FAA-AM-80-15

THE FEDERAL AVIATION ADMINISTRATION"S RADAR TRAINING FACILITY AND EMPLOYEE SELECTION AND TRAINING

James O. Boone
Linda Van Buskirk
JoAnn Steen

Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma

September 1980

Document is available to the public through the
National Technical Information Service,
Springfield, Virginia 22161

Prepared For
U. S. DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
Office of Aviation Medicine
Washington, D. C. 20591

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its content or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-AM-80-15		

4. Title and Subtitle	5. Report Date
THE FEDERAL AVIATION ADMINISTRATION'S RADAR TRAINING FACILITY AND EMPLOYEE SELECTION AND TRAINING	SEPTEMBER 1980
	6. Performing Organization Code
7. Author(s) JAMES O. BOONE, LINDA VAN BUSKIRK, AND JOANN STEEN	8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, Oklahoma 73125	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, SW. Washington, D.C. 20591	13. Type of Report and Period Covered
	14. Sponsoring Agency Code

15. Supplementary Notes

Work was performed under tasks AM-C-79-PSY-70 and AM-C-80-PSY-82.

16. Abstract

The Federal Aviation Administration (FAA) has recently constructed a Radar Training Facility (RTF) in Oklahoma City, Oklahoma, to aid in screening appropriate personnel for work in radar air traffic control (ATC). The approach is based on the idea that limited exposure to simulated radar ATC in a controlled and measured environment will lead to the identification of persons who possess the skills and attributes necessary for success in this type of work. This report describes the results of a study performed at the FAA's National Aviation Facilities Experimental Center (now FAA Technical Center) comparing an over-the-shoulder method of scoring student performance with scoring by computer-derived measures for use in screening at the RTF. Results indicate that the computer-derived measures are far more reliable than over-the-shoulder scoring and the computer-derived measures predict a global rating of potential success in radar ATC at least as well as over-the-shoulder scoring. The implications of the results are discussed in relation to other automated training systems.

17. Key Words	18. Distribution Statement
Radar air traffic control Computer-derived performance measurement Simulation training, Measurement theory Selection and screening Objective performance measurement	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	25	

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

THE FEDERAL AVIATION ADMINISTRATION'S RADAR TRAINING FACILITY AND EMPLOYEE SELECTION AND TRAINING

Successful air traffic control specialists (ATCSs) who have made a transition from manual to automated air traffic control (ATC) appear to prefer the advantages in the automated environment. However, some prospective ATCSs do not perform successfully in radar ATC. Successful employment in the radar environment requires that a person possess certain aptitudes. It is in the interest of the Federal Aviation Administration (FAA) and the prospective ATCSs to determine as soon as possible if the prospective ATCS possesses the aptitude necessary to successfully operate in the radar ATC environment. The philosophy of the FAA in regard to this selection process is that the best way to measure aptitude is to place the prospective ATCS in a radar simulation laboratory and perform a systematic, objective appraisal of the person's potential. To this end the FAA has constructed a Radar Training Facility (RTF) at the FAA Academy in Oklahoma City, Oklahoma. The training/screening process involves a mini-radar training program with rigorous assessment which occurs over a 4- to 5-week period. During this period, the trainee receives basic radar training sufficient to allow systematic evaluation of his or her performance. Those who demonstrate potential to become successful ATCSs are retained and those who do not are screened from the program. To explain this system, the RTF background, RTF positions, system operation, and the evaluation process are described in detail below.

The original simulators used in FAA ATC training were "patches" developed for the operational automated field systems. The "patches" permitted flexible training at designated positions without interfering significantly with the operational positions. Experiences with these prototype simulators resulted in at least two major notions related to using simulation for radar training. First, the value of computer-driven simulation for training purposes was firmly established. Second, several problems associated with using operational field systems in a training mode were identified. A 1965 Institute for Defense Analysis (IDA) study on the training of air traffic controllers discussed some of these problems and suggested that a standardized computer-driven program should be established by the FAA to provide basic radar training. The IDA study further suggested that the radar training should be pass/fail to select out those persons who did not demonstrate the potential to perform proficiently in a radar environment.

In July 1976, engineering requirements were completed by the FAA for a radar training system. During that same month the FAA Administrator approved the procurement and construction of the RTF to be located at the FAA Academy in Oklahoma City.

In October 1977, the FAA completed a program implementation plan that outlined the development and implementation of the RTF. The contract for the development of the computer-driven simulator training system was awarded to Logicon, Tactical and Training System Division, San Diego, California, in January 1978. Groundbreaking for the construction of the new RTF at the FAA Academy was held on December 22, 1977. The new facility was built and accepted by the FAA in January 1980, and the training system developed by Logicon Corporation was accepted in April 1980.

RTF Training System and Laboratory Configuration.

The primary objective of the RTF is to closely duplicate the specialized operational environment existing at automated terminal and en route facilities as well as have the capability of synthesizing and presenting a wide variety of air traffic control situations. These situations would be based on a reference data base created through scenario programs with a full range of control necessary to establish a realistic simulation of actual aircraft traffic under a variety of conditions.

To accomplish this objective, four independent laboratories are utilized. Figure 1 describes how the laboratories are configured.

Positions. There are Trainee positions and Supervisory and Support positions/stations corresponding to each radar training sector. At a "position," the operating personnel have input/output (I/O) equipment access to the system with associated voice communications. A "station" has no I/O equipment access but is equipped with voice communications for monitoring, instructing, and supervisory functions.

Trainee Position.

1. Radar Control Position (R). The R controller positions (six in each lab) have a display console, (PVD) for en route and (DEDS) for terminal. They have associated voice communications. The displays and voice communications are similar to those at field facilities. Displays include maps, weather, aircraft position symbols, alphanumeric readouts, and other digital and symbolic data.

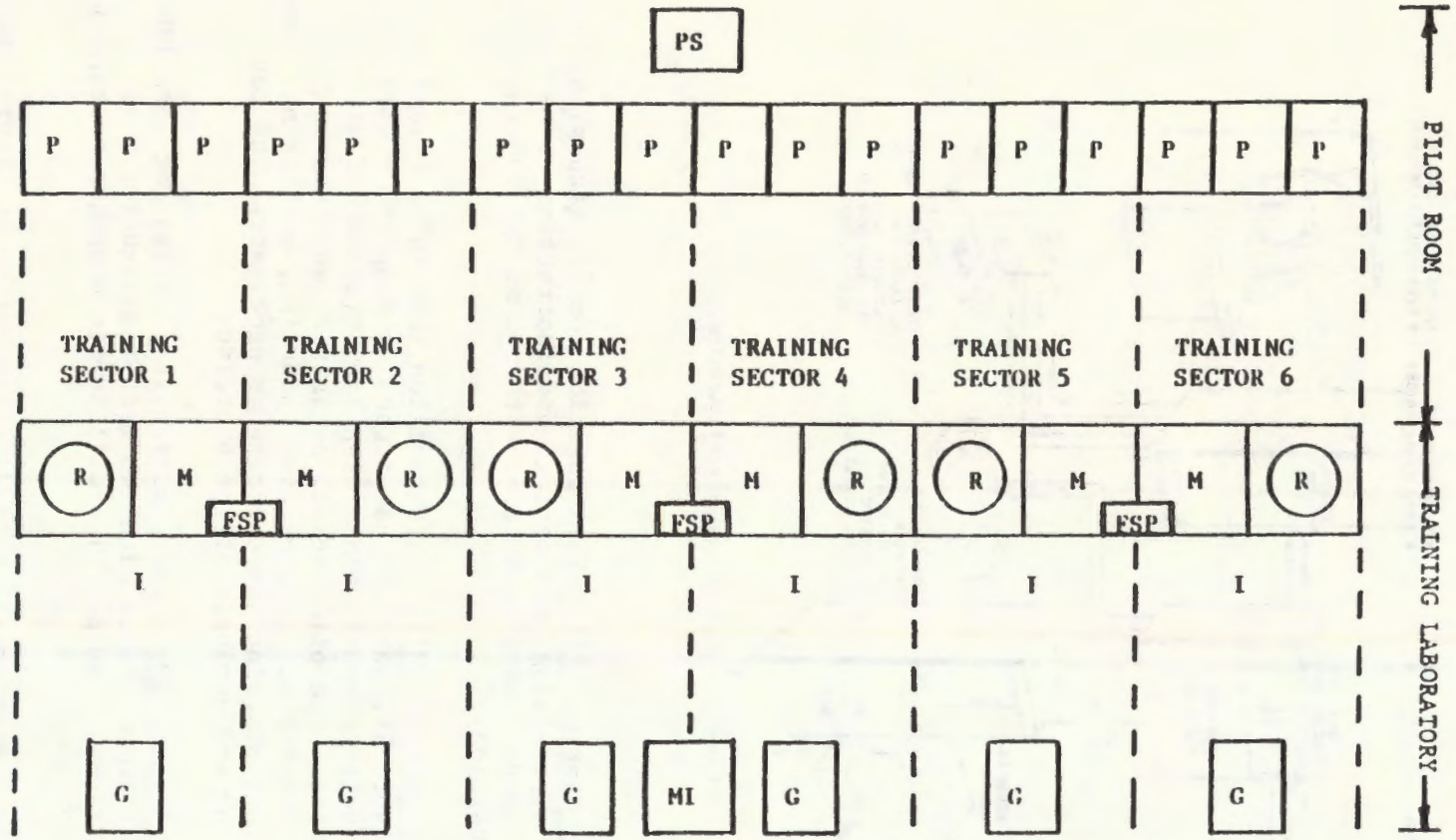
2. Nonradar Controller Position (HO/D). The D controller for en route and the HO position for terminal (six in each lab) have the capability of making and accepting handoffs. This position also permits training for manual or nonradar control by using flight progress strips generated by the flight strip printers.

3. Pilot Position (P). Three pilot positions are associated with each sector (18 in each lab). These positions are in a separate room. Each position operator performs at a console with a tabular display and keyboard for data entry with associated voice communications. These operators simulate aircraft pilots during the exercise by actual responses to ATC clearances/instructions.

4. Ghost Position (G). This position is associated with each R and/or HO/D position. There are six ghost positions in each lab. The position console and display are identical to those of the pilot position. The ghost position operator adds realism to the exercise by performing related functions of adjacent centers, terminals, flight service stations, and positions/sectors. Functions include initiating handoffs, accepting handoffs, and generally ghosting functions of other facilities/sectors.

Supervisory and Support Positions/Stations.

1. Instructor Station (I). An instructor station is provided at each sector (six in each lab). The instructor has voice communication with each student and monitors the overall exercise from behind the trainee positions.



P - Pilot Console
G - Ghost Console
R - Radar Trainee Console
M - Manual Controller Trainee Position
Enroute "D" Controller or
Terminal "HO" Controller

MI - Master Instructor Position
PS - Pilot Supervisor Station
I - Instructor Station
FSP Flight Strip Printer

Figure 1. Laboratory configuration.

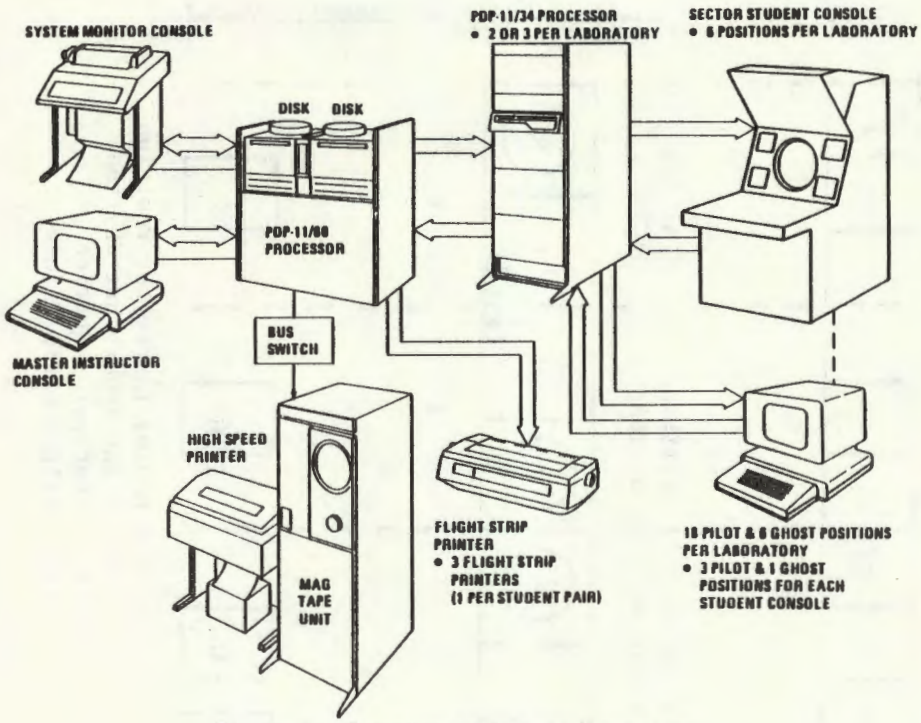


Figure 2. Computer system configuration.

2. Pilot Supervisory Station (PS). This position (one in each pilot room) has voice communications for supervising, monitoring, and instructing operation of pilot positions as well as for coordinating activities with the master instructor station and the system monitor position.

3. Master Instructor Station (MI). This position (one in each lab) controls the exercise within the lab. The position has a tabular display, a data entry keyboard, and associated voice communications with each trainee and with each operator of ghost, instructor, and pilot positions in the lab. The master instructor station will permit setting clock time, starting, monitoring, freezing, backing up, replaying, and restarting the exercise as necessary. The position also provides for data recording and analysis of the exercise.

4. System Monitor Position (SM). One position is provided for each lab. The position will have voice communications with two master instructor positions and two pilot supervisor positions. The position will permit computer operation and operational and maintenance monitoring.

Figure 2 describes the system configuration for operating the positions and stations in each laboratory. The training sectors are controlled by a Digital Equipment Corporation (DEC) PDP 11/60 computer with a PDP 11/34 computer serving as an interface between the PDP 11/60 and the operating positions.

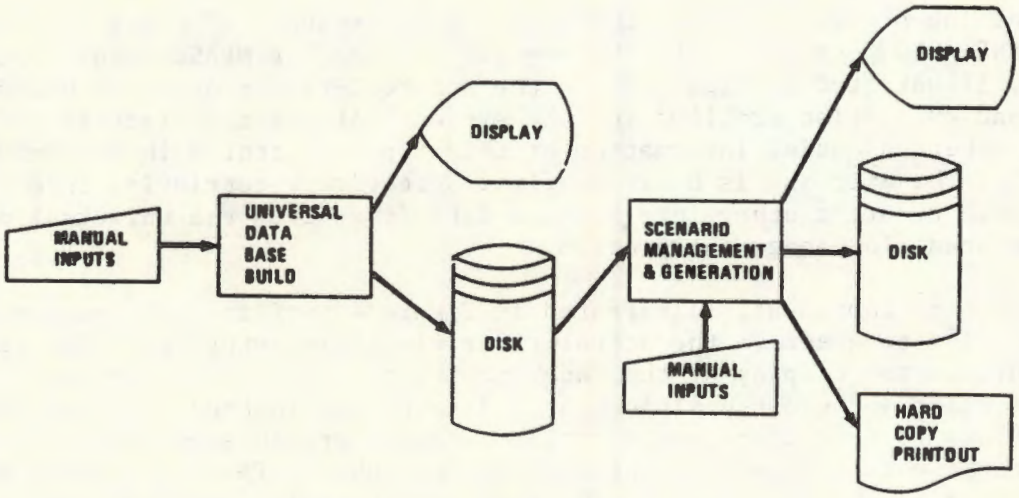


Figure 3. Components of scenario generation.

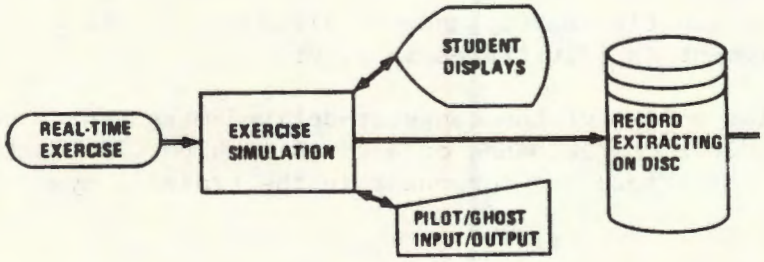


Figure 4. Components of the real-time training system.

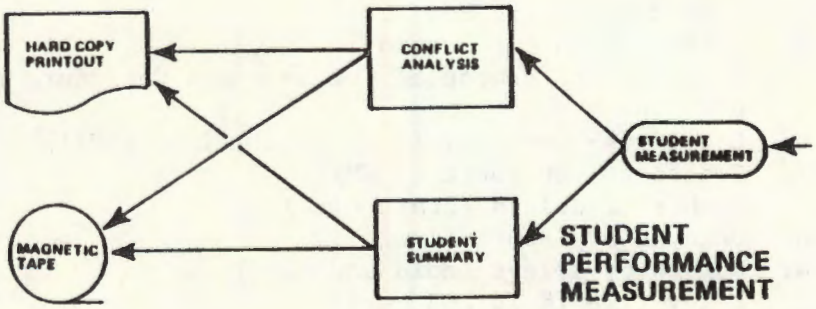


Figure 5. Components of the student performance measurement.

The training process involves three sequential systems of operation: (1) SCENARIO GENERATION --→ (2) REAL-TIME --→ (3) PERFORMANCE MEASUREMENT. Scenario generation, illustrated in Figure 3, is the non-real-time process of building exercises and evaluation problems for the system. Aircraft characteristics, flight plans, and other essential information of this type are stored in the Universal Data Files (UDF). The exercise is built by first selectively retrieving intermediate files and then creating other intermediate data files from the universal data base through the scenario management program.

The real-time component, illustrated in Figure 4, utilizes the scenario management files to generate the actual radar simulation exercise. The real-time component drives the display at the radar position. Aircraft movement is controlled through the pilot and ghost positions according to the instructions the operators of those positions receive from the controller trainee or, in some cases, from a scenario prompt which appears on the cathode-ray tube (CRT) at the pilot or ghost positions. During the operation of the real-time training exercise, all actions taken during the exercise are recorded.

At completion of the exercise, the computer will analyze the recorded actions to determine violations of separation standards and to quantify other pertinent performance information, such as delay times, in order to evaluate the student's ability to move air traffic "safely and expeditiously." The process of student performance measurement is illustrated in Figure 5.

Table 1 contains a list of the computer-derived measures to be employed in evaluating the students' performance on a given problem. The primary focus of this paper is the student performance component in the training system.

TABLE 1. List of RTF Measures

1. Number of aircraft in the sample
2. Ideal aircraft time-in-system (based on filed flight plan)
3. Ratio of the ideal aircraft time in system and the number of aircraft in the sample
4. Number of completable flights
5. Data period duration
6. Number of arrivals
7. Number of departures
8. Arrival/departure ratio
9. Arrival rate scheduled per hour and departure rate scheduled per hour
10. Conflicts--terminal (3 nautical miles (NMI))
11. Conflicts--en route (5 NMI)
12. Number of delays (start time)
13. Delay time (start time)
14. Number of delays (hold and turn)
15. Delay time (hold and turn)
16. Number of delays (arrival)
17. Delay time (arrival)
18. Number of delays (departure)

(CONTINUED ON NEXT PAGE)--

TABLE 1 (continued)--

19. Delay time (departure)
20. Number of delays (total)
21. Delay time (total)
22. Aircraft time-in-system (real)
23. Number of aircraft handled
24. Number of completed flights (total)
25. Number of arrivals achieved
26. Arrival rate achieved per hour
27. Number of departures achieved
28. Departure rate achieved per hour
29. Number of air-ground contacts
30. Air-ground communications time
31. Number of altitude changes
32. Number of heading changes
33. Number of speed changes
34. Number of path changes (altitude, heading, and speed)
35. Number of handoffs

Background in Performance Measurement.

The earliest studies in air traffic control which used some form of automated measurement were conducted by a Civil Aeronautics Administration (CAA) group in Indianapolis, Indiana, at the Technical Development Center (TDC) with support from the Franklin Institute of Philadelphia (18,28, 51, 65). The "dynamic simulator" used at the TDC consisted of a translucent screen on which maps could be projected with motor-driven light projectors capable of projecting a spot of light and moving it across the screen to simulate radar echoes from an aircraft. Personnel acted as pilots by moving the aircraft across the screens according to the control messages they received over a telephone line. The setup resembled the radar Plan Position Indicators (PPI) used in air traffic control (15). Research at this facility spanned from 1950 to 1959, at which point it was moved to the National Aviation Facilities Experimental Center (NAFEC; renamed FAA Technical Center in May 1980).

The research at TDC covered topics in air traffic control such as (i) airport design, (ii) approach systems, (iii) ATCS workload, (iv) data acquisition, and (v) decision making. Reports on the studies contained quantitative data on (i) number of separation violations, (ii) number of aircraft delayed, (iii) average delay per aircraft, (iv) altitude changes, (v) number and length of communications, and (vi) number of missed approaches (10,11,12,13,14,67,68,69).

Concurrent with the TDC studies, a series of 19 simulation-based experiments were conducted in air traffic control at the Ohio State University's Aviation Psychology Laboratory under the direction of Paul M. Fitts (27). The studies were performed between 1954 and 1961 and involved measurement of controller performance. In 1954 Hixson et al. (34) developed an electronic radar target simulator for air traffic control studies. As a part of the development, Hixson made performance measurements on the accuracy of "headings," "airspeed," "turn rate," and "attitude" for each target generated. A camera was mounted on the display indicator and the path of the aircraft was recorded. Calculations were then computed from the

recordings to measure the accuracy of the simulation. These measures were used to determine the accuracy of the system operation.

Later studies at Ohio State University involved more direct measurement of ATCS performance and were conducted on a variety of topics such as (i) pattern-feeder controllers, (ii) individual differences among subjects, (iii) display variables, (iv) workload variables, and (v) procedural variables (9,35,37,39,40,41,42,43,44,45, 46,47,48,49,52,57,58,59,60,61,66). Several different types of measures were used to assess subject and system performance. These include measures of (i) overall flight time, (ii) percent delay time, (iii) fuel consumed, (iv) missed approaches, (v) separation errors, (vi) time intervals between landings and departures, (vii) time and frequency of communications, (viii) delay time in responding to emergency situations, and (ix) traffic load, i.e., number of aircraft in the problem and number of aircraft handled.

During the 1960's research involving ATCS performance was done by at least three groups: (i) the MITRE Corporation, (ii) the Systems Development Corporation, and (iii) NAFEC. Between 1961 and 1963 the MITRE Corporation conducted six studies in air traffic control. The six studies covered topics in (i) high altitude air traffic control, (ii) beacons and automatic tracking, (iii) display clutter on the CRT, (iv) multisector coordination, (v) handoff procedures between en route and terminal, and (vi) conflict resolutions (33,36). The studies were performed by computer-generated simulation where "canned" scenarios were constructed and then run in real time. Automated measures taken in the studies included: (i) traffic load, (ii) teletype usage, (iii) frequency of various displays, (iv) flight plan deviations, and (v) conflicts.

In 1961 the System Development Corporation began a series of studies in air traffic control. The studies were performed by computer-generated simulation and sufficient information from each program to subsequently reproduce the problem was stored on mag tape. Studies were conducted on topics such as: (i) spacing of aircraft, (ii) geographic point of aircraft entry, (iii) heterogeneity of aircraft, and (iv) procedural variations (2,3,4,5,6,7,8,25,29,30,54,55,56). The stored data from the problems made possible an extensive list of postexercise measures. These included (i) safety violations, (ii) percent of time aircraft in holding pattern, (iii) percent of aircraft held, (iv) difference between actual flight time and time by the shortest available path, (v) the ratio of iv and v, (vi) mean time spacing between successive aircraft, (vii) aircraft waiting time before departure, (viii) delay time holding, (ix) fuel consumption, (x) variability in aircraft arrival time, (xi) number of radio communications, (xii) average communication time, (xiii) average number of communications per aircraft, (xiv) total communication time, (xv) number of controller data entries, and (xvi) number of clearance points an aircraft passed.

Perhaps the most extensive research during the 1960's involving performance measurement occurred at the FAA NAFEC facility. As previously mentioned, the simulator at the TDC was moved to NAFEC and used until about 1962. Between 1960 and 1962 NAFEC also had a Model A and Model B simulator installed. The simulators generated radar echoes on a CRT. Pilots were also employed to move the echoes around on the CRT. Later, a Sigma 5 computer was introduced which extended NAFEC's simulation capabilities. A sampling of the research topics covered included: (i) dual approaches, (ii) combining approach facilities, (iii) equipment arrangements,

(iv) traffic flow patterns, (v) final approach spacing, (vi) display usage, (vii) airspace jurisdiction, (viii) helicopter movement, (ix) supersonic control procedures, (x) airport site selection, and many more (1,20,26,32,38,50,53,64,65, 66,70,71). Measures employed in the studies consisted of (i) delay time, (ii) number of vectors, (iii) number of holds, (iv) conflictions, (v) aircraft time in the system, (vi) interval between arrivals, (vii) communication workload, (viii) number of departures and arrivals, (ix) the ratio of departures and arrivals, (x) missed approaches, (xi) total aircraft handled and several others.

During the latter part of the 1960's and into the 1970's, there was a shift in emphasis in performance measurement at NAFEC. While most of the prior NAFEC research had employed the measures noted above to evaluate various equipment, procedures, or configurations, research interests shifted to using automated performance measurement to evaluate how well the ATCS was performing. Thus, a 1969 report by Buckley et al. (19) applied the measures not to evaluate systems, but to assess the performance of individual ATCSs. That initial study examined individual differences in ATCS performance as a function of aging. It was followed up by two reports (21, 22) applying the same measurement approach but with the measures programed into a digital simulator. The latter two studies, termed PROBE tests, further supported the basic performance measurement rationale, and demonstrated the possibility both of developing parallel problems and of identifying a consistent ATCS profile across sections.

With the introduction of the RTF, it was decided that the feasibility of using computer-derived measures to evaluate ATCS student performance should be studied. The present system of student evaluation consists of an over-the-shoulder observation of students by expert air traffic controllers with recent field experience. Scores are comprised on the basis of a composite of instructor ratings (Instructor Assessment) and a count of errors committed while controlling simulated aircraft (Problem Average). A study was designed, employing the computer-driven ATCS simulation lab at NAFEC, to study the possibility of using automated measuring devices as a substitute for the Problem Average portion of the composite score. The purpose of the study was twofold: (i) To make a preliminary assessment of the feasibility of using computer-derived measures (CDM) to evaluate student laboratory performance and (ii) to improve the over-the-shoulder evaluation procedure for student laboratory evaluation.

Methods. To accomplish these goals, 48 students, 24 en route and 24 terminal, were transported to NAFEC to receive radar training and evaluation at the Dynamic Simulation Facility. The students were evaluated over-the-shoulder by an instructor and the problems were recorded by computer on mag tape and later reduced to a set of computer-derived measures (see Table 2 for a listing of the measures used).

Five problems in increasing complexity were administered to each student. Each instructor had an opportunity to observe each student at least once. On problems 4 and 5, randomly selected students were evaluated over-the-shoulder simultaneously and independently by two instructors. An index of agreement (reliability) was computed on the simultaneous evaluations by forming a ratio of the number of agreements over the total number of error conditions recorded by the two instructors. An initial laboratory evaluation manual and a laboratory evaluation form were developed by consensus of the instructors in each option prior to the study; the

TABLE 2. A Listing of the Computer-Derived Measures and Their Corresponding Reference Numbers Employed in the NAFEC Study

1. Conflicts (5-mile separation)
2. Conflicts (3-mile separation)
3. No. Start Point Delays
4. Start Point Delay Time
5. No. Turn and Hold Delays (turns longer than 100 seconds)
6. Turn and Hold Delay Time
7. Aircraft Time-in-System
8. No. Aircraft Handled
9. No. Completed Flights (transfers to 130.5 must be from ghost position)
10. No. En Route Departures (Code 2)
11. No. Terminal Arrivals (Code 3)
12. No. Terminal Departures (Code 4)
13. No. Air-to-Ground Contracts (subject only)
14. Air-to-Ground Communications Time
15. No. Altitude Changes (pilot keyboard messages)
16. No. Heading Changes (pilot keyboard messages)
17. No. Speed Changes (pilot keyboard messages)
18. No. of Handoffs From Feeder Position to Subject
19. Handoff Delay Time
20. No. Beacon Re-Idents

manual and lab form were modified during the study based on new agreements formed after reviewing the disagreements on the laboratory evaluation forms.

After each student was evaluated on the individual problems, each instructor provided an overall, global rating, stating the student's potential to become a full performance level (FPL) radar controller. The rating was a 5-point global scale, (1) definitely will not become FPL, (2) maybe (doubtful) FPL, (3) minimally acceptable FPL, (4) good FPL, and (5) definitely excellent FPL. The global rating was based on the instructor's observations of the student operating the radar problems.

Analyses included the following: To determine the feasibility of using computer-derived measures, those measures were used in a regression equation to predict the (i) problem average (PA), (ii) instructor assessment (IA), and (iii) total score on the over-the-shoulder evaluation. Further, the individual problem scores from the over-the-shoulder evaluation were used in a regression equation to predict the global rating score for each student. A regression analysis was performed using the CDM and IA regressed on the global rating to compare with the PA and IA on the global rating. The indices of agreement reliability for the simultaneous over-the-shoulder evaluations were also computed and listed by problem and option. A reliability index (intraclass correlation) was also performed on the global rating data. Profiles across students and across instructors were computed by stratifying the errors on the lab forms according to error categories identified by a group of controllers who reviewed the worksheets. The frequencies of the errors were then summarized by category (Table 3 contains a listing of the over-the-shoulder measures). An orthogonal, varimax factor analysis was also calculated to group the measures in multidimensional space and to compare the underlying dimensions of the error categories in the over-the-shoulder and computer-derived measures.

TABLE 3. A Listing of the Over-the-Shoulder Measures

and Their Corresponding Reference Numbers

I. SYSTEM ERRORS (S)

1 - Vertical
2 - Lateral, long
3 - Terrain
4 - Airspace outside radar coverage

II. SYSTEM DEVIATIONS (D)

1 - Airspace (lateral)
2 - Altitude (facility)
3 - Altitude (aircraft data block) (min. separation used, no alt. verification)

III. PROCEDURE (P)

1 - Keep them high
2 - Speed control
3 - Bad vector
4 - Delay
5 - L.O.A. (letter of agreement)
6 - Holding-EAC/EFC
7 - WAFDF (wrong altitude for direction of flight)
8 - Needless altitude change
9 - Radar contact not given to ACFT
10 - No reason for vector
11 - Traffic
12 - Position of ACFT. Incorrect or not given.
13 - SID (change in route)
14 - Missed approach instructions
15 - Remarks
16 - Improper coordination
17 - Beacon code
18 - Point out
19 - Route
20 - Altitude
21 - Transfer control
22 - Change of destination
23 - Change of ACFT. Status (VFR/IFR)
24 - Altitude verification
25 - Clearance

IV. OTHERS (O)

1 - Phraseology
2 - Strip marking
3 - Altimeter not issued
4 - Overrestriction
5 - Improper feedback of wrong information
6 - Data block update within sector
7 - Board management

The reliability coefficients for the over-the-shoulder problem averages were computed as previously described. The global rating and instructor assessment reliabilities are intraclass correlations across all instructors for each student. The reliability coefficients are important for several reasons. In the case of the over-the-shoulder evaluation, it indicates the proportion of times that two instructors agreed on a particular error marked against the student's grade. Disagreements occurred in two ways: The instructors recorded the same event as an error but differed in the type of error they called it, or one instructor recorded an error for an event while the other instructor either failed to see or did not judge it to be an error. It can be readily noted that the instructor assessment is more reliable than the problem average. The reliability of the problem average is important since the validity of a measure cannot exceed its reliability. Consequently, it is very important to standardize any portion of the grading procedures that requires instructor judgments.

TABLE 4. Reliability Coefficients for the Over-the-Shoulder

Evaluation and Q-Sort by Option

	Problem Average	Instructor Assessment	Total Score	Global Rating
Terminal	.326	.582	.433	.234
En route	.294	.561	.427	.266

Model 1 (Table 5) demonstrates the ability of the computer-derived measures to duplicate the problem average in the over-the-shoulder evaluation. The Beta weights indicate the relative importance of each of the computer measures in the duplication process. The "R," multiple correlation, ranges from -1.0 to +1.0 and is a measure of the overall fit of the model. A .5212 is a moderate to good value; however, the value could increase considerably if the unreliability in the problem average were minimized.

Models 2 and 3 (Tables 6 and 7, respectively) demonstrate how well the computer-derived measures duplicate the instructor assessment and the total score. The increase in "R" for instructor assessment is probably due to a better reliability in the instructor assessment.

Model 4 (Table 8) demonstrates in the Beta weights a tentative schema for weighting the lab problems to form a composite lab score. The information provided by the problems is highest in problems 4 and 5. The maximum amount of information peaks at problem 4. Thus, a five-problem lab grading procedure offers the most information, but a four-problem procedure would be an efficient manner of maximizing information in the shortest time frame. The relative weightings for five problems would be 10, 10, 15, 40, and 25, and for four problems would be 15, 15, 30, and 40.

Models 4 and 5 (Tables 8 and 9, respectively) demonstrate how well the computer-derived measures can be used in place of the problem average in predicting the global rating. The multiple "R" drops from .4493 to .4299, an insignificant decline. For practical purposes, the computer-derived measures can be used in place of the problem average in forming an overall grade. This approach would have at least one very strong advantage. The computer-derived measures are completely

TABLE 5. Regression of Computer-Derived Measures (CDM)

on the Over-the-Shoulder Problem Average (PA)

Model 1
Predictors = 1-20
R = 0.5212

V	BETA
1	0.1147
2	0.0365
3	0.2637
4	0.0128
5	0.1704
6	0.0126
7	0.0298
8	0.1649
9	0.1791
10	0.8536
11	0.0586
12	0.6821
13	0.2704
14	0.3552
15	0.2906
16	0.1167
17	0.0542
18	0.2582
19	0.1593
20	0.0507

TABLE 6. Regression of Computer-Derived Measures on the Over-the-Shoulder Instructor Assessment

Model 2
Predictors = 1-20
R = 0.5302

V	BETA
1	0.1547
2	0.0390
3	0.3446
4	0.0157
5	0.1669
6	0.0100
7	0.2337
8	0.1343
9	0.2099
10	0.8121
11	0.0000
12	0.7387
13	0.1292
14	0.1970
15	0.3665
16	0.2153
17	0.0169
18	0.1143
19	0.1602
20	0.0041

TABLE 7. Regression of Computer-Derived Measures on the Over-the-Shoulder Total Score

Model 3
Predictors = 1-20
R = 0.5247

V	BETA
1	0.1942
2	0.0387
3	0.3533
4	0.0107
5	0.1370
6	0.0546
7	0.3451
8	0.0039
9	0.2177
10	0.5529
11	0.0790
12	0.5079
13	0.1628
14	0.1861
15	0.3735
16	0.3786
17	0.0335
18	0.0842
19	0.1308
20	0.0808

TABLE 8. Regression of 5 (PA+IA) Probs on Global Rating

Model 4
Predictors = 1-5
R = 0.4493

V	BETA	B
Prob 1	0.0928	0.0062
Prob 2	0.0742	0.0043
Prob 3	0.1376	0.0096
Prob 4	0.3029	0.0147
Prob 5	0.1923	0.0090

REG. CONST. = 1.8253

TABLE 9. Regression of CDM+IA on Global Rating

Model 5
Predictors = 1-5
R = 0.4299

V	BETA	B
Prob 1	0.1851	0.0007
Prob 2	0.3511	0.0012
Prob 3	0.8663	0.0017
Prob 4	0.0515	0.0001
Prob 5	0.6531	0.0012

REG. CONST. = 6.2114

reliable whereas the problem average is considerably unreliable. Combining the highly reliable computer-derived measure with the moderately reliable instructor assessment creates a problem average reliability of approximately .750, which is a significant improvement over the previously reported .433 and .427 for terminal and en route, respectively.

The evidence from Models 1-5 suggests that the computer-derived measures are useful and valuable contributions to the assessment process. The validity of the measures is not established by this study; however, using the computer measures in place of the over-the-shoulder problem average increases the reliability significantly, and reliability is the upper bound for validity.

The factor analyses offer a means to (i) identify cluster areas where general measures are incurred by students and (ii) provide a comparative basis for the underlying structures of the two grading systems. The factor analyses point out at least two major differences in the two evaluation models: delays and system deviations. Delays are difficult to determine over-the-shoulder and no method was available to measure system deviations in the computer-derived measures. A next step would be to attempt an optimal combination of over-the-shoulder measures and computer-derived measures to be averaged with the instructor rating.

Conclusions.

It was concluded from the regression models that the computer-derived measures predict a global rating criterion of potential ATC on-the-job success at least as well as the over-the-shoulder evaluation (Models 1, 2, and 3). Further, it was found that the over-the-shoulder evaluations are not as reliable as the computer-derived measures (Table 3). Since reliability is in general the upper bound for validity, using computer-derived measures would enhance the probability for higher validity. The computer-derived measures, it appears, can be substituted for the over-the-shoulder ratings and used to form a composite laboratory score. Model 4 demonstrates that four or five problems should be employed in forming the laboratory composite with unit weights of (i) 15, 15, 30, and 40, or (ii) 10, 10, 15, 40, and 25, respectively. These results have implications for other training programs where expert observations and ratings are employed for scoring. This study indicates that computer scoring can provide a more objective and reliable measurement, and this increased reliability provides potential for enhancing a program's validity. Further research in this area should include extended analysis of factor structures of the two measurement techniques (Table 10) in an attempt to reach an optimal scoring schema using both computer scoring and expert observation. Future research should also include a long range validity study to determine which measurement technique is more valid in predicting on-the-job success.

TABLE 10. Factor Analyses of the Over-the-Shoulder and Computer-Derived Measures

Computer-Derived Measures

Measure	*Loading
FACTOR 1	(Conflicts)
1. Conflicts	.7843
FACTOR 2 (A/C Workload)	
1. Start Point Delays	.4214
2. No. A/C Handled	.8762
3. No. Completed Flights	.7304
4. No. Arrivals	.6209
5. No. of Departures	.5347
FACTOR 3 (Delays)	
1. Turn and Hold Delays	.4574
2. A/C Time in System	.6302
3. Hand-Off Delay	.4039
FACTOR 4 (Communications)	
1. Air-to-Ground Contacts	.8253
2. No. Beacon Re-Idents	.4928
3. No. Heading Changes	.4291
FACTOR 5 (A/C Direction Vectoring)	
1. No. Speed Changes	.7126
2. No. Altitude Changes	.5284

*Only loadings of .400 or better were retained.

Over-the-Shoulder

Measure	*Loading
FACTOR 1	(Conflicts)
1. System Error	.8188
2. System Error	.6441
3. System Error	.6088
4. System Error	.6836
FACTOR 2 (System Deviations)	
1. System Deviation	.5144
2. System Deviation	.4163
3. System Deviation	.5883
FACTOR 3 (Departures)	
1. Keep Them High	.7902
FACTOR 4 (A/C-Vectoring)	
1. Bad Vector	.5514
2. No Reason Vector	.5108
3. Holding	.4232
FACTOR 5 (Arrivals)	
1. Missed Approach	.6057
FACTOR 6 (A/C Direction)	
1. Improper Coord.	.5511
2. Routing Error	.4294
3. Position A/C Incorrect	.4129
4. Altitude Verification	.5423
FACTOR 7 (Communications)	
1. Traffic	.4374
2. Remarks	.4863

REFERENCES

1. Arad, B., B. T. Golden, J. E. Grambart, C. E. Mayfield, and H. R. vanSaun: Control Load, Control Capacity and Optimal Sector Design. System Design Team, Washington, D.C., and Research Division, National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Report No. RD-64-16, 1963.

2. Alexander, L. T.: Terminal Air Traffic Control System. SYSTEM DEVELOPMENT CORPORATION MAGAZINE, Santa Monica, California, July 1962.

3. Alexander, L. T.: Terminal Air Traffic Control Follow-on Research. System Development Corporation, Santa Monica, California, Report TM-639/007/00, 1963.

4. Alexander, L. T., and M. Ash: Terminal Air Traffic Control: A Laboratory Model for Man-Machine System Research. System Development Corporation, Santa Monica, California, Report SP-1016, 1962.

5. Alexander, L. T., and A. S. Cooperband: A Laboratory Model for Systems Research: A Terminal Air Traffic Control System. System Development Corporation, Santa Monica, California, Report TM-639, 1961.

6. Alexander, L. T., and A. S. Cooperband: Schematic Simulation: A Technique for the Design and Development of a Complex System. HUMAN FACTORS, 6:87-92, 1964a.

7. Alexander, L. T., and A. S. Cooperband: The Effect of Rule Flexibility on System Adaptation. HUMAN FACTORS, 6:209-220, 1964b.

8. Alexander, L. T., and E. H. Porter: Terminal Air Traffic Control and Problems of System Design. System Development Corporation, Santa Monica, California, Report TM-639/008/00, 1963.

9. Alluisi, E. A.: Human Engineering Aspects of Air Traffic Control Systems. Ohio State University Research Foundation, Columbus, Ohio, Final Quarterly Report, 1956.

10. Anderson, C. M., and C. E. Dowling: Evaluation by Simulation Techniques of a Proposed Traffic Control Procedure for the New York Metropolitan Area. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 245, 1954.

11. Anderson, C. M., T. E. Armour, et al.: Dynamic Simulation Tests of Baltimore Friendship Airport at Increased Traffic Densities. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, 1957.

12. Armour, T. E., A. N. Johnson, T. K. Vickers, and R. S. Miller: Simulation Tests of the Factors Affecting IFR Traffic Capacity at Chicago O'Hare Airport. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 341, 1958.

13. Baker, R. E., A. L. Grant, and T. K. Vickers: Development of a Dynamic Air Traffic Control Simulator. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 191, 1953.

14. Berkowitz, S. M., and R. R. Doering: Analytical and Simulation Studies of Several Radar-Vectoring Procedures in the Washington, D.C. Terminal Area. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Franklin Institute Laboratories Report No. 222, 1954.

15. Berkowitz, S. M., and E. L. Fritz: Analytical and Simulation Studies of Terminal-Area Air Traffic Control. Summary of Joint FIL-TDEC Simulation Activities in Air Traffic Control. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Franklin Institute Laboratories, Report No. F-2384, 1955.

16. Berkowitz, S. M., E. L. Fritz, and R. S. Miller: Summary of Joint FIL-TDC Simulation Activities in Air Traffic Control. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 297, 1957.

17. Bottomley, D., R. E. Hansen, T. R. Johnson, H. T. Rohland, S. B. Rossiter, and E. H. Wright: Dynamic Simulation Study and Evaluation of a Proposed Air Traffic Procedural Plan and Control Equipment for the Washington, D.C. Area. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Bureau of Research and Development Report No. 101-112V, 1962.

18. Brinton, J. H., Jr., and R. S. Miller: Summary of Joint FIL-FAA Research in Air Traffic Control During Period From September 1958 to January 1961. Franklin Institute Laboratories, Philadelphia, Pennsylvania, Report No. F-A2221, 1961.

19. Buckley, E. P., W. F. O'Connor, and T. Beebe: A Comparative Analysis of Individual and System Performance Indices for the Air Traffic Control System. Federal Aviation Administration National Aviation Facilities Experimental Center, Atlantic City, New Jersey, NA-69-40, 1969.

20. Buckley, E. P., and R. H. Rood: Tower Training Using Radar Simulators. Federal Aviation Administration National Aviation Facilities Experimental Center, Atlantic City, New Jersey, NA-77-12-LR, 1977.

21. Buckley, R. P., and R. H. Rood: CPM PROBE Experiment on Performance Information Feedback. Federal Aviation Administration National Aviation Facilities Experimental Center, Atlantic City, New Jersey, NA-77-18-LR, 1977.

22. Buckley, E. P., K. House, and R. Rood: Development of a Performance Criterion for Air Traffic Control Personnel Research Through Air Traffic Control Simulation. Federal Aviation Administration National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Report No. FAA-RD-78-71, 1978.

23. Chapman, R. L., W. C. Biel, J. L. Kennedy, and A. Newell: The Systems Research Laboratory and Its Program. Rand Corporation, Santa Monica, California, Report RM-890, 1952.

24. Cooperband, A. S., and L. T. Alexander: The Detection of Compound Motion. System Development Corporation, Santa Monica, California, Report SP-1964/001/00, 1965.

25. Cooperband, A. S., L. T. Alexander, and H. S. Schmitz: Test Results of the Terminal Air Traffic Control Laboratory System. System Development Corporation, Santa Monica, California, Report TM-639/004/00, 1963.

26. Faison, W. E., and A. L. Sluka: Dynamic Simulation Studies of Pictorial Navigation Displays as Aids to Air Traffic Control in a High-Density Terminal Area and a Medium-Density Terminal Area. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Bureau of Research and Development Report, 1961.

27. Fitts, P. M., and M. I. Prosner: Human Performance. Belmont, California: Brooks/Cole Publishing Company, 1967.

28. Franklin Institute: Pilot Experiments Concerning Air Traffic Control Decision Making (internal document prepared for the FAA Bureau of Research and Development). Philadelphia, Pennsylvania, 1960.

29. Grant, E. E., S. L. O'Connell, K. L. Stoker: The Human Factors Laboratory. System Development Corporation, Santa Monica, California, Report TM-561, 1960.

30. Harman, H. H.: The Systems Simulation Research Laboratory. System Development Corporation, Santa Monica, California, Report TM-498, 1960.

31. Hagay, J. A.: The Development of a Procedure for Evaluating the Proficiency of Air Route Traffic Controllers. Civil Aeronautics Administration Division of Research, Washington, D.C., Report No. 83, 1949.

32. Henry, J. H., M. E. Kamrass, J. Orlansky, T. C. Rowan, J. String, and R. E. Reichenbach: Training of U.S. Air Traffic Controllers. Federal Aviation Administration Office of Personnel and Training, Washington, D.C., Report R-206, 1975.

33. Hett, W. D., W. D. Coulopolos, W. Wolff, and R. A. Kowalski: Package D Testing With Air Movements Data Only (DAMDOT). MITRE Corporation, Bedford, Massachusetts, Final Report No. TM-3339, 1962.

34. Hixson, W. C., G. A. Harter, C. E. Warren, J. D. Cowan, Jr.: An Electronic Radar Target Simulator for Air Traffic Control Studies. Aero Medical Laboratory, Wright Air Development Center, Wright-Patterson AFB, Ohio, TR 54-569, 1954.

35. Howell, W. C., R. T. Christy, and R. G. Kinkade: System Performance Following Radar Failure in a Simulated Air Traffic Control Situation. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report TR 59-573, 1959.

36. Jacobs, J. F.: Practical Evaluation of Command and Control Systems. MITRE Corporation, Bedford, Massachusetts, Report MTP-7, 1965.

37. Johnson, B. E., A. C. Williams, Jr., and S. N. Roscoe: A Simulator for Studying Human Factors in Air Traffic Control Systems. University of Illinois National Research Council Committee on Aviation Psychology, Urbana, Illinois, Report No. 11, 1951.

38. Jolitz, G. D.: Evaluation of a Mathematical Model for Use in Computing Control Load at ATC Facilities. FAA Systems Research and Development Service, Atlantic City, New Jersey, Report No. RD-65-69, 1965.

39. Kidd, J. S.: A Comparison of Two Methods of Controller Training in a Simulated Air Traffic Control Task: A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 58-449, 1959, 1961.

40. Kidd, J. S.: A Comparison of One-, Two-, and Three-Man Control Units Under Various Conditions of Traffic Input Rate. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 59-104, 1959, 1961.

41. Kidd, J. S.: A Summary of Research Methods, Operator Characteristics, and System Design Specifications Based on the Study of a Simulated Radar Air Traffic Control System. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 59-236, 1959.

42. Kidd, J. S.: Some Sources of Load and Constraints on Operator Performance in a Simulated Radar Air Traffic Control Task. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 60-612, 1961.

43. Kidd, J. S., and J. J. Hooper: Division of Responsibility Between Two Controllers and Load Balancing Flexibility in a Radar Approach Control Team. A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 58-473, 1959.

44. Kidd, J. S., and R. G. Kinkade: Air Traffic Control System Effectiveness as a Function of Division of Responsibility Between Pilots and Ground Controllers: A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 58-113, 1958.

45. Kidd, J. S., M. W. Shelly, G. Jeantheau, and P. M. Fitts: The Effect of Enroute Flow Control on Terminal System Performance: A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 57-663, 1958.

46. Kidd, J. S., R. G. Kinkade, F. C. Ichniowski, W. G. Widhelm, and S. Urback: Overview of Project No. 102-8X. Aircraft Armaments, Inc., Cockeysville, Maryland, Report ER-3238, Vol. I, 1963.

47. Kidd, J. S., W. N. Widhelm, F. C. Ichniowski, and R. G. Kinkade: Method Development for ATC System Simulation Research. Aircraft Armaments, Inc., Cockeysville, Maryland, Report ER-3238, Vol. II, 1963.

48. Kraft, C. L.: A Broad-Band Blue Lighting System for Radar Approach Control Centers: Evaluation and Refinements Based on Three Years of Operational Use. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 56-71, 1956.

49. Kraft, C. L., and P. M. Fitts: A Broad-Band Lighting System for Radar Air Traffic Control Centers. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 53-416, 1954.

50. McKenzie, R. E.: An Exploratory Study of Psychophysiological Measurements as Indicators of Air Traffic Control Sector Workload. FAA National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Project 157-524-03R, 1966.

51. Miller, R. S.: Summary of Joint FIL-TDS Research in Air Traffic Control During Period From April 1957 to September 1958. Franklin Institute, Philadelphia, Pennsylvania, 1958.

52. Muller, P. F., Jr., R. C. Sedorsky, A. J. Slevinsky, E. A. Alluisi, and P. M. Fitts: The Symbolic Coding of Information on Cathode Ray Tubes and Similar Displays. Aero Medical Laboratory, Wright-Patterson AFB, Ohio, Project No. 7192, Contract No. AF 33(616)-43, 1955.

53. Paul, L. E., and E. P. Buckley: Human Factors Evaluation of a Large Screen Radar Display. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Report No. RD-66-105, 1967.

54. Ratner, R. S., J. O. Williams, M. B. Glaser, and S. E. Stuntz: The Air Traffic Controller's Contribution to ATC System Capacity in Manual and Automated Environments, Vol. I (Summary reports). Stanford Research Institute, Menlo Park, California, Report No. FAA-RD-72-63, 1972.

55. Ratner, R. S., J. O. Williams, M. B. Glaser, and S. E. Stuntz: The Air Traffic Controller's Contribution to ATC System Capacity in Manual and Automated Environments, Vol. II (Appendices). Stanford Research Institute, Menlo Park, California, Report No. FAA-RD-72-63, 1972.

56. Ratner, R. S., J. O. Williams, M. B. Glaser, and S. E. Stuntz: The Air Traffic Controller's Contribution to ATC System Capacity in Manual and Automated Environments, Vol. III (Terminal operations). Stanford Research Institute, Menlo Park, California, Report No. FAA-RD-72-63, 1972.

57. Schipper, L. M., J. Versace, C. L. Kraft, and J. C. McGuire: Human Engineering Aspects of Radar Air Traffic Control. Aero Medical Laboratory, Wright Air Development Center, Wright-Patterson AFB, Ohio, Project No. 7192, Contract No. AF 33(616)-43, 1956.

58. Schipper, L. M., and J. Versace: Human Engineering Aspects of Radar Air Traffic Control: I. Performance in Sequencing Aircraft for Landing as a Function of Control Time Availability. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report No. TR 56-57, 1956.

59. Schipper, L. M., J. Versace, C. L. Kraft, and J. C. McGuire: Human Engineering Aspects of Radar Air Traffic Control: II and III. Experimental Evaluations of Two Improved Identification Systems Under High Density Traffic Conditions. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report TR 56-58, 1956.

60. Schipper, L. M., J. Versace, C. L. Kraft, and J. C. McGuire: Human Engineering Aspects of Radar Air Traffic Control: IV. A Comparison of Sector and In-Line Control Procedures. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report TR 56-69, 1956.

61. Schipper, L. M., J. S. Kidd, M. W. Shelly, and A. F. Smode: Terminal System Effectiveness as a Function of the Method Used by Controllers to Obtain Altitude Information: A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report TR 57-278, 1957.

62. Schipper, L. M., C. L. Kraft, A. F. Smode, and P. M. Fitts: The Use of Displays Showing Identity Versus No-Identity: A Study in Human Engineering. 1957.

63. Slattery, H. F.: Air Traffic Control Simulation Program Conference, Report of the Chairman. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Research and Development Service, 1965.

64. Sluka, A. L.: Dynamic Simulation Studies of Pictorial Navigation Displays as Aids to Air Traffic Control in a Low-Density Terminal Area and in an Enroute Area. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, Final Report, 1963.

65. Test and Experimentation Division: A Report on Dynamic Simulation Tests and Study of the Bureau of Air Traffic Management Plan for the Positive Control of Air Traffic on an Area Basis in the Chicago and Indianapolis Air Route Traffic Control Center Areas. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, 1960.

66. Versace, J.: The Effect of Emergencies and Communications Availability With Differing Entry Rates: A Study in Human Engineering Aspects of Radar Air Traffic Control. Wright Air Development Center, Wright-Patterson AFB, Ohio, Report TR 56-70, 1956.

67. Vickers, T. K.: Development of Traffic Control Procedures for Tactical Airlift Operation. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 235, 1954.

68. Vickers, T. K.: Simulation Tests for Army Air Traffic Control. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 298, 1957.

69. Vickers, T. K.: The Use of Simulation in ATC Systems Engineering. Civil Aeronautics Administration Technical Development and Evaluation Center, Indianapolis, Indiana, Report No. 410, 1959.

70. Vickers, T. K.: Air Traffic Control Simulation Program Conference, Report of the Chairman. Federal Aviation Agency National Aviation Facilities Experimental Center, Atlantic City, New Jersey, 1965.

71. Vickers, T. K., and R. S. Miller: Recent Developments in the Simulation of Terminal Area and Enroute Area Air Traffic Control Problems. IRE TRANSACTIONS ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS ANE-3:51-55, 1956.