DOT/FAA/AM-02/18

Office of Aerospace Medicine
Washington, DC 20591

The Moderator Effects of Taskload on the Interplay Between En Route Intra-Sector Team Communications, Situation Awareness, and Mental Workload

Larry L. Bailey
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

Ben F. Willems
William J. Hughes Technical Center
Federal Aviation Administration
Atlantic City International Airport, NJ 08405

October 2002

Final Report

This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

U.S. Department of Transportation
Federal Aviation Administration

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-02/18		

4. Title and Subtitle	5. Report Date
The Moderator Effects of Taskload on the Interplay Between En Route Intra-Sector Team Communications, Situation Awareness, and Mental Workload	October 2002
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Bailey, L.L.[1], and Willems, B.F.[2]	

9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)
[1]FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125	[2]FAA William J. Hughes Technical Center Atlantic City International Airport Atlantic City, NJ 08405	11. Contract or Grant No.

12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

This report was performed under Task HRR518.

16. Abstract

Recently, the Federal Aviation Administration (FAA) began a program of research to investigate the role that intra-team communication plays in helping radar air traffic control teams coordinate their individual efforts. Based on the literature of controller and pilot communications, it was hypothesized that as taskload increased, communications would increase in order to maintain situational awareness. Furthermore, it was hypothesized that there would be an inverse relationship between taskload and situational awareness. METHOD. Using a high-fidelity air traffic control (ATC) simulator, ten 2-person teams, consisting of certified ATC specialists, performed routine ATC tasks within a single factor (low and high workload) repeated measure design. Performance was videotaped and the frequency of intra-team communications was counted. Post scenario perceptions of taskload and situational awareness were assessed using a version of the NASA Taskload Index (TLX) and a 4-item scale developed at the William J. Hughes Technical Center, respectively. RESULTS. Bivariate correlations of intra-team communications (c), taskload (t) and situational awareness (s) were separately analyzed for low and high workload. Because the sign of the correlations were established a priori, a one tailed test of significance was used with $p < .10$ as a test of significance. Results for the low workload condition were $r_{c,t} = -.14$ (ns), $r_{c,s} = .38$ (ns), and $r_{t,s} = -.62$**. Correlations for the high workload conditions were $r_{c,t} = .51$*, $r_{c,s} = .63$*, $r_{t,s} = -.30$ (ns). CONCLUSIONS. Under high workload conditions, as perceptions of taskload increased, there was a corresponding increase in the frequency of intra-team communications. The data suggest that the increase in communications is used to maintain situational awareness. This conclusion supports the a priori hypotheses. However, under low workload conditions, the data fail to support the hypotheses, with the exception that perceptions of situational awareness decreased as taskload increased.

17. Key Words	18. Distribution Statement
Air Traffic Control, Communication, Situation Awareness, Mental Workload, Teamwork	Document is available to the public through the National Technical Information Service; Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	19	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

The Moderator Effects of Taskload on the Interplay Between En Route Intra- Sector Team Communications, Situation Awareness, and Mental Workload

Three years ago, researchers from the William J. Hughes Technical Center (WJHTC) and the Civil Aerospace Medical Institute (CAMI) began a collaborative program of research in support of the National Airspace System (NAS) modernization effort. The goal of the modernization effort is to increase system capacity (i.e., increase traffic flow) without compromising safety (FAA, 1999). The modernization of the NAS involves, among other things, providing air traffic control specialists (ATCSs) with decision support tools (DSTs) that take advantage of recent advances in information technology. These advances in turn will have an impact on the roles and responsibilities of the air traffic control (ATC) workforce (FAA, 1999; Stefani, 2001).

The CAMI-WJHTC collaboration conducted a series of experiments to serve as a baseline for understanding how NAS modernization may affect the performance of ATCSs at en route centers, and specifically the performance of en route sector teams (ESTs). In this paper, we examined the effects of taskload (defined by the average number of aircraft) on the interplay between intra-team communication, perceived workload, and situation awareness. Prior research on intra-EST communication focused on describing the topic of communications, the grammatical form in which it was expressed, and the mode of communications (Peterson, Bailey, & Willems, 2001; Bailey, Willems, & Peterson, 2001). With this study, we begin to develop a model of how the communication process relates to EST performance. Specifically, we examined how taskload affects the relationship between intra-team communication, perceived workload, and situation awareness.

This study supports the 2001 recommendations contained in an audit of Free Flight Phase I Technologies conducted by the Department of Transportation Assistant Inspector General for Auditing, Alexis Stefani (2001). In her memorandum to the FAA, Ms. Stefani reiterated the need to

> conduct human factors work that examines the combined impact of new technologies (such as conflict alerts, electronic flight data, and enhanced weather information) on controllers, including "human in the loop" simulations. Key issues to be researched include safety, workload, situation awareness, and teamwork (p. xviii).

In the next section, we provide the reader with background information on how this study addresses that need.

BACKGROUND

In this section, a brief overview is presented of the constructs examined in the experiment.

En route Sector Teams

ESTs are formed whenever traffic flow increases beyond what is considered safe for a single ATCS. The simplest and most common EST consists of a radar controller (R-side) and a data controller/radar associate (D-side). The primary responsibility of the R-side is to ensure aircraft separation. An ATCS accomplishes this through the issuance of ATC instructions to aircraft pilots. The D-side's primary responsibilities are two fold: (1) to coordinate the transfer of control of aircraft with other ATCSs in adjacent sectors, and (2) to provide assistance to the R-side, often as an "extra pair of eyes." The D-side provides this assistance by scanning the radar screen to ensure proper aircraft separation. Additional roles and responsibilities of radar control teams appear in FAA Order 7110.65M (FAA, 2001).

For team members to coordinate their individual efforts, Salas, Stout, and Cannon-Bowers (1994) argue that they first must operate from a common perspective. This common perspective or shared mental model represents a common set of expectations concerning the meaning of task cues, compatible assessments of the situation, and common expectations of additional task requirements. Shared mental models allow EST members to know how their individual actions affect and are affected by other team members. It is through the development of shared mental models that team members develop a common awareness of a given situation.

Situation Awareness, Intra-Team Communications and Team Performance

Situation awareness is a construct that represents a person's knowledge of the current and future status of a dynamic environment. Endsley (1988) more formally defines situation awareness as the "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future (p. 97)." Although individuals develop their own awareness of a given situation, when individuals interact, their situation awareness overlaps. Other researchers have called this overlap shared situation awareness (Salas, Stout, and Cannon-Bowers, 1994; Bowers, Braun, & Kline, 1994).

Shared situation awareness provides the basis for the D-side to serve as a "second pair of eyes" for the R-side. For example, the D-side may hear the R-side asking a pilot to descend to a specific altitude. After the pilot verbally acknowledges the reception of the message, the R-side then directs his attention elsewhere. The D-side then notices that the aircraft has failed to descend to the stated altitude. When the R-side completes the task at hand, the D-side may simply point to the errant aircraft and the R-side acts accordingly without saying a word to the D-side.

As the above example illustrates, intra-EST communications are of short duration and may be verbal and/or non-verbal. Previous studies of en route intra-radar team communications demonstrated that R-side and D-side exchanges were infrequent yet important. The dominant topics of communication concerned (1) <u>traffic situations</u> involving a specific aircraft, (2) headings and/or amendments to <u>routing,</u> not in relation to traffic situations, and (3) aircraft <u>altitude</u> changes (Peterson, Bailey, & Willems, 2001; Bailey, Willems, & Peterson, 2001). Most communications were expressed verbally in the form of questions, answers and statements. Furthermore, Bailey et al. (2001) reported in a simulated ATC environment, more communication exchanges about traffic were recorded under higher taskload conditions as compared with lower taskload conditions. Under both conditions, communications were tactical in nature and predominately involved a specific aircraft, in a specific situation, and over a short period. In other words, the R-side and D-side use communications to maintain situation awareness.

Based on the above studies one might conclude that situation awareness-related communications in general would lead to better team performance. From the literature, this appears to be the case. Orasanu (1990) analyzed the communication of aircrews during a simulated mission that required crews to diagnose a problem with the aircraft and to make subsequent changes to their flight plan. Crews with higher situation awareness communications performed better. Following the work of Orasanu (1990), Mosier and Chidester (1991) reasoned that the link between aircrew communication and situation awareness would be most evident during emergencies. In their study, they examined the information exchange of aircrews during two simulated emergencies. They also found that the number of situation awareness-related communications predicted aircrew performance.

Although a link exists between intra team communication, situation awareness and team performance, the link may not always be evident. For example, using the same communication scheme as in her earlier study, Orasanu and Fischer (1991) examined the relationship between communication and flight performance across two separate aircraft: the Boeing 737 and the Boeing 727. The results demonstrated that the frequency of situation awareness communications differentiated between good and poor performing teams in the B737, but not for the B727.

One factor that affects the relationship between intra-team communication, situation awareness, and team performance is the operators' mental workload. Mental workload refers to the mental processing demands placed on an individual by objective task demands, called workload drivers. In an average ATC environment, the single most important task demand is the average number of aircraft within a controller's airspace (Wickens, Mavor, & McGee, 1997). Other examples of ATC workload drivers are sector complexity, multi-tasking, variations in weather patterns, and the use of new technology that interferes with habitual patterns of work.

The precise relationship between controller mental workload, intra-team communication, situation awareness, and taskload remains unclear. This is due in large part to a lack of multivariate studies. In addition, variations in airspace configuration and the average number of aircraft handled make it difficult to generalize from one study to the next. Still, one of the more common findings across studies is that as the average number of aircraft increase, there is a corresponding increase in controller-pilot communications, mental workload, and a decrease in situation awareness (Morrison & Wright, 1989; Morrow, Lee, & Rodvold, 1993).

Although the above studies depicted the relationships as linear, in all likelihood, the relationships are curvilinear with constraints forced by operational realities. Consider the relationship between intra-team communication and the average number of aircraft. As the number of aircraft increases there is an increase in the number of communication exchanges between the R-side and aircraft pilots. However, the relationship between communications and taskload does not stay linear. As Jorna (1991) reports, controllers who spend more than half of their time communicating with pilots report having difficulty maintaining adequate traffic awareness. Thus, there appears to be some upper limit to the amount of communication that can transpire before situation awareness begins to decline.

Curvilinear relationships are likely to appear in operational environments. Figures 1 and 2 provide some theoretical examples. As Figure 1 shows, with an increase in the average number of aircraft, there is a corresponding increase in intra-team communications (segment ab). However, this relationship doesn't continue along the positive trend. Instead with another increase in the average number of aircraft, intra-team communications flattens (segment cd) and then begins to decline (segment ef).

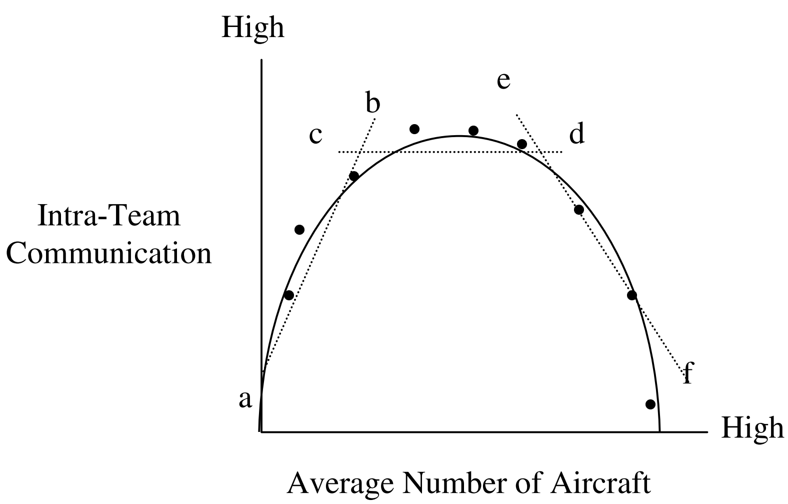


Figure 1. Theoretical relationship between the average number of aircraft and the frequency of intra-team communication.

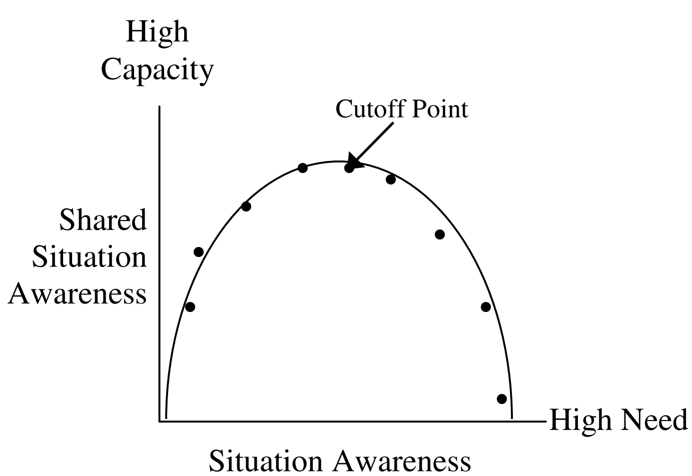


Figure 2. Theoretical relationship between situation awareness and shared situation awareness.

One reason for the drop in intra-team communications is explained in the theoretical relationship illustrated by Figure 2. Intra-team communications primarily serve to develop and maintain a shared awareness/understanding of traffic situations. As the need for situation awareness increases, there is a corresponding increase in the capacity necessary to share the awareness. The trend continues to some cutoff point where the capacity flattens before it begins to decline. With further increase in the need for situation awareness, the capacity for sharing the awareness diminishes. As previously stated, the diminishing capacity is due to the R-side's preoccupation with managing sector traffic. However, in real life the air traffic system is designed to prevent situations from arising where controller teams no longer can manage traffic in a safe and expeditious manner (FAA, 2001). For example, a third controller may be added to the sector team, or a busy sector may be partitioned into smaller sectors.

Technology vs Human Centered Innovations

To help reduce ATC workload and improve the situation awareness capacity of ATCSs, a number of decision support tools (DSTs) are in various stages of development and implementation. The DSTs are part of the FAA's modernization program designed to improve the capacity of the NAS, while at the same time ensuring the safety of flight operations (FAA, 1997). Examples of DSTs include computer modeling to determine optimal flight paths, conflict avoidance tools, and current time weather displays. By incorporating information from adjacent sectors, ATCSs can implement traffic solutions that not only resolve conflicts within the sector, but also prevent a conflict from happening within an adjacent sector.

Although technology is driving the FAA's ATC modernization program, human centered interventions also play a role. Of particular importance to this paper is the concept of EST reconfiguration (Thompson & Viets, 2001; Latron, et al., 1997; and Micro Analysis & Design, Inc. & System Resources Corp, 2000). Rather than having the R-side and D-side working as dyads, the reconfigured EST envisions multiple R-side controllers supported by a common D-side who operates as a multi-sector planner (MSP). In contrast to the short-term focus of the current D-side, the MSP would perform strategic planning tasks across sector boundaries. This, in effect, would minimize the number of conflicts that the respective R-sides would have to resolve and thus enable a more efficient traffic flow through their sectors.

Despite the interest in reconfiguring ESTs, few empirical studies exist that demonstrate the system benefits (e.g., fuel burn and operational errors) of incorporating the MSP into the team. To fill this void, researchers at WJHTC and CAMI designed an experiment to investigate how taskload and various EST configurations affect system performance in a simulated en route ATC environment.

CAMI's role in the experiment was to focus on R-side and D-side communication exchanges to: (1) determine the impact that the experimental manipulations had on intra-team communications, and (2) investigate the relationships between intra-team communication, perceived workload, and situation awareness. The aim of the latter was to develop a model suitable for inclusion in future "human in the loop" simulations targeted at examining the affect that NAS modernization has on controller performance.

Hypotheses

Although we recognize the potential curvilinear relationships of intra-team communications, shared situation awareness, mental workload, and taskload, we do not have an *a priori* understanding of the curvilinear effects of the experimental conditions. We, therefore, assume that the taskload (i.e., average number of aircraft) will not exceed the capacity of the participants to operate as R-side and D-side teams and that all relationships will follow the linear trends. Using the literature previously reviewed, the following hypotheses guided this research:

Hypothesis 1. More intra-team communication will occur under higher taskload conditions, as compared with lower taskload conditions.

Hypothesis 2 R-side/D-side teams will experience greater mental workload under higher taskload conditions, as compared with lower taskload conditions.

Hypothesis 3. R-side/D-side teams will perceive greater difficulty maintaining situation awareness under higher taskload conditions, as compared with lower taskload conditions.

Hypothesis 4a. Situation awareness will be positively related to the amount of intra-team communication.

Hypothesis 4b. Mental workload will be positively related to the amount of intra-team communication

Hypothesis 4c. Situation awareness will be inversely related to mental workload.

METHOD

Participants

Ten 3-person teams, consisting of certified ATCSs from several en route centers participated in a one-week experiment entitled "Study of an ATC Baseline for the Evaluation of Team-configuration (SABET)." The teams consisted of an R-side and D-side pair (the focus of this study) and a single R-side, who worked alone and did not communicate with the other members. Participants received their regular salary and government per diem throughout the duration of the experiment. Following two days of training, we randomly assigned participants to an experimental condition.

Equipment

The SABET experiment used the high-fidelity ATC simulator at the Research and Development Human Factors Laboratory at the WJHTC. The ATC equipment was functionally equivalent to the workstations used by the R-side and D-side ATCSs at en route centers. Included were a radarscope, a full flight strip bay, a Display System Replacement (DSR) keyboard, and a trackball. A high-resolution (2,000 by 2,000 pixel) Sony monitor presented the traffic data. We recorded R-side and D-side communication exchanges, and transferred them to MPEG-files. We then copied the files onto CD-ROMs for use in communication coding.

Stimulus Material

An ATC supervisor on detail to the WJHTC developed ATC scenarios in generic airspace for use in the training and experimental conditions. Each scenario was 40 minutes long. For the training condition, scenarios placed participants under a moderate taskload defined as the amount of air traffic that could be comfortably handled by an R-side/D-side team as perceived by a typical ATC supervisor. For the experimental conditions, we developed lower and higher-taskload scenarios. The lower-taskload scenarios were defined as the least amount of air traffic under which a typical ATC supervisor added a D-side to assist the R-side controller. The standard for developing higher taskload scenarios was the greatest amount of air traffic that a typical ATC supervisor allowed an R-side / D-side team to manage. Based on these standards, participants in the lower, moderate, and higher taskload conditions received a constant flow of 20, 25, and 30 aircraft, respectively.

Team Configuration

Each three-person team performed under three kinds of team configuration. We counter balanced assignment to control for order effects. The first configuration was the standard R-side / D-side EST. One R-side operated a sector by him/herself and the other sector was managed by an R-side / D-side pair (the focus of this study). The second configuration employed the multi-sector planner in the same room with two R-sides. The person occupying the multi-sector planner position was the same D-side controller as in the first configuration. The third configuration was like the second, with the exception that the muli-sector planner was in a different room from the two R-side controllers. Further details about the design of the team configurations and the corresponding traffic flows appear in Willems (2001).

Measures

We collected a variety of human performance measures in the broader SABET experiment. Measures pertaining to the focus of this study included an objective measure of communications, along with subjective measures of workload and situation awareness. We collected both subjective measures in survey format during the end of trial de-briefing.

Communication

We recorded all intra-team communications on videotape. A preliminary review of the data indicated that only the R-side and D-side communications in the first configuration were of sufficient frequency to warrant further analysis. We transferred all task-related R-side/D-side communications to CD-ROMs for coding by two ATC subject matter experts (SME) using the FAA's Controller-to-Controller Communication and Coordination Taxonomy (C^4T). As appendix A shows, the C^4T has three communication categories: the topic of communication, which is situation awareness-related; the grammatical form of communication (e.g. question, answer); and the mode of communication (e.g. verbal, nonverbal). Thus, the C^4T captures the "what" (topic) and "how" (form and mode) of communication. For further information on the development and operational validation of the C^4T, refer to Peterson, Bailey, and Willems (2001).

Both SMEs were retired ATCSs, and each had accumulated over 15 years of controlling air traffic. Previous work as SMEs involved using an operational incident typology for classifying ATC operational errors. We trained coders on the use of the C^4T and checked for agreement before independently coding R-side and D-side communication exchanges. We created team communication scores by summing the number of R-side and D-side transactions.

Workload.

Perceptions of mental workload were assessed using the NASA Taskload Index (TLX) developed by Hart & Staveland (1988). Six items comprise the TLX. These include a subjective appraisal of (1) mental demand, (2) physical demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration level. The items were defined and presented in a questionnaire format as suggested by Nygren (1991). In addition, the evaluation of one's performance (item 4) was dropped from the analysis as recommended by Bailey and Thompson (2001). Participants used a 10-point scale (1 = extremely low, 10 = extremely high) to indicate their perceived workload. The R-side and D-side scores were averaged to create a team score.

Situation Awareness.

Perceptions of situation awareness were assessed using a four-item scale developed at the WJHTC. The items included an assessment of situation awareness: (1) overall, (2) for current aircraft locations, (3) for projected aircraft, and (4) for potential violations. As noted by Endsley (1994), when subjective measures of situation awareness are assessed, they tend to measure the participants' confidence in their situation awareness rather than their actual situation awareness. Items were presented in a questionnaire format using a 10-point scale (1 = extremely low, 10 = extremely high). The average of the R-side and D-side scores was used to create a team score.

Training

Participants received two days of training on the airspace and traffic flow. After completing the familiarization phase of training, participants controlled traffic in six 40-minute air traffic control training scenarios. Each scenario represented a moderate taskload. On average, this meant that there were 25 aircraft on the radar screen at any given time. Data from the training phase were not recorded.

Design and Procedures

Communication exchanges of team members were assessed within a one factor repeated measures design. The two levels of taskload were lower and higher, as described earlier in the Method section.

Two ATC SMEs received 2 hours of training on the use of the C^4T. We then presented a CD of one of the experimental training sessions and asked the SMEs to reach agreement on the coding of each communication exchange. Following the coding, we debriefed the two SMEs to determine if there were any problems concerning the use of the C^4T. We then randomly assigned the SMEs to two sets of CDs containing different experimental sessions. Each SME coded different data sets, except for the inter-rater reliability checks conducted at the start of the coding project, midway through the project, and at project completion. We assessed inter-rater reliability under higher-taskload conditions to ensure that a sufficient number of communication events were present for a proper assessment.

RESULTS

All statistical analyses were conducted using the Statistical Package for Social Sciences (SPSS) for Windows, Version 10.0.

Inter-rater Reliability

Coefficient Kappa was used as measure of inter-rater agreement. The formula for Kappa is:

$$k = \frac{Po - Pc}{1 - Pc}$$

where P_o is defined as the proportion of observed agreement among raters, and P_c is the proportion of agreement expected by chance. Kappa ranges from a value of 0 to 1, indicating no agreement and perfect agreement, respectively. Kappa values of .70 or greater are considered suitable for this program of research.

Table 1 shows the results of the analyses. Kappas for topic were within the acceptable range with perfect agreement achieved midway and at project end. Coders had more difficulty reaching agreement in identifying the grammatical form of communications. The difficulty was associated with confusion over coding an event as a statement vs an implied question or an implied answer. As with the grammatical form, Kappas for the mode of communication were in the acceptable range with the exception of the comparison made at the start of the project.

Communication Descriptives

Table 2 shows the means and standard deviations of the number of communication events for the lower and higher-taskload conditions. Data presentation is organized around the topic, grammatical form, and mode of communication. The same data are graphically displayed in Figures 3-5. Note that the top three topics of communications are about route of flight, altitude, and traffic situations involving a specific aircraft. In addition, most communications are statements that are expressed verbally without a nonverbal component.

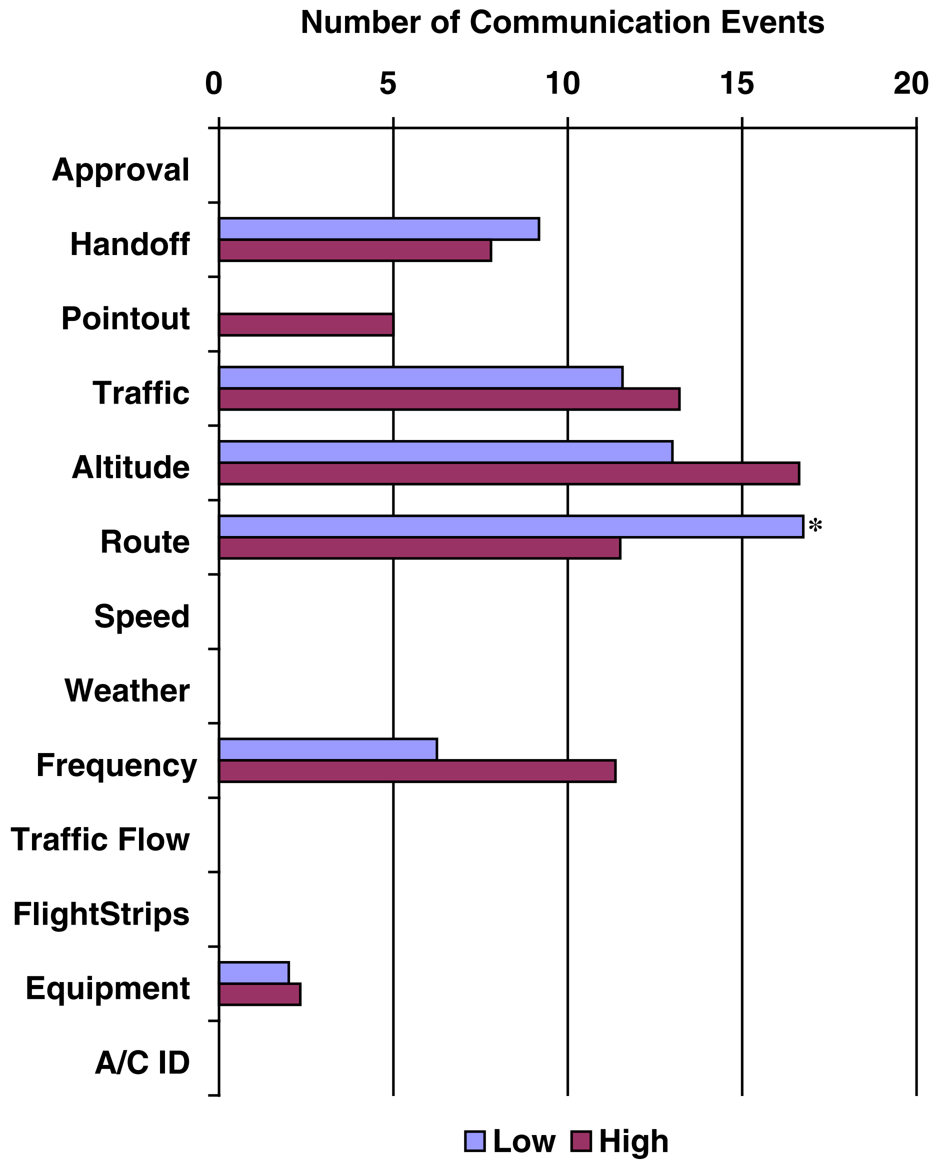
Collapsing across all conditions, Table 3 displays percentage comparisons between R-side and D-side communications. The percentage of R-side and D-side communications is balanced across the top three topics, which include route of flight, altitude, and traffic situations involving a specific aircraft. Compared with the R-side, a greater percentage of D-side communications are statements. In contrast to the D-side, the R-side asks proportionally more questions. Whereas the R-side primarily relies on verbal communications, the D-side uses both verbal and nonverbal

Table 1
Measure of Inter-rater Agreement Based on the Kappa Statistic

	Kappa Coefficient		
Time in project	Topic	Grammatical Form	Mode
Start	.80	.53	.50
Midway	1.00	.94	.89
End	1.00	.73	.82

Table 2
Descriptive Statistics of Topic, Form, and Mode of Communication for Lower and Higher Taskload*

	Lower			Higher		
Measure	n	M	SD	n	M	SD
Total Communication	8	52.38	34.56	9	55.11	41.39
Topic of Communication						
Approval	0			0	0	0
Handoff	7	9.17	11.67	6	7.83	9.47
Pointout	0			1	5.00	
Traffic	7	11.57	11.80	9	13.22	8.24
Altitude	7	13.00	8.89	8	16.63	16.39
Route of Flight	8	16.75	11.39	8	11.50	14.40
Speed	0			0		
Weather	0			0		
Frequency	8	6.25	5.23	8	11.37	9.38
Traffic Flow	0			0		
Flight Strips	0			0		
Equipment	1	2.00		3	2.33	1.53
Aircraft ID	0			0		
Grammatical Form of Communication						
Question	8	10.75	6.30	9	11.78	8.54
Answer	8	12.38	14.81	9	11.67	7.89
Statement	8	23.50	19.57	9	28.89	24.89
Command	7	4.71	3.50	8	3.13	3.00
Communication Mode						
Verbal	8	31.88	24.82	9	31.44	27.52
Mixed Verbal Nonverbal	8	20.13	10.88	9	25.33	15.31



*Statistically significant at $p < .10$.

Figure 3. Mean comparisons of topic of communication for lower and higher taskload.

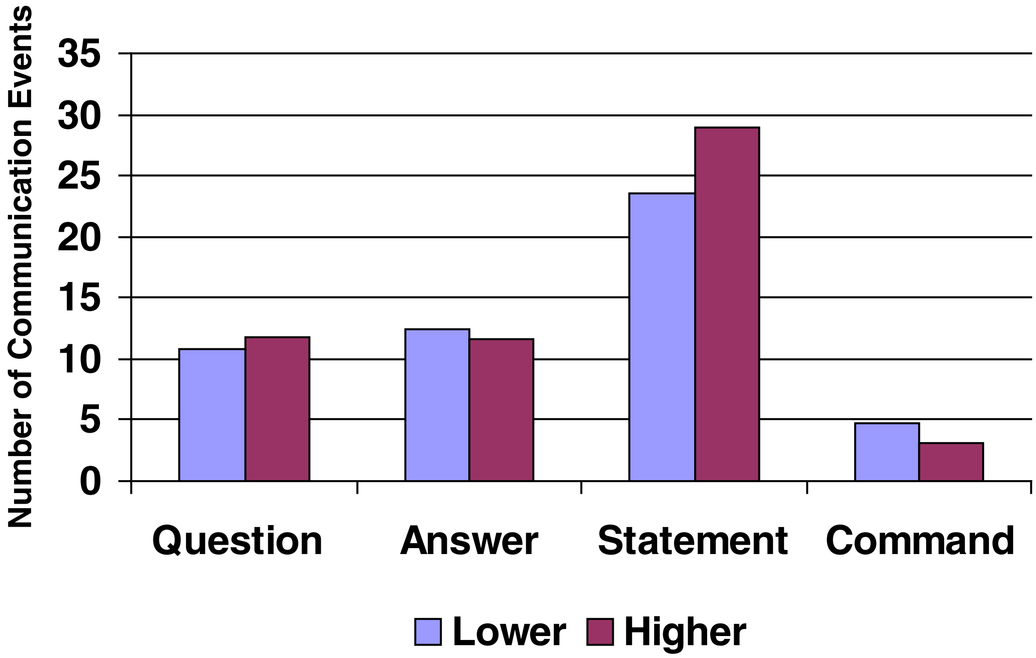
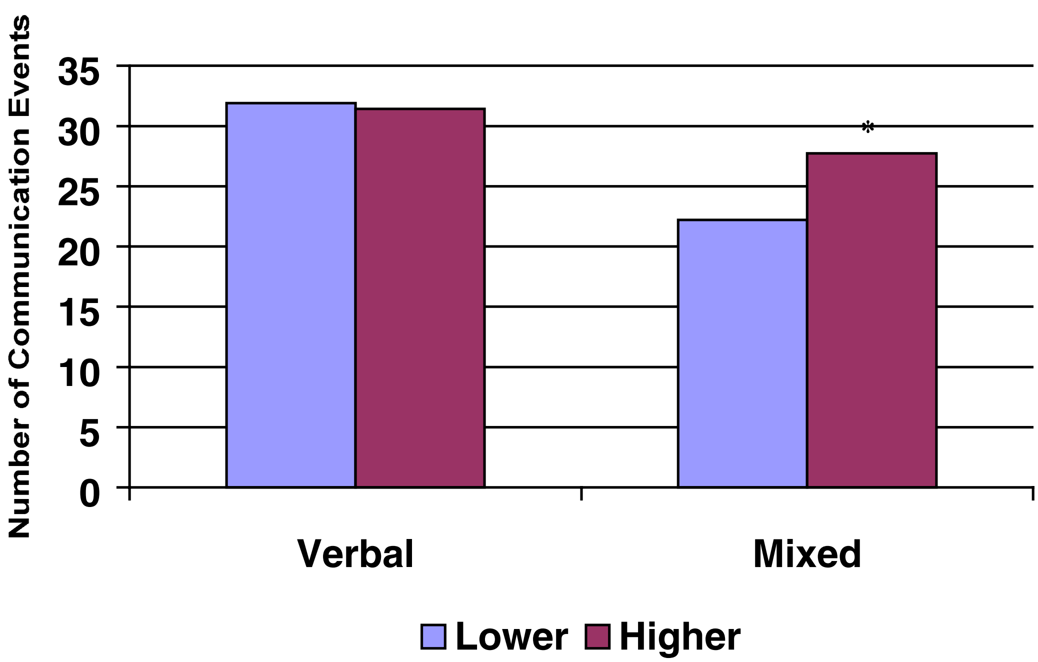


Figure 4. Mean comparisons of the grammatical form of communication for lower and higher taskload.



*Statistically significant at $p < .05$

Figure 5. Mean comparisons of the mode of communication for lower and higher taskload.

Table 3
Contrasting Percentage Comparisons of R-side and D-side Communications in Two Laboratory Settings*

	Current Study		Bailey, Willems, & Peterson (2001)	
	R-side%	D-side%	R-side%	D-side%
Communication Topic				
Route of flight	29.7	26.3	13.1	11.7
Altitude	25.2	24.8	16.0	21.1
Traffic	21.1	18.6	53.7	51.2
Hand-off	14.1	11.2	2.9	1.8
Frequency	8.4	18.6	3.5	2.7
Weather	0.0	0.0	0.0	0.0
Point-out	0.0	0.0	0.0	0.0
Traffic flow	0.0	0.0	0.0	0.0
Flight Strips	0.0	0.0	0.7	1.0
Equipment	1.4	0.4	4.5	4.9
Speed	0.0	0.0	4.4	2.8
Approval	0.0	0.0	1.0	1.8
Column Percent	99.9	99.9	99.8	99.0
Communication Format				
Statement	36.5	60.5	58.0	77.3
Question	34.1	9.2	22.9	11.3
Answer	22.5	25.5	18.3	10.4
Command	7.0	4.8	0.8	1.0
Command Answer	0.0	0.0	0.0	0.0
Column Percent	100.1	100.0	100.0	100.0
Communication Mode				
Verbal	75.2	45.1	93.9	69.0
Verbal & Nonverbal	22.6	45.5	5.0	24.7
Nonverbal	2.2	9.4	0.5	2.8
Equipment			0.0	0.1
Equipment & Verbal			0.6	3.4
Equipment & Nonverbal			0.0	0.1
Column Percent	100.0	100.0	100.0	100.1

*Caution should be used when comparing percentages across studies. Due to individual differences and scenario demands, the number of communication events across studies are not the same.

communications. Data from a similar experiment conducted by Bailey et al (2001) are presented as a reference point. Observed differences between the two studies should be viewed with caution due to sample variations and because the two studies used a different set of scenarios. Because of this, it is the trends that emerge from the comparison that are important, rather than the specific percentages.

Hypothesis Testing

During the experiment, problems occurred with the video recording of three experimental sessions. This resulted in the loss of data for some of the measures reported in this section. Because statistical power was low, all tests of significance were conducted at the level of $p < .10$. This is consistent with practices used in small group research (Stevens, 1996, p. 4).

Hypothesis 1 stated that more intra-team communication would occur under higher taskload conditions, as compared with low taskload conditions. There was inconclusive statistical support for this hypothesis. Although the higher-taskload condition produced a greater number of total communication events (see Table 2), the difference was not statistically significant, $t(6) = -.63$, $p = .55$. Analyses of individual C^4T sub categories, revealed two significant differences. The higher-taskload condition exhibited a greater amount of communications involving both verbal and nonverbal cues, $t(6) = 2.45$, $p = .05$. Under the lower-taskload condition, there were a greater number of communications about the route of flight, $t(6) = -1.90$, $p = .10$. This difference, however, was in the opposite direction as that hypothesized.

Hypothesis 2 indicated that R-side and D-side teams would experience greater mental workload under higher, as compared with lower taskload conditions. The results supported this hypothesis. Under higher taskload conditions, teams perceived that they were under greater workload ($M = 14.34$, $SD = 2.6$) as compared with working under lower taskload conditions ($M = 12.00$, $SD = 2.65$). The difference between the two conditions was statistically significant, $t(9) = -2.81$, $p = .05$.

Hypothesis 3 asserted that R-side/D-side teams would perceive greater difficulty maintaining situation awareness under higher taskload conditions, compared with lower taskload conditions. This hypothesis was supported by the results. While working under higher taskload conditions, teams perceived that their situation awareness was not as high ($M = 12.47$, $SD = 2.53$) as it was while working under lower taskload conditions ($M = 15.03$, $SD = 2.04$). The difference between the higher and lower taskload conditions was statistically significant, $t(9) = 3.02$, $p = .01$).

Hypotheses 4a - 4c addressed the interrelationships between perceptions of situation awareness, workload, and intra-team communications. The results are presented as a correlation matrix in Table 4. The numbers above the diagonal show the relationships under higher taskload conditions. The numbers below the diagonal show the relationships under lower taskload conditions. Given that the hypotheses indicated a direction for all relationships, a one tail test of significance was used. For ease of reporting, in the following section, the higher taskload relationships are presented first and then followed by the lower taskload relationships.

Table 4
Pearson r Correlations Under Lower and Higher Taskloads*

	Intra-Team Communications	Shared Situation Awareness	Mental Workload
Intra-Team Communications	__	.63 (p = .03)	.51 (p = .08)
Shared Situation Awareness	.38 (p = .18)	__	-.30 (p = .20)
Mental Workload	-.14 (p = .37)	-.62 (p = .03)	__

* Upper half of matrix is for higher taskload condition. Lower half of matrix is for lower taskload condition.

Higher Taskload

Hypothesis 4a and 4b achieved statistical support. Correlations of situation awareness perceptions and perceptions of workload with the amount of intra-team communication were .63 (p =.03), and .51 = (p = .08), respectively. Hypothesis 4c, however, did not achieve statistical support. Although the correlation between situation awareness perceptions and perceptions of workload was -.30, the value did not reach statistical significance (p = .20).

Lower Taskload

Situation awareness perceptions were hypothesized in 4a to be positively related to the amount of intra-team communication. Despite a positive correlation of .38, the hypothesis failed to achieve statistical support, p = .18. Hypothesis 4b, which predicted that perceptions of mental workload would be positively related to the amount of intra-team communication, also failed to achieve statistical support (r = -.14, p = .37). However, statistical support was achieved for hypothesis 4c which stated that situation awareness perceptions would be inversely related to perceptions of mental workload (r = -.62, p = 03).

Comparing Correlations

Although there were observed differences in the relationships exhibited while performing under lower and higher taskload conditions, these differences were not statistically significant. Using procedures outlined in Cohen (1988), the correlation coefficients were transformed to Fisher z equivalents so that an effect size index (q) could be computed for determining the sample size necessary to reach statistical significance at p = .10. As shown in Table 5, the computed sample sizes ranged from a low of 14 to a high of 46. In a similar study reported in Bailey et al. (2001), the authors noted that sample sizes in that range were cost prohibitive (as was also true for this study). To address the issue of small sample size, the authors suggested comparing results across similar studies to determine trends. It is in that spirit that the following section draws conclusions and discusses the results.

CONCLUSION AND DISCUSSION

The trend that is emerging in our research on intra-EST communications is that most communication exchanges are directed at maintaining situation awareness and pertain to the route of flight, aircraft altitudes and traffic situations involving a specific aircraft. R-side and D-side controllers differ not in the topics of their communication but rather in the grammatical form used and the mode by which communication occurs. Whereas D-side controllers primarily issue statements that draw attention to a situation, the R-side tends to generate more of a mixture of asking questions as well as issuing statements. Voice is the dominant mode of communication for both R-side and D-side controllers. However, the D-side also

Table 5

Pearson r, Fischer z Transformations, Effect Size Index, and Sample Size Determination

Relationship	Pearson r_L* (Fisher z_L)	Pearson r_H* (Fisher z_H)	Effect Size Index (q) $q = (\lvert z_L - z_H \rvert)$	Sample Size (p = .10)
Intra-Team Communication Vs Shared Situation Awareness	.38 (.40)	.63 .74	.35	46
Intra-Team Communication Vs Mental Workload	-.14 (-.14)	.51 (.56)	.70	14
Shared Situation Awareness Vs Mental Workload	-.62 (-.73)	-.30 (-.31)	.42	32

* Subscripts L and H connote lower and higher taskload conditions, respectively.

relies heavily on a mixture of verbal and non-verbal communications. The latter has implication when considering the impact of NAS modernization.

In the current en route environment, R-side and D-side controllers sit side by side. The D-side can glance over at the R-side's radar screen, thus serving as an "extra pair of eyes." If need be, the D-side can point to a particular location on the radar screen in order to draw the R-side's attention to a traffic situation. However, some aspects of NAS modernization might change that. Plans are underway to switch the role of the D-side from one of providing R-side tactical support to one of providing strategic support (Thompson & Viets, 2001; Latron et al., 1997, and Micro Analysis & Design, Inc., & System Resources Corp, 2000). Plans vary from providing the D-side with additional DSTs while working either in close proximity to the R-side to operating in a separate location. What impact might these changes have on the habitual patterns of intra-EST communications that have evolved over decades of use? Although technology no doubt will provide additional means of communications, what gains or losses might there be if the D-side could no longer read the R-side's body language or draw attention to a traffic situation simply by pointing? Questions such as these emphasize the need to continue developing a baseline understanding of the current use of communications between R-side and D-side controllers, both in the field and in laboratory conditions.

Moving beyond simply describing communication events, in this study, we began to model the inter-relationship between intra-EST communication, workload and situation awareness, and how that relationship is affected by the number of aircraft under the control of the R-side. Manipulations in the average number of aircraft (lower vs. higher) affected the amount of intra-team communications, mental workload, and participants' confidence in their situation awareness. Higher averages of aircraft produced more intra-team communications (not statistically significant), greater mental workload, and decreased the participants' confidence in their situation awareness.

The inter-relationship between intra-team communication, perceived workload, and situation awareness appeared to be differentially affected by the average number of aircraft. With lower numbers of aircraft, the relationship between situation awareness and intra-team communication appeared to be weaker, compared with that observed while performing with higher numbers of aircraft. Furthermore, there appeared to be an absence of a relationship between intra-team communication and mental workload. Taken together, the results suggested that at the lower level taskload, the R-side and D-side did not operate fully as an integrated team. In other words, the R-side did not require the assistance of the D-side but took the help when the D-side offered.

The relationships under higher taskload conditions were more consistent with the initial hypotheses. Intra-team communications was substantially related to mental workload and situation awareness. This indicated that the R-side required the assistance of the D-side to manage the traffic flow. Hence, the demands of the higher taskload condition required greater intra-team coordination/communication between the R-side and D-side.

Much of the discussions comparing the moderator effects of taskload need to be verified in studies with larger sample sizes. Nevertheless, the results of this study suggest that intra-team communications, mental workload, and situation awareness are inter-related and are differentially affected by changes in taskload such as the average number of aircraft being controlled. Experiments and/or field studies will be of limited value if they only examine the respective constructs in isolation or over a limited range of taskload conditions. Thus, it is recommended that a multivariate assessment of intra-team communications, workload, and situation awareness be included in future "human-in-the-loop studies" designed to assess the affect of NAS modernization on controller performance.

REFERENCES

Bailey, L., Thompson, R. (2001). The TLX: One or more constructs. Paper presented at *the Eleventh International Symposium for Aviation Psychology.*

Bailey, L., Willems, B., & Peterson, L. (2001). *The effects of workload and decision support automation on en route R-side and D-side communication exchanges*. (DOT/FAA/AM-01/20). Department of Transportation, Federal Aviation Administration, Office of Aerospace Medicine, Washington, D.C.

Bowers, C., Braun, C., & Kline, P. (1994). Communication and team situational awareness. In R. Gilson, D. Garland, & J. Koonce (Eds.), *Situational awareness in complex systems* (pp. 304-11). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Cohen, J. (1988). Statistical power analysis for behavioral sciences. Hillsdale, NJ: Erlbaum.

Endsley, M., (1994). Situational awareness in dynamic human decision making: Measurement. In *Situational Awareness in Complex Systems (*79-97). Dayton Beach, FL: Embry-Riddle Aeronautical University Press.

Endsley, M. (1988). Design and evaluation for situational awareness enhancement, in *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA: Human Factors Society, pp. 97-101.

Federal Aviation Administration (1997, September). *Air traffic services concept of operations for the national airspace system in 2005*. Washington, DC: Author.

Federal Aviation Administration (1999). *National airspace system architecture, version 4.0.* Washington, DC: Author.

Federal Aviation Administration (2001). *Air traffic control* (*Order 7110.65M)*. Washington, DC: Author.

Hart, S. & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock, & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-83). New York: North-Holland

Jorna, P.G.A.M. (1991). Operator workload as a limiting factor in complex systems. In J. Wise, V. Hopkin, & M. Smith (Eds.). *Automation and Systems Issues in Air Traffic Control* (pp. 281-92). Berlin: Springer Verlag.

Latron, P., McGregor, R., Geissel, M., Wassmer, E., & Marsden, E. (1997). En-route multisector planning procedures (DOC 97-70-15). Bruxelles, Belgium: Eurocontrol.

Micro Analysis & Design, Inc. & System Resources Corp. (2000). *Research task order 34. Intersector planning: En route controller roles and responsibilities. Final report.* Moffet Field, CA: NASA Ames Research Center.

Morrison, R. & Wright, R. H. (1989). ATC control and communications problems: An overview of recent ASRS data. In R. S Jensen (Ed*.) Proceedings of the Fifth International Symposium on Aviation Psychology*, 901-07.

Morrow, D., Lee, A., & Rodvold, M. (1993). Analyzing problems in routine controller-pilot communication. *International Journal of Aviation Psychology* 3, 285-302.

Mosier, K. & Chidester, T. (1991). Situation assessment and situation awareness in a team setting. *Proceedings of the 11th Congress of the International Ergonomics Association*, 798-800.

Nygren, T. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 33(1)*, 17-31.

Orasanu, J. (1990). Shared mental models and crew performance. Cogniive Science Laboratory Technical Report No. 46. Princeton, NJ: Princeton University.

Orasanu, J.& Fischer, U (1991). Information transfer and shared mental models of decision making. In R. Jensen (Ed*.). Proceedings of the Sixth International Symposium on Aviation Psychology*, 272-77.

Peterson, L., Bailey, L., & Willems, B. (2001). *Controller-to-controller communication and coordination taxonomy (C4T)*. (DOT/FAA/AM-01/19). Department of Transportation, Federal Aviation Administration, Office of Aerospace Medicine, Washington, D.C.

Salas, E., Stout, R., & Cannon-Bowers, J. (1994). The role of shared mental models in developing shared situational awareness. In R. Gilson, D. Garland, & J. Koonce (Eds.), *Situational awareness in complex systems* (pp. 297-304). Dayton Beach, FL: Embry-Riddle Aeronautical University Press.

Stevens, J. (1996). Applied multivariate statistics for the social sciences. Mahwah, NJ: Erlbaum.

Stefani, A. (2001). *Memorandum report on free flight phase 1 technologies: Progress to date and future challenges* (AV-2002-067). Department of Transportation, Office of the Inspector General, Washington, DC.

Thompson, K., & Viets, K. (2001, February). *An overview of a possible sector capability evolution for the air route traffic control center (ARTCC)* McLean, VA: MITRE Center for Advanced Aviation System Development.

Wickens, C., Mavor, A., & McGee, J. (Eds). (1997). Flight to the future: Human factors in air traffic control. Washington, DC: National Academy Press.

Willems, B.F. (2001). *Test plan for a study of an ATC baseline for the evaluation of team-configurations (SABET)*(Draft). Atlantic City International Airport, NJ: Federal Administration William J. Hughes Technical Center.

APPENDIX A.

Controller-to-Controller Communication and Coordination Taxonomy

Communication Topic	Definitions and Examples
Approval	Communications about intersector control/approval requests (“Get me control for descent on that aircraft.” “APREQ N1234 climbing to FL330.”).
Handoff	Communications relating to the transfer of radar identification of a particular aircraft (“Handoff N1234.” “Did you handoff N1234?”).
Point Out	Communications relating to the transfer of radar identification of a particular aircraft when radio communications will be retained (“Point out N1234 to 22.”).
Traffic	Communications about a traffic situation involving a specific aircraft. Includes conflict, spacing, other protected air space or terrain and the resolution of that situation (“Are you watching that aircraft?”).
Altitude	Communications about altitude not in relation to traffic (“N1234 is requesting flight level 220.”).
Route	Communications regarding headings and/or amendments to route, not in relation to traffic situations (“N1234 is on a 330 heading.” “Next sector, 27, wants N1234 over WEVER.”).
Speed	Communications about speed not in relation to traffic situations (“These three aircraft are slowed to 250 knots.”).
Weather	Communications about weather display or weather updates (Often communicated nonverbally by passing written information: “Sector 22 says continuous moderate turbulence above FL290.”).
Frequency	Communications about an aircraft’s radio communications transfer or frequency assignment (“Have you switched N1234 yet?” “Tell them to switch to N1234.”).
Flow Messages	Communications about traffic flow restrictions not referring to a specific aircraft (“The next sector is requesting 25 miles in trail.”) (due to radar outage).

Controller-to-Controller Communication and Coordination Taxonomy
(Continued)

Communication Topic	**Definitions and Examples**
Flight Strips	Communications about flight progress strips ("Where is that strip?") Often communicated nonverbally.
Equipment	Communications about any ATC hardware (The radar is out of service.").
Aircraft ID	Communciations involving identifying a specific aircraft (Who was that who called?" "That was N1234 who called.").

Grammatical Form	**Definitions**
Question	A direct inquiry about the state or status of sector events.
Answer	A response to a direct or implied question
Statement	Providing information, without being asked, about the state or status of sector events.
Command	A direct order to perform a specific act

Communication Mode	**Definitions**
Verbal	Use of voice only communication.
Nonverbal	Use of only body movement communication.
Mixed	Communication that contains both a verbal and non verbal component.
Electronic	Communication that is electronically transferred.