DOT/FAA/AM-00/25

Office of Aviation Medicine
Washington, D.C. 20591

# The Effects of Performance Feedback on Air Traffic Control Team Coordination: A Simulation Study

Larry L. Bailey
Richard C. Thompson
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

July 2000

Final Report

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

U.S. Department of Transportation
**Federal Aviation Administration**

DTIC QUALITY INSPECTED 4

20000921 079

# N O T I C E

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents thereof.

**Technical Report Documentation Page**

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| DOT/FAA/AM-00/25 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| The Effects of Performance Feedback on Air Traffic Control Team Coordination: A Simulation Study | July 2000 |
| | 6. Performing Organization Code |
| 7. Author(s) | 8. Performing Organization Report No. |
| Bailey, L.L., and Thompson, R.C. | |
| 9. Performing Organization Name and Address | 10. Work Unit No. (TRAIS) |
| FAA Civil Aeromedical Institute<br>P.O. Box 25082<br>Oklahoma City, OK 73125 | 11. Contract or Grant No. |
| 12. Sponsoring Agency name and Address | 13. Type of Report and Period Covered |
| Office of Aviation Medicine<br>Federal Aviation Administration<br>800 Independence Avenue, SW<br>Washington, D.C. 20591 | 14. Sponsoring Agency Code |

15. Supplemental Notes

This work was performed under Task AM-B-99-HRR-518

16. Abstract

This study examines the efficacy of team members observing a computer playback of their performance as a training tool in building effective air traffic control (ATC) teams. Participants performed various simulated radar-based ATC tasks under varying levels of aircraft density. The results suggest that observing a computer playback of one's team performance enabled team members to gain a system's perspective of how their performance both affects and is affected by others. This perspective enabled team members to better coordinate their individual efforts, which thereby led to improvements in team cohesion and in the percentage of aircraft that reached their destination within the time constraints of a given scenario. However, as aircraft density increased, these improvements began to diminish. Additional specialized training may be necessary for ATC teams to improve their coordination during periods of high workload.

| 17. Key Words | 18. Distribution Statement |
|---|---|
| Air Traffic Control, Training, Teamwork | Document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161 |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 13 | |

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

# The Effects of Performance Feedback on Air Traffic Control Team Coordination: A Simulation Study

## INTRODUCTION

Crew resource management (CRM) refers to the effective coordination of individual crew members' efforts during the course of performing their mission. Although CRM training originally was targeted at improving aircrew coordination, its audience now has expanded to include aircraft dispatchers, flight attendants, maintenance personnel, and air traffic control specialists (Federal Aviation Administration, FAA, 1998a). Recently, CRM principles applied in the air traffic control (ATC) setting have received increasing attention, due in part to the technological changes associated with the modernization of the National Airspace System (NAS) (FAA, 1998b). For example, changes in radar displays, flight strips, and automated updates of flight plans will not only affect how individual air traffic control specialists (ATCSs) manage the traffic flow within a given sector but will also influence how ATCSs interact with each other and other NAS users. Thus, ATC-CRM training should not only address the current NAS structure but also prepare ATCSs for future developments. However, if ATC-CRM training is to help facilitate the transition from the current to the new NAS, it is important that the training delivered is of proven quality and effectiveness.

One of the techniques used in delivering skill-based CRM training is to provide crews with a video playback of their performance conducted under mission simulated conditions (FAA, 1998a). This training technique is based on the premise that by observing a video playback of their own performance and discussing the results, crews will be able to identify critical factors affecting crew coordination and will adjust their behavior accordingly (Hartel and Hartel, 1995; Prince, Childester, Bowers, and Cannon-Bowers, 1992). However, this training technique has not been empirically evaluated to determine if it *does* produce a positive effect on teamwork, and, if so, what aspect(s) of teamwork is/are being affected.

In this paper, we report on the results of an ATC simulation experiment that examined the effect that observing a computer generated playback had on building an effective ATC team. Specifically, we examined the differential impact that observing a playback of one's team vs. one's own performance had on a number of dependent variables. This included measures of: a) team cohesion, b) shared mental model, c) ATC system outcomes, defined as the percentage of aircraft that reached their destination, the amount of aircraft delay, and the number of safety errors, and d) subjective workload.

Research on teamwork suggests that effective team training techniques facilitate the development of teamwork by building greater team cohesion and task coordination. The latter then leads to increases in productivity (Feldman, 1968; Salas, Dickinson, Converse and Tannenbaum, 1992). However, this increase in productivity does not necessarily mean that team members are working harder. To the contrary, research suggests that effective teams are able to produce more because less time is spent on addressing interpersonal problems and problems associated with coordinating individual efforts (Morgan, Glickman, Woodard, Blaiwes, and Salas, 1986). Furthermore Cannon-Bowers, Salas, and Converse. (1993) suggest that effective task coordination may be accomplished through the development of a shared mental model of teamwork. Included in the mental model are team members' expectations concerning the time-sequencing of events, the tasks to be performed, and how individual efforts will be coordinated. For a more complete review of the literature on shared mental models the reader is referred to Klimoski and Mohammed (1994) and Kraiger and Wenzel (1997).

Although the above description depicts a causal pathway, this was not tested in the study reported in this paper. Instead, each of the above constructs was independently examined in a 2 x 3 analysis of variance cross sectional experimental design. In the first factor, two levels of training were manipulated, system oriented vs. individual oriented. In the second factor, three levels of aircraft density were manipulated (low, medium, and high) to simulate changes in objective workload. Based on the literature previously reviewed, the following hypotheses were derived:

**Hypothesis 1.** Participants trained to assume a system perspective of teamwork, defined as monitoring the performance of the entire team, will have a greater sense of team cohesion, compared with those trained to focus on individual performance.

**Hypothesis 2.** Participants trained to assume a system perspective of teamwork will have more fully developed shared mental models, compared with those trained to focus on individual performance.

**Hypothesis 3.** Participants trained to assume a system perspective of teamwork will outperform participants trained to focus on individual performance, as measured by a reduction in safety errors, aircraft delay time, and an increase in the percentage of aircraft reaching their destination.

**Hypothesis 4.** Participants trained to assume a system perspective of teamwork will experience less subjective workload, compared with those trained to focus on individual performance.

## METHOD

**Participants.** Two hundred forty adults (53% female) between the ages of 18 and 30 participated in simulated radar-based ATC tasks using four-person teams. Participants were recruited through a local temporary help provider and were paid $10/hour.

**Equipment.** A high psychological fidelity, multi-sector, personal computer-based, ATC team training device was used to simulate radar-based ATC tasks. The training device consisted of four computer workstations networked in an interlocking pattern, as shown in Figure 1. This configuration enabled four-person teams to perform ATC tasks under varying levels of aircraft density. Figure 2 shows a screen shot of one of the workstations.

**Stimulus material.** Three 28-minute scenarios were developed and calibrated to create three levels of aircraft density based on the number of aircraft presented to the team over time. In the low, medium, and high-density scenario conditions, aircraft were presented so that, on average, each sector would be managing 1.9, 2.7, and 3.6 aircraft, respectively. Aircraft originally appeared in an inactive state and were activated at the discretion of the participant in the originating sector. Once activated, the aircraft had to travel through three sectors before landing at an airport in the fourth sector. Further information about the scenario development is reported in Bailey, Broach, Thompson, and Enos (1999).

**Measures.** Three measures of team performance and three measures of individual performance were assessed.

### Team Measures.

*Team Cohesion* was assessed as a psychological construct (James, James, and Ashe, 1990) using a 15-item scale modeled after Feldman's (1968) three-factor model of cohesion: a) interpersonal liking, b) task coordination, and c) normative integration. Participants used a five-point Likert-scale (1 = strongly disagree, 5 = strongly agree) to indicate their responses.

An index of *shared mental model* completeness was developed using the Kendall's *W* statistic. Kendall's *W* measures the degree of agreement (0 = no agreement, 1 = perfect agreement) among raters (in this case team members) relative to a set of variables. Team members were asked to determine the relative importance of nine task-specific activities evenly distributed over three categories: a) safety, b) efficiency and effectiveness, and c) coordination. Each of the nine variables was independently compared to the other variables for a total of 36 pair-wise comparisons. For each comparison, participants chose the task variable that was most important to their success in the scenario they had just completed. A numerical value ranging from zero to eight was then assigned to each variable based on the number of times it was endorsed. The degree of team member agreement was then computed using Kendall's *W*. Higher Kendall *W* values were used to represent a more complete shared mental model among team members. Due to the nature of the index, this measure can only be assessed at the team level of analysis.

*Percentage of aircraft reaching their destination* was computed by dividing the number of aircraft that landed at the designated airports by the total number of aircraft eligible to land at a given airport. Higher percentages of aircraft landings indicate higher degrees of team performance. This measure is an indicator of team performance due to the way scenarios were developed. All aircraft fly through three sectors before landing in the fourth. Consequently, each

team member's score on this measure represents 25% of his or her own performance and 75% of other team members' performance.

### Individual Measures

*Aircraft delay time* was computed based on the mean of activation delay and destination delay time. Activation delay represented the time it took (in seconds) for an aircraft to be activated once it appeared on the computer screen. Destination delay referred to the amount of time aircraft deviated from optimal flight paths within a given sector. This measure is also an indicator of individual performance.

*Safety errors* represented a composite of three domains: a) separation errors, which were the number of times participants violated the five mile horizontal separation rule, b) procedural errors, which were related to the number of times participants violated speed and altitude restrictions, and c) number of aircraft "crashes." This measure is primarily an indicator of individual performance, as each participant was in control of aircraft within his/her sector.

*Subjective workload* was assessed using a variation of the National Aeronautic and Space Administration Task Load Index (TLX; Hart and Staveland, 1988). In the TLX, subjective workload is viewed as a multidimensional construct involving one's subjective appraisal of: a) mental demand, b) physical demand, c) temporal demand, d) performance, e) effort, and f) frustration level. These dimensions were defined and presented as single items in a questionnaire format. Participants used a 21-point scale (1 = low, 21 = high) to indicate their responses. The ratings were then averaged to produce an overall measure of subjective workload.

**Procedures.** Participants were first screened to determine their eligibility for participation. To be considered, candidates had to pass a 10-minute test based on simulated ATC tasks. The tasks required that subjects use a point and click method with a mouse to issue changes in aircraft direction, speed, and altitude. Thirty minutes of training were provided prior to testing. During the testing phase, candidates were given up to three chances to receive a passing score. Those who failed to achieve a passing score received payment for their time and were excluded from further participation in the study.

Qualified participants were randomly assigned to one of 60, four-person teams with the goal of having two males and two females on each team. Each team was assigned to one of six experimental conditions based on the type of training received (two levels) and the level of aircraft density used in the experiment (three levels).

Teams participated in two hours of training and one hour of experiments, for a total of three hours. During the two hours of training, teams were assigned to one of two training conditions. The first condition (self) emphasized the development of individual sector management strategies. This was accomplished by explaining that participants did not need to be concerned about how other team members were performing. All that was necessary was to follow the rules and concentrate on their individual performance. After each training exercise participants were presented with a computer replay of their performance and were asked to examine it to determine ways to improve their individual performance. No opportunity was presented for team members to exchange information with each other.

In the second training condition (team), the development of team management strategies was emphasized. This was accomplished by emphasizing that participants were part of a team, and each member needed to think about how his or her performance both affected and was affected by others. After each training exercise, participants were presented with a computer replay of their team's performance and were asked to examine it, as a group, to determine ways for their team to improve its overall performance. A facilitator was provided to ensure that the group focused their discussions on issues related to their team performance.

## RESULTS

A series of 2 X 3 analyses of variance (ANOVAs) were used to examine the effects that training (self vs. team) and aircraft density (low, medium, and high) had on the separate measures of: a) team cohesion, b) the shared mental model index, c) the percentage of aircraft that reached their destination, the amount of aircraft delay, the number of safety errors, and d) subjective workload. These results are presented in Table 1. Table 2 presents the supporting descriptive statistics that are graphically displayed in Figures 3-6.

**Table 1.**

Analysis of Variance Summary Table

| Dependent Measure | Independent Measure | df | F | p |
|---|---|---|---|---|
| Team Outcomes | | | | |
| Team Cohesion | | | | |
| | Training | 1 | 7.16 | .01 |
| | Aircraft Density | 2 | 3.75 | .03 |
| | Training X Aircraft Density | 2 | 4.46 | .01 |
| Shared Mental Model* | | | | |
| | Training | 1 | .64 | .43 |
| | Aircraft Density | 2 | 2.38 | .10 |
| | Training X Aircraft Density | 2 | .40 | .67 |
| Percentage to Destination | | | | |
| | Training | 1 | 1.17 | .28 |
| | Aircraft Density | 2 | 96.01 | .00 |
| | Training X Aircraft Density | 2 | 8.61 | .00 |
| Individual Performance | | | | |
| Aircraft Delay Time | | | | |
| | Training | 1 | .86 | .35 |
| | Aircraft Density | 2 | 72.21 | .00 |
| | Training X Aircraft Density | 2 | 2.17 | .12 |
| Safety Errors | | | | |
| | Training | 1 | .22 | .64 |
| | Aircraft Density | 2 | 37.34 | .00 |
| | Training X Aircraft Density | 2 | .20 | .82 |
| Subjective Workload | | | | |
| | Training | 1 | 0.09 | .76 |
| | Aircraft Density | 2 | 26.46 | .00 |
| | Training X Aircraft Density | 2 | 2.33 | .10 |

**Table 2.**

Descriptive Statistics for Dependent Measures by Training Condition

| Aircraft Density | Measure | Individual Focus n | Individual Focus M | Individual Focus SD | Team Focus n | Team Focus M | Team Focus SD |
|---|---|---|---|---|---|---|---|
| Low | | | | | | | |
| | Team Cohesion | 40 | 3.55 | .69 | 40 | 3.98 | .54 |
| | Shared Mental Model* | 10 | .31 | .18 | 10 | .29 | .14 |
| | % to Destination | 40 | 68.74 | 25.37 | 39 | 89.31 | 18.53 |
| | Aircraft Delay Time | 40 | 16.41 | 14.5 | 36 | 12.54 | 8.35 |
| | Safety Errors | 40 | 6.93 | 4.46 | 38 | 6.42 | 3.37 |
| | Subjective Workload | 40 | 9.42 | 3.00 | 40 | 8.44 | 3.51 |
| Medium | | | | | | | |
| | Team Cohesion | 39 | 3.59 | .40 | 44 | 3.50 | .64 |
| | Shared Mental Model* | 10 | .31 | .16 | 11 | .39 | .18 |
| | % to Destination | 40 | 50.44 | 26.53 | 44 | 43.88 | 26.12 |
| | Aircraft Delay Time | 40 | 36.51 | 25.85 | 44 | 53.22 | 41.93 |
| | Safety Errors | 40 | 11.18 | 8.22 | 44 | 11.52 | 9.06 |
| | Subjective Workload | 40 | 11.31 | 3.05 | 44 | 11.94 | 2.66 |
| High | | | | | | | |
| | Team Cohesion | 39 | 3.43 | .58 | 36 | 3.68 | .49 |
| Scenario | Measure | N | U | SD | N | U | SD |
| | Shared Mental Model* | 10 | .39 | .15 | 9 | .43 | .11 |
| | % to Destination | 30 | 29.17 | 21.41 | 32 | 25.00 | 10.58 |
| | Aircraft Delay Time | 30 | 83.23 | 50.25 | 31 | 82.96 | 41.39 |
| | Safety Errors | 30 | 19.00 | 11.12 | 32 | 17.66 | 9.46 |
| | Subjective Workload | 40 | 11.83 | 3.43 | 36 | 12.14 | 2.59 |

* Team level of analysis

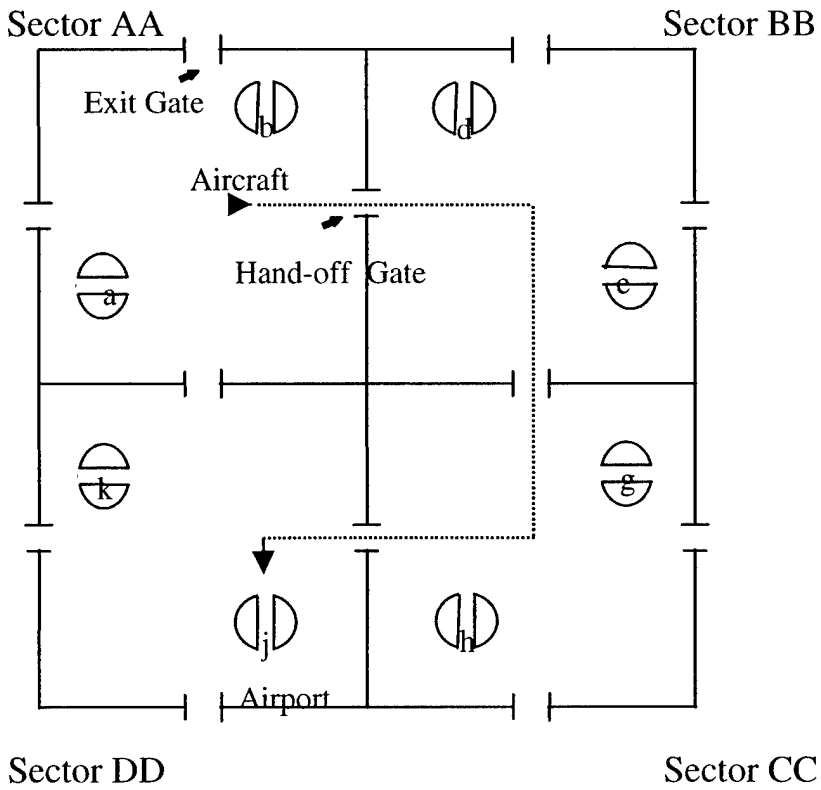


**Figure 1.** Sector configuration showing airport layout (lower case alpha), exit gates (at the perimeter), hand-off gates (at the interior), aircraft (arrow), and route of flight (dashed line).

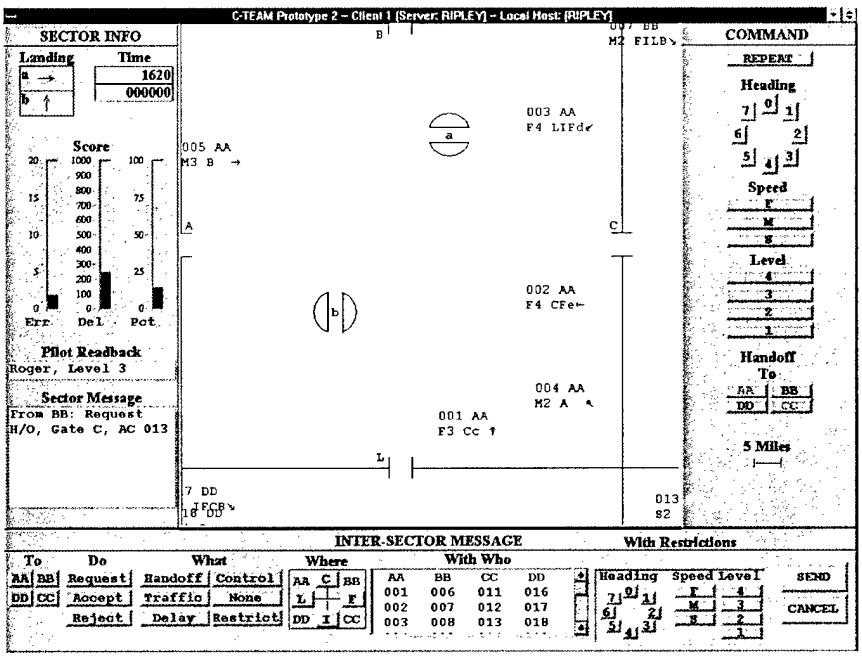


**Figure 2.** Screen shot from a CTEAM workstation showing sector airspace (middle), sector information (top left), aircraft commands (top right), and the controller-to-controller communication panel (bottom).

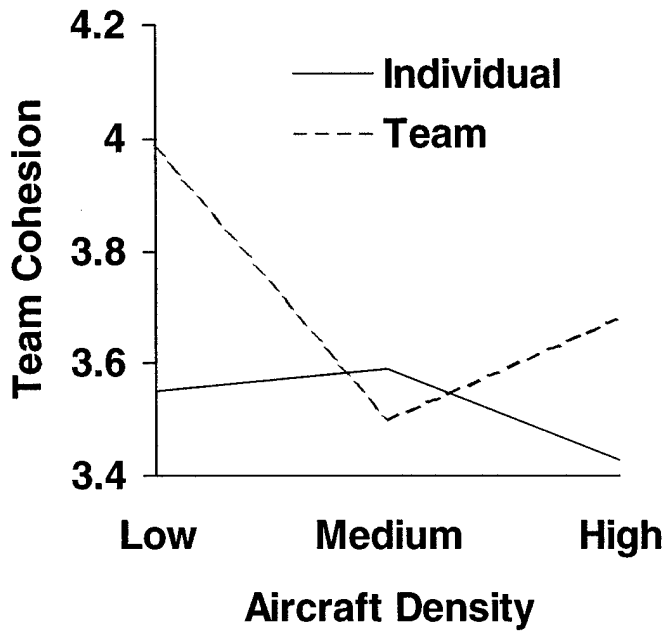


**Figure 3.** Team Cohesion as a Function of Aircraft Density and Training

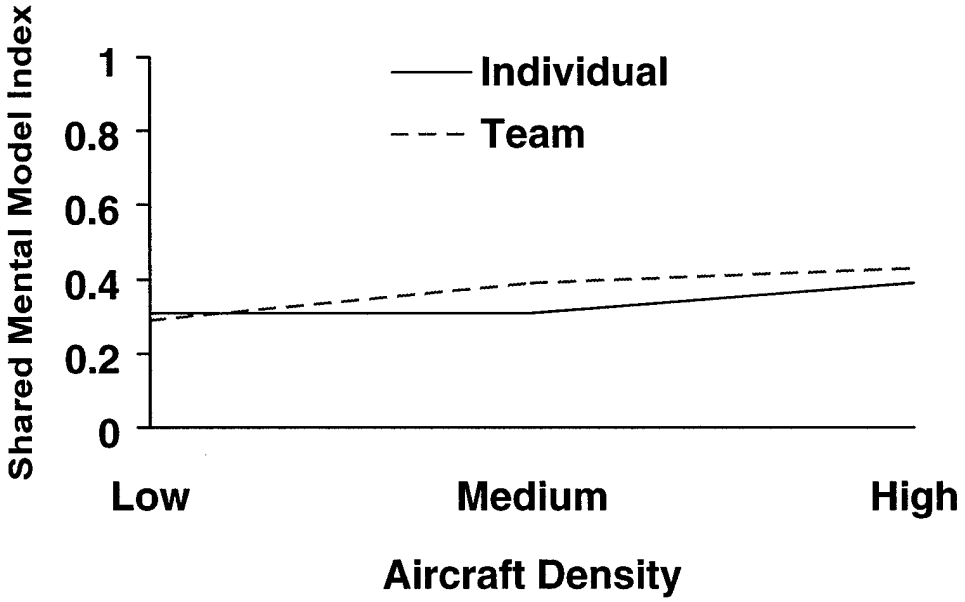


**Figure 4.** Shared Mental Model Index as a Function of Aircraft Density and Training

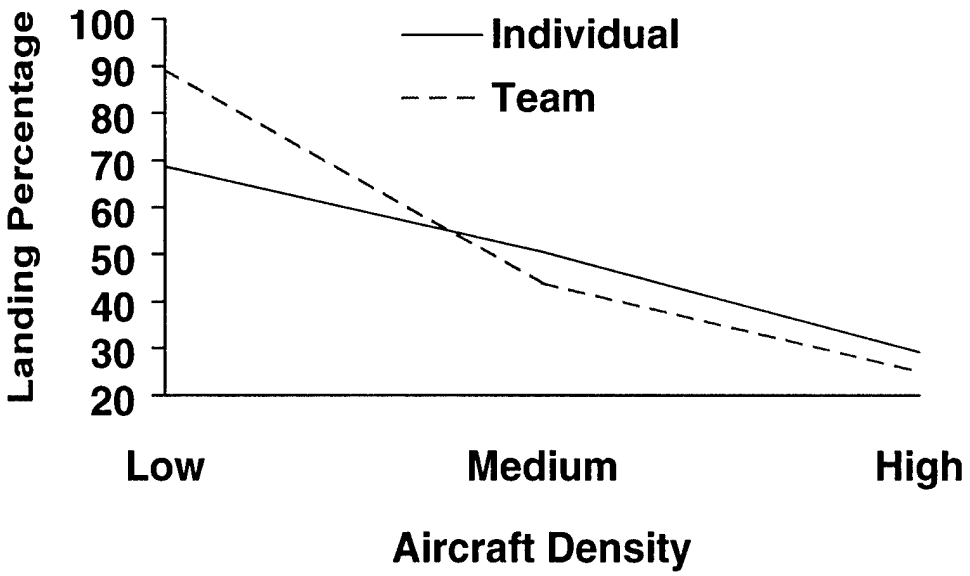


**Figure 5.** Percentage of Aircraft Landing as a Function of Aircraft Density and Training

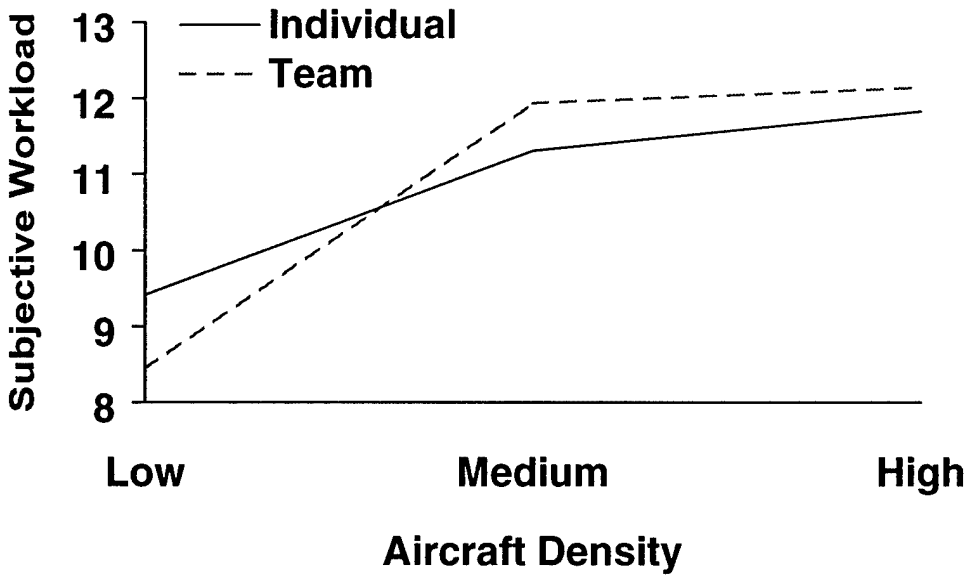


**Figure 6.** Subjective Workload as a Function of Aircraft Density and Training

Hypothesis 1 stated that training from a team perspective would result in greater team cohesion compared with training that focused on individual performance. This hypothesis was partially supported by a significant training by aircraft density interaction observed for team cohesion, $F(2, 232) = 4.46$, $p < .01$ (see Table 1). As Figure 3 shows, the greatest level of team cohesion for the team-training group occurred under low aircraft density conditions. The team-training group also reported a greater sense of team cohesion under high aircraft density conditions. However, this difference did not achieve statistical significance.

Hypothesis 2 predicted that effective team training would result in a more fully developed, shared mental model (i.e., higher values of Kendall's *W*) among team members compared with training that emphasized individual performance. There was no statistical evidence to support this hypothesis. However as Figure 4 shows, with increasing aircraft density there appears to be a tendency toward the formation of a more complete shared mental model among team members. A subsequent correlational analysis confirmed the statistical significance of this observation with a Pearson r of .28 ($p < .05$).

Hypothesis 3 stated that team-oriented training would produce greater improvement in system outcomes, compared with training that focused on individual performance. There were no statistically significant differences between training conditions observed for the amount of aircraft delay or the number of safety errors (see Table 1). However, due to a significant training by aircraft density interaction, partial support was provided for the percentages of aircraft that reached their destination, $F(2, 219) = 8.61$, $p < .01$. An examination of Figure 5 reveals that this interaction was due to higher percentages of aircraft reaching their destination for the team training group under low aircraft density conditions.

Hypothesis 4 predicted that those who received team-oriented training would experience less subjective workload, compared with those who received training that focused on individual performance. As Table 1 shows, there was no statistical support for this hypothesis. Figure 6 suggests that the team-training group appeared to experience less subjective workload under low aircraft density conditions, and slightly more subjective workload under medium and high aircraft density conditions.

## CONCLUSION AND DISCUSSION

This study examined whether observing a computer generated playback of one's team performance would lead to: a) greater team cohesion, b) more complete shared mental models, c) better air traffic control system outcomes, and d) a reduction in subjective workload. Groups who received team feedback had statistically greater team cohesion and more aircraft landings under low aircraft density conditions, compared with groups who received feedback just on their individual performance. There were no statistically significant group differences observed under medium and high aircraft density conditions. Also, contrary to expectations, more complete shared mental models were observed in the high density conditions regardless of how teams were trained.

The results of this study have several implications for ATC researchers, trainers, and supervisors.

First, from an ATC-CRM training perspective, it is interesting to note that teamwork improved under low aircraft density conditions simply from observing and discussing the results of a computer generated playback of one's team performance. These results suggest that visual feedback, which allows individuals to observe how their performance affects, and is affected by others, provides enough stimulation for team members to improve their coordination. Although the utility of using video playbacks in team training has a long history in CRM training, this appears to be the first study that has empirically measured its training efficacy.

Second, from a psychometric standpoint, the results identified two potentially useful measures of effective teamwork: one subjective (team cohesion), and one objective (percentages of aircraft reaching their destination). Both measures were sensitive to the experimental manipulation of team training. Moreover, only the measures of teamwork and not the measures primarily focused on individual performance, amount of aircraft delay, number of safety errors, and subjective workload, were sensitive to the effects of observing the video playback of one's team vs. one's own performance. Although team cohesion has long been used as a measure of teamwork (Salas et al., 1992), there has been a lack of objective measures of ATC teamwork reported in the scientific literature (Hartel and Hartel, 1995).

Third, as new technologies are integrated into the NAS, it is important for managers and policy makers to understand how ATC coordination is affected by varying levels of workload, such as aircraft density. As reported, ATC team coordination in a simulation environment was significantly improved only under low aircraft density conditions. It is possible that, with a stronger training manipulation, such as conducting more structured team debriefings, similar improvements could have been achieved at higher levels of aircraft density. However, it may be that there is also a limit on the amount of work that a team is able to effectively coordinate. Perhaps as aircraft density increased, the corresponding task demands forced individuals in both training conditions to focus more closely on their individual performance despite the presence of more complete shared mental models of task priorities. That is, there just was not enough time to effectively coordinate team actions. The increase in subjective workload trends observed in Figure 6 suggests that this indeed may be so. Although both suggestions are a subject for future research, the latter suggestion is especially relevant for those concerned about improving ATC coordination: *while learning or using newly acquired skills, as individual taskload increases, various forms of ATC team coordination may decrease.*

## REFERENCES


Bailey, L., Broach, D., Thompson, R., & Enos, R. (1999). *Controller teamwork evaluation and assessment methodology (CTEAM): A scenario calibration study.* (DOT/FAA/AM-99/24). Washington, DC: Department of Transportation/Federal Aviation Administration.

Cannonon-Bowers, J., Salas, E., & Converse (1993). Shared mental models in expert decision making teams. In N. Castellan, Jr. (Ed.), *Current issues in individual and group decision making* (pp.221 – 246). Hillsdale, NJ: Lawrence Erlbaum Associates.

Federal Aviation Administration (1998a). *Crew resource management training* (Advisory Circular 120-51A). Washington, DC: author.

Federal Aviation Administration (1998b). *Air traffic services concept of operations for the national airspace system in 2005* (draft report). Washington DC: author.

Feldman, R. (1968). Interrelationships among three bases of group integration. *Sociometry, 31*, 30-46.

Hart, S. G. & Staveland, L. E. (1988). Development of a NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Handcock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-83). Amsterdam: North-Holland.

Hartel, C., & Hartel, G (1995). *Controller resource management - What can we learn from aircrews* (DOT/FAA/AM-95/21). ). Washington, DC: Department of Transportation/Federal Administration (ADA297386).

James, L., James, L., & Ashe, D. (1990). The meaning of organizations: The role of cognition and values. In B. Schneider (Ed.), *Organizational Climate and Culture* (pp. 40-84). San Francisco: Jossey-Bass.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20*, 403-37.

Kraiger, K., & Wenzel, L. (1997). Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness. In M. Brannick, E. Salas, & C. Prince (Eds.). *Team performance assessment and measurement* (pp. 63 – 84). Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, B., Jr., Glickman, A., Woodard, E., Blaiwes, A., & Salas, E. (1986). *Measurement of team behaviors in a navy environment* (NTSC TR-86-014), Orlando, FL: Naval Training Systems Center.

Prince, C., Childester, T., Bowers, C., & Cannon-Bowers, J. (1992). Aircrew coordination - Achieving teamwork in the cockpit. In R. Swezey & E. Salas (Eds.), *Teams: Their training performance* (pp. 329-54). Norwood, NJ: Ablex.

Salas, E., Dickinson, T., Converse, S., & Tannenbaum, S. (1992). Toward and understanding of team performance. In R. Swezey & E. Salas (Eds.), *Teams: Their training performance* (pp. 3-29). Norwood, NJ: Ablex.