At the operational warfighting level, Objective Force units as part of joint teams will conduct operational maneuvers from strategic distances (U.S. Army White Paper: Concepts for Objective Force).

The Problem
Team communication data can provide a rich data set for analyzing team performance. However, hand coding for content can take a very long time and can be subjective.

Goal
Predict team performance based on automated analyses of communication among team members.

Approach
Use Latent Semantic Analysis, a computational model of language and memory, to code/measure and content of communication team tasks.

Evaluate effectiveness of these techniques for predicting team performance and decision making.

Evaluate the combination of these techniques with other methods of cognitive modeling for improving measurement performance.

The CERTT Lab
- Research laboratory to develop, apply, and evaluate measures of team cognition.
- Hardware and software support synthetic team tasks.
- Can be configured to simulate colocated or distributed task environments.
- Current Configuration: Unmanned Air Vehicle Control.

- An automated measurement and recording system captures team behaviors:
  - audio & video streams
  - task performance indices
  - communication flow
  - computer events
- Post-processing routines summarize the data.
- Experimenters observe team behavior remotely via:
  - audio & video monitors
  - performance indices
  - shared displays
- Experimenters record observations using time-stamped annotation software.
- Participants complete on-line measures that elicit task and team knowledge:
  - factual test
  - structured interviews
  - concept ratings (individual & teams levels)

How do we automatically determine what makes a good/poor team?
Analyze semantic components of communication
Use these semantic components to predict knowledge and communications skills and predict team performance.

Latent Semantic Analysis (LSA) is a:
- Psychological Theory:
  - of the acquisition, induction, and representation of knowledge
- Model:
  - how people learn the meanings of words
- Model:
  - a mathematical system for computational modeling of cognitive processes
- Tool:
  - Artificial Intelligence (Machine Learning) system for matching words/texts at a semantic level
- LSA learns the relationships between text documents and their constituent words (term) when trained on very large numbers of background texts (thousands to millions)
- LSA learns how to group documents and terms that are similar in a "Semantic Space"
- Documents can be domain knowledge, writing samples, e-mail files, course materials, personnel records, etc.
- LSA judgments of similarity agree well with human judgments

Measuring Team Communication with LSA

A good team
- High agreement: 5.00
- High disagreement: 1.00
- Low agreement: -5.00
- Low disagreement: -1.00
- +ve: team is performing well
- -ve: team is performing poorly
- 0: the team isn’t being assessed

A poor team
- High agreement: 0.00
- High disagreement: 0.00
- Low agreement: -0.00
- Low disagreement: -0.00
- +ve: team is performing relatively well
- -ve: team is performing relatively poorly
- 0: the team isn’t being assessed

How do we automatically determine what makes a good/poor team?
Analyze semantic components of communication
Use these semantic components to predict knowledge and communications skills and predict team performance.

Experiments
The Data
- 67 Transcripts from 11 teams, 7 missions
  - XML tagged
  - 2700 minutes of spoken dialogue
  - 20,545 separate utterances (turns)
  - 232,000 words (660 k bytes of text)
- Logs of the speaker, listener, and duration of each communication from each participant

Experiment 1: Predicting team performance from dialogues as a whole
Goal:
Predict objective individual and team performance measures based on transcript of team communication.

Approach:
- Match team dialogue patterns and content against database of prior dialogues.
- Assign a predicted team score based on similarity of a transcript to other team transcripts with known performance measures.

Result:
Correlation between LSA-derived measures of communication to Team Performance measures r=.76 (p<.01)

Implication:
We can automatically predict how well a team is performing based on analysis of their communication.

Experiment 2: Automatic Tagging of Transcripts
Goal:
Tag utterances from transcripts for types of dialogues
- Frequency and sequences of tags can predict team performance

Approach:
- For each utterance, find the most semantically similar one in a database of tagged utterances.
- Suggest when trainer needs to intervene
- Analyze rich source text to that utterance.

Results:
- Human-human reliability: 0.71 (c-value of agreement)
- LSA-Human reliability: 0.59
- LSA-Syntax reliability: 0.63
- LSA-based confidence measure of tagging can help improve performance to be even closer to that of human performance

Implications:
We can automatically tag transcripts and use results to predict team performance.
Tagging by computer takes seconds compared to minutes to hours by human coders.
Performance is not quite as human-human reliability, but can provide an acceptable level of accuracy to provide fast predictions.

Conclusions
- Objective Force and FC8 put greater emphasis on distributed team communication.
- Monitoring and assessing team performance will become more critical in these distributed contexts.
- Secure level is too low for obtaining diagnostic information for training, design, and selection.
- Semantic and statistical analyses of team dialogues can reveal the effectiveness of a team.
- Permits automatic analyses of the content of team dialogues.
- This approach can be applied to any domain in which there is team dialogue.

Current Extensions
- Creation of new laboratory for studying team communication in a simulated SASSO context underway at NMSU using DDD.
- Development of a web-based communication analysis system for use by ARL CTA member underway.
- Can use the approach to determine what makes a good team.
- Match team dialogue patterns and content against database of prior dialogues.
- Predict individual and team performance.
- Detect “unusual” events.
- Automatically code utterances for types of dialogues.
- Matching individuals, skills, training material to teams.

Applications of LSA
- Individual/Team Assessment
  - Evaluating teams and individuals through measurement of communication
  - Identifying critical events
  - Suggest when trainer needs to intervene
  - Suggest ways to improve team communication
- Matching individuals, skills, teams, and training material

Additional Applications
- Integrated into existing cognitive models (ACT-R)
- Automated essay scoring/knowledge assessment embedded within training (scaffolding)
- Automated tools for aiding collaborative learning environments
-Within and Cross-Language filtering/Retrieve/Topic-detection systems
-Lessons Learned systems

Acknowledgements
This research was completed in collaboration with the CERTT laboratory including, Nancy Cooke, Steven Shope, Preston Kiekel, and Jamie Gorman.

Peter W. Foltz & Melanie J. Martin
New Mexico State University
pfolez@nmsu.edu