

Automating Measurement of Team Cognition through Analysis of Communication Data

Preston A. Kiekel¹, Nancy J. Cooke¹, Peter W. Foltz¹, and Steven M. Shope²

Department of Psychology, New Mexico State University, Las Cruces, NM 88003¹
Sandia Research Corporation, 4200 Research Drive, MSC3-ARP PO Box 30001, Las Cruces, NM 88003²

Abstract

In this paper we propose a general methodological approach for semi-automatically assessing team cognition using communication data. The approach rests on four premises: 1) analyzing communication data is a means of assessing team cognition, 2) both substantive content and physical quantity measures of communication are needed, 3) sequential flow methods are especially helpful to effectively make use of communication data, and 4) analysis of both data types can be automated with contemporary tools, both statically and sequentially. We begin by illustrating the first three points. Next we briefly review commonly employed methods of communication analysis, and note the difficulties of analyzing such data. We then suggest that appropriate automatic methods for analyzing communication data are becoming increasingly available, and we give some examples of our approach. Finally, we conclude by discussing the implications of this approach, especially for team training and groupware design.

1. INTRODUCTION: A HOLISTIC DEFINITION OF TEAM COGNITION

For the purposes of this paper, any small group of people collaborating on a task constitutes a team. Though teams have been a ubiquitous component of most organizations for some time, research on team cognition is relatively new. Team cognition can be defined as the team analog of individual cognition, i.e. the thoughts and knowledge of the team. This is distinct from “social cognition,” which revolves around individuals’ cognitions about other people. Certainly, social cognition is relevant to team cognition, but team cognition is also the interaction of team members as they collaborate. As such, it must be somewhat distinct from individual cognition, which has a physical mechanism for storage and processing. There is no single brain and body to contain the thoughts and knowledge of a team, yet there is an emergent property when a group of people collaborate (Steiner, 1972).

Usually, this emergent property is treated as the sum of individual cognition, and this has value. However, the fact that current measures of team cognition focus on the similarities among the individuals’ cognition (Langan-Fox, Code, & Langfield-Smith, 2000), makes a paradigmatic assertion that the tasks in which we are most interested are those in which there is little individuation of subtasks. For tasks that require specialization of individuals, such as in an operating room, should our definition of the team’s cognition still focus only on similarities among individual cognition (Cooke et al., 2000)? Just as an individual has contradictory and dynamic thoughts, so individual team members have different ideas and knowledge regarding a task. So it is perfectly reasonable to define team cognition at the holistic level, where the team is the unit, rather than at the level of averaged individual cognition.

If teams are to be the unit of analysis for this holistic definition, we will need to measure behaviors exhibited by the team as a whole. Just as we use think aloud protocols to examine what an individual is saying to herself when she thinks, so we can use the communication inherent in the collaborative process to assess team cognition in a holistic way. In this respect, team cognition is easier to measure than individual cognition; teams need not interrupt their process in order to “think aloud,” since they are always “thinking aloud” in some sense. With a newly formed team, team cognition begins as the sum of individual cognition. Then, as the team “thinks” (interacts), dynamic changes occur in the “team mind” as a natural result of the interaction. Effects of this process on performance depend on the type of task (Steiner, 1972). Thus, an analysis of team communication provides a window through which to view team cognition.

2. TYPES OF COMMUNICATION DATA AND COMMONLY EMPLOYED METHODS

Communication measures can be characterized as “physical” (i.e. relatively low-level measures such as duration of speech, e.g. Watt & VanLear, 1996) and “content” (what is actually being said). Both can be useful in characterizing team cognition. For instance, major drops in communication frequency over time may be driven by implicit learning of the task, leading to less need for explicit discussion. Likewise, content data might show that a team talks about mechanics of the software early in the session, but later only talks about more substantive issues.

Though physical measures can often be taken in real time by human or machine recorders, content measures are typically taken on transcripts made from taped interaction. After transcription, a coding scheme is employed

that classifies the utterances into meaningful categories (Emmert, 1989). Analyses can be either time-static (e.g. code counts), or sequential. Codes can themselves be sequences, such as “argument followed by name-calling.”

Statements-- the basic verbal units teams use to define a common frame of reference-- are themselves driven by previous statements. Thus, we cannot readily isolate a statement without its context. This gives an advantage to sequential analyses over static. However, both types of analysis are important for assessing team cognition. As an example of a useful static analysis, it would be quite helpful to know that 98% of a team’s communication is argument and 0% is planning. The approach advocated here includes both physical and content data, and emphasizes sequential flow analysis over static measures. This yields a 2*2 table of possible data types (see Table 1).

Hand coding can be time consuming, for both content and physical measures. For physical hand coding, speech duration can be captured by recording onset/offset times, at the rate of one hour per hour of interaction. For content coding, Emmert (1989, p. 244) cites a study that required 28 hours of transcription and encoding per hour of tape. We propose more automatic methods to circumvent such time requirements, and also to increase reliability.

Table 1. Types communication analysis, with an example of each.

	Static	Sequential
Content	Number of arguments	Number of arguments followed by insults
Physical	Total seconds spoken	Number of Person A speaking after Person B

3. AUTOMATIC ANALYSIS OF COMMUNICATION DATA

To collect physical data, we developed software that records quantity of verbal communication as a $K^2 \times N$ communication log matrix of dichotomous values. K is the number of team members, and N is the number of time intervals (e.g. seconds) across which the communication spans. All possible pairs of the K speakers account for the K^2 columns. At each time interval, a measure is automatically taken of which team members are talking, and to whom. This creates the N rows in the communication matrix. The result enables rapid analyses of sequential flow.

To code content data, we use Latent Semantic Analysis (LSA) (Landauer, Foltz, & Laham, 1998). LSA is a computational linguistic technique which can measure the semantic similarity between strings of text. Its “knowledge” of the language is based on a semantic model of domain knowledge acquired through training on a corpus of domain-relevant text. Through a statistical analysis of how words occur across contexts (e.g., paragraphs), LSA generates a high-dimensional semantic space, in which each original word, as well as larger units of text (utterances, paragraphs, documents), are represented as vectors in the space. The derived vectors for words and utterances can be correlated by taking the cosine between their vectors. This permits the matching of strings based on semantic relatedness, rather than just direct keyword overlap.

LSA can be applied to communication data in a variety of ways, with differing degrees of automation. A completely automatic means of categorizing content is to use LSA to generate a correlation matrix of every utterance with every other utterance, then cluster highly correlated utterances. Each original utterance can be classified according to the cluster of which it is a part. A less automatic method is to develop a coding scheme, and create text strings that are prototypically representative of each coding category. Then, each utterance can be classified according to which of the prototypical categories it correlates with the best. Even less automatically, raters can code a subset of the dialogues, and use LSA to compare each new utterance to one of the pre-classified utterances. LSA can further be used to automatically analyze the coherence, quality and amount of information flowing between speakers, based on analyzing utterances. This permits a wide range of measures that can be correlated to team performance.

4. EXAMPLES

Our approach incorporates several sequential data analysis methods to further refine the data, both in content and in physical dimensions of communication. Tools include lag sequential analysis (Bakeman & Gottman, 1997), graphical display methods, ARIMA models (Suen & Ary, 1989) and fourier analysis (Vallacher & Nowak, 1994), network models such as PRONET (Cooke, Neville, & Rowe, 1996), etc. For instance, we have analyzed team discourse using communication log data, LSA applied to the transcribed dialogue, and PRONET as a graphical, sequential data compression tool. The sequel describes an example, using LSA to assess communication content.

4.1 Content data: static and sequential analyses

Dyads collaborated for one hour to write an essay on censorship and pornography. We used LSA to correlate each utterance with every other utterance. The correlation matrix was then submitted to factor analysis and cluster analysis to give a notion of what sentences were related. We boiled an hour’s worth of discussion down to 13 statement types. In a sense, showing that these specific statements were typical of the communication is a static analysis of content. Further analysis would include counts of how frequently each statement type was uttered, by whom, how much information in multidimensional LSA space is conveyed by each speaker, etc.

We went on to find which statements tend to follow each other by using PRONET (Cooke, Neville, & Rowe, 1996). PRONET is a sequential analysis that relies on the network modeling tool, Pathfinder (Schvaneveldt, 1990). Transition probability matrices among a set of nodes are input to the Pathfinder algorithm, and a network representation of prominent pairwise connections is generated. Here, each statement was coded by classifying it as an instance of a statement type. Statement types were nodes in a lag-1 transition matrix.

Figure 1 shows the output. One possible interpretation is that this dyad's "thought processes" are focused on retaining agreeability. Strong assertions, such as the PornBad statement type, tend to be followed by weaker, more clearly acceptable statements (e.g. "children need to be protected from pornography"), or by clarifying, opinion seeking statements (e.g. "what do you think?"). Note also that specific questions tend to precede questions regarding punctuation, implying that more specific clarifications tend to motivate the team to write.

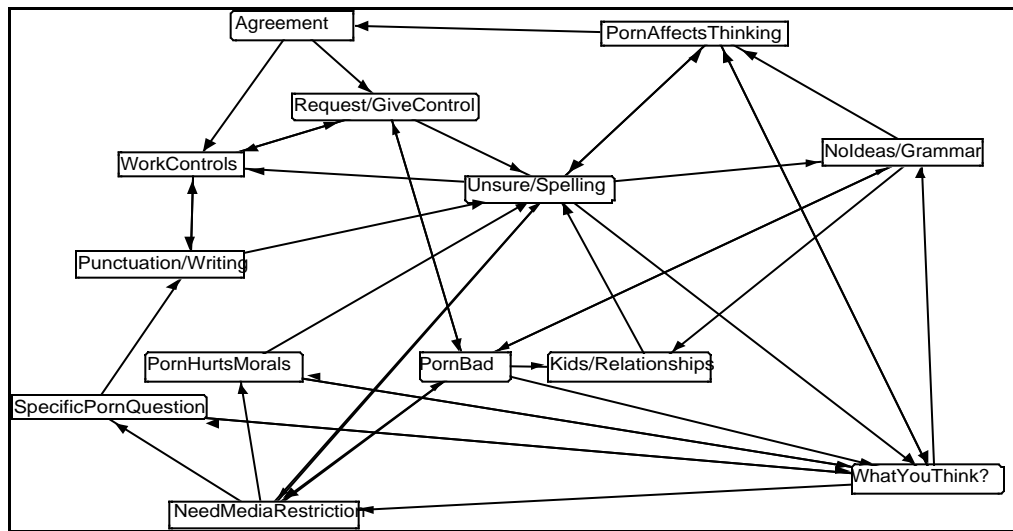


Figure 1. PRONET representation of LSA-classified statement types during a discussion on censorship and pornography. Nodes are statement types. Arrows indicate prominent lag-1 transitions.

Examination of the network helps to create interesting hypotheses about the team's process. It is important to validate the interpretations implied by this type of analysis with other measures. In the next example, the team process and cognition uncovered by the methods we describe was validated by human observers viewing videotaped dialogue. In addition, converging validity was determined against other measures of team cognition, and predictive validity was assessed using measures of team process, performance, and situation awareness.

4.2 Physical data: sequential analysis

This time we use a sequential analysis of physical data taken from the communication log. Each of the three team members had a specialized role: AVO, PLO, or DEMPC. Teams flew a simulated plane for 10 missions. Six events were defined, one for each team member beginning or ending a speech sequence. These events were treated as nodes in a lag-1 transition probability matrix, fed into PRONET (see Figure 2 for output).

Networks for missions 1, 2, and 4 are identical. DEMPC appears to be the focal point between AVO and PLO. At mission 5 this flow pattern is interrupted, and appears relatively chaotic. In particular, the connections between PLO and DEMPC are gone. In subsequent networks, there is never again a completed trapezoid between PLO and DEMPC. In fact, the two are never connected at all, except at mission 9. We chose this team because the PLO and DEMPC had a fight during mission 5. What we find interesting here is that, using the PRONET method, with only lag-1 transitions for this example, we can clearly portray this team's pattern of interaction. Information flow is severely hampered between two team members because of a conflict in team process, and after that, the team's pattern is altered. In essence, the way the team "thinks" is changed by this shock to the system.

4.3 Physical data: static analysis

We conclude with an example of a semi-static analysis of physical data taken from the same team's communication log output. Our goal was to identify the pattern of communication dominance among team members. We separated the mission into segments of one minute duration, and formed a multinomial model of how much each team member spoke during that minute. These models were then used to generate expected values for every other minute, whose observed deviation was tested with a χ^2 . We test every minute against every other minute for similarities in the dominance pattern. Those segments whose differences could not be detected with a χ^2 were pooled to form a new model. The process of pooling and testing was continued until all "similar" minutes

were pooled, and only “distinct” dominance patterns remained. For this static analysis example, we retained counts of distinct patterns the team exhibited during each mission. In addition, flow between minutes could be examined by sequential methods.

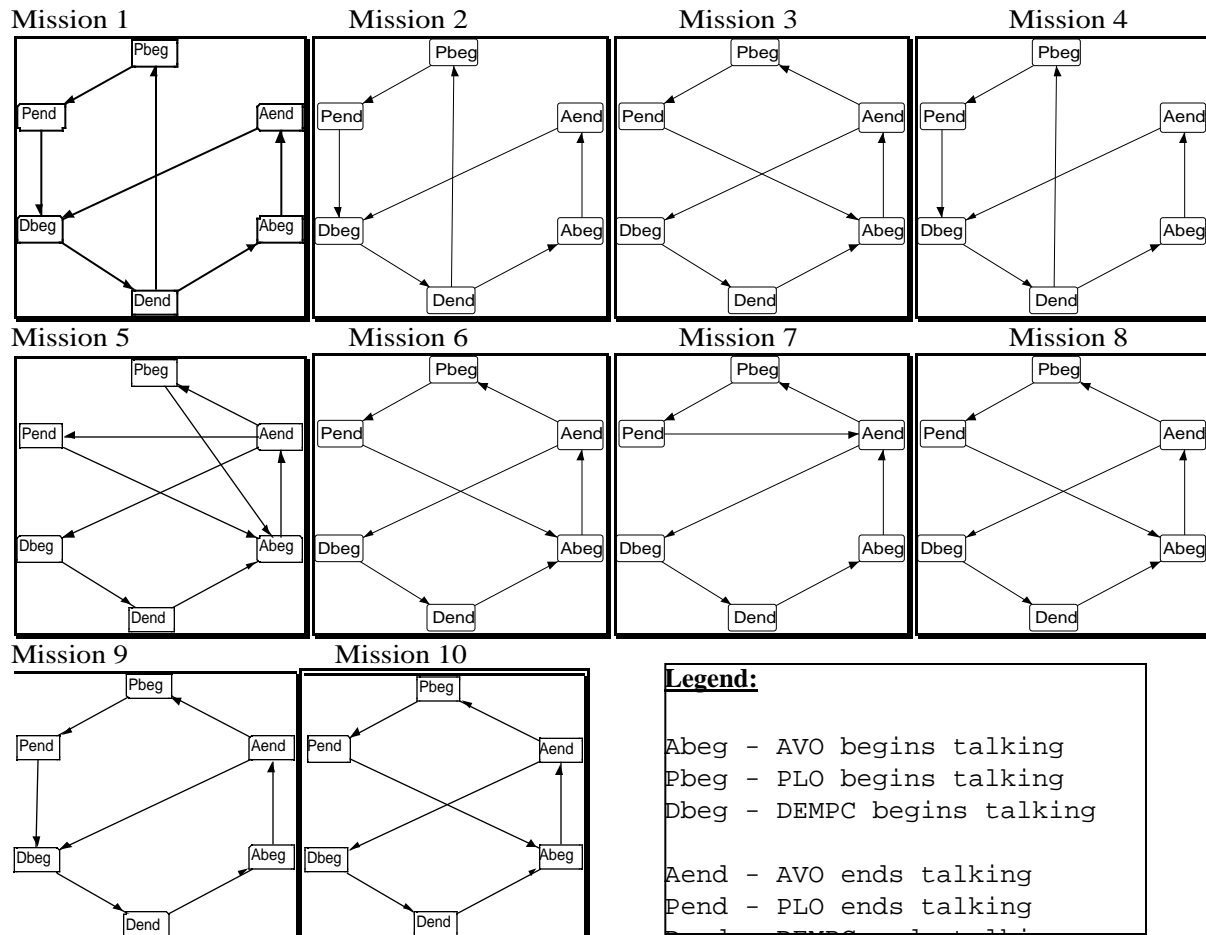


Figure 2. PRONET representation of shift in team communication pattern across 10 missions.

The number of statistically distinct communication patterns in any mission is a measure of communication stability for that mission. We used this to estimate how well the team had an established process for the passing of knowledge. The more patterns exhibited during a mission, the less stable the team’s communication, and so the less stable we predict their team knowledge to be.

We plotted reverse-scaled z-scores of the number of distinct patterns and distinct patterns per minute against z-scores for performance (see Figure 3; performance data were missing for the 10th mission) and situation awareness (not shown). Except at mission 1, where communication patterns were stable in spite of presumably low team knowledge, general shifts in performance and situation awareness correspond to the number of distinct patterns.

5. CONCLUSION AND IMPLICATIONS

A general principle of our approach is to take communication data and reduce it to its core components. Then look for shifts in the communication patterns. Next, tie these pattern shifts to known events and internal team trends. This process is essentially an application of the dynamical systems paradigm (e.g. Vallacher & Nowak, 1994). Employed in an exploratory way, it can lead to general predictive principles of team cognition. These principles can then be applied to other teams by relying solely on the communication data to assess team cognition.

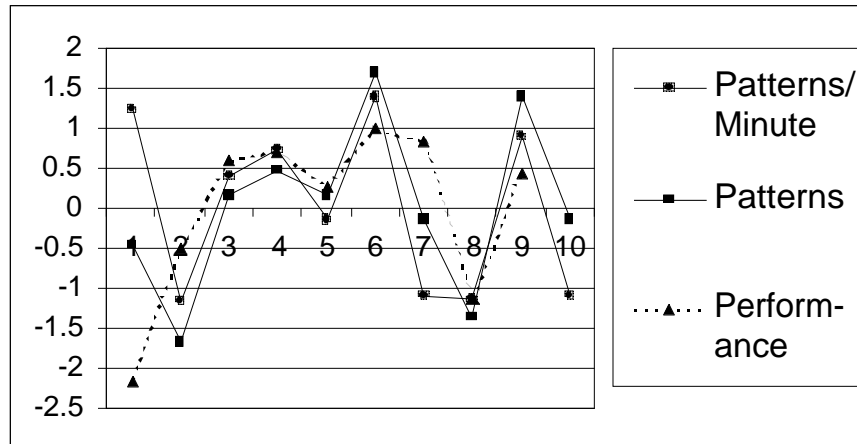


Figure 3. Z scores for number of communication patterns per mission and performance scores across 10 missions.

The efficiency of these methods has not been explicitly computed at this time. But we know that time spent recording who is talking to whom at what time, is eliminated by use of the communication logger. Furthermore, time spent coding text can be eliminated by the use of LSA. With text-based computer mediated communication, even transcription becomes unnecessary.

There are a number of implications that automated, sequential measurement tools have for design and for cognitive modeling. First, these methods will allow for rapid and on-line assessment of team cognition. This is important for evaluating and designing training programs and system interfaces. Moreover, poor process detection can automatically trigger interventions on the part of the system. Second, computerized tools such as PRONET and LSA allow the assignment of numeric values to variables in models of team cognition, facilitating the development of relatively specific predictive models.

6. REFERENCES

- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). Cambridge: Cambridge University Press.
- Cooke, N. J., Neville, K. J., & Rowe, A. L. (1996). Procedural network representations of sequential data. *Human-Computer Interaction, 11*, 29-68.
- Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. (2000). Measuring team knowledge. *Human Factors, 42*, 151-173.
- Emmert, B. J. L. (1989). Interaction analysis. In P. Emmert and L. L. Barker (Eds.), *Measurement of Communication Behavior*, pp. 218-248. New York: Longman.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes, 25*(2 & 3), 259-284.
- Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors, 42*, 242-271.
- Schvaneveldt, R. W. (Ed.) (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.
- Steiner, I. D. (1972). *Group Process and Productivity*. New York: Academic Press.
- Suen, H. K. & Ary, D. (1989). *Analyzing Quantitative Behavioral Observation Data*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vallacher, R. R., & Nowak, A. (Eds.) (1994). *Dynamical Systems in Social Psychology*. San Diego: Academic Press.
- Watt, J. H. & VanLear, C. A. (Eds.) (1996). *Dynamic patterns in communication processes*. Thousand Oaks, CA: Sage.

7. ACKNOWLEDGEMENTS

This work was supported by ONR Grant No. N00014-00-1-0818.