



**APTIMA**  
HUMAN-CENTERED  
ENGINEERING

# **A Methodology to Predict Specific Communication Themes from Overall Communication Volume for Individuals and Teams**

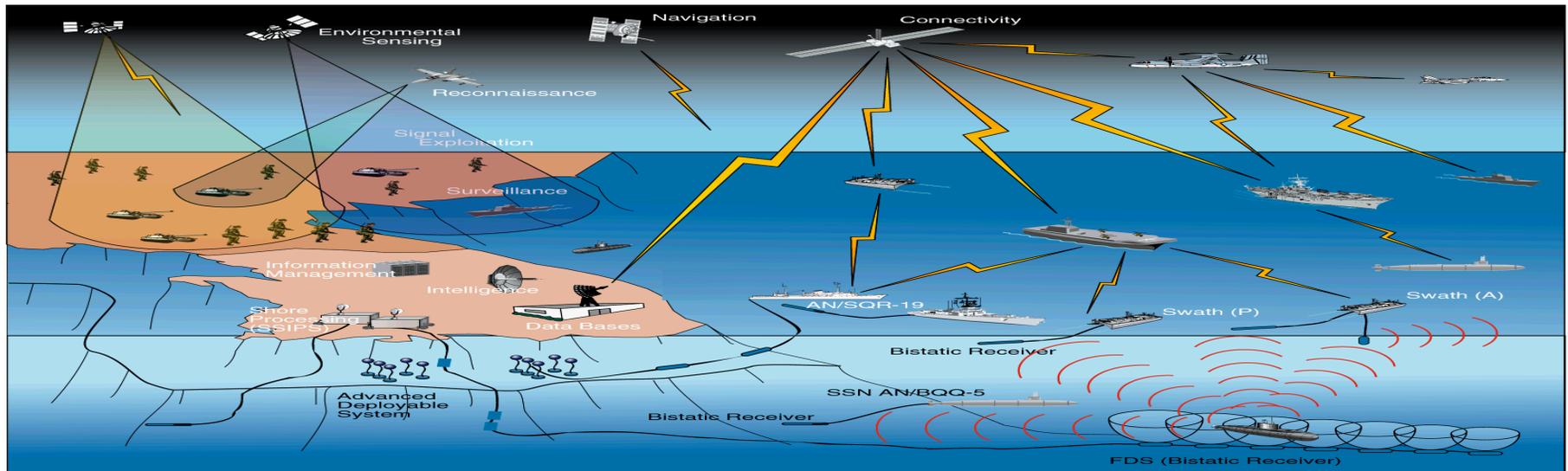
**The Command & Control Research & Technology  
Symposium, San Diego, CA**

**20 June 2006**

**Elliot E. Entin  
Shawn A. Weil**

**Sponsored in part by the Office Of Naval Research,  
Contract No. N00014-99-C-0255, COTR Gerald Malecki**





- Communications – key to the Network Centric world and the “life’s blood” to organizations and teams
  - Coordinate actions such as the reallocation of assets, the redistribution of workload, and joint processing of tasks
  - Major conduit for the sharing of information and identifying the need for change



- Various protocols exist to code and analyze communication
  - Most procedures are difficult and time consuming
    - Counts of utterances, though straightforward, do not provide a meaningful window into team processes
    - At a detailed semantic level it can take months to code
- Intermediate level
  - Adopt small finite number of categories
  - Capture speaker and recipient





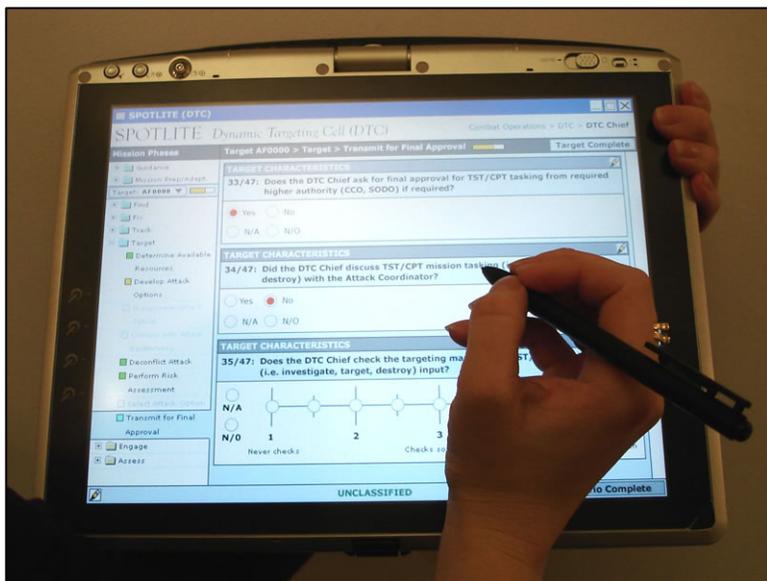
# Communication Categories

## Two primary axes: Request & Transfers, Information & Action

- Information requests
- Information request about a mission task
- Information request about an asset
- Action and task request
- Resource utilization request
- Coordination request
- Information transfers
- Information transfers about a mission task
- Information transfer about an asset
- Action and task transfers
- Resource utilization transfers
- Coordination transfers
- Acknowledgements



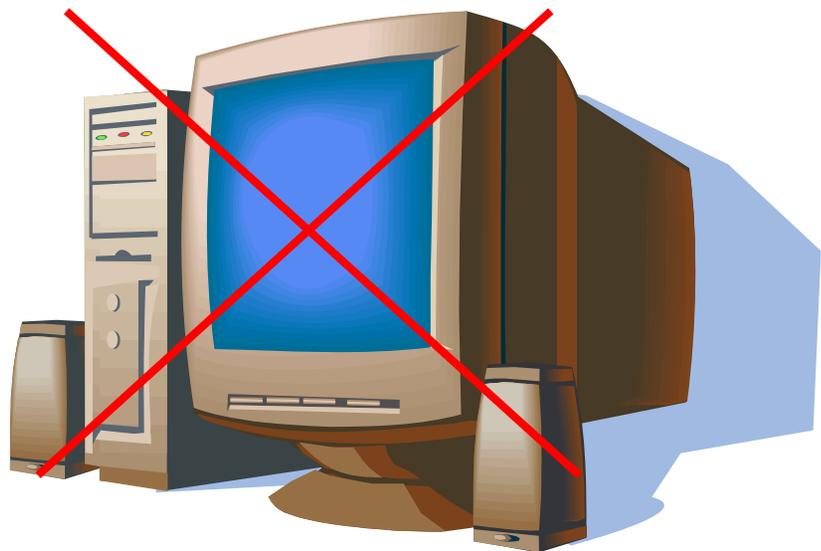
# Coding Communication In Real-time



- **Coding done at an intermediate level of detail**
  - Incorporates both semantic and quantitative aspects of the communication stream
- **Computer Assisted Coding**
  - Touch sensitive tablet
  - Specifically designed software
  - Two interfaces
  - Application produces two output files
- **Training**
  - 25-40 hours to become proficient
- **Adequate procedure, but still time consuming and complex**



# Automated Methods for Communications Coding?



- No software application or package appears to be up to the challenge
  - Some applications can count the number of utterances (communication volume)
- If I know communication volume, can I know more?
  - Could I build a model or family of models to predict some about pattern of communications within a team?
  - **Maybe**



# Methodology: Overview

- Randomly select a 20% to 25% sample of trials from all trials conducted
- Code communications manually for this random sample
- Construct a linear regression model using total number of utterances to predict one or more communication categories
  - This results in a **part-whole correlation** and necessitates a **correction** to the coefficient to control for inflation
  - One such correction is described by Guilford (1965, p.350)
  - **Corrected part-whole correlation is used to build regression model**



- To test the veracity of the procedure we **used archival data**
  - 32 trials from Diedrich et al. (2003)
  - **Four communication categories selected**: information requests, information transfers, coordination requests, and coordination transfers
    - Found to be important to team dynamics across several studies
- **At random, eight trials were selected, representing a 25% sample**
- Regression models were constructed using communication volume to predict the four selected communications categories using the whole sample and the 25% sample
  - Use regression models derived from 25% sample to predict scores for the four communication categories for each individual in the whole sample
- Compared patterns resulting from the 25% sample regression models to the whole sample





# Results: Whole Sample

| Communication Categories | Mean  | Std. Dev. | N  | Correlation | Adjusted Correlation | R Square | b     | a     | Regression |
|--------------------------|-------|-----------|----|-------------|----------------------|----------|-------|-------|------------|
|                          |       |           |    |             |                      |          |       |       | F          |
| Information requests     | 9.44  | 10.68     | 32 | 0.48*       | 0.44*                | 0.2      | 0.05* | -5.8  | 8.60*      |
| Information transfers    | 30.94 | 22.95     | 32 | 0.53*       | 0.45*                | 0.2      | 0.11* | -2    | 11.51*     |
| Coordination requests    | 17.44 | 12.31     | 32 | 0.71*       | 0.68*                | 0.46     | 0.09* | -9.41 | 29.12*     |
| Coordination transfers   | 14.75 | 11.35     | 32 | 0.69*       | 0.66*                | 0.43     | 0.08* | -9.28 | 26.24*     |

\* Significant  $p < .01$

- A summary of the regression analyses based on the whole sample
  - All the zero order correlations between the four communication categories and the total volume of communication are significant
  - **All the regression models are significant**
    - The models account for between 20% to 45% of the variance.
    - Inspections of the residuals revealed no significant outliers or influential data points



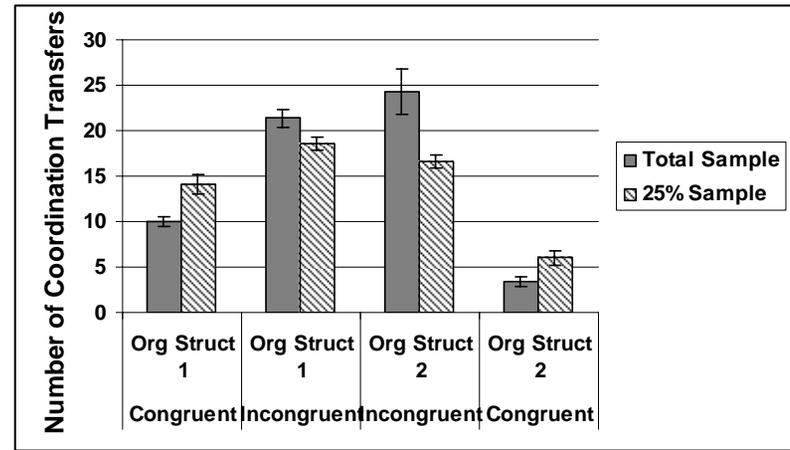
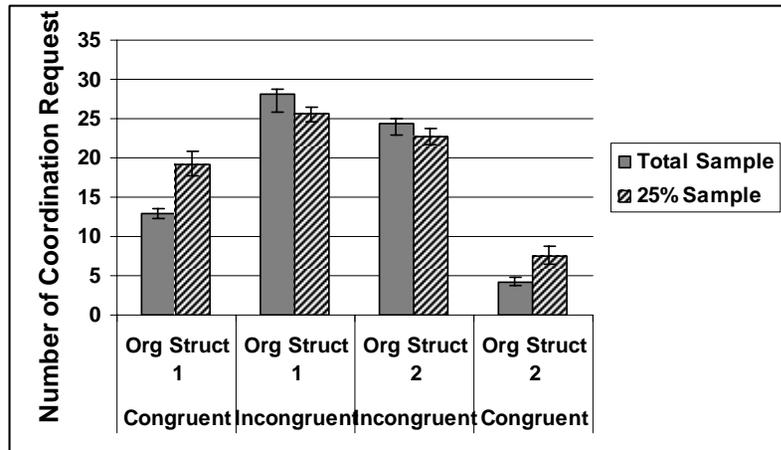
# Results: 25% Sample

| Communication Categories | Mean  | Std. Dev. | N | Correlation | Adjusted correlation | Rsquare | b     | a      | Regression |
|--------------------------|-------|-----------|---|-------------|----------------------|---------|-------|--------|------------|
|                          |       |           |   |             |                      |         |       |        | F          |
| Information requests     | 8.63  | 14.4      | 8 | 0.69        | 0.65                 | 0.42    | 0.09  | -14.5  | 4.99       |
| Information transfers    | 24.13 | 29.5      | 8 | 0.78*       | 0.70*                | 0.48    | 0.20* | -26.53 | 9.01*      |
| Coordination requests    | 14.38 | 12.76     | 8 | 0.84*       | 0.82*                | 0.68    | 0.10* | -11.55 | 13.90*     |
| Coordination transfers   | 10.88 | 9.63      | 8 | 0.79*       | 0.77*                | 0.59    | 0.07* | -7.41  | 9.10*      |

- All correlations derived from the random 25% sample were somewhat larger than that seen for whole sample,
  - With reduced sample size only three out of four of the correlations proved significant
  - **The R squares for the 25% random sample varied between .42 and .68**
    - Three out of the four regression models were significant
    - An examination of the residuals revealed only one significant outlier for information request; all else appear nominal



# Testing the Veracity of the 25% Regression Model



- To assess the utility of the communications scores derived from the 25% sample regression models, the predicted coordination request & coordination transfer scores were compared to their counterpart derived from the whole sample
  - The patterns of both the predicted and whole sample scores are quite similar
  - ANOVA performed with predicted scores yielded significant, but weaker effect size results compared to the whole sample scores



- Our findings imply that adequate predictive regression models can be constructed for communication categories using the communication volume for the whole sample
  - Moreover an adequate regression models can be constructed from a 25% randomly drawn sample
- Predicted scores from the 25% sample regression models showed similar communication patterns found for scores derived from the whole sample
- There are limitations – this procedure is no panacea
  - Not amenable to small samples
  - Communication volume may not be strongly related to communication categories in every situation
- When the procedure is applicable the savings are substantial and may lead to more research with verbal communication & a better understanding of team dynamics