NASA/TM–1998-208750

Human Factors Assessment: The Passive Final Approach Spacing Tool (pFAST) Operational Evaluation

Katharine K. Lee
Ames Research Center, Moffett Field, California

Beverly D. Sanford
Sterling Software, Inc., Redwood City, California

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

October 1998

Acknowledgments

The authors would like to acknowledge the contributions of the DFW Passive FAST Assessment Team who participated in the operational evaluation of Passive FAST, as well as the many months of testing leading up to the evaluation. We would also like to extend our thanks to the DFW TRACON, who was always very accommodating and patient with our testing procedures. We would also like to recognize the tremendous support of William Eudaley (Sterling Software, Inc.) who, as part of our development team and the CTAS representative at the DFW TRACON, was a great source of information and supported us in collecting data. Finally, we would also like to express our appreciation to Santosh Mathan (Sterling Software, Inc.) for assisting in the analysis of the coordination data; Deborah Medved (Sterling Software, Inc.) for data entry assistance; and the other members of the Passive FAST test team: Tom Davis, John Robinson, Doug Isaacson, Shawn Engelland (all of NASA), and Stephane Couillaud (Raytheon STX).

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
(703) 487-4650

Contents

1.0 Summary 1

2.0 Introduction 1

 2.1 The Introduction of Automation into a Complex Environment 2

 2.2 Human Factors Assessment Framework 3

3.0 Methods 3

 3.1 Questionnaire Data 4

 3.2 Controller Observation Data 5

4.0 Results 6

 4.1 Test Period Characteristics 6

 4.2 Demographics 7

 4.3 Usability 7

 4.4 Suitability 9

 4.5 Acceptance 15

5.0 Discussion 17

 5.1 Usability 18

 5.2 Suitability 18

 5.3 Acceptance 20

 5.4 Lessons Learned: Constraints of Field Testing 21

6.0 Conclusion 21

Appendix A. Questionnaires and Rating Scales Used in the Operational Field Evaluation 23

Appendix B. Controller Acceptance Rating Scale (CARS) Use Guidelines 31

Appendix C. Coding Categories for Arrival Position Observations 33

References 37

Human Factors Assessment: The Passive Final Approach Spacing Tool (pFAST) Operational Evaluation

KATHARINE K. LEE AND BEVERLY D. SANFORD*

Ames Research Center

1.0 Summary

Automation to assist air traffic controllers in the current terminal and en route air traffic environments is being developed at Ames Research Center in conjunction with the Federal Aviation Administration. This automation, known collectively as the Center-TRACON Automation System (CTAS), provides decision-making assistance to air traffic controllers through computer-generated advisories. One of the CTAS tools developed specifically to assist terminal area air traffic controllers is the Final Approach Spacing Tool (FAST), which was tested extensively both in simulation and in the field. In 1996, FAST underwent an operational evaluation at the Dallas/Ft. Worth, Texas, Terminal Radar Approach Control (TRACON) facility. Engineering results showed increases in throughput and runway balancing efficiency.

Human factors data collected during the test describe the impact of the automation upon the air traffic controller in terms of perceived workload and acceptance. The human factors results showed that controller self-reported workload was not significantly increased or reduced by the FAST automation; rather, controllers reported that the levels of workload remained primarily the same. Controller coordination and communication data were analyzed, and significant differences in the nature of controller coordination were found. Acceptance ratings indicated that this new system was acceptable.

This report discusses the human factors data that were collected during the 1996 FAST Operational Field Evaluation and describes the controller-reported levels of acceptance, usability, and workload in the operational environment. The lessons learned from the perspective of human factors in the field testing process will also be discussed, along with comments on the development of future air traffic control automation.

* Sterling Software, Inc., Redwood City, California.

2.0 Introduction

Automation tools to assist air traffic controllers in the current terminal and en route air traffic environments is being developed at Ames Research Center in conjunction with the Federal Aviation Administration (FAA). This set of tools is collectively known as the Center-TRACON Automation System (CTAS), which provides decision-making assistance to air traffic controllers through computer-generated advisories. CTAS is distinctively human-centered and works to optimize arrival traffic flow for both the en route and terminal area environments (ref. 1). CTAS is comprised of several tools; three of these—the Traffic Management Advisor (TMA), the Descent Advisor (DA), and the Final Approach Spacing Tool (FAST)—have all undergone thousands of hours of controller-in-the-loop simulation testing and in the past several years have been the focus of extensive field testing. The tools have been developed at the field sites in Dallas/Ft. Worth, Texas, and Denver, Colorado. The focus of this paper is the human factors results from the operational field evaluation of the terminal area tool, FAST. Further information regarding the development and testing of TMA and DA can be found in other publications (refs. 2–8).

FAST provides advisory information to the air traffic controllers in the terminal area, also known as the Terminal Radar Approach Control, or TRACON. The FAST advisory information, as initially conceptualized, included turn, heading, and speed clearances, as well as runway assignments and sequence numbers (ref. 9). The advisories were integrated into the arrival controllers' radar displays by adding runway assignments and sequence numbers to the full datablock (FDB) and by providing symbology to indicate locations where speed clearances and turns should be initiated. In the early development of FAST, as with the other CTAS tools, controllers from the field sites participated in simulations and provided feedback into the development process. The controllers indicated that displaying all five types of advisories together on their monochrome radar displays produced excessive clutter. For this as well as other concerns, the

FAST functionality was split into "passive" and "active" phases (ref. 10). The passive phase includes the runway and sequence number advisories. The active phase adds the turn and speed advisories. Passive FAST (pFAST) was developed first, and recently completed an operational field evaluation at the Dallas/Ft. Worth TRACON (DFW). Active FAST is currently under development at Ames Research Center and is scheduled to begin simulation testing near the end of 1998.

The engineering specifications, methods, and results of the pFAST field evaluation are reported in several publications (refs. 11–13). Overall, an increase in throughput and runway balancing efficiency was shown, coupled with benefits demonstrated for Tower operations (ref. 11). But as Hopkin (ref. 14) has stated, for a system to be successfully developed for air traffic control (ATC), significant benefits must be provided to the air traffic controller or air traffic facility. Thus, it is important to fully understand levels of perceived workload and the aspects of the automation that influence controller acceptance. The evaluation and assessment of these issues fall under the domain of human factors, an important part of CTAS development which contributes to the characteristically human-centered design of CTAS as a whole.

Hopkin's statement is a reminder that engineering and human factors should work together to develop ATC automation. If development were not coupled in this way, it would be possible to create ATC automation aids that increase traffic handling capacity, but as a by-product also increase controller workload, stress, and required coordination. Such systems would ultimately be doomed to failure because of unjustifiable demands upon both the facility and the controllers, which could easily lead to an unsafe situation. By the same token, it would be possible to create a very usable human-computer interface with many of the latest interface design innovations, but which lacks significant, sophisticated advances "under the hood." Such a system would also fail because the interface alone cannot guarantee that the user will be able to effectively gather and process information, and the system may do nothing to reduce or mitigate workload or stress.

The CTAS tool development process has successfully coupled engineering and human factors efforts. This report will first describe previous ATC automation development, then the framework for the pFAST operational evaluation. Then methods used in the operational evaluation and detailed results and discussion are provided. Preliminary results have appeared in other publications (refs. 11 and 15), but are discussed here in significantly more detail.

2.1 The Introduction of Automation into a Complex Environment

The ATC environment provides many unique challenges to the introduction of new systems. As the first responsibility of the air traffic control system is safety, anything that is attached to the ATC environment must not compromise safety. In addition, the ATC environment has highly specialized constraints on lighting, displays, radar interface, and procedural and personnel requirements. Because there has been little change in the U.S. ATC facility equipment in the last 20–30 years, new software automation must work within existing FAA guidelines and procedures that may not be easily altered. In the TRACONs throughout the United States, for example, the typical controller display is a large, monochrome radar scope with the aircraft information presented via alphanumeric data tags associated with alphanumeric position symbols. This graphical user interface is unable to present menus, windows, and other such features which are considered conventional components of current human-computer interfaces. As a result, recent software development approaches regarding human factors issues may not be appropriate, and may need to be modified to meet the requirements of the specialized ATC environment.

The CTAS software development process utilizes procedures that are common to industry software development, such as rapid prototyping, change tracking, and verification and validation (M. Eshow, personal communication, 1997). These procedures have worked well within the development of CTAS because they enable iterative development and testing, and allow for user feedback before full implementation. Consequently, safety concerns and other problems can be resolved and demonstrated to users early, thus enhancing user confidence in the system. In addition, users have direct involvement with all aspects of the development process: the software changes, the testing, and the interaction with the developers themselves. Extensive simulations are conducted before the system is introduced into the field, and sometimes in the early stages of field deployment and testing. Human factors assessment is integrated throughout CTAS development to measure the impact on the controller, as well as to identify where engineering benefits may fall short in terms of user acceptance.

Previous development of ATC automation has met with mixed success. In the United States, the Advanced Automation System, or AAS, was slated to produce the next advances in ATC automation. However, the AAS development experienced many problems, stemming from issues such as its lack of iterative prototyping and delayed involvement of controllers in system evaluations (ref. 16). Human factors expertise was not incorporated in the requirements specification process, and human factors issues were limited to interface concerns. Consequently, a workable ATC automation system was not produced.

In contrast, European ATC automation development has met with better results. For example, the German research organization, Deutsche Forschungsanstalt für Luft und Raumfahrt (DLR), has developed advanced automation for German air traffic control. DLR-Braunschweig has implemented the Computer Oriented Metering Planning and Advisory System (COMPAS) to provide a strategic arrival planning system for both terminal area and en route controllers (ref. 17). This system underwent simulator evaluations, followed by operational testing, several years ago in the Frankfurt Control Center. The development of COMPAS has incorporated human factors issues in its design, and had the goal of matching a controller's mental model of the air traffic situation (ref. 18) to the development of the automation.

2.2 Human Factors Assessment Framework

The human factors operational evaluation of pFAST was built upon previous human factors evaluations of TMA and DA (refs. 15 and 19), as well as COMPAS. The general approach included developing an understanding of the existing operational environment and the tasks for which the controllers, area supervisors, and traffic management coordinators (TMCs) are responsible. Significant interaction between the researchers and controllers was required. This interaction helped both researchers and controllers to define the operational tasks and the testing objectives, while respecting the boundaries and needs of both groups during testing activities. In addition, these interactions contributed to refinement of data collection procedures and interpretation of results.

The usability, suitability, and acceptance concepts defined by Harwood (ref. 20) were used to organize the data collection efforts. Together, these results provide a fairly complete picture of the human factors impact of pFAST on the arrival controller. The data collection focused on each of these three areas, with observations and rating scales used to assess each category of information. These areas are defined below.

- Usability: perceptually based aspects of the human-computer interface, including the interaction with the interface (such as keystrokes, pointer movement, and other equipment manipulation).
- Suitability: information content and representation for the users' tasks; the support of the users' tasks and the workload level that results.
- Acceptance: a final "verdict" on the overall system, reflecting usability and suitability of the system, as well as job satisfaction, demonstrable performance, and esteem (ref. 19).

3.0 Methods

The operational assessment of pFAST took place over a period of six months. The test was conducted during arrival traffic rushes spanning the entire spectrum of traffic patterns at the DFW facility. Engineering data such as throughput, in-trail separation on final approach, and adherence to the sequence and runway advisories were collected; these findings are described in references 11–13. The engineering team was stationed in a room adjacent to, but separate from, the operational TRACON. In this separate area, the engineering data were collected, and the overall system was monitored during operational use of pFAST.

The human factors team conducted their data collection activities on the operational floor. Their role was to observe operations, collect data, and limit their interaction with the controllers, except to be available to answer questions about pFAST. The human factors team also occasionally provided feedback between the operational floor and the engineering team.

Data collection in the field, especially over a several-month effort, is subject to numerous constraints. There is no opportunity to exercise experimental control over traffic conditions, and test personnel must adhere to operational restrictions. It was clearly understood by all test personnel involved that operational demands took priority over any type of evaluation activity. Therefore, the human factors team curtailed their data collection activities whenever there were excess demands on space or personnel on the operational floor. Likewise, severe weather, training requirements, or other operational constraints on a few occasions led the facility representative to completely cancel evaluation sessions.

The controllers used pFAST advisories during 25 arrival rush periods across 7 different rush times. Baseline observation data were collected during 12 rush periods. There were 5 rushes in which pFAST was in operation for only part of the rush. These partial data are not included in the present report.

The pFAST advisories, which consisted of runway assignments and sequence numbers to the assigned runway, were incorporated into the existing Full Digital Automated Radar Terminal System (ARTS) Displays (FDADs) utilized by the TRACON arrival controllers. The advisories were added to the FDBs of the arrival aircraft (fig. 1). Controllers were required to make a few additional keyboard entries to input runway changes and accept runway advisories when they differed from default runway assignments. This was the extent of any additionally required physical manipulation of the equipment when using pFAST.

As shown in figure 1, information on the pFAST FDBs contained timeshared information on the second line; in one mode, the default runway assignment and the aircraft type are displayed. In the second mode, the aircraft's altitude and speed are displayed. On the third line, the aircraft's sequence number to the runway allocated by pFAST is displayed, together with the pFAST runway advisory, but only if the runway advisory differed from the default runway assignment. In figure 1, for example, the pFAST runway advisory is to 17L, and the default runway assignment is 17C. Until the controller acknowledged the pFAST runway advisory (through a keyboard entry), the 17L advisory continued to be displayed in the third line of the FDB. If pFAST's runway advisory did not differ from the default runway assignment, there would be no additional runway information in the third line of the FDB. The sequence number displayed in the third line is for the pFAST-advised runway. If the controller chose not to direct the aircraft to the pFAST-suggested runway, another entry could be made to indicate the controller's runway assignment, and the sequence number would update accordingly.

The pFAST Assessment Team (who participated in the operational evaluation) was composed of a group of eight controllers and one area supervisor. The Assessment Team had been involved in the development of pFAST for over a year prior to the operational evaluation. Consequently, they were trained to use pFAST and were familiar with its operation. All of the human factors data were collected from this pool of controllers, with the exception of two substitute controllers who participated when there was a staffing shortage. The substitute controllers were chosen by the Assessment Team and were briefed on the operation of pFAST prior to their participation in the operational evaluation.

The test plan was reviewed by representatives from the National Air Traffic Controllers Association (NATCA) who were involved in CTAS development. The human factors data consisted of questionnaires, operational observations, and in-depth debriefings. The procedures and questionnaires were developed with the aid of the Assessment Team controllers to ensure that the observation methods would not be intrusive to live operations and that the questionnaires were understandable and meaningful.

3.1 Questionnaire Data

There were several different questionnaires used in the operational evaluation. A demographics questionnaire was administered once. The other questionnaires, which examined usability, suitability, and acceptance issues, were administered after each test rush. Baseline questionnaire data were not collected as the data collection process was not finalized sufficiently ahead of time. The rating scales are listed below. Copies of the rating scales are provided in Appendix A.

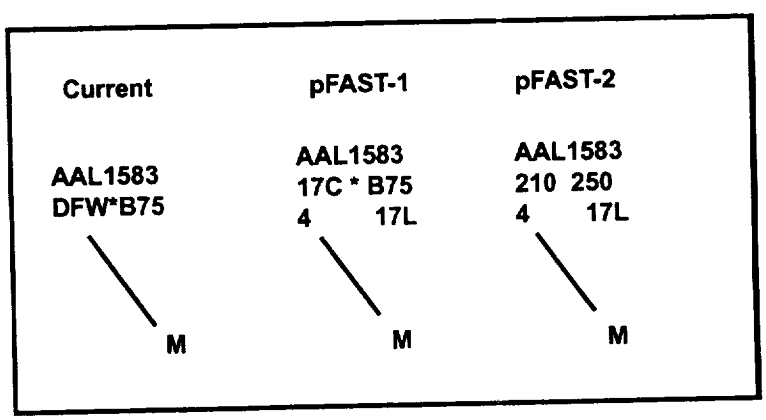


Figure 1. pFAST information added to the FDAD flight datablocks. Information displayed in Line 2 alternates ("timeshares") the presentation of two groups of information.

Line 1: ACID
Line 2: Current: Airport Destination, Aircraft Type
pFAST-1: Runway Assignment, Aircraft Type
pFAST-2: Altitude, Speed
Line 3: Sequence Advisory, pFAST Runway Advisory

3.1.1 Overall Workload and Workload Contributors

The workload ratings were collected using two different scales. First, a scale modeled after the NASA-TLX (ref. 21) was used to provide workload ratings along a 0 to 10 point range and included questions regarding mental demand, time pressure, performance support (provided by the pFAST advisories), overall effort, and the satisfaction versus frustration experienced. These workload ratings did not include the paired comparisons that are used with administering the original TLX. In addition, the physical demand rating from the original TLX was not used; in early testing, controllers reported that the physical demand rating was not a relevant question.

A second scale was used by controllers to rate a list of possible workload contributors on a range of 1 to 4, indicating how each of the items contributed to their overall workload.

3.1.2 In-depth Rush Information

Approximately once per day when pFAST was tested, the controllers were asked to provide more in-depth information regarding one of the rushes. Separate questionnaires were presented which included questions regarding controlling strategy, perceived coordination, and perceptions of how the Center handled the traffic flow to the TRACON (the Center feed).

3.1.3 Acceptance Ratings

The controllers provided a direct acceptance rating using the Controller Acceptance Rating Scale (CARS) (ref. 10) after each test rush. The CARS is adopted from the Cooper-Harper Scale for pilot evaluation of aircraft handling qualities (refs. 22 and 23). The Cooper-Harper scale has been used for pilot evaluation since its development in the late 1960s, becoming a worldwide standard (ref. 24). The test subject uses the Cooper-Harper scale by following a decision-tree structure and answering a series of dichotomous (yes or no) questions. Based on the responses, a numerical rating on a scale of 1 to 10 is selected. The Cooper-Harper rating falls into one of four possible rating groups: controllability, tolerability, satisfaction, and desirability. For each complete rush in which the pFAST advisories were shown, the controllers provided acceptance ratings. A description of the CARS and the criteria used in the acceptance ratings can be found in Appendices A and B.

The CARS was developed specifically for the assessment of CTAS automation and reflects the structure of the Cooper-Harper scale. The CARS is reoriented from the Cooper-Harper scale such that a rating of 1 reflects a lower, more undesirable rating, and a rating of 10 reflects a higher, more desirable rating. The CARS' physical appearance is also structured such that the decision-making process proceeds from the top of the diagram and moves down. The descriptive anchors for each rating on the scale reflect the ATC environment, and pFAST automation specifically (see Appendices A and B for examples of the CARS form and the guidelines that were used in the pFAST test).

The use of a Confidence Rating (a rating of A, B, or C), as with the Cooper-Harper scale, is maintained in the CARS design. The Confidence Rating is an expression of how much information the rater had to assess the system. It is important to reinforce that the Confidence Rating is not used to express the rater's confidence in the system itself.

3.2 Controller Observation Data

During both baseline and pFAST test conditions, observations were recorded by two human factors engineers at two positions along the arrival wall: one between the two parallel finals and one on the busy side of the rush (typically this was the East side of the arrival wall). Figure 2 describes the location of the controller and observer positions.

West side operations were located on the left of the arrival wall, and East side operations were located on the right. The two feeder positions (Feeder West, or FW, and Feeder East, or FE) were assisted by handoff positions (designated by "h" preceding the feeder name). The feeder controllers were responsible for controlling the traffic that arrives from the Center and merging different streams of traffic (which may be separated by altitude as well as arrival fix) into single streams towards the runways. In the DFW airspace configuration during the operational test, the FW controller was responsible for merging traffic arriving over both West arrival fixes, and the FE controller was responsible for merging traffic arriving over both East arrival fixes.

The final controllers were responsible for controlling the traffic handed off from the feeder controllers and directing the aircraft to their final approach courses. AR2 and AR1 (the parallel final controllers) were responsible for working the two parallel runways. Either the Meacham North (MN) or the Dallas South (DS) position was responsible for the diagonal runways, 13R (South flow) and 31R (North flow), respectively. The MN and DS positions were not co-located on the arrival wall, and observations were not collected from these positions (though questionnaire data were collected).

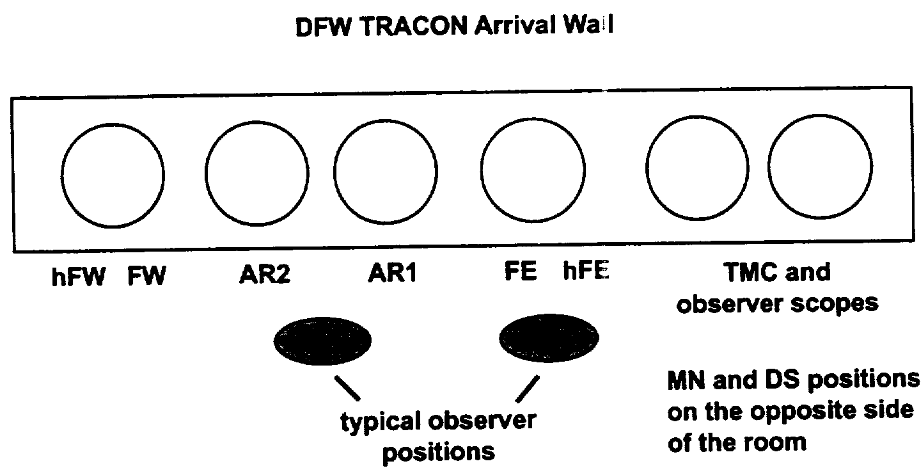


Figure 2. Controller and observer positions during the operational evaluation.
hFW = handoff, Feeder West — *FW = Feeder West*
hFE = handoff, Feeder East — *FE = Feeder East*
AR1, AR2: parallel finals
MN = Meacham North — *DS = Dallas South*

Basic characteristics of each observed rush, such as airport configuration, weather conditions, and changes to staffing, were noted by the human factors engineers. Coordination between the area supervisor and the TMCs and between the area supervisor/TMCs and the Tower and the Center was also noted. Specific observations were concentrated on coordination between the arrival controllers along the arrival wall, and, where possible, coordination with the Center. Controller coordination was defined as an instance of any verbal or nonverbal contact that was related to controlling traffic. The observations from the two observer positions were merged into a single transcript for each rush period observed. Any observation events that were incomplete, or unrelated to the traffic situation, were not included in the analysis.

The two human factors engineers who recorded the observations assigned the codes to each observed event by consensus. The coordination events in the transcript were assigned codes from 9 general topic areas: Runway, Sequence, TRACON situation, Aircraft Status, Coordination, Weather, Traffic Management Issues, Communication Issues, and Equipment Problems. Within each of the 9 major categories were a range of 2 to 6 subcategories. A total of 33 subcategories were available. A full text of the coding categories and the rules for assigning the codes is provided in Appendix C.

From these data collection materials, the usability, suitability, and acceptance areas were assessed in the following manner:

- Usability: primarily questionnaire data pertaining to issues of keyboard and slewball use, ability to detect the advisories themselves, the update rate of the advisories on the displays, equipment problems, and related communication problems.
- Suitability: questionnaire data pertaining to overall perceptions of workload, strategies in traffic control, the helpfulness of the advisories, and coordination and communication between the various ATC personnel.
- Acceptance: questionnaire data regarding specifically how acceptable the overall system was and comments from the controllers with regard to their areas of concern that influenced their acceptance ratings.

4.0 Results

The results are described in a general framework of usability, suitability, and acceptance, with the exception of sections 4.1 and 4.2, which describe test period characteristics and demographics information.

4.1 Test Period Characteristics

The DFW airport operates primarily in either North flow or South flow, which means that the traffic arrives and departs either landing towards the North or towards the South. South flow is the predominant airport configuration. The airport configuration defines the landing direction as well as which runways are in operation. During the testing period, it was possible to have, at most, three runways in the DFW airport configuration (two parallel runways and one diagonal runway). Since October 1996, the DFW airport has added another parallel runway. For the purposes of this paper, references to airport configuration refer to the landing direction and a three-runway operation. All three runways were in use whenever human factors data were collected. Six of the 25 total test rushes were in North flow.

One-way analyses of variance (ANOVAs) were used to compare North versus South flow questionnaire data. With the exception of one question, regarding the amount of perceived coordination between the arrival controllers, there were no significant differences between North and South flow responses. Consequently, all of the questionnaire data are considered together, regardless of airport configuration.

Passive FAST was tested during seven different rush periods. These time periods were (in local time): 8:00 AM, 9:30 AM, 11:00 AM, 2:00 PM, 3:30 PM, 5:00 PM, and 8:00 PM. The majority of the questionnaire data (nearly 62%) came from the 8 AM, 9:30 AM, and 11 AM rushes. For the purposes of the analysis, the data were treated all together, regardless of the time of the rush. This was due to the relatively small amount of data available, and its unequal distribution across the different rushes.

4.2 Demographics

Seven members of the Assessment Team filled out a general demographics questionnaire. Their ATC experience ranged from 9 to 19 years, with a mean of 13.3 years. DFW is a level 5 facility, which is the highest level in the FAA classification of facilities based on their hourly traffic density (ref. 25). The controllers were asked to indicate the number of years they spent at a level 5 facility. The reported range of years at a level 5 facility was 4 to 9 years, with a mean of 5.9 years. The range of years of experience at DFW TRACON was 3 to 8 years, with a mean of 5 years.

The controllers were also asked about their experience with computers as a whole. None of the controllers reported working with personal computers at work, on a day-to-day basis. Three of the seven controllers reported having a personal computer in their homes.

4.3 Usability

Because the use of the FDADs restricted how the advisories would be presented to the controllers, there were relatively few changes to the controller interface (see fig. 1). It was expected that the usability issues would be confined to the ability of the controllers to visually detect, and respond to, the advisory information, and to make the necessary inputs to interact with the system when changes to the advisories were required.

Questionnaire responses comprise the majority of the usability data. Questions pertaining to the pFAST advisories included using the equipment (making handoffs, using the keyboard and slewball, and making runway assignment changes), equipment problems, stability and update rate of the advisories, how much controller communication and coordination was required, and the use of the sequence numbers in coordination. Each of these results will be presented in detail in the sections below. Several of these questions were phrased in terms of how the usability item contributed to the controller workload. This is different from the suitability issues, in that the usability questions are not concerned with the information content of the features.

4.3.1 Using the Equipment

As shown in figure 3, giving handoffs, receiving handoffs, and using the ARTS keyboard and slewball were all rated as <u>minimally</u> to <u>not at all</u> contributing to the controllers' workload. Making runway assignment changes overall was also rated as <u>minimally</u> to <u>not at all</u> contributing to the controllers' workload. The keyboard entry requirements as a whole were rated as a little less demanding than normal keyboard entry requirements.

Feeder controllers are largely responsible for establishing the aircraft sequences; generally, the final controllers themselves make few changes to the traffic plan. This is reflected in the results shown in figure 4; the feeder controllers rated the keyboard entry requirements significantly more demanding than the final controllers ($F(1,42) = 6.406$, $p < 0.02$). The feeder controllers rated the keyboard entry requirements as about the same as they currently experience. These results also suggest that the keyboard entry requirements that are imposed by pFAST do not add significantly to controller workload.

Of all the controller positions, the hFE controllers rated the keyboard entries as most demanding. In general, all of the East side controllers rated the keyboard entries as significantly more demanding than the West side controllers. This is likely due to the nature of the rush patterns at DFW; as the predominance of data collected was in the morning hours, the rushes were mostly from the East. Under South flow configurations, rushes were generally busier for the East side due to the heavier traffic levels and the fewer available arrival runways on the East side of the airport.

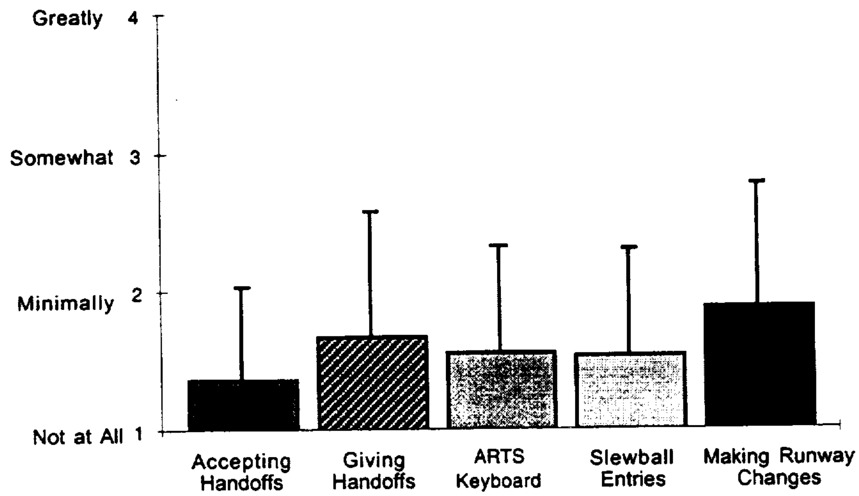


Figure 3. Usability items' contribution to overall workload.

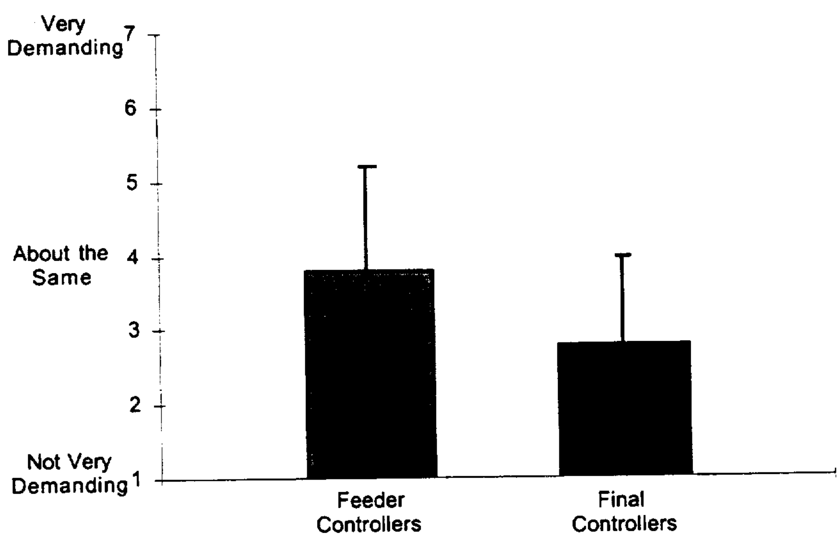


Figure 4. Keyboard entry requirements ratings using pFAST.

4.3.2 Equipment Problems

There were occasional equipment problems during the operational test. One problem was that the FDAD at the AR1 position was unable to display the pFAST advisories for certain runs. A second problem was interference created by inadvertent entries from other FDADs. The controllers rated these occurrences as <u>minimally</u> to <u>not at all</u> contributing to their workload.

4.3.3 Advisory Stability

Advisory stability is defined as the pFAST advisories not changing frequently on the controllers' FDADs. The pFAST advisories did not generally change past a certain "freeze" location unless a runway change was made by a controller or area supervisor using the ARTS keyboard. Exceptions to this did occasionally occur; most notable were sequence advisories between two aircraft in which one aircraft was turning. Sometimes the turn would cause the advisories to switch between the two. The sequence would correct itself once the turn was detected or completed. When runway advisories were changed, there was sometimes a perceptible delay as the pFAST software recomputed the advisories and the updated information was displayed on the FDADs. This delay was usually on the order of a few radar sweeps, and some controllers commented that some runway assignments changed later than expected.

Overall, the controllers reported no obvious stability problems for the runway and sequence advisories. They rated the update as occurring neither very well nor very poorly (fig. 5). Controllers were asked to rate how the wait for the update contributed to their workload; they rated this delay as minimally to not at all contributing to their workload. These results suggest the controllers were expecting some amount of update-related delay, but what they experienced was not excessive. It is a potential area of concern because the feedback is not instantaneous and the delay is noticeable. However, given the current hardware constraints on the display of pFAST, some update delay may be unavoidable.

4.3.4 Coordination and Communication

Coordination and communication were measured both through observations and through ratings. The ratings results describe these data in terms of frequency. The controllers rated the amount of communication that they had with the aircraft under their control. On average, the controllers reported talking to each aircraft a range of 2 to 5 times. The reported average over all of the controllers was 3.8 times (SD = 0.80). None of the controllers reported having to talk to any aircraft more frequently due to the pFAST advisories.

A single sample of the actual communication between each arrival controller and each aircraft that was worked was taken from one busy North flow rush. From this sample, it was calculated that across all the positions, controllers communicated with each aircraft an average of 3.74 times. This single sample is not adequate to suggest how reliable the controllers are about predicting the frequency of communication with the aircraft, but with further analysis of such data, the actual radio communication impact of using pFAST can be determined. Such data will be analyzed and discussed in a future report.

The controllers were also asked to rate the level of coordination required (with other controllers and facilities) during the test. They reported that the level of coordination that was required was not in excess of what they normally experienced.

4.3.5 Use of the Sequence Advisories in Coordination

The controllers reported referring to the sequence advisories rarely to sometimes when coordinating with other controllers. The average response was 2.30 (SD = 1.37) on a scale of 1= rarely to 7 = often.

4.4 Suitability

Objective workload measures, such as throughput and runway balancing, indicate the impact of automation on the work environment, but do not provide adequate information about controller workload, or coordination required between controllers. Therefore, suitability

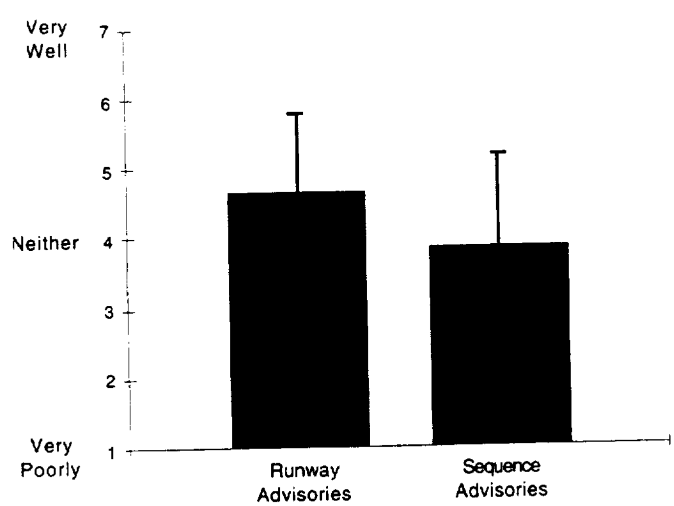


Figure 5. How well the advisories updated in response to changes.

questions are used to assess how the system provides assistance to the controller in performing her/his job. Suitability issues concern the ability of pFAST to support controlling strategies and planning. To meet its intended functionality, pFAST must provide accurate and useful information. The major issues of interest within the category of suitability are workload and coordination/communication. Workload has been a key concern of all parties involved in the development of pFAST.

The workload data are examined by considering all the controller positions equally; the data are not analyzed separately (East versus West side controllers, or feeder versus final controllers). Again, this was done because of the relatively small sample size and a restricted amount of data available in the different conditions.

4.4.1 Workload

In the beginning of the pFAST operational evaluation, the traffic into DFW TRACON arrived at a "free-flow" traffic rate. A traffic rate, or airport acceptance rate, reflects a number of arriving aircraft per given time period (typically, an hour). A free-flow rate is one that essentially allows traffic from the Center to enter the TRACON with no restrictions (such as metering) on the number of aircraft. This was done in part to exercise the limits of pFAST (by feeding as much traffic as they could into the TRACON). One possible covariate in the analysis of the workload questionnaire data was the decision to stop allowing the traffic to free-flow into the TRACON. This decision was made approximately three months into the operational evaluation and was based on two main factors: (1) the enhanced capacity with pFAST had already been demonstrated, and (2) the Center traffic feeds were, at times, too inconsistent during the peak flow periods.

After the decision to stop allowing free-flow rates under pFAST testing conditions, the traffic fed by the Center was limited to a rate of 102 aircraft per hour. An analysis of the questionnaire data was conducted to contrast the ratings before and after free-flow rate conditions. No significant differences between the runway advisory agreement before and after free-flow conditions were found. Consequently, the remainder of the data described below combines both traffic rate levels in the analyses.

4.4.1.1 Overall Workload

The areas of workload described in the following section include workload scale (TLX-modeled) questions, controlling strategies (including planning activities), and sequence and runway advisory usage and support.

As described earlier, the workload scale used to measure overall workload incorporated categories of mental demand, time pressure, performance support, overall effort, and satisfaction versus frustration. The workload scale utilized a 0 to 10 range, with 0 representing the lowest score (lowest workload, most favorable rating) and 10 representing the highest score (highest workload, least favorable rating). Figure 6 depicts the mean workload ratings from the workload scale.

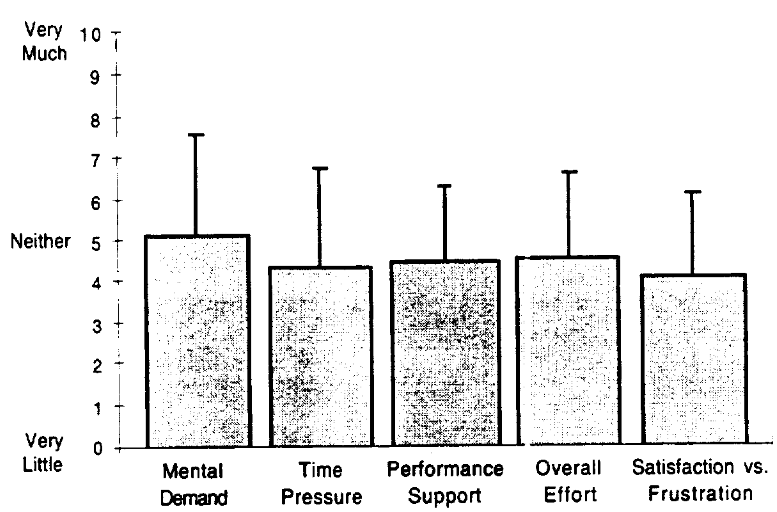


Figure 6. Workload scale ratings.

As can be seen from the graph, all of the responses are clustered around the middle of the scale. This suggests that pFAST did not increase controller workload. There is also no dramatic reduction in controller workload.

4.4.1.2 Controlling Strategies

It was of interest to determine how pFAST advisory information was incorporated into the controllers' tasks, as well as to determine how pFAST might be selectively utilized.

The helpfulness of the sequence numbers in terms of providing a common reference point was rated from 1 to 7, where 1 represented not at all useful and 7 represented very useful. The mean rating was 2.66 (SD = 1.35).

The controllers were asked to rate the amount of effort required to use the pFAST advisories. The mean response was 4.29 (SD = 0.77), which was slightly above the middle anchor of about the same towards the made it much easier end of the scale.

Controllers were asked if they followed the advisories more at some times than at others; one-third of the responses to this question were "yes." The reasons given for how the controllers followed the advisories were contradictory, however; some of the controllers reported greater advisory use during **lower** traffic conditions, and some reported greater advisory use during **higher** traffic conditions.

Controllers were also asked how pFAST advisories affected their ability to control traffic in their sectors. The mean response, on a scale of 1 to 7, was 4.43 (SD = 0.67). Controllers reported that, overall, pFAST had no effect on their ability to control traffic in their sectors.

The controllers were asked if they felt that they had to compensate for the pFAST advisories by changing what they would normally do. One-third of the responses were "yes." There was no additional elaboration on this result, however.

4.4.1.2.1 Sending and Receiving Aircraft "Over-the-Top." Sending aircraft over the top of the airport is a procedure that may arise because of the pFAST runway advisories. As mentioned earlier, under South flow, the East side controllers direct traffic to primarily one runway and the West side controllers direct traffic to primarily two runways. Consequently, when the bulk of the traffic is arriving from the East, pFAST may suggest runway advisories that would involve sending aircraft over-the-top, which would likely produce better runway balancing, and help to off-load the East side controllers. However, sending aircraft over-the-top may not always be the easiest task for a controller.

Figure 7 depicts the controller ratings of sending and receiving aircraft over-the-top. Sending aircraft over-the-top was rated as somewhat to minimally contributing to the overall workload, and receiving aircraft over-the-top was rated as minimally to not at all contributing to the overall workload. These are moderately positive results which suggest that the added tasks of changing runway assignments to the opposite side of the airport and requiring aircraft to be vectored over-the-top do not significantly impact the controllers' workload. Neither the controller who must initiate an over-the-top instruction nor the controller who receives aircraft from over-the-top are significantly impacted by this task.

4.4.1.2.2. Advisory Agreement. The controllers were asked how much they agreed with the runway and sequence advisories (fig. 8). Their reported agreement with the runway advisories was between sometimes and often. Their reported agreement with the sequence advisories was just above the middle-response of sometimes. It should be made clear that agreement with the advisories was not necessarily synonymous with adherence to the advisories, which was determined from the engineering data and is reported in Robinson et al. (ref. 13) and Isaacson et al. (ref. 12). While the controllers may have performed at a 95% adherence to the pFAST advisories, they may not have agreed with the advisories 95% of the time. In other words, the controllers could work the traffic in accordance with the pFAST advisories, but not agree with some sequences or runway advisories. Unless a particular advisory was unworkable from the controller's viewpoint, the adherence to the pFAST advisories in general was likely to be high.

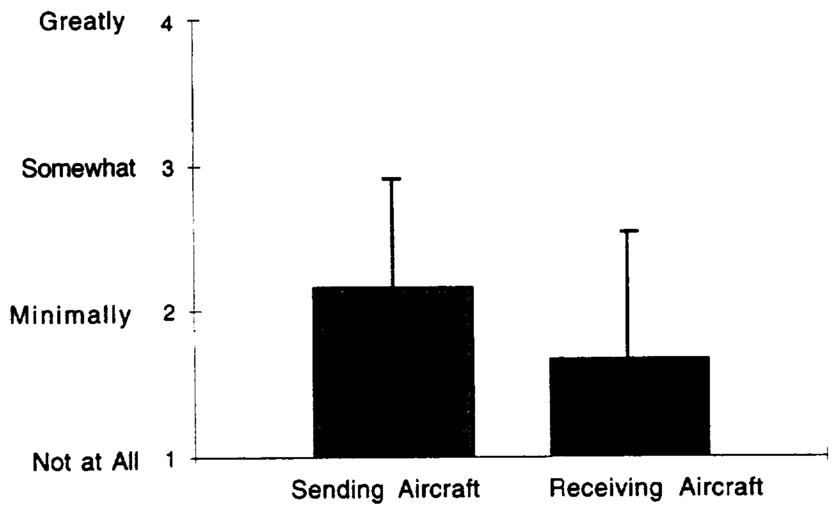


Figure 7. Controlling over-the-top aircraft and the contribution to overall workload.

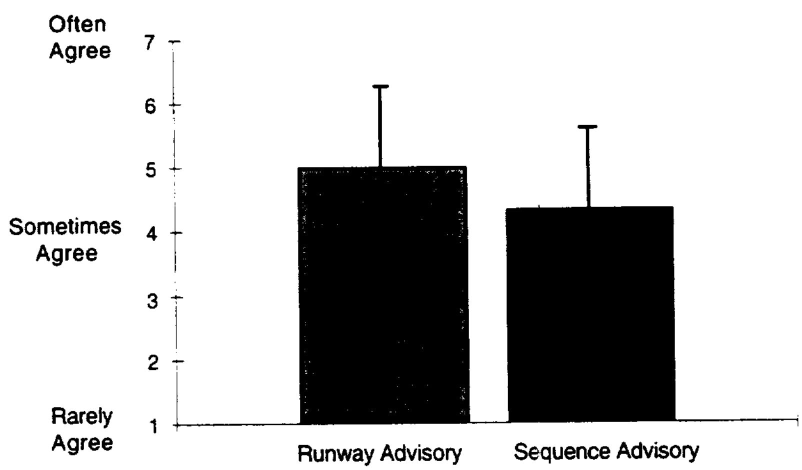


Figure 8. Agreement with the advisories.

4.4.1.2.3 Workload Impact of Planning and Following Advisories. The ratings for planning and following the runway and sequence advisories are shown in figure 9. Controllers rated both advisories between somewhat and minimally contributing to their overall workload.

4.4.2 Coordination/Communication

4.4.2.1 Observation Data

The transcript data were used to describe the impact of pFAST on controller coordination and communication. Available baseline observations were compared with field evaluation observations. It should be noted that baseline observations were gathered both before the operational testing and within the overall time frame during which the pFAST testing took place (but when the advisories were not being presented). The data were collapsed across both North and South flow, and the number of instances of each code was tabulated. The baseline data consist of a larger pool of controllers; in addition to the pFAST Assessment Team, other controllers who were not trained on pFAST were observed.

4.4.2.2 Most Frequent Coordination Categories

Over both baseline and test conditions, the five most frequently discussed categories were pFAST/ARTS-related issues, point-outs, handoff issues, runway assignments, and aircraft altitude changes. These categories are described in table 1.

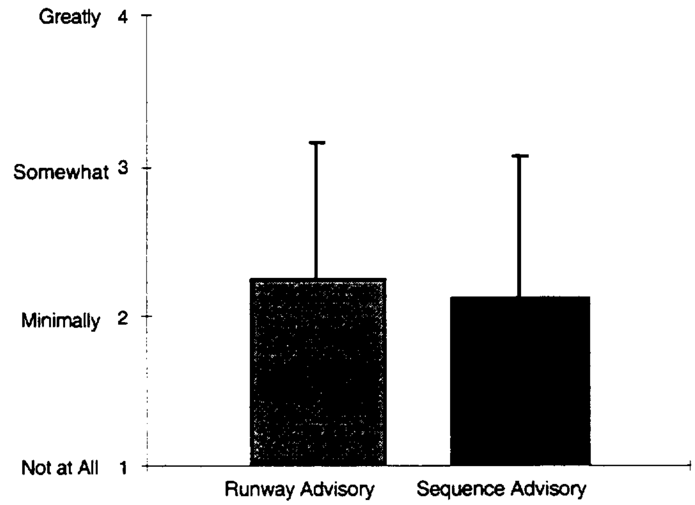


Figure 9. Advisory contribution to workload.

Table 1. Five most frequent coordination categories

Category Name	Description
pFAST/ARTS-related Issues	• Keyboard entry procedures required for pFAST-related inputs, as well as display issues related to pFAST • pFAST being turned on or off, or problems with the display of pFAST information (due to the ARTS interface)

Point-outs	• Aircraft requiring:
	    * Special handling
	    * Crossing through airspace that was not normally assigned to such aircraft
	    * APREQs (approval requests, especially from airports internal to the TRACON)
	• Utilizing another controller's airspace, but retaining communication/control of the aircraft
	• Often nonverbal

Handoff Issues	• Asking for handoffs • Frequency changes • Ownership
Runway Assignments	• What the runway assignments were • Changes to runway assignments
Aircraft Altitude Changes	• Expedited descents • Coordination based on altitude • Inquiring about aircraft altitudes

4.4.2.3 Baseline versus Test Coordination Comparison

The baseline and test conditions were compared and statistically significant differences in coordination were found in the categories of Runway Assignment, Sequence, Spacing, Point-outs, and Status Check. Figure 10 depicts the means and standard deviations of the baseline data compared to the pFAST test data. Table 2 lists the results of the statistical tests.

In four of these categories—Runway Assignment, Sequence, Spacing, and Status Check—the pFAST test conditions demonstrated more coordination per rush regarding these topics than the baseline conditions. The Runway Assignment category, as described in table 1, related to runway assignments or changes to the runway assignments. The Sequence and Spacing categories both concern the sequence advisories. The sequence category specifically refers to which aircraft are to follow which other aircraft and the sequence advisory itself. The spacing category refers to accommodating the sequence through changes to the existing spacing. The Status Check category was assigned to discussions referring to the current state of the traffic situation in qualitative terms, such as "Is everything going all right?" and comments from area supervisors checking on the workload of the controllers.

The point-outs category was the only coordination category which demonstrated a significant trend in the opposite direction. Point-outs are defined as coordination with another position so as to utilize another controller's airspace, but retaining communication and control (M. Prichard, personal communication, 1997). There was significantly more point-out coordination observed in the baseline than in the test condition. However, the controllers' mean ratings of point-outs contributing to workload fell in the range of minimally to not at all under the test conditions.

Tables 3 and 4 list the five most frequent categories of discussion in the baseline versus test conditions. The mean frequency (and standard deviation) of instances of coordination per rush is presented.

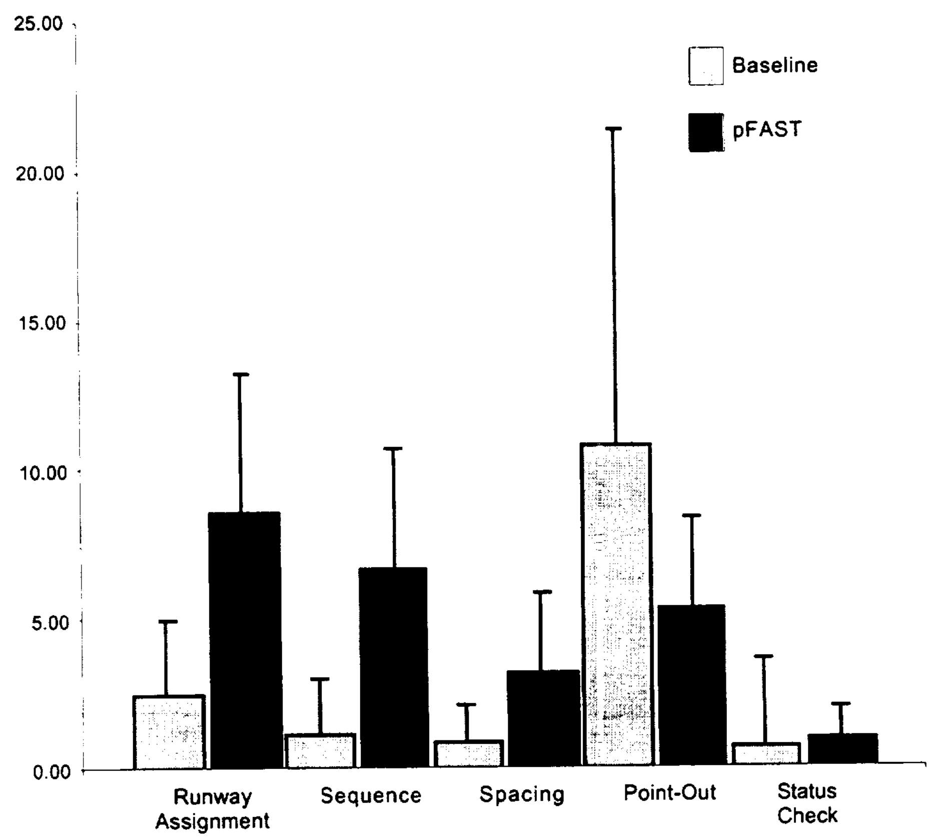


Figure 10. Baseline versus pFAST coordination comparison.

Table 2. Baseline versus pFAST coordination comparison

Category	Statistical Results
Runway Assignment	$F[1,32] = 14.97, p < 0.001$
Sequence	$F[1,32] = 16.72, p < 0.001$
Spacing	$F[1,32] = 7.43, p < 0.05$
Point-outs	$F[1,32] = 5.62, p < 0.05$
Status Check	$F[1,32] = 9.87, p < 0.05$

Table 3. Most common categories of coordination under baseline conditions

Category	Mean (SD)
Point-outs	10.90 (10.52)
Altitude Changes†	4.90 (3.45)
Handoffs†	4.20 (3.52)
Heading Changes	4.10 (2.85)
Runway Assignment†	2.50 (2.59)
*Weather**	*2.50 (4.12)*

†Categories common to both baseline and test conditions.

*The weather category result may be misleading, as weather conditions were more uniform during the pFAST test than during baseline observations.

Table 4. Most common categories of coordination under pFAST test conditions.

Category	Mean (SD)
Runway Assignment†	8.58 (4.65)
ARTS Problems	7.50 (5.41)
Handoffs†	7.21 (4.19)
Sequence	6.75 (4.09)
Altitude Changes†	5.74 (4.45)

†Categories common to both baseline and test conditions.

As shown in tables 3 and 4, there were three categories whose coordination frequency was common to both baseline and test conditions: altitude changes, runway assignments, and handoffs. There were more frequent altitude change discussions in the test condition than in the baseline condition. In addition, there was more frequent coordination regarding handoffs in the test condition than in the baseline condition. Runway assignments were discussed in both conditions, but as described above, were discussed significantly more in the test condition.

If the top five categories are an indication of discussion per rush, it appears that the frequency of discussion under pFAST conditions is higher and more evenly distributed for the top five categories. In baseline conditions, with the exception of point-outs, there is relatively infrequent discussion about the other four categories.

4.4.2.4 Center Comments

Some positive comments were collected from Ft. Worth Center, after the operational testing was completed (due to constraints on researcher staffing, no formal assessment was made at the Center during the pFAST test). One Center controller who was interviewed reported noticing turbo props being assigned to runway 18R, which he found unusual. This controller also reported that he noticed his holding was reduced by about 20% during the pFAST test. It should be pointed out that this is just one controller's observation and reflects just one aspect of delay reduction.

4.5 Acceptance

Usability and suitability results ultimately help to determine the overall acceptance of the system. In addition to providing usability and suitability measures, the controllers provided a direct rating of acceptance using the CARS. Prior to the beginning of the pFAST field evaluation, the CARS was used in simulation testing (ref. 10). Further, the pFAST Assessment Team helped provide the specific definitions of the CARS anchors, including defining adequate versus desired performance.

4.5.1 Numerical Ratings

The controllers' overall CARS rating across all the test rushes was 7.82 (SD = 1.10). This rating, rounded to 8, is associated with the following description of the system: "Mildly unpleasant deficiencies. System is acceptable and minimal compensation is needed to meet desired performance."

As discussed above, a portion of the test rushes occurred under free-flow acceptance rate conditions. The increased airport acceptance rate could have affected controller acceptance of pFAST, as a higher traffic level would presumably create more workload. Figure 11 shows the CARS ratings under free-flow and under more restrictive airport acceptance rates. There was no statistically significant difference between the two sets of CARS ratings.

The CARS ratings were significantly correlated with agreement with the runway advisories and how often the sequence numbers were considered to be in error. The higher the agreement with the runway advisories, the higher the CARS rating ($r = 0.502$, $p < 0.01$). The more often a sequence number error was noted, the lower the CARS rating ($r = -0.424$, $p < 0.02$).

The CARS ratings were also significantly correlated with the amount of effort required to accomplish the controlling tasks, using the advisories. The more the advisories were rated as making the work easier, the higher the CARS rating ($r = 0.55$, $p < 0.001$).

Finally, the CARS ratings were also significantly correlated with final controller ratings of their traffic feed. The more the final controllers felt that pFAST made their traffic much easier to manage and control, the higher the CARS rating ($r = 0.702$, $p < 0.002$).

4.5.2 Comments

In addition to the numerical and confidence ratings, the controllers were asked to provide comments on their CARS rating forms that would help clarify their ratings. Forty-five percent of the CARS data collected did not include comments. The lack of formally reported comments is due to two major factors. First, there were extensive debriefing sessions following the test rushes, often providing an opportunity for the controllers to report their opinions. Second, testing periods sometimes occurred with limited downtime in between the rushes. As the controllers were required to fill out, at minimum, three different surveys following each rush period, they were likely to only provide comments on the CARS form when they experienced problems that they wanted to highlight. As a result, it should be noted that positive comments were provided during debriefings, but were not always written down on the CARS form.

The comments that were reported on the CARS forms were summarized into six major categories, as shown in figure 12. The six categories were Sequence advisories, Runway advisories, ARTS problems, Traffic Load, Positive Comments, and Other. The Other category included comments regarding general questions about pFAST, the update rate of the advisories, external forces on the performance of pFAST (such as the Center feed or weather problems), and the effects of a lack of familiarity with pFAST. The controller comments were not categorized according to the severity with which a controller assigned a particular topic, so the tabulation of these comments reflected a continuum of minor disagreements with advisories to major philosophical differences with how the traffic should be controlled.

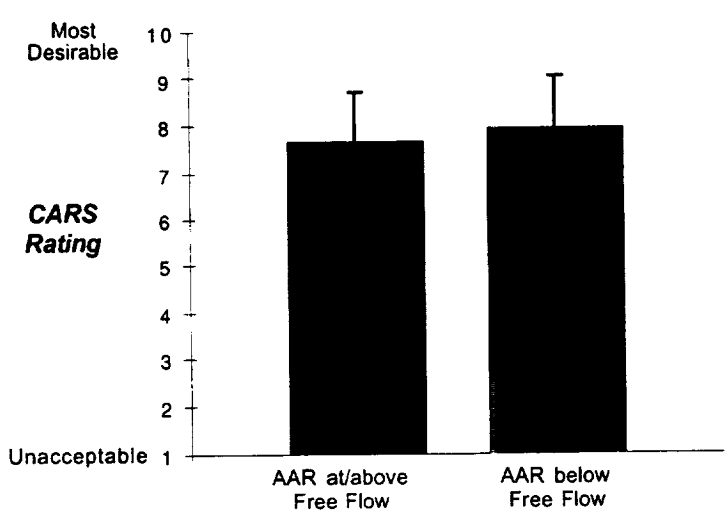


Figure 11. Free-flow AAR and CARS ratings.

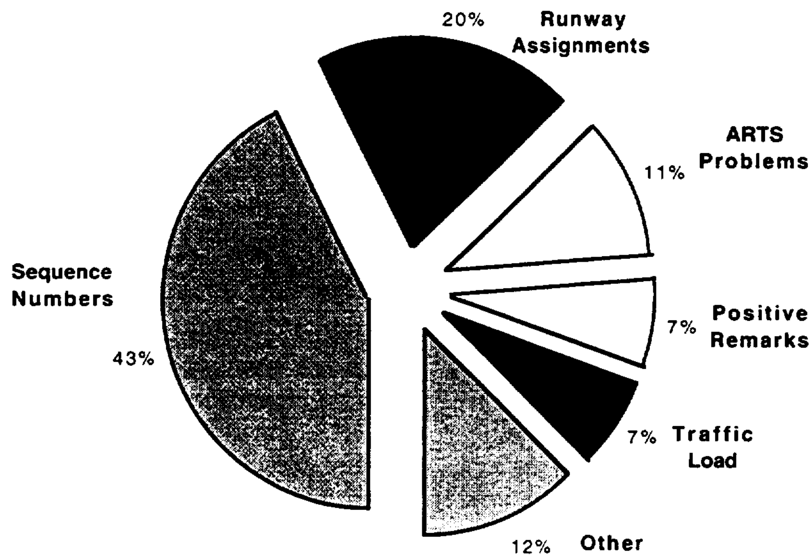


Figure 12. Top six categories of CARS comments.

As shown in figure 12, the majority of the comments (over 40%) were concerned with the sequence advisories. These comments related to overtakes and general disagreement with some sequences. The next-most-common comments related to the runway advisories where the controllers identified such issues as runway balancing and difficulty in achieving the over-the-top runway assignment.

Seven percent of the comments were purely positive in nature; for example, a controller expressing the opinion that the system ran very well with pFAST advisories.

The remaining two major categories reflect the difficulty in evaluating the system, or conditions affecting the acceptance rating: ARTS problems (11% of the comments) and Traffic Load (7% of the comments). The ARTS problems reflected issues unrelated to CTAS operations, such as the lack of advisory information at a position (due to equipment problems), the improper display of advisories, the "slinky effect" (which denoted a noticeable display lag between the FDB movement and advisory movement on the FDAD), and situations in which it was not possible to "quick-look" a controller's advisories from another controller position. The traffic load comments related to the traffic load being either too high or too low at that particular controller position for the controller to feel that s/he could make a sound evaluation.

5.0 Discussion

The pFAST operational test was conducted during a variety of rush periods over several airport configurations. The human factors data that were gathered contribute to the understanding of the impact of pFAST on the air traffic controller and the tasks for which the controller is responsible.

The pFAST operational evaluation results can be compared and contrasted with results obtained in the operational evaluation of COMPAS by DLR researchers. Although the COMPAS tool differs significantly from pFAST, and there are inherent differences in the operating procedures and facilities into which COMPAS was deployed versus pFAST, it is useful to examine what factors contributed to the success of COMPAS. The DLR researchers found general controller acceptance of COMPAS which they attributed to less required vectoring (more direct clearances), better coordination between en route and terminal environments, and a decrease in the minimum separation distance over time.

Similar engineering results to the COMPAS results were achieved in the pFAST operational evaluation (ref. 11). An analysis conducted prior to the pFAST field evaluations suggested that reduced spacing between arrivals on final approach could also be anticipated in pFAST that would contribute to an overall increase in efficiency of operations (ref. 26). The controller-rated acceptance of pFAST can be attributed to both the functional engineering benefits that were achieved and the positive human factors results discussed below.

5.1 Usability

From the usability perspective, controller ratings indicated that the additional inputs required to manipulate some of the pFAST advisories did not significantly increase workload. At best, the runway advisories were acceptable enough to require few corrections, or at worst, did not impact controller workload significantly when changes were indeed required.

When changes to the runway advisories were required, the greatest concern that the controllers voiced had to do with the delay from waiting for changes to update. While the delay was not rated as excessive, it is a potential area of concern which relates to the interface between CTAS and existing FAA hardware systems. Some observable lag between inputs and feedback may be unavoidable, but may also be reduced or alleviated with future equipment upgrades. It will be critical for the operational system to provide adequate training to the controllers so that they expect a lag time, and are able to work with it and distinguish genuine update time from a delay that might signify other problems with the system.

From a communication and coordination standpoint, the usability results showed that the amount of communication required due to the use of the pFAST advisories was not more than normal. This shows that pFAST is not creating additional interactions with other controllers or with the aircraft. Further examination of the communication data, such as determining the types of commands that controllers issued and contrasting such data under baseline with pFAST operations, would be useful in describing the impact of pFAST on controller communications. Such data analysis is forthcoming. The COMPAS testing determined that fewer heading changes were issued in the terminal area, and more direct vectoring was observed when COMPAS was in use (ref. 27). The controllers did not comment about this under the pFAST conditions, but this may be an area worth investigating as the communications data are analyzed. Again, the very different control environments would likely contribute to differences in results, but the COMPAS results are instructive in suggesting likely effects of ATC automation tools.

Unexpectedly, the controllers reported that the sequence advisories were not that useful when coordinating with other controllers. This result is somewhat contradicted by two other findings: first, observations determined that there was significantly increased discussion about the sequence advisories under test conditions, compared to available baseline data; second, controllers seemed very concerned about the sequences when they were asked to rate system acceptance. Robinson et al. (ref. 13) have suggested that the sequence advisories provide an additional benefit to controllers by indicating a gap in the sequence and show where a hole in the traffic stream should be maintained. The human factors data show that the controllers clearly were paying attention to the sequence advisories, but perhaps they did not consider their discussion about maintaining a sequence to be pertinent to the actual sequence advisories themselves.

5.2 Suitability

The suitability results show that pFAST is able to provide assistance to the controller by supporting controlling strategies and planning. Workload is a key measure in this analysis. The workload results reflect how usability elements contribute to the overall workload experienced. A highly positive workload result would have been an indication of a dramatic reduction in workload; a highly negative workload result would have been an indication of a dramatic increase in workload. The workload ratings suggest that pFAST had little to no effect on workload levels. The "non-effect" can be seen as a positive result, however, demonstrating that pFAST did not detract from operations. Improvements in throughput and runway spacing were achieved without adversely impacting the controller's workload.

Additional positive results can be seen in the comparison of free-flow and below free-flow traffic rates. No difference in overall workload between the two traffic levels was found, suggesting that pFAST can be helpful under highly challenging traffic loads without increasing workload. It should be noted, however, that in the future, free-flow operations may require some modifications to Center/TRACON traffic management coordination and procedures.

Controlling strategies were for the most part unaffected, though there was some discrepancy over whether the advisories were followed more closely at selective times. Following the advisories more when there was low traffic suggests that the controllers were paying attention and evaluating the advisories, and that they did so when they had time. Following the advisories more when there was high traffic suggests that the controllers had enough trust in the system to use the advisories even when they did not have adequate time to fully consider each advisory. While these responses seem to conflict, it should not be ruled out that different controllers will rely upon pFAST differently. Since both responses were obtained, it is reasonable to assume that pFAST will be used in both ways.

Another controlling strategy, sending and receiving aircraft over-the-top, was a likely source of increased effort, but was not rated as a significant contributor to workload.

This strategy could be an issue that is resolvable with experience; in initial (simulation) testing of pFAST, it appeared to be a more significant issue than it actually became during the operational test.

Overall, the controllers did not report that pFAST affected their ability to control traffic in their sectors; this result suggests that pFAST did not interfere with controllers' day-to-day responsibilities, and allowed them to continue to achieve safe and expedient traffic flows.

The controllers did not report tremendous agreement with the advisories themselves; their mean responses fell between "sometimes agree" and "often agree." However, the engineering data show very positive results for adherence to both runway and sequence advisories during the operational test (ref. 12). It is possible that the controllers felt that the advisories needed to be "perfect," thus their ratings may reflect their tendency to characterize a less-than-perfect test rush as problematic. This would contribute to their agreement ratings being less positive than the adherence results.

The controllers and the engineering team differed in their definition of perfect advisories. From the controller's perspective, a perfect rush likely reflected a condition in which the advisories matched her/his view of the traffic situation; this does not account for pFAST's knowledge of traffic outside of the controller's perception. Furthermore, it is unrealistic to expect that pFAST advisories would always perfectly match each controller's preferences. In contrast, a perfect rush in terms of the flow efficiency measured by the engineers was one in which delay was minimized. To attempt to issue advisories that always minimized delay could have produced a traffic scenario that might have been more difficult (or impossible) for the controller to accomplish (whether in terms of ability or comfort level). Thus, Robinson et al. (ref. 13) noted that it was more important to prevent the occurrence of poor advisories rather than to strive for issuing a series of perfect advisories. By occasionally presenting advisories that were less than optimal (engineering-wise), it was possible to achieve greater controller agreement and allow the controllers to work with the advisories. The balance between the optimization of the advisories and the workability of the advisories will always be an issue in the development of automation aids.

The sequence and runway advisories have been treated together in the human factors data analysis. It should not be assumed that their impact is necessarily equivalent, however. Disagreements with the sequence advisories did appear to be more noticeable, and created more concern than runway advisory disagreements. It is possible that an incorrect sequence is more obvious than an incorrect runway assignment. In addition, an incorrect sequence is something that must be corrected. A runway assignment can be a source of disagreement, but may still be correct and must be assigned because there is no other choice.

The coordination data provided some of the most interesting results. As Hopkin (ref. 14) has stated, coordination (between controllers) helps ensure safe aircraft handling. While the controllers did not report any significant increase in controller-to-aircraft or controller-to-controller coordination, some changes in coordination were observed between baseline and pFAST conditions. Runway assignments, sequences, and spacing were discussed with significantly greater frequency under the pFAST conditions. This result is somewhat expected, as the new information provided to the controller, as well as the testing environment itself, would likely promote discussion about the advisories. Increased discussion regarding status checking was also found under pFAST conditions relative to baseline, but could be an artifact of the operational evaluation itself. It is likely that the testing environment prompted the controllers and supervisors to increase their monitoring and awareness of operations in order to identify problems.

The most interesting coordination finding was the significant decrease in point-out activity under operational test conditions relative to baseline. Twice as many point-outs occurred under baseline conditions as occurred under pFAST conditions. Point-outs are common coordination activities between controllers and, as described above, are used to retain control over an aircraft, but to utilize the airspace of another controller. Reducing the number of point-outs could allow controllers to spend more time separating aircraft and monitoring the aircraft, rather than being concerned with coordination (M. Prichard, personal communication, 1997). It could also allow controllers to coordinate regarding other aspects of the traffic control process; perhaps more advance planning could be accomplished given more time to evaluate the traffic situation, therefore resulting in controllers using each other's airspace less than they would have to otherwise. Alternatively, point-outs could be reduced out of necessity as there was increased discussion regarding the advisories. However, if this were the case, the controllers should have indicated concerns over coordination. In contrast, the controllers did not report any difficulties with the amount of coordination that they experienced, and did not feel the amount of coordination required was increased by the use of the pFAST advisories. The point-outs themselves were not reported to contribute, on average, more than minimally to the overall workload.

It is important to note that the coordination discussed here does not just refer to that which involves the arrival controllers, but could reflect coordination between the arrival controllers and the Center or elsewhere in the TRACON. This is a positive finding that should be verified through further study during actual implementation of pFAST. When such an assessment is attempted, there should be discussion with the Center controllers to see if they also notice changes in the frequency of coordination with the TRACON. While information about Center operations was obtained from one Center controller, indicating a reduction in holding, it should be noted that this was one controller's assessment, and that holding at other sectors might not have been as great, if holding would have happened at all (holding rarely occurs simultaneously at all sectors). It is worth pointing out, however, that reduced delay at one sector translates into benefits for all sectors and the overall traffic flow, though these benefits may be manifested in different ways.

TMA was used for Center metering on two days at the end of the pFAST field evaluation. This TMA-influenced traffic feed has not been analyzed to see if there were any detectable differences in the human factors data. Such an analysis would be very useful in helping to determine how two of the CTAS tools work together. However, the data could also be confounded with the fact that TMA was used to meter at the end of the pFAST assessment, when issues such as training and system familiarity might also contribute to any significant differences that would be detected.

There was no attempt to analyze the impact of workload upon the controller in terms of the traffic complexity. Indeed, the time of day and rush periods define the traffic complexity that is experienced per rush. Bruce et al. (ref. 28) found that traffic complexity in an en route environment was a significant predictor of performance pressure, and it is likely that the same would be true in the terminal environment. Because of the small sample sizes in the different rush periods, we are not confident of drawing conclusions about the impact of traffic complexity upon controller time pressure or other aspects of workload. This would be an area for further investigation in the future, however.

5.3 Acceptance

The acceptability of the overall pFAST system was measured through the CARS and controller comments. The CARS was an easy-to-use, simple system for gathering acceptance ratings, and the ratings did not significantly differ between free-flow and below free-flow traffic rates. Further, the CARS ratings were correlated with advisory agreement, amount of effort used to accomplish controlling tasks, and how well pFAST made the traffic easy to manage and control. Comments provided from the CARS forms again reflect the predominant concern about the sequence advisories over all other topic areas.

In CTAS development, the CARS has only been used in the assessment of the pFAST automation; consequently, there are concerns about the interpretation of the CARS results and the scale needs to be further validated. Some validation issues include verifying that providing means and standard deviations in describing the ratings is appropriate; incorporating the confidence ratings in the interpretation of the results; and interpreting groups of controller results, where the actions of one controller directly impact the actions of another. Because Mitchell and Aponso (ref. 24) have determined that reporting means and standard deviations in using the Cooper-Harper scale is appropriate, the CARS data are also reported here using means and standard deviations; it is recognized, however, that further analysis is still necessary to explore this issue.

It should be acknowledged that even positive acceptance ratings themselves do not provide a full indication of how the entire facility is likely to react towards the eventual deployment of pFAST. There are still issues regarding job satisfaction and other elements of acceptance that are not easily quantified. It was clearly demonstrated in the operational evaluation that the performance of pFAST was acceptable within the boundaries of the testing environment. How the tool itself will be accepted by the controllers at DFW, as well as in a national deployment, can be influenced by many other factors.

Controllers who have not been so deeply involved in the development of pFAST are likely to be concerned that this new automation will change the nature of the controller's job quite substantially. This is a typical concern with automation that is not unique to air traffic control. Indeed, this is an issue that will be faced as new automation is developed that provides even more assistance to the controller. Currently, pFAST only suggests runway assignments and the landing sequence. Controller concern may grow as Active FAST begins to suggest headings, altitudes, or other control functions to manage the traffic. When considering these issues and concerns, it is important to keep in mind that the nature of the traffic environment itself is evolving to a condition in which such assistance may be essential to the controlling task.

Wickens et al. (ref. 29) have described the concept of automation as "a device or system that accomplishes (partially or fully) a function that was previously carried out (partially or fully) by a human operator." Their discussion points out that the definition of automation will change over time with the continued advances in technology and the continued interaction of the human user. Therefore, what was once considered a primary controller function is likely to change, and what the current controllers might consider an invasion of their job responsibilities will change as more automation is integrated into their work environment.

The impact on overall acceptance could be that some controllers will be less interested in their jobs as automation assistance is increased. Some of the typical planning and strategizing functions, which may be what currently makes the job rewarding, will certainly decrease or be removed. However, the integration of automation will likely pose new and different challenges, creating an environment that will still appeal to controllers, although perhaps in different ways than before.

5.4 Lessons Learned: Constraints of Field Testing

As outlined in this report, the field evaluation environment produced challenges in the assessment of pFAST that would probably not have been faced if such an assessment had been conducted in a laboratory setting. For example, there was no control over the airport configurations and AARs, which would have simplified the data analysis. There were also no guarantees about the staffing of the controller positions; while the majority of the rushes were staffed by the Assessment Team controllers, some substitute controllers participated when Assessment Team members were not available. Because only a small group of the overall facility had been chiefly involved in pFAST development, the facility at large did not have a good understanding of pFAST. As a result, misconceptions could occur suggesting that pFAST was causing problems, even when pFAST was not being used.

Concerns might also be raised about the fact that pFAST was developed and tested by the same group of controllers. It could be argued that the Assessment Team was not able to give the most balanced appraisal of pFAST capabilities, though in actuality, the Assessment Team members were always very frank in their evaluation of the system. However, if time and resources on the part of the facility had permitted, it would have been useful to train a new group of controllers on the use of pFAST and have this new group involved in the testing and validation process, and help to provide a wider range of experience and skill levels to the data collected.

The test periods themselves were limited by facility concerns about the traffic, or conditions of severe or unpredictable weather. The human factors team's concerns about the negative impact of the questionnaire-gathering process reduced the overall amount of questionnaire data that would have ideally been collected. The human factors team was also unable to directly observe or measure the impact of pFAST on the Tower or the Center. Finally, the problems that come from research in general, such as unexpected loss of data, and the unavailability of data, also occurred.

Some of these challenges were not anticipated in the development process leading up to the pFAST evaluation and led to more scrutiny and analysis of the data. The human factors data gathered during the pFAST evaluation are "noisy" as a consequence, and care should be exercised in extending the interpretation of the results to other similar tools or ATC environments.

There are several lessons learned from the pFAST field evaluation that could be instructive in future field evaluations and field deployment. Some of these recommendations are as follows:

- Plan to collect data systematically at the upstream and downstream facilities (Center and Tower, for pFAST).
- Plan alternate data collection in situations where the automation must be shut off due to unanticipated traffic or weather.
- Work with the facility to recognize when problems are due to the test and not to other unrelated hardware or software problems.
- Attempt to reduce the fatigue involved with repeatedly asking the same questions using the same questionnaires: perhaps streamline the questionnaires themselves mid-way through the test to eliminate questions that have not shown meaningful results; devise ways to achieve the same objectives as those intended by questionnaires to create more variety; focus questionnaires more narrowly to target fewer areas of interest.

6.0 Conclusion

The development of new ATC automation tools must provide demonstrable benefits for controllers. Such benefits should, at least, be in the form of accurate and useful information. From an overall system perspective, operations should demonstrate improvement, such as increased throughput and enhanced efficiency. Additionally, a benefits assessment must examine the system's impact on the controller and the controller's job. Positive benefits to the controller would be in the form of reduced or maintained controller workload, no unanticipated or unreasonable increase in controller responsibilities (such as increased frequency of inter-facility coordination or communication with aircraft), and continued job satisfaction (where the challenge of the daily tasks is at least maintained). Overall acceptance, which will determine how much a new system is utilized, will depend directly on these elements.

The human factors data from the pFAST field evaluation provide a complement to the overall engineering data that were collected. The engineering data show benefits of runway balancing and throughput. The human factors data describe the outcome of these benefits on the controllers themselves. Because of the heightened throughput and more efficient runway balancing during the pFAST field evaluation, it would not have been surprising if controllers reported increased workload. The human factors data instead bear out a different conclusion: despite the increased number of aircraft controlled during the field evaluation, the controllers did not report any significant increase in mental demand, time pressure, or overall effort. While controllers did not rate pFAST as improving their performance, they reported no detriment to their job satisfaction. The perceived workload remained at about the level to which the controllers say they are accustomed. These findings can be viewed as very positive.

Further, the additional information provided by the pFAST advisories and the increased discussion regarding the advisories were not found to significantly impact controller workload. Point-outs were reduced during pFAST test rushes compared to baseline data. Reduced point-outs suggest that the controller is able to concentrate on the key tasks of monitoring and controlling aircraft, and possibly coordinate regarding other aspects of traffic control. It is also possible that the pFAST operations lead to less frequent use of another controllers' airspace. This could provide benefits for not only the arrivals to the major airport within the TRACON, but for other airports within the TRACON, departures, and Center operations. Further studies should investigate the impact of pFAST upon other sectors, and try to determine if the reduced point-outs can translate into more time for the controllers to engage in planning activities, or other traffic control-related tasks.

The human factors data also show that the controllers were primarily concerned with the accuracy of the sequence advisories. They appeared to comment most frequently on sequence advisory problems, but their adherence to the advisories themselves was very high. The expectation that the advisories should be perfect is unrealistic, and may be something that will change with continued use of the system. Also, as CTAS is able to be implemented on faster hardware and the interface between CTAS and existing ATC hardware is improved, the update lags that were observed may be significantly reduced.

The ultimate success of pFAST as demonstrated by the operational evaluation is due to the successful incorporation of pFAST into ATC operations; this was partially aided by the long history of controller involvement in the design and testing of pFAST. As Hopkin (ref. 14) has suggested, controllers need to understand new systems in order to effectively utilize and integrate them into their existing knowledge and experience. The development of pFAST employed a strategy of closely coupling the human factors engineers, developers, and controllers. The human factors involvement in the development process contributed to identifying controller needs and determining if those needs were met. The trust of the air traffic controllers and their willingness to test pFAST operationally were results of this design approach. Without controller understanding and support of the system, benefits might never have been identified.

The attention that has been paid to the human factors issues has helped to define CTAS and ensure that it will meet controller needs. The human factors findings from the pFAST operational evaluation help to validate the processes which guided pFAST (and CTAS) software development and demonstrate how benefits are achieved not only in terms of overall airport throughput and efficiency, but in terms of the impact upon the controller. The positive human factors findings increase the confidence in the operational deployment of pFAST by making sure that key issues from the controllers' perspective have been examined.

Appendix A. Questionnaires and Rating Scales Used in the Operational Field Evaluation

Controller Workload Assessment (Modified NASA-TLX)

Controller Workload Assessment

Please rate along the scales below the following attributes of the last traffic period you just experienced:

MENTAL DEMAND

Please rate the amount of **mental and perceptual activity required** (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.):

Very Much — Neither Very Much nor Very Little — Very Little

TIME PRESSURE

Please rate the amount of **time pressure** you felt due to the rate or pace at which events occurred, while using the Passive FAST advisories:

Very Much — Neither Very Much nor Very Little — Very Little

PERFORMANCE SUPPORT

Please rate how much **the Passive FAST advisories, as well as the displays, slewball, ARTS keyboard, and the radar information** **assisted you** in accomplishing your tasks:

Provided Very Much Support — Neither Very Much nor Very Little — Provided Very Little Support

OVERALL EFFORT

Please rate the overall **amount of mental *and* physical effort required** to accomplish your level of performance using the Passive FAST advisories:

Very Much Effort Required — Neither Very Much nor Very Little — Very Little Effort Required

SATISFACTION vs. FRUSTRATION

Please rate the overall **degree to which you felt secure, gratified, content, and relaxed *versus* the degree to which you felt discouraged, irritated, stressed or annoyed** during the rush, using the Passive FAST advisories:

Very Satisfied and Not Frustrated — Neither Satisfied nor Unsatisfied, nor Frustrated — Very Unsatisfied and Frustrated

Contributors to Workload

Workload Topics

Please indicate how each of the following tasks contributed to your workload during the past rush, according to the following scale:

The task contributed ...

1	2	3	4	X
Greatly	Somewhat	Minimally	Not at All*	Did Not Experience**

...to my overall workload during this past rush.

** Please rate a **4** if you experienced a particular item, but it did not impact your overall workload.*

*** Please rate an "X" or leave the space blank if you **did not** experience a particular item listed below.*

	Description	Rating
1.	Traffic Load	
2.	Pilot responses—errors and delays	
3.	Aircraft/pilot procedural violations	
4.	Giving handoffs	
5.	Planning or following the sequence	
6.	Overshoots	
7.	Number of altitude changes issued	
8.	Pilot routing/altitude errors	
9.	Number of vectors/routing changes issued	
10.	Using the slewball	
11.	Planning or following the runway assignment	
12.	Coordinating a staggered approach	

	Description	Rating
13.	Aircraft flight characteristics	
14.	Lack of a handoff controller	
15.	Making runway assignment changes	
16.	Using the ARTS keyboard	
17.	Equipment problems	
18.	Accepting handoffs	
19.	Pointing out aircraft	
20.	Dealing with go-arounds or missed approaches (if any)	
21.	Sending aircraft over the top	
22.	Receiving aircraft over the top	
23.	Waiting for the sequence to update	
24.	The type of feed from the Center	

25. How would you describe **the overall amount of traffic** (the traffic load) that you experienced in this rush?

Low Moderate High

Post-Rush Questionnaire

Post Rush Questionnaire **OP-RMT**

Overall Impressions

1. How **typical** was the **amount of traffic** you experienced for this rush, compared to similar rushes at the same time and under the same configuration conditions?

Much Lower than Usual	About the Same	Much Higher than Usual

---

2. Approximately how often did you feel there **should have been a runway assignment change** from the one that Passive pFAST assigned?

Rarely	Sometimes	Often

---

3. Approximately how often did you feel **there should have been a sequence number change**?

Rarely	Sometimes	Often

4a. Were the **sequence numbers** ever clearly in **error** (e.g., reversed sequences for aircraft established on final, duplicate sequence numbers)?

yes **no**

4b. **If yes,** approximately **how often** did this occur?

Rarely	Sometimes	Often

5a. Did you make any **changes** to the **runway assignments**?

yes **no**

5b. **If yes, how well** did Passive pFAST **update** to meet your runway assignment change(s)?

Very Poorly	Neither Very Well nor Very Poorly	Very Well

6a. Did you make any **changes** to the **sequence**? **yes** **no**

6b. **If yes, how well** did Passive pFAST **update** to meet your sequence change(s)?

Very Poorly | Neither Very Well nor Very Poorly | Very Well

7. How **stable** were the **runway advisories** in your position?

Very Unstable | Very Stable

8. How **stable** were the **sequence numbers** in your position?

Very Unstable | Very Stable

9. Overall, how did you feel that the use of the **Passive pFAST advisories** affected the **amount of effort** you needed to accomplish your task?

Made it Much Harder | About the Same | Made it Much Easier

10. How **demanding** were the **keyboard entry** requirements?

Very Demanding | About the Same | Not Very Demanding

11. How did you think that **Passive pFAST** affected your **ability** to control the traffic in your sector?

Made it Much More Difficult to Control | Had no Effect | Made it Much Easier to Control

12. Did you have to **compensate** for the Passive pFAST **advisories** by **changing** what you would normally do?

yes **no**

13a. Do you think that you **followed** the Passive pFAST advisories more closely **at certain times of the rush than others**?

yes **no**

13b. If yes, when did you follow the advisories **the most**?

____
____
____

13c. If yes, when did you follow the advisories **the least**?

____
____
____

Controlling Strategies

1a. Did you use **double basing** in this rush?

yes **no**

1b. **If yes, how well** did the Passive pFAST advisories support you in doing this?

|___|___|___|___|___|___|

Made it Much Harder	Had No Effect	Made it Much Easier

2a. Did you send aircraft **over the top of the airport** in this rush?

yes **no**

2b. **If yes, how well** did the Passive pFAST advisories support you in doing this?

|___|___|___|___|___|___|

Made it Much Harder	Had No Effect	Made it Much Easier

---

3. Did you have to control any **go-arounds** or **missed approaches**?

yes **no**

Communication and Coordination

For the following questions, please **do not include** communication or coordination that was due to go-arounds or missed approaches.

1. Approximately how many times did you talk to each aircraft? _______ times

2a. Did you have to communicate with the aircraft more frequently because of the Passive pFAST advisories?

yes **no**

2b. **If yes,** what did you have to communicate with the aircraft about?

____

____

3. How many times did you coordinate with the following positions?

a. Other ARs or Feeders	________times	d. Departures	________times
b. Area Supervisor/TMC	________times	e. Satellites	________times
c. Center	________times	f. Other: ______________	________times

4. Aside from go-arounds/missed approaches, did you notice **greater** coordination than normal between yourself and (please circle):

a. Other ARs or Feeders	**yes**	**no**	d. Departures	**yes**	**no**
b. Area Supervisor/TMC	**yes**	**no**	e. Satellites	**yes**	**no**
c. Center	**yes**	**no**	f. Other: ______________	**yes**	**no**

5. When you coordinated with other arrival controllers, or with other personnel, how often did you refer to the **sequence numbers**?

Rarely Sometimes Often

6. How useful were the sequence numbers in helping you to coordinate with other personnel by providing a common reference?

Not at all Useful Somewhat Useful Very useful

Feeder and Handoff-Feeder Controllers only

1. How satisfied were you with the feed that you got from the Center?

Not Very Satisfied	Neither Satisfied nor Unsatisfied	Very Satisfied

2a. Were there any problems with the feed that you got from the Center?

yes **no**

2b. **If yes,** what problems did you have?

3. How satisfied were you with the traffic that you fed to the finals?

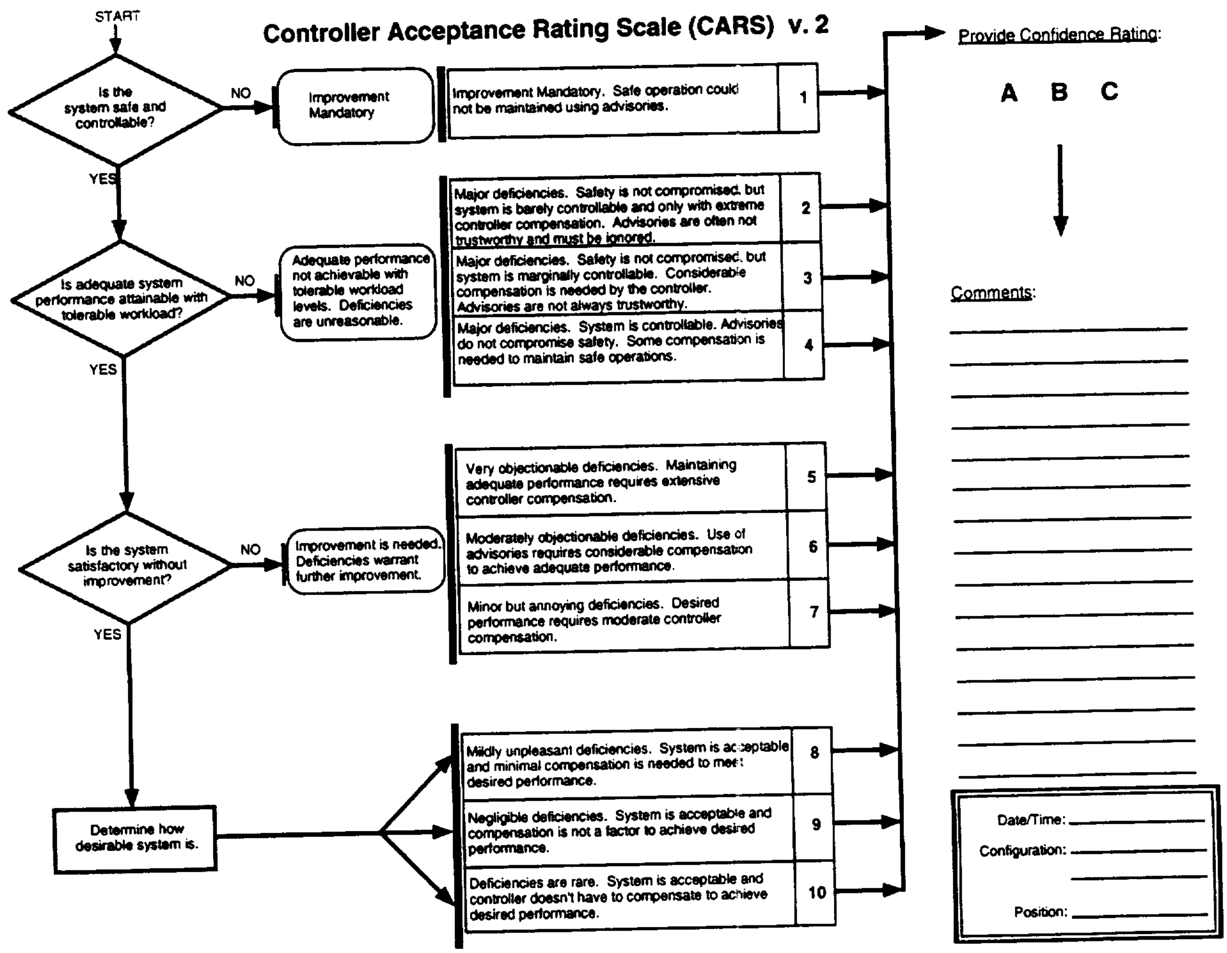
Not Very Satisfied	Neither Satisfied nor Unsatisfied	Very Satisfied

Final Controllers only

1. How did you think that Passive pFAST affected the traffic you were fed?

Made it Much More Difficult to Manage/Control	Had no Effect	Made it Much Easier to Manage/Control

Controller Acceptance Rating Scale



Appendix B. Controller Acceptance Rating Scale (CARS) Use Guidelines

Controller Acceptance Rating Scale (CARS) Guidelines

CARS will be used by the controllers to rate the acceptability of Passive pFAST during the Operational Assessment at DFW. To use this scale consistently, some basic definitions are defined below.

System

The system is taken to mean **everything** being rated: the controller's performance, the performance of Passive pFAST (runway advisories and sequence number advisories), and the performance of the ARTS.

Pilot performance should **also** be considered in the system rating. A couple of conditions to keep in mind when evaluating the overall system, and considering pilot performance:

1. If pilot response is exceptionally bad (for example, not very responsive) over several aircraft, then this could lead to a poorer picture of how well the overall system could perform. This should be reflected in the confidence rating. But to the extent that Passive pFAST was affected by bad pilot response, that should be considered in the numerical rating.

2. If pilot response is bad, but Passive pFAST seems to react especially poorly or especially well in incorporating the pilot situation, then this should be considered in the numerical rating.

Performance

The following are characteristics of adequate and desired performance.

ADEQUATE PERFORMANCE	DESIRED PERFORMANCE
The system performs at least as well as the current system performs.	**The system performs above and beyond the current system performance levels.**
• System performs much as it does currently.	• System performs better than it does currently.
• Runways balanced as well as they are currently.	• Runways well-balanced, ahead of when it is normally expected.
• Coordination between controllers is similar to what currently is required.	• Coordination between controllers is reduced.
• Reduced "guesswork" about where aircraft could be going.	• Does away with guesswork about where aircraft could be going.
	• Less sequence swapping close in.
• Advisories can be reasonably followed.	• Advisories are realistic in taking into account aircraft performance.
• Runway assignments are good, sequence numbers are OK (not "great").	• The system behaves predictably; it reacts approximately the same way under the same conditions.
• Runway assignments 90% accurate.	• Runway assignments 90-100% accurate.
• Sequence numbers 50% accurate.	• Sequence numbers 75-80% accurate.
	• Workload is well-balanced.
• Meeting the advisories doesn't result in excessive pressure.	• Meeting the advisories doesn't increase pressure.

Confidence

The **Confidence Rating** should describe confidence in the rating itself. It is not a rating of how confident one is about CTAS or Passive pFAST. For example, the confidence rating does not answer the question, "How confident am I that CTAS is good ATC software?"

Instead, it answers the question, "How confident am I that the rating I just made is an accurate one, reflecting the overall system performance, based on the amount of information I had available to me?"

The confidence rating should reflect the amount of information you think you had available to you in making your overall rating. It should also reflect problems that you encountered that are not necessarily an indication of how the software performed. As in the example above, a pilot that is especially unresponsive and uncooperative which results in a difficult traffic situation could mean that any problems encountered in the traffic situation could be due to more than just Passive pFAST; the pilot response is also a factor. How much a factor is reflected in the confidence rating.

There are 3 levels of Confidence rating:

A. High Confidence

You were able to account for the traffic events that occurred. You are very certain what problems or benefits could be due to Passive pFAST, the traffic situation, etc., and can therefore provide a rating that really reflects how well Passive pFAST performed.

B. Moderate Confidence

You were able to account for some of the traffic outcome. You are somewhat certain what problems or benefits could be due to Passive pFAST, the traffic situation, etc. There is some uncertainty about how well Passive pFAST performed, given the overall situation. You have some reservations about the accuracy of your numerical (CARS) rating.

C. Low Confidence

It was difficult to account for the traffic outcome. There is a great deal of uncertainty about the performance of Passive pFAST, and how you were able to work within the whole system. You have many reservations about the accuracy of your numerical (CARS) rating because of external factors that you can't adequately account for.

Appendix C. Coding Categories for Arrival Position Observations

Code	Topic	Subtopics/Explanation
	Runway	
01	Runway	• what is the runway assignment/change to the runway assignment
02	Over the top	• whether an aircraft is going over the top of the airport to a runway on the opposite side from where it originated
03	Runway Balancing	• discussion of how the runways are balanced • specific discussion of the traffic load with reference to the runways
	Sequence	
11	Sequence	• who follows whom • sequence number
12	Spacing	• filling a hole • general spacing comments, such as the spacing needed to accommodate departures
13	Blow By	• includes overtakes
	TRACON Situation	
21	Traffic Load	• specific to the amount, distribution, and level of traffic • the amount of traffic they are experiencing • the amount of traffic they are expecting • airport acceptance rate
22	Final Length	
23	Center feed	• how the Center is feeding them • coordinating with the Center regarding the feed • does not include coordinating with Center about ownership • delay discussions • holding
24	Airport Configuration	• asking about the current airport configuration • discussing a change to North flow or South flow
	Aircraft Status	
31	Speed	• asking another controller about an aircraft's speed • discussing what speed to take an aircraft
32	Altitude	• asking about an aircraft's altitude • coordinating based on altitude • expedite descent
33	Heading	• specific, explicit heading discussions • asking about where an aircraft is going with reference to a heading • discussing routing of an aircraft • includes discussion of aircraft going through the final
34	Priority Aircraft	• Emergency • aircraft equipment or mechanical problems that render it difficult or impossible to do what ATC instructs • does **not** include larger equipment problems related to VORs, DMEs, etc.
35	TCAS	• comments on an aircraft equipped with TCAS, or TCAS warning
36	Inquiring	• trying to determine the aircraft status (it is currently unknown)
37	Missed Approach	• go-arounds

Coordination

41	Point Out	• warning about an aircraft going through someone else's airspace • asking for an approval to go through someone else's airspace • includes physically pointing to the ac on the display
42	Ownership/Handoff	• asking for an aircraft to be handed off • asking if an aircraft was handed off • indicating ownership of an aircraft • whether or not one is talking to an aircraft • wrong handoff • frequency change related to ownership
43	APREQ	• approval request • departures internal to TR airports

Weather

51	Weather	• ATIS comments • Noting changes in the weather • altimeter setting • winds
52	Stagger/Simuls	• coordination specifically referencing staggering or simuls
53	Visual Conditions	• VFR or IFR conditions • Discussing ILS

Traffic Management Issues

61	AS/Controller coaching	• providing suggestions about how to run the traffic • inquiring about staffing needs • asking for information from the AS
62	Staffing	• Discussing needing handoff controllers • Need for final monitors • Briefing the next controller
63	Status check	• asking **how** things are going (qualitative statement) • comment on **how** the overall rush is going • controller performance

Communication

71	Communication Problem	• stuttering speech • not hearing a comment correctly or asking for clarification about what was said (but not with regards to clearing up a clearance or command) • speech formalisms • apologies
72	Correction to issued command	• could be prompting on the part of the handoff controller to tell the radar controller to correct what's been said • could reflect comments about a misheard or misreadback clearance

Hardware

81	Display problems/issues	• problems related to just the display and not with pFAST • scope problems; comment on not getting advisories, if it's just this particular scope

Hardware, cont'd

82	pFAST/ARTS (problem)	• anything pFAST-ARTS related:
		    • slinky effect
		    • TATCA Data Unavailable
		    • ZZZ's
		    • pFAST being turned on or off, or otherwise not working
		• comments about **data entry** (what keys are required to make a runway assignment change)

83	Frequency problem	• jammed frequency • wrong frequency
85	Equipment Malfunction	• larger problems with DME outage, VOR problems

99	Not Codable	• not understandable, based on the transcript, what exactly was the topic

Coding Rules

What to Code:

- Code only when the arrival controllers, area supervisor, and/or TMC are involved in the communication.
- Code even when there is only one of the TRACON arrival controllers/area supervisor/ TMC as a party in the communication.
- Code even when the communication itself is incomplete, as long as the participants are valid (in such cases, a 99 code is generally assigned).

What not to Code:

- Don't code communications between the controllers/area supervisor/TMC and the NASA test team.
- Don't code actions, such as controllers plugging in or getting up, unless some kind of communication is attached to this action.
- Don't code communications to the aircraft.

References

1. Erzberger, H.; Davis, T. J.; and Green, S. M.: Design of Center-TRACON Automation System. In: Proceedings of the 56th Symposium on Machine Intelligence in Air Traffic Management, Berlin, Germany, 1993, pp. 52-1–52-14.

2. Denery, D. G.; and Erzberger, H.: The Center-TRACON Automation System: Simulation and Field Testing. NASA TM-110366, 1995.

3. Erzberger, H.; and Nedell, W.: Design of Automated System for Management of Arrival Traffic. NASA TM-102201, 1989.

4. Harwood, K.; and Sanford, B.: Denver TMA Assessment. NASA CR-4554, 1993.

5. Hoang, T.; and Swenson, H. N.: The Challenges of Field Testing the Traffic Management Advisor in an Operational Air Traffic Control Facility. AIAA 97-3734, 1997.

6. Swenson, H. N.; Hoang, T.; Engelland, S.; Vincent, D.; Sanders, T.; Sanford, B.; and Heere, K.: Design and Operational Evaluation of the Traffic Management Advisor at the Fort Worth Air Route Traffic Control Center. Presented at the 1st USA/Europe Air Traffic Management Research and Development Seminar, Saclay, France, June 17–19, 1997.

7. Green, S. M.; and Vivona, R. A.: Field Evaluation of Descent Advisor Trajectory Prediction Accuracy. AIAA 96-3764, 1996.

8. Green, S. M.; Vivona, R. A.; and Sanford, B.: Descent Advisor Preliminary Field Test. AIAA 95-3368, 1995.

9. Davis, T. J.; Erzberger, H.; Green, S. M.; and Nedell, W.: Design and Evaluation of an Air Traffic Control Final Approach Spacing Tool. J. Guidance, Control, and Dynamics, vol. 14, 1991, pp. 848–854.

10. Lee, K. K.; and Davis, T. J.: The Development of the Final Approach Spacing Tool (FAST): A Cooperative Controller-Engineer Design Approach. Control Engineering Practice, vol. 4, no. 8, August 1996.

11. Davis, T. J.; Isaacson, D. R.; Robinson, J. E. III; den Braven, W.; Lee, K. K.; and Sanford, B.: Operational Test Results of the Final Approach Spacing Tool. In: Proceedings of the IFAC 8th Symposium on Transportation Systems '97, Chania, Greece, June 16–18, 1997.

12. Isaacson, D. R.; Davis, T. J.; and Robinson, J. E. III: Knowledge-Based Runway Assignment for Arrival Aircraft in the Terminal Area. AIAA 97-3543, 1997.

13. Robinson, J. E. III; Davis, T. J.; and Isaacson, D. R.: Fuzzy Reasoning-Based Sequencing of Arrival Aircraft in the Terminal Area. AIAA 97-3542, 1997.

14. Hopkin, V. D.: Air Traffic Control. In: E. L. Wiener and D. C. Nagel (eds.), Human Factors in Aviation, San Diego: Academic Press, 1988.

15. Harwood, K.; Sanford, B. D.; and Lee, K. K.: Developing ATC Automation in the Field: It Pays to Get Your Hands Dirty. ATC Quarterly Journal, vol. 6, no. 1, 1998, pp. 45–70.

16. Small, D.: Lessons Learned: Human Factors in the AAS Procurement. MP 94W000088, MITRE, McLean, VA, 1994.

17. Volckers, U.; Brokof, U.; Dippe, D.; and Schubert, M.: Contributions of DLR to Air Traffic Capacity Enhancement Within a Terminal Area. AGARD 56th Conference of GCP, Paper No. 24, Berlin, Germany, 1993.

18. Winter, H.: The Role of Planning Systems in Future Air Traffic Management. In: The COMPAS System in the ATC Environment, Scientific Seminar of the Institute for Flight Guidance, 1990.

19. Harwood, K.; and Sanford, B.: Evaluation in Context: ATC Automation in the Field. In: J. Wise, V. D. Hopkin, and P. Stager (eds.), Human Factors Certification of Advanced Aviation Technologies, Daytona Beach, FL: Embry-Riddle Aeronautical University Press, 1994.

20. Harwood, K.: Defining Human-Centered System Issues for Verifying and Validating Air Traffic Control Systems. In: J. Wise, V. D. Hopkin, and P. Stager (eds.), Verification and Validation of Complex and Integrated Human Machine Systems, Berlin: Springer-Verlag, 1993.

21. Hart, S. G.; and Staveland, L. E.: Development of a NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: P. A. Hancock and N. Meshkati (eds.), Human Mental Workload, Amsterdam: North-Holland, 1988, pp. 139–183.

22. Cooper, G. E.; and Harper, R. P.: The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. NASA TN D-5153, 1969.

23. Harper, R. P.; and Cooper. G. E.: Handling Qualities and Pilot Evaluation. J. Guidance, vol. 9, 1986, pp. 515–529.

24. Mitchell, D. G.; and Aponso, B. L.: Reassessment and Extensions of Pilot Ratings with New Data. AIAA 90-2323-CP, 1990.

25. U.S. Department of Transportation, Federal Aviation Administration: FAA Order SW 3510.10E, 1993.

26. Ballin, M. G.; and Erzberger, H.: Benefits Analysis of Terminal-Area Air Traffic Automation at the Dallas/Fort Worth International Airport. AIAA 96-3723, 1996.

27. Schick, F. V.: Evaluation of the COMPAS Experimental System. In: The COMPAS System in the ATC Environment, Scientific Seminar of the Institute for Flight Guidance of DLR, Braunschweig, Germany, September 12–13, 1990.

28. Bruce, D. S.; Freeberg, N. E.; and Rock, D. A.: An Explanatory Model for Influences of Air Traffic Control Task Parameters on Controller Work Pressure. In: Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, Seattle, WA, 1993.

29. Wickens, C. D.; Mavor, A. S.; and McGee, J. P.: Flight to the Future: Human Factors in Air Traffic Control, Washington, DC: National Academy Press, 1997.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
October 1998

3. REPORT TYPE AND DATES COVERED
Technical Memorandum

4. TITLE AND SUBTITLE
Human Factors Assessment: The Passive Final Approach Spacing Tool (pFAST) Operational Evaluation

5. FUNDING NUMBERS
538-18-24

6. AUTHOR(S)
Katharine K. Lee and Beverly D. Sanford*

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Ames Research Center
Moffett Field, CA 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER
A-9900077

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/TM–1998-208750

11. SUPPLEMENTARY NOTES
Point of Contact: Katharine K. Lee, Ames Research Center, MS 210-9, Moffett Field, CA 94035-1000
(650) 604-5051
*Sterling Software, Inc., Redwood City, California

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified — Unlimited
Subject Category 03 Distribution: Standard
Availability: NASA CASI (301) 621-0390

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Automation to assist air traffic controllers in the current terminal and en route air traffic environments is being developed at Ames Research Center in conjunction with the Federal Aviation Administration. This automation, known collectively as the Center-TRACON Automation System (CTAS), provides decision-making assistance to air traffic controllers through computer-generated advisories. One of the CTAS tools developed specifically to assist terminal area air traffic controllers is the Passive Final Approach Spacing Tool (pFAST). An operational evaluation of pFAST was conducted at the Dallas/Ft. Worth, Texas, Terminal Radar Approach Control (TRACON) facility. Human factors data collected during the test describe the impact of the automation upon the air traffic controller in terms of perceived workload and acceptance. Results showed that controller self-reported workload was not significantly increased or reduced by the pFAST automation; rather, controllers reported that the levels of workload remained primarily the same. Controller coordination and communication data were analyzed, and significant differences in the nature of controller coordination were found. Controller acceptance ratings indicated that pFAST was acceptable. This report describes the human factors data and results from the 1996 Operational Field Evaluation of Passive FAST.

14. SUBJECT TERMS
Air traffic control, Human factors, Terminal area

15. NUMBER OF PAGES
41

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		