom (b, bottom). (glitch at Mission 6).

## ProNet Results

Several ProNet sequence counts were higher in Co-located teams, but this is somewhat obfuscated by the fact that Co-located teams made more utterances in general,  $t(16.23) = 3.82, p = .001$ . This does not impact inter-mission contrasts. Co-location effects on AiD dropped between each mission, from 4 down to 7 (respectively,  $t(89) = -2.13, p = 0.036$ ;  $t(89) = -2.64, p = 0.010$ ;  $t(89) = -2.25, p = 0.027$ ). PiA co-location differences also dropped from Missions 4 down to 7 (respectively,  $t(89) = -1.82, p = 0.072$ ;  $t(89) = -3.07, p = 0.003$ ;  $t(89) = -3.37, p = 0.001$ ).

Turning to glitch effects, between Missions 5 and 6 (the glitch mission), DAcycles decrease ( $Wald(1) = 3.15, p = .076$ ), DPcycles increase ( $t(89) = 1.82, p = .073$ ), and PAcycles increase ( $t(89) = 2.11, p = .038$ ). Also, PDcycles decrease ( $Wald(1) = 3.30, p = .069$ ).

## DISCUSSION and CONCLUSION

ProNet chain counts were particularly good at discriminating glitch effects, because they identify specific sequences of communication events. Apparently to compensate for the glitch, DEMPC spoke more to PLO and less to AVO, creating an alternate communication route. PLO spoke more to AVO, apparently passing information from DEMPC. Contrary to our expectation, PLO appears to have passively received DEMPC's input, since PDcycles decreased.

Results of other methods were entirely consistent with these interpretations. It is generally found that the glitch was linked to weakened co-location effects, and all teams behaved more like distributed teams. CHUMS measures showed that 1) distributed teams had more models, and hence less stable communication, 2) the glitch was correlated with more distributed-like (i.e. less stable) communications, and therefore 3) teams created additional communication patterns during the glitch. Dominance measures implied that, though DEMPC is less able to exert predictive speech patterns in distributed teams, this disadvantage disappears during the confusion of the glitch. Also, during the glitch, Co-located teams become more AVO-dominated, and PLO-reactive, the reverse of the general pattern. This suggests that PLO is passively replying to the added information.

In addition to concurrent validity, some of the findings lend external validity to the measures, by replicating findings in the literature. For instance, on

ProNet-based interruption counts, the effects of the medium (i.e., co-located vs. distributed) diminish with experience (e.g. Walther, 1996).

These findings reveal the power of communication measures for assessing team cognition. The ComLog file from which these measures were derived is a simple record of when team members are speaking. Yet, these data lend themselves to a recombination into three quite distinct measures of team cognition, each tapping into a different construct: stability, dominance, and information passage. Moreover, using these measures, we were able to decipher high-level findings, such as observing how teams work around a communication glitch in order to retain information passage. With additional research, automatic measures of team cognition should be able to yield very powerful diagnosis and prediction of team behavior.

#### **ACKNOWLEDGEMENTS**

This work was supported by ONR Grants N00014-00-1-0818 and N000140310580.

#### **REFERENCES**

Budescu, D. V. (1984). Tests of lagged dominance in sequential dyadic interaction. *Psychological Bulletin*, 96, 402-414.

- Cooke, N. J., Neville, K. J. , & Rowe, A. L. (1996). Procedural network representations of sequential data. *Human-Computer Interaction*, 11, 29-68.
- Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., Cooke, N. J. (2003). Evaluation of Latent Semantic Analysis-based measures of team communications content. *Proceedings of the 47<sup>th</sup> Annual Human Factors and Ergonomics Society Meeting*.
- Kiekel, P. A., Cooke, N. J., Foltz, P. W., Gorman, J & Martin, M. M. (2002). Some promising results of communication-based automatic measures of team cognition. *Proceedings of the Human Factors and Ergonomics Society 46<sup>th</sup> Annual Meeting*.
- Kiekel, P. A.; Cooke, N. J.; Foltz, P. W.; & Shope, S. M. (2001). Automating Measurement of Team Cognition Through Analysis of Communication Data. In M. J. Smith, G. Salvendy, D. Harris, and R. J. Koubek (Eds.) *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality*, pp. 1382-1386. Mahwah, NJ: Lawrence Erlbaum Associates.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, 3-43.