

# Analysis of Semantic Content and Its Relation to Team Neurophysiology during Submarine Crew Training

Jamie C. Gorman<sup>1</sup>, Melanie J. Martin<sup>2</sup>, Terri A. Dunbar<sup>1</sup>,  
Ronald H. Stevens<sup>3</sup>, and Trysha Galloway<sup>3</sup>

<sup>1</sup> Texas Tech University, Lubbock, TX, USA

<sup>2</sup> California State University—Stanislaus, Stanislaus, CA, USA

<sup>3</sup> The Learning Chameleon Inc., Culver City, CA, USA

jamie.gorman@ttu.edu

**Abstract.** A multi-level framework for analyzing team cognition based on team communication content and team neurophysiology is described. The semantic content of team communication in submarine training crews is quantified using Latent Semantic Analysis (LSA), and their team neurophysiology is quantified using the previously described neurophysiologic synchrony method. In the current study, we validate the LSA communication metrics by demonstrating their sensitivity to variations in training segment and by showing that less experienced (novice) crews can be differentiated from more experienced crews based on the semantic relatedness of their communications. Cross-correlations between an LSA metric and a team neurophysiology metric are explored to examine fluctuations in the lead-lag relationship between team communication and team neurophysiology as a function of training segment and level of team experience. Finally, the implications of this research for team training and assessment are considered.

**Keywords:** Latent Semantic Analysis, Team cognition, Team communication, Team neurophysiology, Teamwork.

## 1 Introduction

A team is an interdependent group of two or more people who work together for a fixed amount of time to achieve a common goal [1-2]. Across a wide range of work environments, including business, military, medical, academic, and culinary settings, there are many common goals that are either too physically or cognitively demanding to be achieved by individuals working alone. To meet such goals, tasks must be performed in real time by people working together as a team. This paper focuses on the communicative and neurophysiological aspects of team cognition as crews work together to solve navigation problems and coordinate solutions in a submarine crew training environment.

## 1.1 Levels of Analysis in Team Cognition

Team cognition has been defined theoretically as either the shared declarative knowledge of team members, the *shared cognition perspective* [e.g., 3-5], or as the dynamic interactions (e.g., communications) between team members, the *interactive theory of team cognition* [6]. Inspired by the interactive theory of team cognition, we take a multi-leveled approach to studying team cognition as it unfolds across segments of submarine crew training. In this research, our overarching focus is on how team communication and team neurophysiology function as related metrics for team cognition, though each is concerned with a different level of analysis.

At a more “micro” level of analysis, we have analyzed neurophysiological patterns of team members as they work together to acquire team skill [7]. We argue that solely focusing on individual neurophysiology does not capture the dependencies that develop across team members as they acquire overt interaction patterns. The team neurophysiology paradigm developed by Stevens and colleagues [7-8] addresses this by allowing one to examine distributions of neurophysiological patterns as they develop across team members. We have also developed a variety of approaches for studying team cognition at a more “macro” level by focusing on overt interaction patterns during team skill development [9]. Those overt measures of team cognition include behavioral [e.g., 10] and communication-based [e.g., 11] metrics. The major theme of this paper is to continue to extend our communication-based metrics of team cognition into a submarine crew training environment, while beginning to develop a framework to support a joint team communication—team neurophysiology paradigm for understanding team skill development.

The team neurophysiology methods considered in this paper are described elsewhere [7-8]. Therefore, we want to devote much of this paper to describing our analysis of team cognition during submarine crew training using semantic content analysis of team communication. Hence, we devote the following section to team communication analysis.

## 1.2 Analyzing Team Communication Using Latent Semantic Analysis

Latent Semantic Analysis (LSA) is a mathematical/statistical method for representing and analyzing semantic knowledge within a domain [12]. LSA is based on the theory that knowledge is reflected in the contextual usage of words within meaningful discourse [13]. LSA takes as its input a raw corpus of text and represents the corpus as a matrix of unique terms (e.g., words) by documents (e.g., paragraphs). LSA assumes that lower-dimensional (latent) semantic factors account for the frequency of co-occurrence between words and documents in the raw matrix. The space of factors is called the semantic space, and it is constructed through singular value decomposition. The optimal number of dimensions can be determined such that the relationships between words and context results in correct inductions (e.g., synonym matching; missing word replacement).

LSA has been used to successfully distinguish high-performing from low-performing unmanned air vehicle (UAV) teams by comparing their transcribed communications to a UAV semantic space [11]. For the current study, we constructed a

semantic space representing nautical navigation knowledge with which to analyze submarine crew communication. The semantic space (314 factors, 124,326 total words, 6,846 terms, and 5,904 documents) was constructed from a corpus created from submarine crew training transcripts, the Navigational Rules of the Road (COLREGS), Submarine Operations Manual, and the unclassified Doctrine for Submarine Interior Communications.

### **1.3 The Current Study**

The primary goals of the current study are to (a) validate semantic content metrics for identifying critical differences between submarine crew training segments, (b) use semantic relatedness metrics to differentiate between levels of submarine crew experience, and (c) demonstrate the dependent relationship between team semantic content and team neurophysiology in a submarine crew training environment. Finally, we consider the implications of our results for team training and assessment.

## **2 Method**

### **2.1 Participants**

The data used in the current study were collected from Junior Officer Navigation teams enrolled in the Submarine Officer Advanced Candidacy class at the US Navy Submarine School. These teams consisted of six or more crew members, including: Quartermaster on Watch; Navigator; Officer on Deck; Assistant Navigator; Contact Coordinator; and Radar. (Other team members were also present and participated, but were not analyzed using the neurophysiological methods described later.) These teams participated in Submarine Piloting and Navigation (SPAN) simulation sessions during the class. We analyzed seven of these SPAN sessions: Four are from more experienced teams, and three are from less experienced (“novice”) teams. In the statistical analyses we present below, we use a between-subjects variable, Experience, to index Novice vs. Experienced SPAN training sessions.

### **2.2 Training Simulation**

The communication metrics analyzed in the current study were calculated from transcripts of team communication across crews during SPAN training simulations. The SPAN sessions are high-fidelity training simulations that consist of three segments: Briefing; Scenario; and Debriefing [7-8]. During the Briefing segment, the overall goals of the mission are presented and discussed. The Scenario is the dynamically evolving segment of the training, during which teams navigate through a route in a high-fidelity submarine simulation. The Scenario segment requires teams to steer and change course or speed while identifying landmarks and other ships that factor into SPAN. During the Scenario, the team must also periodically take Rounds, during which three navigation points are chosen, and the bearing of each point from the boat

is measured and plotted on a chart. The accuracy and variability of Rounds may serve as a team performance measure in future research, but those data will not be analyzed here. The Debriefing segment is an after-action-review, during which teams discuss what worked and what other options or actions could have been taken during the Scenario. The Debriefing segment provides a teaching experience, where both short- and long-term learning goals are discussed. In the statistical analyses we present below, we use a within-subjects variable, Training Segment, to index the Briefing, Scenario, and Debriefing training segments. In addition to communication metrics, neurophysiology data taken from crew members during the training simulations are used in some analyses presented below.

### 2.3 Measures

Two metrics derived from the geometrical interpretation of the semantic space are (1) the *vector length* of a piece of discourse and (2) the *cosine* between two pieces of discourse. We calculated both of these metrics for seven SPAN transcripts (i.e., the four experienced team sessions and three novice team sessions). These metrics will be used to analyze the semantic content of team communication during SPAN.

The vector length of a piece of discourse (e.g., an utterance; “Recommend steering course 178 to regain track.”) is the Euclidean norm of the vector, created by summing the semantic space vectors of words in the discourse, plotted in the semantic space. The vector length measures the amount of semantic content (cf. knowledge) a piece of discourse contains relative to the domain of discourse, as represented by our SPAN semantic space.

The cosine between any two pieces of discourse (e.g., any two utterances; any two training segments; any two complete transcripts; etc.) is the vector dot product between two vectors plotted in the semantic space. The correlation between two vectors can be shown to be the cosine of the angle joining them (e.g., independent, perpendicular vectors have  $\cos[90^\circ] = 0$ , and they are completely uncorrelated). Hence, the cosine measures the degree of semantic relatedness, or correlation, between any two pieces of discourse.

The team neurophysiological measure we will use (*NS Entropy*) to examine the relationship between communication content and team neurophysiology is derived from the EEG-based Neurophysiological Synchrony (NS) method, which is more fully described elsewhere [7-8]. Using this method, discrete, team-level NS states are sampled at a fixed interval (we used 1 Hz) from continuous EEG streams collected from each team member. The EEG-to-NS mapping is such that each discrete NS state identifies a different distribution of cognitive engagement (or workload; not analyzed here) across team members. As training segments unfold, the team engagement distribution changes, and is captured in a time series of discrete NS states. The set of NS states for SPAN was determined using an artificial neural network approach [7-8], which resulted in a set of 25 discrete NS states.

Though the cardinality of the NS states is fixed, there is no inherent numerical ordering of states. To quantify NS organization, therefore, we calculated the Shannon entropy across NS states using a sliding window of size 100s.

$$NS \text{ entropy} = - \sum_{i=1}^{\#NS \text{ States}} p_i \cdot \log p_i, \quad (1)$$

where  $p_i$  is the relative frequency of NS state  $i$  over a 100s window, was repeatedly calculated as the 100s window slid over the original, discretely-varying NS time series. Using this technique, for an input NS time series of length  $N$ , the output is a continuously-varying NS entropy time series of length  $N - 99$ . In this way, we use the first 100 samples to calculate the first entropy value at time  $t = 100$ , samples 2 through 101 to calculate the second entropy value at  $t = 101$ , and so forth. Using a window smaller than 100s has been found to increase the potential for false (discontinuous) spikes in the NS entropy time series [7]. In terms of team cognition, low entropy may be interpreted as a highly-ordered team neurophysiological state, whereas high entropy corresponds to a more random mix of team neurophysiological states [7].

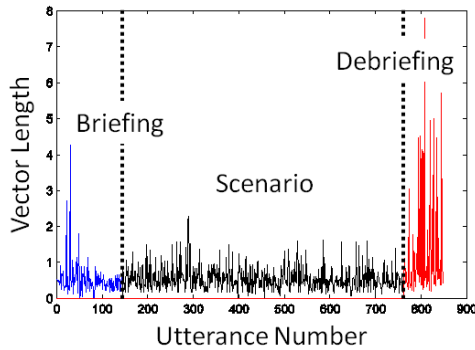
### 3 Results

#### 3.1 Differentiating between Task Phases Using Vector Length

To examine whether Training Segment and Experience underlie communication differences captured by LSA metrics, we first computed mean vector length across utterances for Briefing, Scenario, and Debriefing segments of each transcript and then we analyzed those mean vector lengths using a 3 (Training Segment)  $\times$  2 (Experience) mixed ANOVA. (We also analyzed cosines taken between successive utterances using this approach; however, none of those results were significant.) As illustrated by the vector length data shown in Figure 1, there was a significant main effect of Training Segment,  $F(2, 10) = 15.78$ ,  $p = .001$ ,  $\eta^2 = .76$ . No other omnibus effects were significant. A follow-up Tukey test on Training Segment ( $\alpha_{FW} = .05$ ) revealed that mean Debriefing vector length ( $M = 1.21$ ;  $SD = .27$ ) was significantly greater than both Briefing ( $M = .79$ ;  $SD = .30$ ) and Scenario ( $M = .51$ ;  $SD = .03$ ) mean vector lengths.

Careful reading of the utterances in the transcripts clearly indicated that teams were communicating differently as a function of Training Segment. Specifically, teams communicated with shorter, to-the-point utterances during the Scenario segment and longer, conversational utterances during the Briefing and Debriefing segments. To quantify this observation, we controlled for word count by dividing each utterance's vector length by the number of words in each utterance. The resulting quantity measures the rate of semantic content per word in each utterance: Communication Efficiency = Vector Length / Word Count [11]. To examine whether Training Segment and Experience underlie communication differences captured by the Communication Efficiency measure, we computed mean Efficiency across utterances for Briefing, Scenario, and Debriefing segments in each transcript and then analyzed those mean

Efficiency values using a 3 (Training Segment)  $\times$  2 (Experience) mixed ANOVA. The main effect of Training Segment was significant,  $F(2, 10) = 25.81, p < .001, \eta^2 = .84$ . No other omnibus effects were significant. A follow-up Tukey test on Training Segment ( $\alpha_{FW} = .05$ ) revealed that Communications Efficiency was significantly higher during the Scenario ( $M = .078; SD = .001$ ) than during the Briefing ( $M = .054; SD = .006$ ) and Debriefing ( $M = .039; SD = .003$ ) segments.



**Fig. 1.** Vector length of each utterance for an experienced team separated by Training Segment

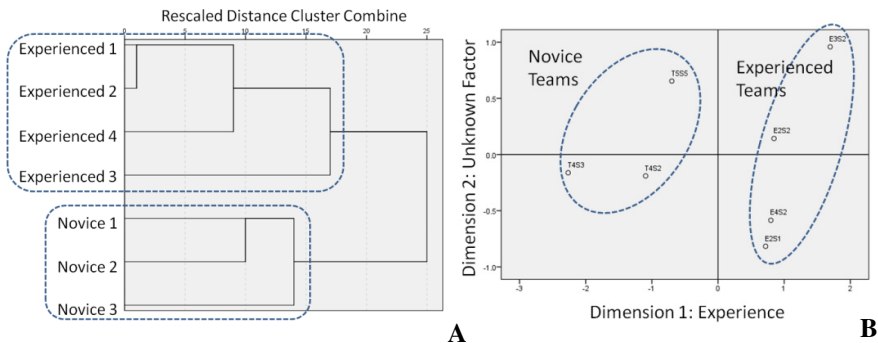
### 3.2 Differentiating between Experienced and Novice Teams Using Cosines

To determine whether experienced teams' semantic content was more similar to each other than to novice teams and that novice teams' semantic content was more similar to each other than to experienced teams, we first calculated the LSA cosine metric between all possible pairs of transcripts as a function of Training Segment. We show the cosine matrix for the Scenario training segment in Table 1. If it is the case that semantic relatedness differentiates between experienced and novice team communication, then the bold values in Table 1 should be larger than the italicized values. We further examined that qualitative grouping using cluster analysis and multidimensional scaling (MDS).

A hierarchical cluster analysis of the Scenario cosine matrix (Table 1) using average between-groups linkage revealed that experienced teams clustered together and novice teams clustered together based on the semantic content of their communications for the Scenario training segment (Figure 2a). Similarly, a two-dimensional MDS solution for the Scenario cosine matrix ( $Stress = .86; R^2 = .96$ ) revealed an "Experience" dimension, with novice teams low and experienced teams high on this dimension, and a second dimension that also appears to differentiate between teams by an as yet unidentified factor (Figure 2b). Hierarchical clustering and MDS conducted on the Briefing and Debriefing cosine matrices revealed that novice teams tended to be more tightly grouped in terms of semantic content than experienced teams during those training segments.

**Table 1.** Cosine Similarity Matrix Computed between All Pairs of Transcripts during Scenario

	Exper. 1	Exper. 2	Exper. 3	Exper. 4	Novice 1	Novice 2	Novice 3
Exper. 1	1.00						
Exper. 2	<b>0.91</b>	1.00					
Exper. 3	<b>0.81</b>	<b>0.85</b>	1.00				
Exper. 4	<b>0.85</b>	<b>0.87</b>	<b>0.82</b>	1.00			
Novice 1	0.82	0.85	0.78	0.81	1.00		
Novice 2	0.76	0.77	0.72	0.77	<b>0.88</b>	1.00	
Novice 3	0.81	0.81	0.79	0.84	<b>0.85</b>	<b>0.83</b>	1.00



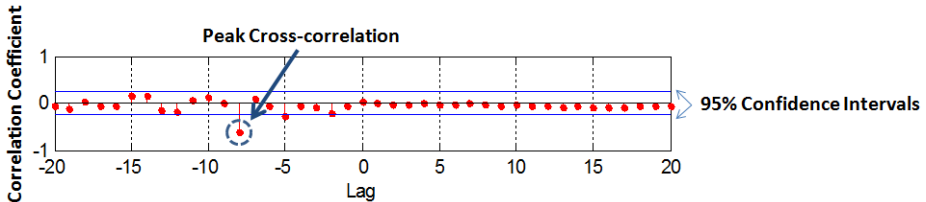
**Fig. 2.** (A) hierarchical clustering of the LSA cosine matrix from the Scenario training segment; (B) multidimensional scaling of the LSA cosine matrix for the Scenario training segment

### 3.3 Cross-Correlations between Semantic Content and Team Neurophysiology

Having established that (a) LSA-based vector length metrics differ as a function of training segment and (b) that LSA-based cosine metrics differentiate between experienced and novice teams, we turn to the question of how these differences in team communication are related to changes in team neurophysiology as a function of Training Segment and Experience.

We calculated the lagged cross-correlation function between LSA vector length of each utterance (Variable 1) and mean NS Entropy during each utterance (Variable 2) for each combination of Training Segment and Experience. The peak cross-correlation between these two variables (e.g., Figure 3) was identified to determine whether semantic content was leading (+ lag) or following (- lag) team neurophysiology and whether that correlation was significantly positive (+ direction) or negative (- direction). Table 2 provides basic information for each of the cross-correlations analyzed in the current study. (Because we were simply concerned with determining whether these variables were cross-correlated for the current study, we do not report or interpret exact values of lags and strength of correlation in this paper.) If the

correlation is significantly negative at a negative lag, as it is for the novice teams during the Briefing segment, then team neurophysiology is leading during that segment of training, such that higher entropy tends to temporally precede lower vector lengths. In terms of team cognition, this suggests that a more random mix team neurophysiological engagement states tends to temporally precede a reduction in the semantic content (cf. “knowledge”) embodied in a team’s communications, at least for some training segments and levels of team experience. Interestingly, we see the opposite cross-correlational pattern emerge for two of the three novice teams during the Debriefing segment. Though the cross-correlational patterns for the experienced teams are more varied, their interpretation can be carried out in the same way as our interpretation of novice teams’ patterns.



**Fig. 2.** Cross-correlation function for a novice team during the Briefing segment

## 4 Discussion

The LSA-based vector length metric of team communication content significantly differed as a function of training segment, and the cosine metric allowed us to differentiate between experienced and less experienced submarine crews. Those results lend support to our expectation that LSA-based metrics of semantic content can successfully distinguish teams of different skill levels and under different task constraints (e.g., planning a scenario vs. actually performing the scenario). The LSA communication efficiency measure also differed as a function of training segment. This metric is a hybrid between semantic content and simple syntactic markers of communication and represents just one of many possibilities for augmenting LSA metrics by adding contextual and syntactic details of team communication. In future research, each of these metrics should be validated against objective team performance measures (e.g., taking Rounds).

The lagged cross-correlations were largely significant and seem to suggest that team communication and team neurophysiology may lead or lag each other at different stages of team training or experience. Interestingly, we saw a mix of positive and negative peak cross-correlations, indicating that at times higher synchronization at the neurophysiological level is associated with increased domain specific semantic content at the verbal communication level and at other times with decreases in domain specific communication content. The prevalence of significant cross-correlation between team neurophysiology and communication content begins to lend support to a joint team communication—team neurophysiology paradigm for understanding team skill development; however, the variety of lead-lag patterns and directions of cross-correlation must be disentangled in future research.



**Table 2.** Lag and Direction of Peak Cross Correlation between LSA Vector Length and NS Entropy as a Function of Training Segment and Experience

Experience	Training Segment					
	Briefing		Scenario		Debriefing	
	Lag	Direction	Lag	Direction	Lag	Direction
Exper. 1	-	+ *	-	+ *	+	- *
Exper. 2	+	+	+	+	+	+ *
Exper. 3	-	- *	-	+ *	-	+ *
Exper. 4	-	- *	-	+ *	+	+ *
Novice 1	-	- *	-	- *	+	+ *
Novice 2	-	- *	-	+ *	-	+ *
Novice 3	-	- *	-	- *	+	+ *

Note. \* This correlation lies beyond the 95% Confidence Interval for no correlation;  $p < .05$ .

Finally, in this paper we have suggested a multi-leveled analysis of team cognition, which has implications for team training and assessment. During team development, the goal of team performance may be reflected in a variety of adjustments in the neurophysiological and overt behavioral patterns exhibited by teams as they learn to work together. In keeping with the interactive theory of team cognition [6], and similar to the theory of embodied cognition [14], neurophysiological patterns may constrain behavior patterns, or vice versa, at critical points during team skill development. As we are beginning to see with cross-correlation analyses, hierarchical patterns of constraint, in the form of lead-lag relationships between team neurophysiology and team communication, may be significantly altered by type of training and level of team experience. Although the exact nature of these developmental transitions remains to be seen; in the future, a joint neurophysiological/communication analysis may be critical for assessing key transitions in neural/cognitive team development.

**Acknowledgments.** This research is supported by Defense Advanced Research Projects Agency Contract W31P4Q-12-C-0166 and National Science Foundation Small Business Innovation Research Grant IIP 1215327. The findings, views, and opinions expressed in this paper are the authors' and do not necessarily represent the official views of any funding agency.

## References

1. Cooke, N.J., Salas, E., Cannon-Bowers, J.A., Stout, R.: Measuring team knowledge. *Human Factors* 42, 151–173 (2000)
2. Salas, E., Dickinson, T.L., Converse, S.A., Tannenbaum, S.I.: Toward an understanding of team performance and training. In: Swezey, R.W., Salas, E. (eds.) *Teams: Their Training and Performance*, pp. 3–29. Ablex, Norwood (1992)

3. Cannon-Bowers, J.A., Salas, E., Converse, S.: Shared mental models in expert team decision making. In: Castellan, N.J. (ed.) *Individual and Group Decision Making*, pp. 221–246. Lawrence Erlbaum Associates, Hillsdale (1993)
4. DeChurch, L.A., Mesmer-Magnus, J.R.: The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology* 95, 32–53 (2010)
5. Langan-Fox, J., Code, S., Langfield-Smith, K.: Team mental models: Techniques, methods, and analytic approaches. *Human Factors* 42, 242–271 (2000)
6. Cooke, N.J., Gorman, J.C., Myers, C.W., Duran, J.L.: Interactive team cognition. *Cognitive Science*. [Electronic publication ahead of print] (2012)
7. Stevens, R.H., Gorman, J.C., Amazeen, P., Likens, A., Galloway, T.: The organizational dynamics of teams. *Nonlinear Dynamics, Psychology and Life Sciences* 17, 67–86 (2013)
8. Stevens, R., Galloway, T., Wang, P., Berka, C.: Cognitive neurophysiologic synchronies: What can they contribute to the study of teamwork? *Human Factors* 54, 489–502 (2012)
9. Cooke, N.J., Gorman, J.C.: Interaction-based measures of cognitive systems. *Journal of Cognitive Engineering and Decision Making* 3, 27–46 (2009)
10. Gorman, J.C., Cooke, N.J., Amazeen, P.G., Fouse, S.: Measuring patterns in team interaction sequences using a discrete recurrence approach. *Human Factors* 54, 503–517 (2012)
11. Gorman, J.C., Foltz, P.W., Kiekel, P.A., Martin, M.J., Cooke, N.J.: Evaluation of latent-semantic analysis-based measures of team communications. In: *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, pp. 424–428 (2003)
12. Landauer, T.K., Foltz, P.W., Laham, D.: Introduction to Latent Semantic Analysis. *Discourse Processes* 25, 259–284 (1998)
13. Landauer, T.K., Dumais, S.T.: A solution to Plato’s problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review* 104, 211–240 (1997)
14. Shapiro, L.A.: *Embodied cognition*. Routledge, New York (2011)