White Paper Report

Experimental Evaluation of the New Automated Team Composition System (ATCS) by Formation of Optimal Teams for an Engineering Research Task

Gershon Weltman, Marvin Cohen, Raj Ratwani, Eduardo Salas[1] and Amos Freedy
Perceptronics Solutions, Inc., Sherman Oaks, California

Abstract

This paper describes an experimental study conducted in partnership with the UCLA School of Engineering and Applied Science and designed to demonstrate and evaluate in a real-world situation the use of the new Automated Team Composition System (ATCS) being developed by Perceptronics Solutions under a US Army Research Institute (ARI) Small Business Innovative Research (SBIR) project. The ATCS *is an innovative PC-based tool that supports rapid team formation in critical incident preparedness and response as well as in other applications. The ATCS uses advanced multiagent system software technology to help command and/or decision making groups solve the daunting organizational problem of optimally assigning personnel to mission-oriented teams. The team performance criteria used by the Multiagent System derive from the analysis of specific mission features as well as from the psychological literature. The present study used archival data from a UCLA (University of California Los Angeles) engineering course in which nominally student teams prepare a research report on a societal problem involving both technical and ethical issues. Data on approximately 700 students and 130 teams were first examined to determine the relationships among student characteristics (including major, gender, ethnicity and essay and test scores) and team performance (as defined by scores on several components of the research report). Analysis of the data revealed a number of significant, marginally significant and trend relationships between team member characteristics and team performance. These results, in combination with other information about course goals, were used to program the ATCS to form optimal 5-person teams with specific role assignments. The ATCS was able within 1 minute to recommend teams from the candidate population that met all the programmed criteria and had strong face validity as well -- a task well beyond human capabilities. The study showed that the ACTS can be readily used in a real-world situation in which rapid team formation is required and where knowledge about the team mission can be combined with knowledge about the factors influencing team performance. Previous ATCS demonstrations have used on-the-spot questionnaires to gather attribute and qualification data about team candidates in anti-terrorist exercises. In the present evaluation we used candidate data from a stored data base, which we see as a major advantage in future applications of the ATCS tools, and also used a non-military, non-security situation to demonstrate the broad applicability of the technology.*

1. Introduction

Military and non-military organizations alike need optimally-constituted teams to carry out immediate missions and long-term objectives. Of particular concern are missions and operations requiring the participation of diverse services, of different agencies and organizations, and possibly also of partners dispersed in different locations [1] [2].

Bureaucratic hierarchies generally do not solve the problem of connecting people with complementary skills and/or mutual interests, and consequently these structures often force organizations to perform inefficiently and ineffectively on problems that require rapid and team composition – such as response to civil emergencies or critical corporate tasks [3]. Often, organizations *have* personnel with the right skill sets to meet a particular challenge or to create innovative solutions to difficult problems, but are unable to carry out the task of efficiently and effectively putting together the optimal mix of personnel. Previous research has shown that properly designed computer support tools can improve human ability to deal with such complex organizational functions. The goal of the Automated Team Composition project is to provide such a team formation tool.

---

[1] Department of Psychology, University of Central Florida

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **24 DEC 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Experimental Evaluation of the New Automated Team Composition System (ATCS) by Formation of Optimal Teams for an Engineering Research Task**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Perceptronics Solutions, Inc,3527 Beverly Glen Ter,Sherman Oaks,CA,91423**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT

This paper describes an experimental study conducted in partnership with the UCLA School of Engineering and Applied Science and designed to demonstrate and evaluate in a real-world situation the use of the new Automated Team Composition System (ATCS) being developed by Perceptronics Solutions under a US Army Research Institute (ARI) Small Business Innovative Research (SBIR) project. The ATCS is an innovative PC-based tool that supports rapid team formation in critical incident preparedness and response as well as in other applications. The ATCS uses advanced multiagent system software technology to help command and/or decision making groups solve the daunting organizational problem of optimally assigning personnel to mission-oriented teams. The team performance criteria used by the Multiagent System derive from the analysis of specific mission features as well as from the psychological literature. The present study used archival data from a UCLA (University of California Los Angeles) engineering course in which nominally student teams prepare a research report on a societal problem involving both technical and ethical issues. Data on approximately 700 students and 130 teams were first examined to determine the relationships among student characteristics (including major, gender, ethnicity and essay and test scores) and team performance (as defined by scores on several components of the research report). Analysis of the data revealed a number of significant, marginally significant and trend relationships between team member characteristics and team performance. These results, in combination with other information about course goals, were used to program the ATCS to form optimal 5-person teams with specific role assignments. The ATCS was able within 1 minute to recommend teams from the candidate population that met all the programmed criteria and had strong face validity as well -- a task well beyond human capabilities. The study showed that the ACTS can be readily used in a real-world situation in which rapid team formation is required and where knowledge about the team mission can be combined with knowledge about the factors influencing team performance. Previous ATCS demonstrations have used on-the-spot questionnaires to gather attribute and qualification data about team candidates in anti-terrorist exercises. In the present evaluation we used candidate data from a stored data base, which we see as a major advantage in future applications of the ATCS tools, and also used a non-military, non-security situation to demonstrate the broad applicability of the technology.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**16**	

The ATCS team formation process makes certain fundamental assumptions about the ingredients of an optimal team based on the most recent organizational research findings [4] [5] [6] [7]; these findings suggest that an optimal team may include:

- ***Qualified Personnel***. Appropriate position and technical experience and/or knowledge.
- ***Good Teamwork Attitudes and Skills.*** Motivation and know-how that can support effective cooperation and coordination.
- ***Social Connections*****.** Team members are centrally connected within organizational networks and teams have high within-team connectivity.
- ***Appropriate Weighting***. Relative importance weights assigned to the various team formation factors – for each team role and for the team as a whole.
- ***Team Formation Constraints*****.** Real-world requirements and preferences that influence both role assignments and overall team formation.

Optimizing over all of these conditions simultaneously and rapidly for a reasonable number of team roles to be filled from a sizable pool of candidates can involve choosing from among *millions* of possible team configurations. Optimizing this task rapidly is beyond human capabilities, and accordingly the job can benefit significantly from sophisticated computational processing. The ATCS computational framework described below provides the needed rapid processing capability by applying advanced algorithms in a form usable by non-experts and supportable on standard PCs.

2. ATCS Description

Figure 1 shows the Automated Team Composition System developed under the present project and a related DARPA STTR project [8, 9]. The main components are described briefly in the following.

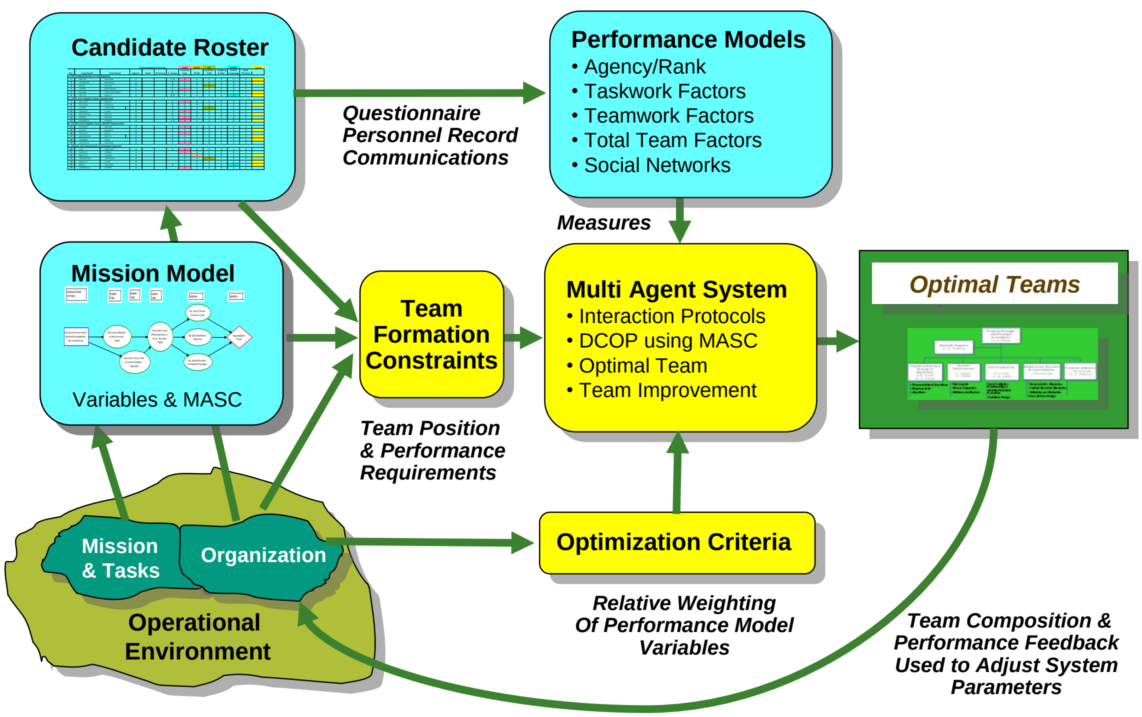

Figure 1 Automated Team Composition System (ATCS)

- **Operational Environment** is composed of the *Organization***,** which can be military or non-military, and the *Mission***,** which is the identified near-terms or long-term objectives and/or goals of the Organization

- **Mission Model** is a representation of the mission in which the key factors are identified and expressed in terms of their relative importance to the mission as a whole
- **Candidate Roster** is the list of personnel from which the team will be formed
- **Performance Model** is a representation of the key factors characterizing the candidates; these may include:
  - *Qualifications,* such as organization, rank, experience and technical knowledge, based on self assessment or actual performance data;
  - *Teamwork Measures*, which are teamwork attitudes and skills also obtained through questionnaire or available data;
  - *Network Measures,* these are social networking measures – such as individual *centrality* within a social network and team *connectivity* of its individual members – which again can be obtained from questionnaire concerning person-to-person connections or directly from communications data such as the number of messages exchanged among candidates.
- **Multiagent System (MAS)** is a computational system consisting of multiple software agents where each agent represents a candidate along with his or her attributes and capabilities, and the agents "negotiate" among themselves by means of an *agent interaction protocol* to solve the *distributed constraint optimization problem (DCOP)* and select the members of the *locally optimal team or teams* [10]. Critical to this process are:
  - **Optimization Criteria,** which are the rules governing how the various attributes, qualification and social networking measures are combined to calculate the overall criterion by which the value of potential teams will be judged as well as the size of the "neighborhood" over which the search for optimal teams will be conducted.
  - **Team Formation Constraints,** which are the rules pertaining to specific team roles or to the overall team makeup, and are respectively defined as "single-agent" or "multi-agent" constraints
- **Optimal Teams** are the teams selected by the system from among the many possible combinations, taking into account the constraints and optimization criteria. The selected teams are presented to the users as recommendations, and may be modified directly or by changing the original constraints and optimization criteria. The optimization process can be adaptively improved over time by observing the performance of the selected teams and using that feedback to adjust the parameters

The ATCS provides a human-factored graphical user interface that makes its complex support tools easily accessible for non-specialized users. For example, familiar drag and drop functions allow users to select rapidly from stored templates of team roles and performance factors, and simple screens allow users to apply factor weights to the roles and entire team, as well as select various degrees of optimization by computation time or search parameters.

A key requirement for a system designed to support critical incident response is compatibility with the National Incident Management System (NIMS). NIMS is "a consistent approach for federal, state, tribal, and local governments to work effectively and efficiently together to prepare for and respond to all hazards including acts of terrorism" (National Response Plan). ATCS is compatible with NIMS and supports its major functions.

Overall, ATCS is oriented to the three main stages in team composition, which are applicable to general organizational requirements as well as to critical incident preparedness and response:

- **Preplanning Stage**, where organizational data are gathered, the several models are created and virtual teams are configured in anticipation of an event; this stage may also highlight organizational deficiencies;
- **Response Stage**, where actual teams are rapidly formed from available personnel just prior to a planned for event or after an unexpected critical incident has occurred; and

- **Post-Event Stage**, where evaluation of actual team performance is used as an aid to organizational improvement.

The ATCS was previously successfully demonstrated during anti-terrorist exercises conducted by the Center for Asymmetric Warfare, a part of the Naval Postgraduate School. In those exercises we used on-the-spot questionnaires to gather attribute and qualification data about the team candidates; and in those demonstrations the main objective was to show that the ATCS could form teams in the required time frame. In the present evaluation we used candidate data from a stored data base, which we see as a major advantage in future applications of the ATCS, and also used a non-military, non-security situation to demonstrate the broad applicability of the technology.

2. Study Design

Target Course and Team Task. The UCLA School of Engineering and Applied Science (SEAS) Engr 183 course fulfils both the School's ethics and writing requirements, and according is taken by about 150 students each of the four quarters it is offered. The course requires the students to form into multidisciplinary teams and perform a typical engineering research project culminating in a Team Research Report. Each quarter the students, with some assistance from the instructors and TAs, form themselves into about 25 to 27 three- to five- person teams. Each team produces its Team Research Report on a topic selected by the team from a suggested list of five or six topics covering current societal problems that have both clear technical and ethical components. The Team Research Reports are graded from 0 to 100 on seven internal criteria (executive summary, background, technical discussion, ethical factors, recommended actions, organization and references) that are also combined to form the overall 0-100 score. Student data are available from the class roster (name, major, gender) and from recorded performance scores on the two individual writing assignments, the midterm and final exams, and the Report components.

Study Scope and Objectives. In the present study we used only 183EW data that is historical (i.e., past quarters only) and anonymous (i.e., no student names connected) to examine the existing correlations between team member characteristics and the performance of the total team in performing this typical engineering task.

Our analysis focused on the period after Engr 183 became a requirement for the entire School of Engineering and Applied Science (this increased the diversity of student majors) and before it also became an Engineering Writing course (the additional TA attention both raised the level and narrowed the range of Team Report scores). Consequently, the team candidate list comprised student from five prior quarters (from Spring '07 through Fall '08); As shown in the table below, it involved approximately 700 students and exactly 130 teams.

Quarter	Students	Teams
1. Fall '08	172	34
2. Spring '08	145	27
3. Winter '08	129	29
4. Fall '07	107	22
5. Spring '07	138	24
Total	691	130

A primary objective was to allow further testing of the ATCS using optimization criteria based on knowledge derived from an existing "personnel" data base. A second objective was to use the empirical correlations for further refinement of the Team Performance Framework, helping shed light, for example, on the relative value of team diversity vs. team homogeneity (based on such measures as individual performance, major, general ethnicity, etc.) for team performance (based on the overall performance score or constituent scores). A third and different objective was to potentially benefit the Engr 183 course itself by helping the instructors to guide students into better performing teams and to provide extra assistance in areas where already-formed teams are likely to need it. The results may also benefit the UCLA SEAS as a whole by providing more information about the factors that contribute to the success of engineering teams, and UCLA or other academic organizations may also in the future be able to make use of the Perceptronics Solutions automated team composition system that results in part from the research results.

Data Anonymity Scheme. The historical data were formatted by the Engr 183 course instructors and delivered under a subcontract to UCLA SEAS from Perceptronics Solutions. Figure 2 shows the preliminary data format and the anonymity scheme. The student names in the Engr 183 Course data were converted to numbers by the course instructors as the first step of the study, and the analysis of individual characteristics to team performance was based entirely on these numbers and did not involve student names or other personal identification. There was be no way for an outsider (including Perceptronics Solutions staff) to trace these data back to a particular student because the quarter in which it was obtained, the specific team topic involved, the student UCLA identification number, and any other identifying material was either redacted, randomized or anonymized. The existing class data containing student names and performance records remained in the current data files accessible only to the course instructors and could not be further released -- in accord with all privacy and security concerns and/or regulations. This scheme was approved by the UCLA Registrars Office and also by the UCLA Internal Review Board (IRB). *To emphasize: The data used was only archival and did not involve or affect any students previously or currently enrolled in the class, and the data was anonymous, student names having been removed in the initial processing step.*

Quarter Redacted | Team # Randomized | Student ID Anonymized | Student # Redacted | Student Final Grade Redacted

Student Data				Individual Performance							Team Performance						
Team	ID	Major	Gender	Particip	Team Present	Essay Case St	Essay Hardin	Midterm Total	Final Part I	Final Part II	Exec Sum	Prob Back	Tech Econ	Ethical Societal	Actions Recom	Org & Present	Report Total
1	11	ELE	M	100	100	95.5	92.5	47	25.5	92	86	85	94	89	90	92	91.2
1	18	MEC	M	100	100	95	89	38	23.1	88	86	85	94	89	90	92	91.2
1	43	MAT	F	100	100	93	90.5	44	22.8	89	86	85	94	89	90	92	91.2
1	100	MEC	M	100	100	93	91	41	21	92	86	85	94	89	90	92	91.2
2	9	ELE	M	100	100	86	83	28	15.6	83	100	98	98	100	94	95	97.2
2	14	ELE	M	100	100	93	89	43	26.4	93	100	98	98	100	94	95	97.7
2	37	MEC	M	100	100	68	70	41	24	83	100	98	98	100	94	95	97.7
2	81	ELE	F	100	100	80	82	41	25.8	89	100	98	98	100	94	95	97.7
2	85	COM	M	90	100	81	80	37	24	92	100	98	98	100	94	95	97.7
2	93	CHM	M	100	100	89	83	46	30	87	100	98	98	100	94	95	97.7
2	98	ELE	F	100	100	99	100	48	29.7	100	100	98	98	100	94	95	97.7
3	7	COM	M	100	100	95	100	45	27.9	94	98	96	100	100	96	98	97.9
3	33	ELE	M	90	85	91	90	43	26.1	84	98	96	100	100	96	98	97.9
3	59	ELE	M	90	85	84	85	34	22.2	87	98	96	100	100	96	98	97.9
3	71	MEC	M	80	85	89	88	29	22.2	66	98	96	100	100	96	98	97.9
4	32	CIV	M	100	95.83	83	89	48	25.5	93	92	89	97	96	88	85	92.2
4	36	AER	M	100	87.5	100	85	46	26.4	90	92	89	97	96	88	85	92.2
4	48	MEC	M	100	87.5	90	90	48	30	91	92	89	97	96	88	85	92.2
4	82	MEC	F	80	95.83	100	78	39	30	96	92	89	97	96	88	85	92.2
4	118	CIV	M	90	83.33	100	95	41	25.8	98	92	89	97	96	88	85	92.2
5	19	COM	M	90	98	94.8	94	45	30	93	100	100	100	100	100	100	100
5	69	CHM	F	100	98	99.5	99	46	25.2	98	100	100	100	100	100	100	100
5	79	COM	M	90	98	89.2	89	50	30	90	100	100	100	100	100	100	100
5	106	COM	M	90	98	86.3	89.6	43	30	100	100	100	100	100	100	100	100
5	107	COM	M	90	98	79	70	43	27	91	100	100	100	100	100	100	100
6	29	COM	M	95	100	92	93	47	30	91	90	95	97	92	95	97	95.3
6	78	COM	M	100	100	93	92	47	22.5	97	90	95	97	92	95	97	95.3
6	80	MEC	M	90	100	82	79	38	21.9	83	90	95	97	92	95	97	95.3

No way to identify individual student From Study ID or grade components

Figure 2 UCLA Study Sample Data and Anonymity Scheme

Data Analysis. We exercised the data base both with some simple "warm-up" analyses and a number of more complex analyses of the correlations among individual student characteristics, team diversity measures and team performance scores.

Our independent variables were based on data covering individual student performance (two essays, a midterm exam, a 2-part final exam), student participation (class attendance and observed team contribution) as well as student characteristics (engineering major, student gender as determined by first name and roster picture, and an informal designation of student ethnicity as determined primarily by his or her name). Coding for ethnicity was as follows:

1. ASI; Asian, includes Chinese, Korean, Japanese, etc. names
2. AME; American/European, includes all English, Irish, Russian, etc. names
3. HIS; Hispanic, includes Mexican and Spanish names
4. INP; Indian and Pakistani names
5. MDE; Middle Eastern, includes Arabic, Israeli and Armenian names
6. AFR; African, includes only African names, African-Americans are included in AME

Coding for academic major was as follows.

Abbreviation	Aggregated	Program Name
CHM		Chemical Engineering
CIV		Civil Engineering
COM		Computer Science
ELE		Electrical Engineering
MEC		Mechanical Engineering
Other	AER	Aerospace Engineering
	BIO	Bioengineering
	CSE	Computer Science and Engineering
	MAT	Materials Science
	DBLE	Double Major

In the case of ethnicity and academic major, low frequency categories were sometimes collapsed to avoid empty cells in factorial analyses. With respect to ethnicity, Hispanics, Indians and Pakistanis, Middle Easterners, and Africans were sometimes combined into the single “Other” category.

Dependent variables were the Team Research Report Scores, including the seven component scores and the overall Report score.

The following presents those results from our analyses that pertain most directly to the identification of relevant measures that might be useful as inputs to the Automated Team Composition System. We exercise some caution in the interpreting these results for a variety of reasons; these include: the relative homogeneity of the student population, ceiling effects on grades, the non-random selection of team members, and the large number of possible effects examined, leaving wide scope for chance. The analyses are accordingly exploratory: We report significant effects ($p \leq 0.05$), marginally significant effects ($.05 < p \leq 0.10$) and non-significant trends ($.10 < p \leq 0.15+$).

3. Analytical Results

Data Base Characteristics. The data base of approximately 700 students contained 80% males and 20% females. There were significantly more Asians (52%) than American/Europeans (28%) or Other (20%). Using the latter three categories in a Chi-squared test, this pattern held without significant variation across quarters. Students were not spread evenly across academic majors. Overall, 23% of students were in Electrical Engineering; 17% in Mechanical Engineering; 13% in Computer Science; 13% in Civil Engineering; and 11% in Chemical Engineering. All other fields, totaling 23%, had less than 10% of the students.

Figure 3 (left) shows the number of teams with specified number of male team members. There were no teams with zero male members**.** Figure 3 (right) shows the number of teams with specified number of female team members. There were a large number of teams with no females; the mean was one female.

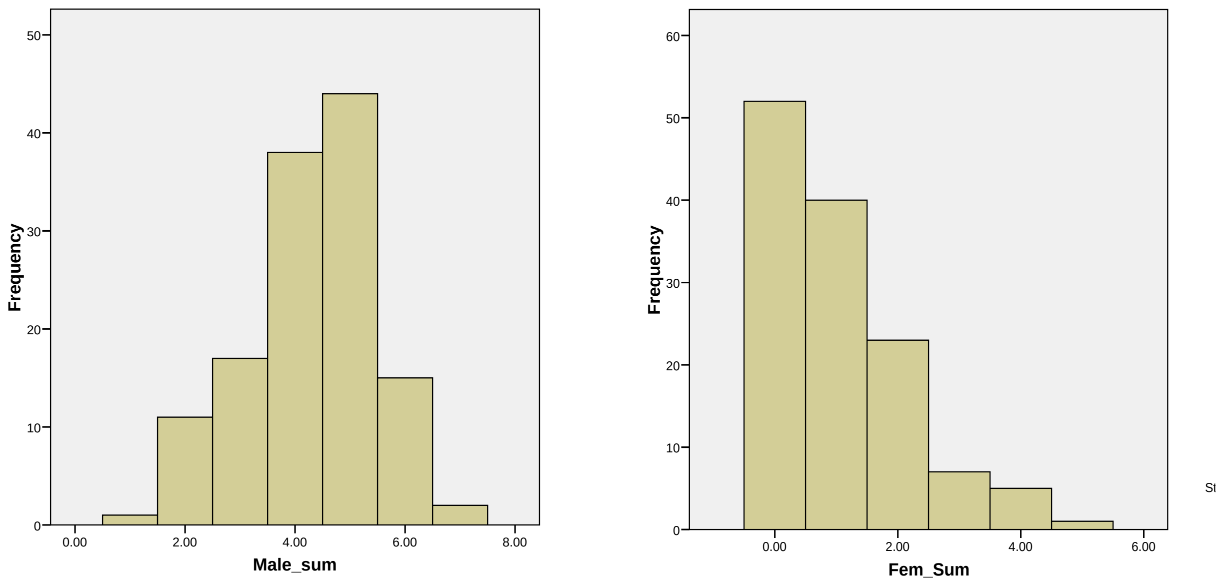

Figure 3 Distribution of male team members (left) and female team members (right).

Team Diversity. A direct way to study the effect of diversity is by a diversity measure. We did this by means of Shannon entropy. It is a widely used measure of dispersion among nominal values, which plays the same role as variance does for dispersion of numerical values. The formula for Shannon's index of diversity is:

$$D = - \sum p_i \ln(p_i)$$

where p_i is the proportion of the population represented by subgroup i. Diversity is at a maximum when each of n subgroups represents 1/n of the population; the maximum in that case is ln(n). Diversity is at a minimum, equal to 0, when the entire population is concentrated in one subgroup.

There was a good amount of variation across teams on ethnicity and academic major as exemplified by Figure 4 showing the distribution of diversity measure for ethnicity and majors..

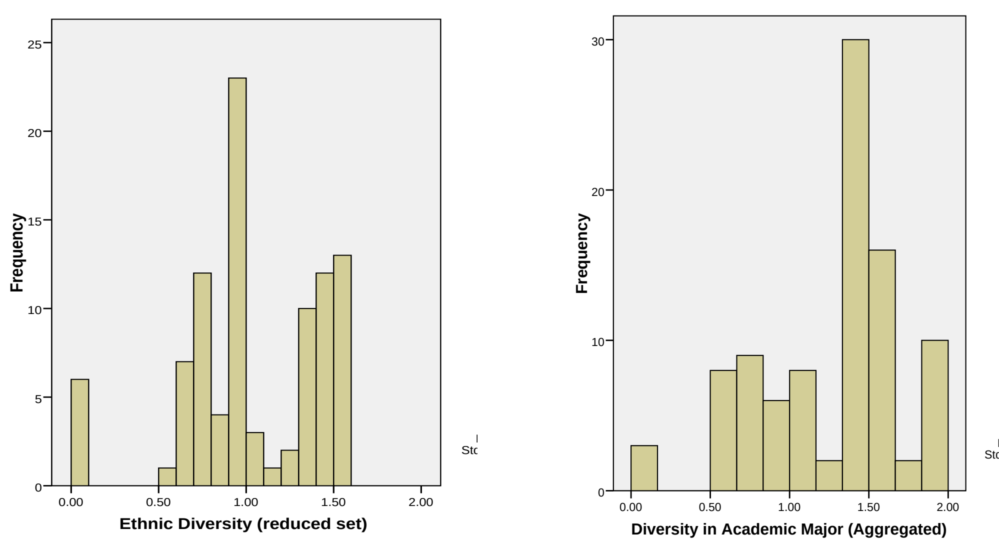

Figure 4 Distribution of Teams by Diversity Measure for Ethnicity (left) and Academic Major (right)

Individual Grade Scores. We found that Individual grades of different types are correlated with one another across students, and moreover, these correlations yield the two-dimensional Euclidean distance solution shown in Figure 5.

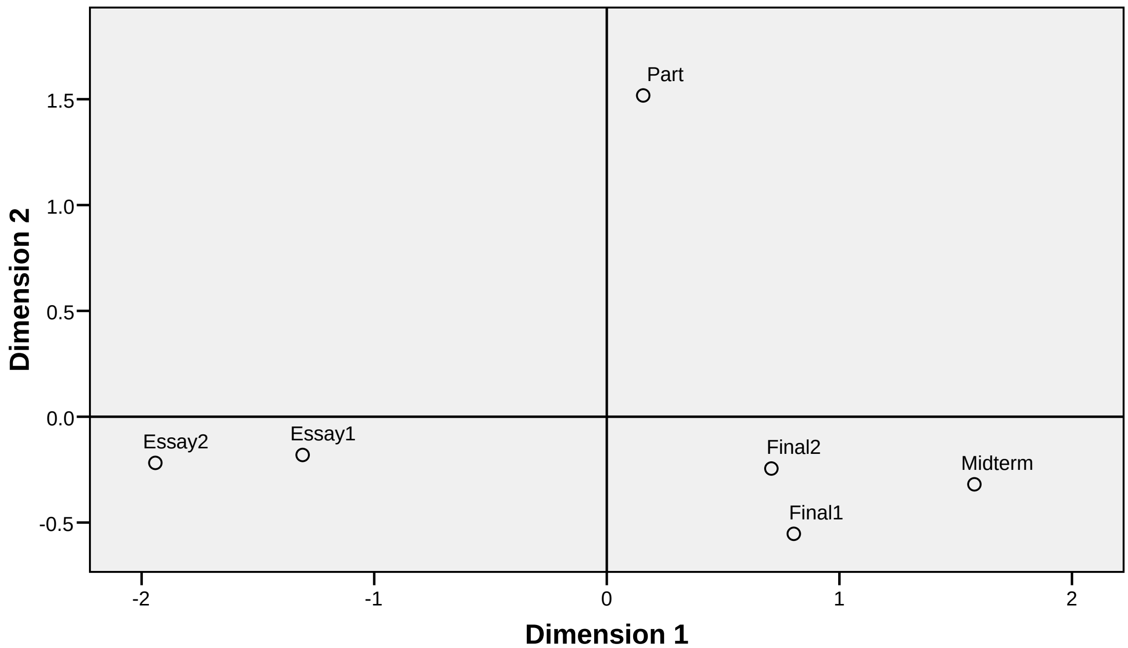

Figure 5 Multi-dimensional scaling in which greater distances represent less correlation among different kinds of individual student scores.

The two-dimensional solution provides a very good fit to the data, as evidenced by low stress and high dispersion accounted for. Moreover, the different grades appear to cluster in a very natural way: With respect to dimension 2, participation is at one extreme and products (essays and exams) at the other. With respect to dimension 1, essay products are on the left and test products on the right. This intuition is confirmed by hierarchical cluster analysis, as shown below. At the level of three clusters, we have distinct clusters for: (1) participation, (2) the three tests, and (3) the two essays.

Influence of Individual and Team Variables on Team Performance. Table 1 on the following page is a summary of the significant, marginally significant and trend effects determined from our various data analyses. The effects themselves are a combination of the contribution to team performance made by single individuals with a particular attribute (e.g., Computer major) and the contribution made by having increasing numbers of individuals with that attribute (e.g. several members with Computer majors). The goal is to give an idea of how the candidates' attributes affect the performance of the teams of which they are a member. Our review of this summary suggests a number of relevant findings.

- **Gender.** There is a clear advantage to having female team members, who contribute positively almost across the board
- **Major.** Some majors have a balance of positive and negative effects, but it appears that Material Science and Mechanical Engineering majors are detrimental to team success, while Aerospace and Bioengineering majors are beneficial over a wide range of Report components.
- **Ethnicity.** There is a tendency for Asian students to add to team performance in the Technical area and the Executive Summary but not in the Ethics or Organization area, while American/European students add in Ethics as well as Technical but not in the Executive Summary. This suggests a cultural difference in the approach to or the importance given to the Ethical considerations. 'Other' ethnicities contribute positively in several areas, but not to the Executive Summary; and there was also a trend for Hispanics to positively influence Background scores.

- **Individual Scores.** Good performance on tests, whether separate or averaged, has its main effect on the Technical section, which seems understandable, but the average essay and test scores have a negative effect on recommended actions, which suggests that students who are good at working with the text and lecture material may not also be good at solving problems not previously encountered. Participation also has a positive effect on the Technical section..

Table 1 Summary of the Data Analyses

Team Member Variables		Team Research Report Variables							
		Exec	Back	Technical	Ethics	Actions	Org	Ref	Report
Gender	Male								[Neg]
	Female	(Pos)	(Pos)				(Pos)	(Pos)	(Pos)
Major	CHM				(Neg)				
	CIV								
	COM	<Neg>		[Pos]		<Neg>		(Neg)	
	ELE							(Pos)	(Pos)
	MAT			[Neg]	<Neg>		(Neg)		
	MEC		[Neg]	[Neg]			[Neg]	[Neg]	
	AER					(Pos)			[Pos]
	BIO		[Pos]	[Pos]	[Pos]				[Pos]
	CSE								
Ethnicity	ASI	(Pos)					[Neg]		
	AME	[Neg]		[Pos]	<Pos>				
	Other	[Neg]	[Pos]				[Pos]		
Individual Scores	Essay 1								
	Essay 2								
	Midterm			<Pos>					
	Final 1			[Pos]				[Neg]	
	Final 2								
	Ave Grade[1]			<Pos>					
	Participation			(Pos)					
Team Scores	Mean Ave Grade			(Pos)		(Neg)			[Neg]
	Mean Participation			[Pos]					
Team Size	--			(Pos)	[Pos]				
Diversity	Ethic								
	Major								

<significant>, p ≤ 0.05; (marginal), 0.05 > p ≤ 0.10; [trend], p > 0.10

- **Team Scores.** The mean of team members' average essay and test scores reflects the individual effects – positive on the Technical section and negative on Recommended Actions. Likewise the mean Participation score has a positive effect on the Technical section
- **Team Size**. Increasing team size had positive effects on the Ethics section as well as on the Technical section.
- **Diversity.** Neither Ethnic nor Major diversity had any significant, marginal or trend effect on any of the team performance measures; i.e., teams with high diversity performed as well as teams with lower diversity.

Discussion. The extensive data analysis summarized in the above table has led us to some (very) preliminary conclusions regarding the three main questions addressed in this study, namely:

- What do we learn about team performance in general?
- How would we use the results to form optimal teams using the ACTS tool?
- How might we use the results to improve instruction in this specific course?

We provide some initial answers in the following.

- **General Team Performance.** Our results from this study suggest that:
  - Team members in an academic setting will self-select (perhaps with some degree of help) for significant levels of diversity both in ethic and technological dimensions.
  - Diversity itself may not improve or degrade performance, but it may have other benefits in the overall task context and with respect to other, broader objectives. In the current analysis, diversity per se is seen to be less important than contributions of specific groups to specific team competencies.
  - Females can be a general asset to team performance in a technical context, and may exert one of the strongest effects.
  - Members from different ethnic backgrounds can contribute differently to specific task elements. Their contributions may reflect their cultural backgrounds.
  - Individual task skills, reflected in this case by academic test scores, may affect only limited team task areas – in this case only the Technical and Recommended Actions portions of the Report.
  - Teams composed of members with high individual performances on constrained tasks such as test taking may do worse on team tasks requiring more creativity or decision making – in this case the Recommended Actions portion of the report.
- **ATCS Team Formation.** Establishing optimization criteria and constraints for the ACTS tool might make use of present results, including the following areas.
  - Possible team role weightings. A Technical Lead might be weighted higher on average test scores; other team roles would likely have different weightings for Essay scores, Exam scores and Participation score.
  - Possible role constraints. Constraints on individual roles might take into account the positive influences of Females, Bioengineer and Aerospace Engineering majors, as well as the negative effects of other major
  - Possible team constraints. These might consider both the results of the correlation analysis and the other objectives of the course, which include having students work on teams of significant diversity.
- **Use of Results for Improved Instruction.** The instructors of Engr 183 as well as of other Engineering course might make use of the results in the following ways:
  - Recognize the student distributions in terms of Ethnicity, Gender and Majors.
  - Identify females as a particularly valuable team asset, which perhaps should be rationed.
  - Enforce the modal team size of 5 as optimal – not too small, not too large – even though some performance improves with increasing team size.
  - Observe that the objective of promoting specific types of team diversity seems to be working.
  - Recognize that individual performance has three clear and somewhat separate dimensions – Essays, Examinations and Participation.
  - Recognize that Asian students in general may be having problem with the Ethics area possibly due to a significant number of recent arrivals to the US, English language problems, or cultural differences, and might require special attention to ensure that the course objectives in this are understood and integrated into the Team Research Report

- And finally, recognize that some Majors seem to be having problems with the more societal aspects of the course – and this might warrant a closer look at what characteristics might be causing this and what could be done to counteract it.

In the following section, we show how we used the analytical results to program the ATCS tool for the rapid formation of optimal teams capable of producing a strong Team Research Report.

4. Team Formation Using the ATCS Tool

The hypothetical real-world problem was to rapidly form *the best team* to produce a Team Research Report having at hand; (1) the general objectives of the course; (2) the 700-student roster and accompanying data; (3) the results of the present correlation analysis; as well as (4) the general team performance research literature and in particular our preliminary formulation of a Team Performance Framework.

Team and Roles. To address this problem we used the same set-up of the ATCS as we had for the Kaimalu O Hawaii (KOH) demonstration conducted in May 2009 with the assistance of the Center for Asymmetric Warfare (CAW) as part of a Port of Hawaii anti-terrorist exercise [9]. In the KOH case we formed a virtual six-person command team to oversee the anti-terrorist operations; in the present case we formed a virtual five-person team to conduct research and produce a Team Research Report as previously analyzed. For the research team we defined the distinct roles shown in Table 2, assuming a specified Lead person for each of several key Research Report components.

Table 2 Team Roles for the Engr 183 Research Report Mission

1	Organization & References Lead (ORL)
2	Executive Summary Lead (ESL)
3	Background & Technology Lead (BTL)
4	Ethics and Societal Lead (BTL)
5	Recommended Actions Lead (RAL)

Team Candidate Characteristics and ATCS Model Factors. In the KOH case the candidates' Individual Characteristics comprised their agency (local police, FBI, DHS, etc.), their position or rank; their Mission Factors came from self-assessment of experience/skills in a number of mission-related technical areas (HAZMAT, civil aviation, rescue operations, etc.); their Teamwork Factors were assessed by a 20-question survey; and their social Network Factors was assessed by self-reported recent communications.

In the present case the candidates' individual characteristics comprised their engineering major, their gender and their ethnicity; their mission-related experience/skill (taskwork) factors came from their test and essay scores; and their teamwork factors came from their participation score. No Network Factors were applied.

For the present case we defined two types of Mission Factors, as follows; these were based on the analytical findings regarding the clustering of performance scores:

1. Research/Exposition Skills = Ave[Essay1 + Essay2]
2. Course Purpose/Content Knowledge = Ave[Midterm + Final1 + Final2]

We defined a weighted multi-attribute Mission Score ($MASC_M$) as follows, and equal weighting was used for the initial trials.

$$MASC_M = w_R[\text{Research/Exposition}] + w_P[\text{Purpose/Content}]$$

We defined a weighted multi-attribute Teamwork Score ($MASC_T$) was in this case simply as:

$$MASC_M = [\text{Participation}]$$

Role and Total Score Weighting. Table 3 table shows how the Mission and Teamwork Factors were weighted for the individual team roles in our trial runs. The weights were based in part on the analytical finding and in part on our assumptions regarding the skills required for each role.

Table 3 Factor Weights by Role for the Engr 183 Team

	Organization & Refs Lead	Executive Sum Lead	Background & Tech Lead	Ethics Lead	Actions Lead
Indiv Research/Exposition	0.15	0.50	0.30	0.20	0.40
Course Purpose/Content	0.35	0.35	0.40	0.60	0.30
Teamwork Skills	0.50	0.15	0.30	0.20	0.30

In the ATCS we define a Total Score used for team optimization by the multiagent system; in the present case this could be expressed as:

$$\text{Total Score} = w_M MASC_M + w_T MASC_T$$

In the present case, we assumed that the Mission Factors were more important than the Teamwork Factors, and so assigned weights of 0.70 and 0.30 to these components, respectively

Role and Team Constraints. We applied the role constraints shown in Table 4 in accord with the associated objectives, based primarily on the relationships between team member characteristics and team performance resulting from the prior data analysis.

Table 4 Role Constraints Used in the Engr 183 Team Formation Trials

Role Constraint	Objective
BTL must be BIO Major	Strengthen Background and Technology Section
BTL must have Purpose/Content > 90.0	Strengthen Background and Technology Section
RAL must be AERO Major	Strengthen Recommended Actions Section
ESL must be Female	Strengthen Ethics and Societal Section

For the sake of these trials, we assumed it would be acceptable to specify a female for a particular role, as well as to specify the total number of females on the team, since females were clearly shown to a valuable asset in the data analysis, and planned representation in this arena is generally approved (even on the Supreme Court). We made the opposite assumption in the case of ethnicity, i.e., we imposed no role or team constraints based on ethnicity, reasoning that it would not be acceptable in this type of situation to specify a particular ethnicity for a role or to specify the number of members of particular ethic groups for the team, even in the interests of ethnic diversity and with regard to our analytical results. (This assumption might be questioned in this and other situations).

With respect to diversity of engineering major, however, we recognized that a major objective of Engr 183, as described in the Course Syllabus, is to give students the experience of working on a technically diverse team. Accordingly, we included in the Team Constraints shown in Table 5 the requirement that there be at most one member from a particular major.

Table 5 Team Constraints Used in the Engr 183 Team Formation Trials

Team Constraint	Objective
At least one Female	Strengthen Total Team Performance
Not more than two Females	Maintain Gender Diversity
Not more than one member per Major	Maintain Major Diversity

Team Composition. We limited the ATCS compute time to 1 minute on the assumption that this would be a reasonable and acceptable run duration, and because we were interested in whether the ATCS could find optimal teams in this relatively short time.

The ATCS was successful in forming an optimal team within one minute from the approximately 700 member candidate pool and the initial role and team constraints. Table 6 shows the team makeup as well as the Optimal Score, which is a relative measure of optimality. The four Role Constraints met are shown in **bold type;** and the team also meets the Team Constraints of not more than two Female Members and complete diversity of engineering majors.

Table 6 Optimal Team Composed from Full Candidate Data Base

Team 1 (Optimal Score=15993)		ID	Major	Gender	Ethnicity	Tests	Essays
1	Org & Ref Lead	581	CHM	Male	ASI	98.5	100
2	Executive Summary Lead	580	ELE	**Female**	ASI	96.0	100
3	Background & Tech Lead	121	**BIO**	Male	AME	**90.8**	94.5
4	Ethics Lead	645	COM	Male	INP	98.9	99.5
5	Recommended Actions Lead	480	**AER**	Female	AME	93.9	100

We also examined the new case of forming two teams simultaneously, where the teams cannot share members, and each team should be as close to optimum as possible. This is a common situation where multiple teams must be formed to perform different parts of an overall mission, and the teams cannot be considered *primary and backup* but must each be as nearly equivalent in capability as possible. The standard unaided method of picking one team and then picking the next team from the remaining candidates does not satisfy the formal conditions of optimality. In this case, we were able to exercise the ATCS in a dual-team optimal mode with the following results. As seen in Table 7, the two optimal teams meet all the role and team constraints, and have a common Optimal Score just slightly lower than the single optimal team above.

Table 7 Two Optimal Teams Composed from Full Candidate Data Base

Team 1 (Optimal Score=15860)		ID	Major	Gender	Ethnicity	Tests	Essays
1	Org & Ref Lead	645	COM	Male	IDP	98.6	99.5
2	Executive Summary Lead	144	CIV	**Female**	AME	97.2	93.5
3	Background & Tech Lead	642	**BIO**	Female	ASI	**97.8**	98.5
4	Ethics Lead	149	ELE	Male	ASI	97.8	98.0
5	Recommended Actions Lead	633	**AER**	Male	AME	83.5	96.0

Team 2 (Optimal Score=15860)		ID	Major	Gender	Ethnicity	Tests	Essays
1	Org & Ref Lead	250	MEC	Male	ASI	99.0	96.5
2	Executive Summary Lead	370	CHM	**Female**	AME	94.2	97.0
3	Background & Tech Lead	531	**BIO**	Male	ASI	**98.6**	90.5
4	Ethics Lead	580	ELE	Female	ASI	96.0	100
5	Recommended Actions Lead	38	**AER**	Male	ASI	93.5	97.0

Interestingly, while we did not impose any role or team constraints based on ethnicity, overall the percentage of Asians (60%), American-Europeans (30%) and Others (10%) in the two selected teams is quite well reflective of the candidate population in which Asians are 52%, American-European are 28% and Others are 20%.

Finally, we examined the case of using only half the candidate data base for team composition, because this increased the number of alternative team configurations that the ATCS could examine during the same computational time. For this trial the approximately 700 member candidate pool was randomly divided into A and B halves. Table 8 shows the results in terms of team member IDs in comparison to two teams formed with the full data base. As seen, the half-DB results share some role assignments with the full-DB runs, but not many. In addition, the optimality scores for the half-DB teams are the same or even slightly higher than those for the full-DB case, which is consistent with the ATCS being able to process more team options when it is working with a smaller candidate pool. It appears that when the candidate pool is this large, dividing it in half randomly does not materially affect team composition.

Table 8 Comparison of Teams from Full and Half Candidate Data Bases

Team Role		Full Data Base		Half Data Base A		Half Data Base B	
1	Org & Ref Lead	581	581	250	645	581	618
2	Executive Summary Lead	580	580	580	580	83	36
3	Background & Tech Lead	642	121	642	642	20	20
4	Ethics Lead	645	645	645	250	618	651
5	Recommended Actions Lead	633	480	633	633	480	480
	Optimal Score	16030	15993	16023	16020	15990	15984

5. Conclusions

End-to-End Demonstration. The ATCS team composition runs successfully concluded the UCLA Engr 183 archival study as a complete end-to-end demonstration of the following major team composition processes:

1. Extracting from an existing personnel data base the information about individual characteristics and performance needed for automated team composition, as would be the case in many envisioned applications;
2. Using knowledge about the factors influencing team performance for a well defined mission along with knowledge about general mission objectives and team performance factors to establish the ATCS model parameters and optimization criteria, as well as the role and team constraints;
3. Using the programmed ATCS to form optimal teams in an acceptable computational time frame using a substantial candidate pool and typical model parameters and constraints.

The teams formed in the UCLA case certainly have great face validity: They have the right people in the right roles, as well as a high degree of diversity. They appear fully representative of teams that would be selected manually, but have the major advantages of known optimality using all programmed criteria, in addition to very rapid formation despite a large candidate pool – features impossible to achieve manually.

Because we used average Essay and Test scores as Mission Factors in our team optimization criteria, the team members selected have almost uniformly high scores for these taskwork variables. There was some suggestion in the statistical analysis (as well as in the literature) that having a team in which all members have a high individual performance level may not yield the best total team performance in all aspects of the mission. Accordingly, another useful team constraint might be to include some number of lower scoring members (i.e., to impose Grade Diversity), even though this is basically counter-intuitive. Likewise, in some situations the use of ethnicity data might be acceptable, and this could add another dimension of optimality to the present results. Our planned experimental research may help clarify this point.

Relationship to Future R&D. The ATCS project contains a number of innovative elements that will influence future research and development in the related research fields. These elements include:

- Structuring the immediate and/or long term mission of an existing organization in terms of a formal decision model in order to quantitatively represent important characteristics as part of the team optimization criteria. This is a significant contribution to practice in both organizational and team performance research domains because it effectively combines methodologies that are not usually, if ever, brought together.
- Enabling the combination of empirical results from team performance studies with automated team composition methodology based on multiagent system technology, where the first serves as an input to the latter. This is a totally new approach and will be a major contribution to research in both areas.
- In the future, using the results of the automated team composition system as a means of evaluating both the contribution of the various factors previously identified with effective team performance and as a means of testing empirically whether teams formed on the basis of previously identified key factors actually perform better in their processes and outcomes.

In essence, we are bringing together two disciplines: one, team performance modeling, including investigation of how teams evolve; and two, proactive computer-supported team composition through the use of advanced multiagent system technology. This is a significant contribution to team performance research and also an incentive to further studies of multiagent system applications. In future planned experimental work we will also encourage research and development by demonstrating solutions to two difficult problems in this field:

- What key measures of team performance can we get hold of practically by using available instruments, existing personnel data bases, acceptable questionnaires, etc.?
- How can we incorporate those measures and instruments into computer-assisted team composition systems that are both useful and usable?
- How do we address human and social issues in the area of computer-assisted team composition and improvement? Problems can arise both with individuals who want to protect their privacy (e.g. their address book, their qualifications, their associations, etc.) and with the “IT mafia” who are afraid of compromising “their” command and control network.

We believe that our solutions, which will be based on continued comprehensive literature research, thorough analyses, and empirical studies involving computer-assisted methodology, will be a major resource for future researchers.

6. References

1. Hackman, J. R. (1990). Work Teams in Organizations: An Orienting Framework. In J. R. Hackman (Ed.), *Groups that work (and those that don't)*. San Francisco, CA: Jossey-Bass
2. Campion, M. A., Medsker, G.J. & Higgs, A.C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*, 823-847.
3. McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. A. Guzzo, & Salas, E. (Ed.), *Team Effectiveness and Decision Making in Organizations* (pp. 9-45). San Francisco, CA: Jossey Bass.
4. Cannon-Bowers, J. A., Tannenbaum, S.I., Salas, E., & Volpe, C.E. (1995). Defining competencies and establishing team training requirements. In R. Guzzo, & Salas, E. (Ed.), *Team effectiveness and decision making in organizations* (pp. 333-380). San Francisco, CA: Jossey-Bass.

5. Driskell, J. E., & Salas, E. (1991). Group Decision Making Under Stress. *Journal of Applied Psychology, 76*(3), 473-478.

6. Driskell, J. E., Goodwin, G.F., Salas, E., & O'Shea, P.G. (2006). What makes a good team player? Personality and team effectiveness. *Group Dynamics: Theory, Research, and Practice, 10*(4), 249-271.

7. Salas, E., Guthrie, J.W., Jr., Wilson-Donnelly, K.A., Priest, H.A., & Burke, C.S. (2005). Modeling team performance: The basic ingredients and research needs. In W. B. Rouse, & Boff, K.R. (Eds.), *Organizational Simulation* (pp. 185-228). Hoboken, NJ: John Wiley & Sons.

8. Freedy, A., Cohen, M., Salas, E., Weltman, G., Tambe, M., Pearce, J., McDonough, J. and Freedy, E. "Automated Team Composition System (ATCS) Using Psychological Team Performance Modeling And Multiagent System Technology," SBIR Phase I Final Report, ARI Contract No. W91WAW-07-P-0021, Perceptronics Solutions, May 9, 2007

9. Weltman, G., McDonough, J., Freedy, E. Horvath, D., Cohen, M. Freedy, A. and Tambe, M. "Collaborative Assistance and Rapid Team Optimization System (CARTOS)." STTR Phase II Semi-Annual Report, DARPA Contract No. W31P4Q-08-C-0104, Perceptronics Solutions, December 31, 2009

10. Pearce, J. & Tambe, M. (2007). Quality guarantees on k-optimal solutions for distributed constraint optimization problem. *Proceedings of the 20th International Joint Conf. on AI, Hyderabad, India, Jan 6-12*

7. Acknowledgements

The research described here was performed under US Army Research Institute (ARI) SBIR Phase II Contract No.W91WAW-09-C-0151; the ATCS technology was partially developed and tested under DARPA STTR Phase II Contract No. W31P4Q-08-C-0104. We gratefully acknowledge the role of our subcontractor the UCLA School of Engineering and Applied Science (SEAS) in providing and formatting the Engr 183 student data, and the assistance at SEAS of Professor Rick Wesel and Dr. Donald Browne, who acted as Principal Investigator and Research Scientist, respectively. We also greatly appreciate the support and guidance of our ARI program manager Dr. Gerald Goodwin.