HIGH FIDELITY SIMULATION TEST OF NEW AIR TRAFFIC CONTROL CONCEPTS

Todd R. Truitt, Ph.D. & D. Michael McAnulty, Ph.D.
FAA William J. Hughes Technical Center
Atlantic City International Airport
New Jersey 08405

Researchers at the William J. Hughes Technical Center examined an Air Traffic Control (ATC) concept that collocates Terminal Radar Approach Control (TRACON) and Air Route Traffic Control Center (ARTCC) operations while expanding the terminal airspace. The concept may provide benefits for intersector coordination, traffic sequencing and spacing, holding, and overall traffic flow. By simulating arrival flows through two en route and two terminal sectors simultaneously, the researchers tested the effects of collocation and an expanded terminal airspace on system efficiency and Certified Professional Controller (CPC) performance, workload, and communication behavior.

Introduction

As the number of air carrier flights continues to climb, there is an ever increasing effect on the efficiency of the National Airspace System (NAS). According to the Federal Aviation Administration (FAA) Research and Development Strategy (2002), the FAA expects that 50% more airline passengers will travel in the year 2013 than in 2001. This predicted movement of 1.1 billion passengers a year requires a safe and highly efficient ATC system. To meet the challenge of a growing air transport industry, the FAA must reduce delays and improve overall system efficiency. Changes in ATC procedures and airspace redesign have not kept pace with the concomitant increase in air traffic volume and complexity. Furthermore, there are no proposals to construct new runways at these airports for the next 10 to 15 years.

We conducted a high fidelity, human-in-the-loop ATC simulation to scientifically compare current procedures with two alternative procedures for directing arrivals into a major New York airport. The experiments examined system performance and various aspects of CPC behavior and performance that occurred during each of three conditions.

Method

Participants

Eighteen CPCs participated in the experiment. Nine were from the New York TRACON (N90), and nine were from the New York ARTCC (ZNY). Two N90 sectors, Yardley/Penns (ARD) and Newark (EWR), and two ZNY sectors, Broadway (74) and Milton (75), comprised the simulated domain. Both ARD and 75 included handoff positions (ARD H and 75H) in addition to the radar positions. We also simulated the adjacent ARTCC facilities of Washington (ZDC) and Cleveland (ZOB).

Research Personnel

One experimenter oversaw the experiment and ensured proper data collection. A second experimenter observed and recorded communication behavior. Four Subject Matter Experts (SMEs) each observed one of the four simulated sectors and made over-the-shoulder ratings. Approximately 15 simulation pilots imitated the adjacent ARTCCs and all pilot activity.

Scenarios

The SMEs constructed nine scenarios based on actual traffic that occurred in the simulated sectors during the year 2001. They created the scenarios so that they were 30% busier than what the CPCs would normally experience at these sectors. We increased the taskload to test the new concepts under likely future traffic loads and to provide the opportunity to offset these high traffic loads with the new procedures. Each scenario was 50 minutes in length and contained a consistent level of traffic from beginning to end.

Equipment

We conducted the experiment at the FAA Research Development and Human Factors Laboratory using Standard Terminal Automation Replacement System and Display System Replacement displays and interfaces. We simulated the radar displays and their functionality with the Distributed Environment for Simulation, Rapid Engineering, and Experimentation ATC simulator. We simulated all aircraft using the New Generation Target Generator Facility. This hardware and software combined to present a high fidelity ATC simulation. Researchers developed these simulation tools at the FAA William J. Hughes Technical Center.

Design and Procedure

The participants controlled traffic at each of three respective terminal or en route positions under three conditions. In the Normal condition, a wall separated the terminal and en route sectors, and participants controlled traffic as they normally would while performing inter facility coordination via landline. In the Collocated condition, we removed the wall and permitted face-to-face (FTF) coordination. In the Terminalized condition, the wall remained down, and the en route sector 74 that fed the terminal sector ARD used terminal separation rules (3 nm rather than 5 nm) to effectively expand the terminal airspace. We counterbalanced the order of participant rotation, conditions, and scenarios appropriately.

After signing an informed consent statement, the participants completed a biographical questionnaire. The experimenters read instructions aloud prior to each scenario. The participants then received a position relief briefing from the SME assigned to their sector. During each scenario, the participants provided on-line ratings of workload every 5 minutes using the Workload Assessment Keypad (WAK). The WAK is based on the Air Traffic Workload Input Technique (Stein, 1985). An SME observed each sector and made over-the-shoulder performance ratings during each scenario by using either the terminal or en route Observer Rating Form (ORF; Sollenberger, Stein, and Gromelski, 1997). We also recorded communication activity by counting the number and duration of ground-ground and ground-air transmissions. During the Collocated and Terminalized conditions, an experimenter recorded FTF communication behaviors using taxonomy based on the Controller-to-Controller Communication and Coordination Taxonomy (C^4T; Peterson, Bailey, & Willems, 2001). After each scenario, the participants made subjective ratings of workload using a modified version of the NASA Task Load Index (TLX; Hart & Staveland, 1988) and responded to a Post-Scenario Questionnaire (PSQ) that asked them about the scenario they had just completed. The participants took a minimum 15-20 minute break between each scenario. After the participants had controlled traffic in each sector position and condition combination, they completed the Post-Experimental Questionnaire (PEQ) to provide their observations about the experiment.

Results

We analyzed the data from the terminal and en route participants separately and report those data accordingly. We used a repeated measures analysis of variance (ANOVA) for each dataset while making Geisser-Greenhouse and Huynh-Feldt adjustments to address the circularity assumption. We explain any significant main effects ($p \leq .05$) using Tukey's Honestly Significant Difference (HSD) test. We present statistics only for significant post hoc tests. In all graphs, the Normal, Collocated, and Terminalized conditions appear from left to right.

System Performance Measures

During each scenario, we recorded the number of aircraft handled by each sector, distance flown by each aircraft, number of landings, number and duration of holds, and number of altitude, heading, and speed changes.

In comparison to the Normal condition, the Terminalized condition showed an increase in the number of aircraft handled, a decrease in the distance flown per aircraft, an increase in the number of landings, a decrease in the number of holds, and a decrease in the duration of holds (see Table 1). The number of control commands (altitude, heading, speed) remained constant across conditions.

Table 1. Means and Standard Deviations for System Performance Measures by Condition

	Normal	Collocated	Terminalized
Number of a/c handled	52.3 (10.8)	53.9 (11.0)	54.1 (12.1)
Distance flown per a/c (miles)	40.9 (15.1)	40.9 (14.3)	40.7 (14.9)
Number of A/c landed	29.9 (6.0)	29.4 (4.9)	32.0 (6.6)
Number of holds	32.4 (14.9)	36.9 (12.1)	28.9 (15.0)
Duration of holds (minutes)	11.6	9.6	9.0

These differences were not statistically significant, but we consider them operationally interesting. They will sum over time and result in increased efficiency for the NAS and savings for the airlines.

Communications

The ground-ground transmissions for the en route Push to Talk (PTT) data include landline transmissions made from sector position 74 to ARD and transmissions made from sector position 75H to ZOB. For the number of en route ground-ground transmissions, there was a significant main effect of Condition, $F(2, 16) = 13.09$. Averaging across both sectors, the number of ground-ground transmissions was highest in the Normal condition and decreased significantly for the Collocated and Terminalized conditions, $HSD(16) = 9.87$ (see Figure 1).

Although not statistically significant, the apparent decrease in ground-ground transmissions for sector position 74 was not surprising because the wall that separated sectors 74 and ARD was not present during these conditions thereby allowing FTF communication. The data trend suggesting a reduced number of transmissions between sector position 75H and ZOB may have resulted from the decreased number of holds that participants performed during the Collocated and Terminalized conditions.

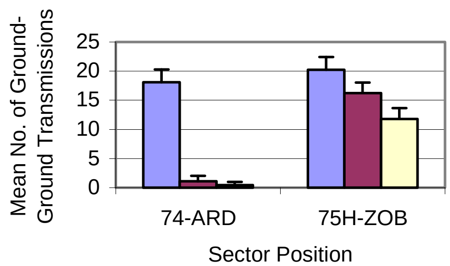

Figure 1. Mean Number of En Route Ground-Ground Transmissions by Sector Position and Condition.

We analyzed the mean duration of en route ground-ground transmissions and found a significant Sector Position by Condition interaction, $F(2, 16) = 4.86$ (see Figure 2).

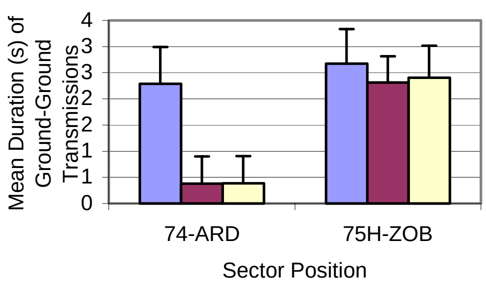

Figure 2. Mean Duration (s) of En Route Ground-Ground Transmissions by Sector Position and Condition.

The post hoc test indicated that the mean duration of transmissions was significantly lower at sector position 74 compared to sector position 75H during the Collocated and Terminalized conditions, $HSD(16) = 1.35$. The significant Sector Position X Condition interaction and post hoc test suggest that Condition affected each en route sector position differently. During these conditions, sectors 74 and ARD were collocated and could use FTF communication. On the other hand, sector position 75H still had to communicate with ZOB via landline.

The en route ground-air transmissions include transmissions made from sector positions 74 and 75 to a pilot. There was a significant main effect of Sector Position, $F(1, 8) = 22.40$. The participants at sector position 75 made more transmissions to pilots than did the participants at sector position 74 (see Figure 3). We did not find any significant differences in the mean duration of en route air-ground transmissions. On average, each transmission made to a pilot took about 3.5 s.

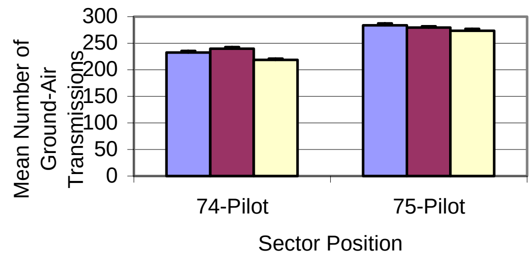

Figure 3. Mean Number of En Route Ground-Air Transmissions by Sector Position and Condition.

The ground-ground transmissions for the terminal PTT data include transmissions made from sector position ARD to 74 and transmissions made from sector position ARD H to ZDC. There were no significant differences. Although not statistically significant, the data trend suggests that the participants at ARD reduced their communication with sector position 74 in the Collocated and Terminalized conditions, but the frequency of their communication with ZDC increased (see Figure 4).

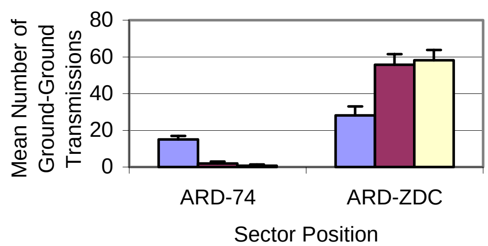

Figure 4. Mean Number of Terminal Ground-Ground Transmissions by Sector Position and Condition.

For the mean duration of terminal ground-ground transmissions, there was a significant Sector Position X Condition interaction, $F(2, 16) = 5.69$. The participants reduced the mean duration of their communications between sector position ARD and 74 in the Terminalized condition as compared to the Normal condition, $HSD(16) = 1.92$ (see Figure 5). The duration of communications between sector position ARD and ZDC appeared to increase but this difference was not significant.

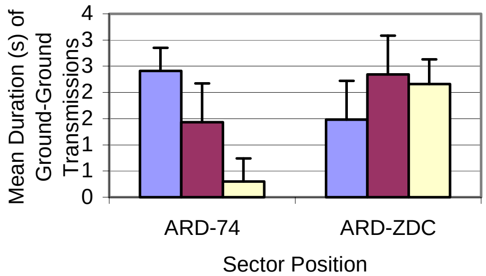

Figure 5. Mean Duration (s) of Terminal Ground-Ground Transmissions by Sector Position and Condition.

The terminal ground-air transmissions include radio transmissions made from EWR and ARD to a pilot. For the mean number of terminal air-ground transmissions, there was a significant main effect of Sector Position, $F(1, 8) = 147.94$. The participants at EWR made more ground-air radio transmissions than did those at ARD. The mean number of ground-air transmissions appeared to be relatively stable across Condition. However, the participants at EWR may have realized a small benefit during the Terminalized condition because they made 34 fewer ground-air transmissions on average as compared to the Normal condition (see Figure 6).

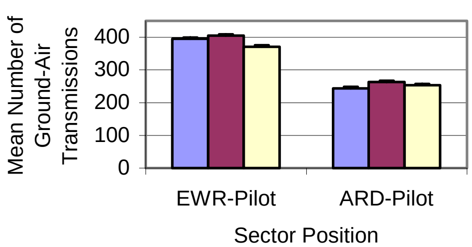

Figure 6. Mean Number of Terminal Ground-Air Transmissions by Sector Position and Condition.

There were no significant effects for the mean duration of terminal air-ground transmissions. The participants' ground-air transmissions took about 3 s on average and remained relatively stable across Sector Position and Condition.

The participants may have compensated for the reduction in the number of landline communications during the Collocated and Terminalized conditions, especially those between ARD and 74, by engaging in more FTF communication. FTF communication between the terminal and en route sectors was possible only in the Collocated and Terminalized conditions because the experimenters removed the wall separating terminal and en route sector positions during these two conditions.

The participants took advantage of the opportunity for FTF communication. In addition to exchanging information, they also acquired information by glancing at one another's radar display. Table 2 shows the mean number of communication behaviors by Condition and Type. We recorded a Glance whenever a participant at sector position ARD looked over at sector 74's radar display or whenever the participant at sector 74 looked over at the ARD display. Verbal communications included any type of ATC-related communication. Nonverbal communications included gestures such as pointing, giving a "thumbs up" or nodding one's head in acknowledgment.

Table 2. Mean Number of Communication Behaviors by Condition and Type

	Glance	Verbal	Non-verbal
Collocated	8.81	12.93	1.07
Terminalized	10.52	13.52	1.22

Overall, the participants took advantage of their collocated situation during the Collocated and Terminalized conditions by engaging in FTF communication instead of making ground-ground landline transmissions to one another. The participants also glanced at each other's radar display to gain information.

Observer Rating Form

The data from the ORF indicated that all of the participants performed well throughout the experiment. The en route participants received the highest mean ratings in the Terminalized condition for 12 different items of the ORF. The SMEs rated them as performing best in the Terminalized condition in terms of taking actions in the appropriate order of importance, preplanning for control actions, marking flight progress strips while performing other actions, prioritizing overall, providing control information overall, showing knowledge of procedures, showing knowledge of aircraft capabilities and limitations, overall technical knowledge, using proper phraseology, communicating clearly and efficiently, listening to pilot read backs and requests, and communicating overall. The ORF did not prove to be as sensitive a tool for measuring the performance of the terminal participants. Although the SMEs' subjective ratings of the terminal participants' performance did not indicate any significant differences, they did not identify any areas of concern, either. The terminal participants' performance was high in all conditions and at all positions.

Workload Assessment Keypad

For the en route participants, the subjective ratings of workload were moderate and did not change across conditions. For the terminal participants, WAK ratings were low to moderate and tended to be lower in the Normal condition and higher in the Terminalized condition. The difference between WAK ratings across conditions was not statistically significant. Figures 7 and 8 show the respective data.

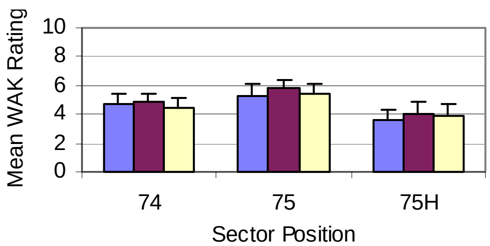

Figure 7. Mean En Route WAK Ratings by Sector Position and Condition.

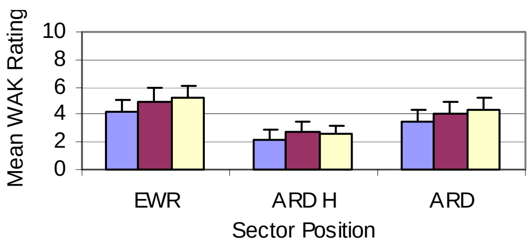

Figure 8. Mean Terminal WAK Ratings by Sector Position and Condition.

NASA-TLX

We analyzed each item of the NASA-TLX separately. The items asked participants about their mental demand, physical demand, temporal demand, effort, frustration, and performance.

For the en route participants, there was a significant main effect of Sector Position, $F(2, 16) = 7.01$, on mental demand. The data trend suggests that the participants perceived mental demand as being lowest at sector position 75H and somewhat higher at sector position 75 (see Table 3).

There was also a significant main effect of Sector Position, $F(2, 16) = 5.29$, on physical demand. It appears that the participants rated sector position 75H as being relatively lower than either 74 or 75. We did not find significant effects of either Sector Position or Condition for the items regarding temporal demand, effort, frustration, or performance for en route participants.

Table 3. Means and Standard Deviations for En Route NASA-TLX Ratings

		Condition		
Item	**Sector**	Normal	Collocated	Term.
Mental	74	5.89 (0.78)	6.56 (0.67)	6.56 (0.71)
	75	7.00 (0.90)	7.22 (0.76)	7.33 (0.78)
	75H	5.11 (0.93)	5.44 (0.90)	5.00 (0.86)
Physical	74	5.78 (0.80)	6.33 (0.83)	6.56 (0.75)
	75	6.33 (0.92)	5.89 (0.91)	6.78 (0.91)
	75H	4.00 (0.86)	5.11 (0.97)	5.00 (0.87)
Temporal	74	5.33 (0.85)	6.11 (0.75)	6.78 (0.80)
	75	6.56 (0.92)	5.89 (0.78)	7.00 (0.91)
	75H	5.22 (0.99)	5.33 (0.96)	5.44 (0.91)
Effort	74	7.44 (0.64)	6.78 (0.57)	7.33 (0.69)
	75	6.44 (0.80)	7.22 (0.68)	7.11 (0.85)
	75H	5.78 (0.90)	6.44 (0.96)	6.22 (0.91)
Frustration	74	3.89(0.93)	3.11 (0.77)	2.56 (0.73)
	75	4.22 (1.00)	4.56 (0.98)	3.67 (0.84)
	75H	3.67 (0.76)	3.78 (1.01)	2.33 (0.76)
Perform.	74	8.00 (0.54)	8.44 (0.64)	8.89 (0.65)
	75	7.78 (0.66)	7.78 (0.76)	8.00 (0.74)
	75H	8.00 (0.74)	7.89 (0.91)	8.67 (0.66)

For the terminal participants, the analyses identified significant differences in all of the items except performance. There was a significant main effect of Sector Position for mental demand, $F(2, 16) = 10.44$. Mental demand was significantly lower for ARD H than for both ARD and EWR, $HSD(16) = 2.54$. Although not statistically significant, the data trend suggests that mental demand may have increased slightly for the Collocated and Terminalized conditions as compared to the Normal condition (see Table 4).

Table 4. Means and Standard Deviations for Terminal NASA-TLX Ratings

		Condition		
Item	**Sector**	Normal	Collocated	Term.
Mental	EWR	4.89 (0.86)	6.44 (1.08)	6.33 (0.88)
	ARD H	2.89 (0.85)	3.56 (0.89)	3.44 (0.95)
	ARD	4.22 (0.72)	5.00 (0.95)	5.11 (0.73)
Physical	EWR	3.44 (0.71)	5.33 (1.08)	5.00 (1.03)
	ARD H	1.89 (0.78)	2.78 (0.92)	2.78 (0.94)
	ARD	3.11 (0.70	4.44 (0.88)	4.11 (0.72)
Temporal	EWR	4.56 (0.79)	6.33 (1.05)	6.22 (0.83)
	ARD H	3.00 (0.79)	3.11 (0.94)	3.22 (0.97)
	ARD	3.33 (0.80)	4.78 (0.88)	5.11 (0.68)
Effort	EWR	6.00 (0.76)	7.11 (0.94)	6.78 (0.86)
	ARD H	3.22 (0.84)	5.44 (0.89)	5.00 (1.01)
	ARD	5.44 (0.87)	6.00 (0.89)	6.67 (0.91)
Frustration	EWR	4.56 (0.89)	5.78 (1.07)	4.78 (1.01)
	ARD H	3.11 (1.03)	3.00 (0.97)	3.22 (1.03)
	ARD	4.67 (0.90)	4.89 (0.88)	5.11 (0.81)
Perf.	EWR	7.33 (0.76)	8.11 (0.82)	8.11 (0.80)
	ARD H	7.44 (0.98)	6.33 (0.80)	7.56 (0.78)
	ARD	7.56 (0.82)	8.22 (0.91)	8.33 (0.98)

There was a significant main effect of Sector Position, $F(2, 16) = 10.02$, and Condition, $F(2, 16) = 5.52$, for physical demand. The post hoc test for Sector Position was marginal, suggesting that physical demand may have been lower at the ARD H sector position than at both ARD and EWR. The post hoc test for Condition was not significant. The data trend suggests that physical demand may have increased slightly for the Collocated and Terminalized conditions as compared to the Normal condition.

There was a significant main effect of both Sector Position, $F(2, 16) = 13.03$, and Condition, $F(2, 16) = 6.48$, for temporal demand. The post hoc test for Sector Position was significant, $HSD(16) = 2.27$, and indicated that temporal demand was lower at the ARD H sector position than at either ARD or EWR. The post hoc test for Condition was not significant. However, the data trend suggests that, at least for the sector positions EWR and ARD, temporal demand may have been higher during the Collocated and Terminalized conditions as compared to the Normal condition.

We also found significant main effects of Sector Position, $F(2, 16) = 5.34$, and Condition, $F(2, 16) = 5.70$, for ratings of effort. The data trend suggests that participants may have perceived the Collocated and Terminalized conditions as requiring more effort than the Normal condition.

There was a significant main effect of Sector Position, $F(2, 16) = 7.23$ for frustration. The data trend suggests that the participants perceived frustration to be lowest at the ARD H sector position.

There were no significant differences in the participants' perception of their performance across either Sector Position or Condition as they rated their performance to be moderately high throughout the experiment.

Post-Scenario Questionnaire

The participants indicated on the PSQ that the reduced lateral separation standards used in the Terminalized condition had a generally positive effect on their ability to control traffic. They were able to adapt to the new procedures and use them effectively despite the fact that they had never used them before. The data from the PSQ showed that the en route participants did not believe that the Collocated or Terminalized conditions affected them negatively. Their overall subjective ratings of their situation awareness were high. The data from the PSQ also showed that terminal participants thought they were best able to move aircraft during the Terminalized condition.

Post-Experimental Questionnaire

The participants reported on the PEQ that both the Collocated and Terminalized conditions affected their communication and control strategies positively. They rated the realism of the simulation as moderate to high and indicated that interference from the WAK device was negligible.

Conclusion

The results indicated that the new concepts, as simulated in the experiment, are feasible. The participants' performance was high in all conditions, and we observed some improvements in the Terminalized condition. The participants used landline communication less in the Collocated and Terminalized conditions but compensated with FTF and non-verbal communication. The terminal participants indicated a slight increase in subjective workload ratings during the Collocated and Terminalized conditions, but overall workload was still at a moderate level. Both participants and expert observers provided positive feedback regarding the perceived benefits provided by the Collocated and Terminalized conditions.

References

Federal Aviation Administration. (2002). *Research and development strategy*. Washington, DC: Office of Aviation Research.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Peterson, L. M., Bailey, L. L., & Willems, B. F. (2001). *Controller-to-controller communication and coordination taxonomy* (C^4T). (DOT/FAA/AM-01/19). Washington, DC: Office of Aerospace Medicine.

Sollenberger, R., Stein, S., & Gromelski, S. (1997). *The development and evaluation of a behaviorally based rating form for assessing air traffic controller performance.* (DOT/FAA/CT-TN 96/16). Atlantic City International Airport, NJ: William J. Hughes Technical Center.

Stein, E. S. (1985). *Air traffic controller workload: An examination of workload probe*. (DOT/FAA/CT-TN 84/24). Atlantic City International Airport, NJ: FAA Technical Center.