To appear in J. Wise, V.D. Hopkin, and P. Stager (Eds.), Verification and validation of complex and integrated human machine systems. Berlin: Springer-Verlag (1993).

Defining Human-Centered System Issues for Verifying and Validating Air Traffic Control Systems

Kelly Harwood

CTA INCORPORATED, Moffett Field, CA, USA

Over the past 40 years, the application of automation to the U.S. air traffic control (ATC) system has grown enormously to meet significant increases in air traffic volume. The next ten years will witness a dramatic overhaul of computer hardware and software in enroute and terminal facilities to accommodate future growth in air traffic activities. From a human factors perspective, notable changes are the new controller workstations or sector-suites which will provide such new features as adjustable consoles, graphic situation displays, and electronic flight strips. This modernization will provide the basis for introducing automated functions that will transition the controller from tactical control to strategic traffic management. The U.S. Federal Aviation Administration (FAA) recognizes the importance of an effective human-system interface to successful operations (Kloster and Zellweger, 1987). Because various phases of test and evaluation are just around the corner for these new system upgrades, questions arise concerning what aspects of the human-system component must be addressed to verify system safety and efficiency. Such questions are not trivial. They strike at the heart of the "omnipresent criterion problem" (Christensen, 1958), that is, the difficulty of defining criterion measures for verifying and validating complex systems.

This paper first discusses the criterion problem, focusing on the unique constraints within ATC. The central argument is that before criteria and measures can be specified, human-centered issues associated with ATC technology upgrades must be carefully determined. An approach is discussed for disclosing such issues drawing on techniques and philosophies from traditional human factors engineering, cognitive systems engineering, and ethnography. The approach is illustrated for the Center Terminal Automation System (CTAS), a set of automation tools, currently under development and evaluation by NASA-Ames in partnership with the FAA. CTAS will assist air traffic personnel in managing arrival traffic flow in the center and terminal environments.

The Criterion Problem

The criterion problem is essentially the "problem of validating [and verifying] procedures and equipment against a goal, purpose, or set of aims" (Fitts, 1951; p. 76). Three key factors confound the problem for ATC.

First, the ultimate goals or criteria for ATC --namely safe, expeditious and orderly flow of traffic--are too general to easily quantify or set some measurable criteria (Fitts, 1951). Moreover these goals impose constraints on one another to achieve stable scoring criteria

(Whitfield and Stammers, 1978; Hopkin, 1980). For example, the controller may request an aircraft to deviate from its current route to maintain the orderly flow of traffic into the terminal area. Here, expediency is sacrificed for orderly flow, but the controller's performance still demonstrates effective judgment and planning.

A second confounding factor is the general lack of knowledge regarding job performance of individual and controller teams in current and future ATC environments. Without such knowledge it is difficult to establish meaningful functional relationships between aspects of system goals --safety and efficiency-- and aspects of controllers' job performance for validating and verifying future systems (Federal Aviation Administration, 1990). This problem is magnified as ATC system complexity increases and different couplings between the controller and machine are produced. Increases in intelligent decision-aiding automation will shift the unit of analysis from controller to cognitive system[1] (Hollnagel and Woods, 1983). Our understanding of what it means for a controller to evaluate a computer-generated conflict resolution or what awareness of the traffic situation entails for cognitive systems must be improved before we can identify meaningful functional relationships between system goals and measures for the human-system component for future ATC systems.

A third factor is the stringent requirement for sensitive criterion measures when transitioning from old to new ATC systems. A cautious transition strategy is necessary for maintaining ATC system continuity and safety. Thus system evolution occurs in small steps, and only minimal changes to the old system are permitted. Gradual evolution seeks to ensure that controllers have at least the same level of functionality as the old system and that critical cues for rapidly analyzing data and making decisions are preserved (Hunt and Zellweger, 1987). The challenge lies in verifying that this requirement is met. On the surface, minimal system changes would seem to imply minimal changes to criterion measures--that is to the functional relationship between system goals and measures. However, criterion measures must not simply tap performance with the new system, but must be sensitive to *consequences* of the new system for controller task performance (Hopkin, 1980). For example, electronic flight strips merely appear to change the medium of presentation. Yet questions have been raised regarding the consequences of electronic flight strips for the controller's understanding and memory of the situation and controlling strategies (Hopkin, 1991). Criterion measures must be sensitive to these potential consequences if informed decisions are to be made about ultimate system safety and efficiency. This requires a thorough understanding of tool use in the current ATC system as well as an understanding of future conditions of tool use with the new system. Our understanding of such consequences is limited.

In summary, the generality of ultimate criteria for ATC systems, our lack of detailed knowledge about controller/team job performance and cognitive systems, and stringent requirements for sensitive criterion measures during system transition confound the criterion problem for validation and verification of new ATC system components. Further, in the absence of clear statements of goals and criteria there is a risk of collecting data that is difficult to integrate into a decision about system safety and efficiency (Van Cott and Kincaid, 1972; Meister, 1985) or risk of taking measurements that are expedient versus appropriate (Parsons, 1972; Hopkin, 1980). Accordingly, what is needed is a clear specification of human-centered issues associated with ATC system upgrades from which criteria and measures may be identified. The following section describes an approach for identifying such issues in an effort to tackle the criterion problem for ATC.

---

[1] A cognitive system includes human operators and machine components. Together these components function adaptively as a system "using knowledge about itself and the environment in the planning and modification of actions." (Hollnagel and Woods, 1983; p. 583).

Approach for Defining Human-Centered System Issues

Human-centered system issues fall into three broad categories:

- technical usability
- domain suitability
- user acceptability.

Others have distinguished previously between two or three of these categories (e.g., Hopkin, 1980; Gould, 1988; Federal Aviation Administration, 1989; Rasmussen and Goodstein, 1988). Technical usability refers to perceptual and physical aspects of the human computer interface such as display formatting, graphics and human-computer dialog as well as anthropometric characteristics of the workstation. Issues in this category address the access and manipulation of data. Occasionally, human factors system evaluations stop here. Yet assessing issues of interface usability does not provide insight into the suitability of the system for the domain. Here, domain suitability must be considered, which refers to the content of information and display representation for domain tasks as well as functionality and decision-aiding algorithms. Issues in this category address the appropriateness of information and functions for supporting the cognitive requirements of the domain. It is possible for a system to be usable but not suitable for domain tasks, and thus both must be considered.

User acceptability is obviously enhanced by the ease of use and suitability of the system for supporting cognitive task requirements. Yet user acceptance also depends upon job satisfaction. It is generally acknowledged that the role of the air traffic controller will evolve from tactical control to traffic monitoring and management with the increased application of information technology and intelligent automation. Accordingly, criteria for controller selection will change, and with this so too will the definition of what is satisfying and motivating about the job. However, for the next 15 years at least, the "transition population" of controllers must be considered. Hopkin (1980; 1992) has argued that issues of controller job satisfaction, esteem, and individual merit in the context of technology upgrades are generally overlooked, but may possibly have serious consequences for ultimate system safety and efficiency. Attention must thus be devoted to disclosing issues associated with the impact of new technology on ATC job satisfaction.

Technical usability is characterized as a bottom-up, technology driven process, while domain suitability is characterized as top-down and problem-driven (cf Rasmussen and Goodstein, 1988). User acceptability is influenced by the usability and suitability of the system. Figure 1 shows the relationship between these three human-centered system categories. "U", "S", and "A" indicate regions that correspond respectively to technical usability, domain suitability, and user acceptability. These regions are important considerations for system validation and verification.

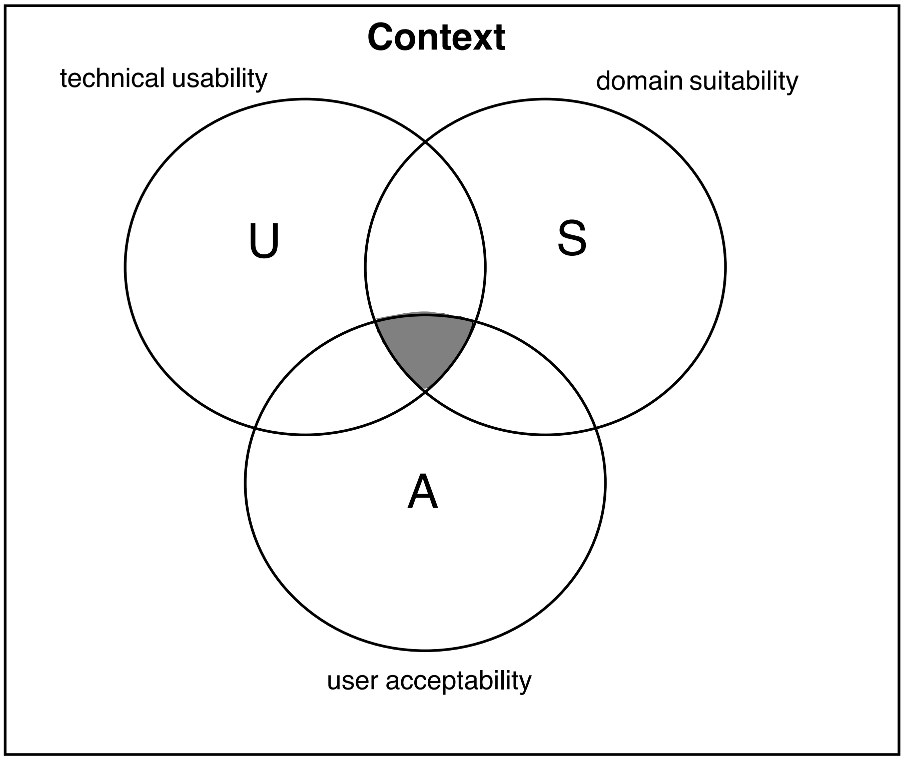

Figure 1. Relationships between technical usability, domain suitability, and user acceptance

Targeting the combination of technical usability, domain suitability, and user acceptability is necessary for system evaluations to provide meaningful input to decisions on system deployment (shaded region, Figure 1). Focusing on only one or two categories may be misleading. For example, a system may be usable and suitable, but if the system disrupts aspects of the job that are satisfying, then user acceptance may suffer (U overlaps S but not A). Likewise, a system may provide effective support for domain tasks and retain the human as the final decision-making authority, but use colors and character sizes that are difficult to discriminate under operational viewing conditions (S overlaps A but not U). System validation must consider all three aspects of the user's experience with the system.

Human-centered system issues are meaningless if derived without consideration of the context of the domain --ATC, nuclear power processes, flightdeck. Approaches for identifying issues must be contextually based; that is, based on an understanding of the physical characteristics of the environment (lighting, workplace layout), task domain (goals/functions of the domain), and work activities (social aspects of coordination; job satisfaction). The importance of context is depicted in Figure 1 by enclosing the figure in a box labelled "context."

Considerable efforts in the fields of human factors engineering, cognitive engineering, and usability engineering have been devoted to building a knowledge base from which principle driven approaches for designing complex systems may be derived. Aspects of these approaches are essential for validating and verifying the technical usability, domain suitability and user acceptability of a complex system. This paper discusses these techniques and approaches for defining human-centered issues for validation and verification of ATC systems and uses the CTAS Traffic Management Advisor (TMA) as an illustrativeexample.

CTAS

The continuing growth of air transport activities is challenging the capacity of terminal areas and airport facilities. In response to increased terminal delays and airspace congestion, the FAA Terminal ATC Automation (TATCA) program has initiated an effort to develop automated systems for assisting controllers in handling larger volumes of departing and landing traffic in major terminal areas. The FAA is currently evaluating CTAS, a prototype system, developed by NASA-Ames, to help controllers manage the flow of arrival traffic in the terminal area.

CTAS is an integrated set of automation tools, designed to provide decision-making assistance to both center and terminal controllers via planning functions and clearance advisories. CTAS consists of three sets of tools: the Traffic Management Advisor (TMA), Descent Advisor (DA), and Final Approach Spacing Tool (FAST). TMA generates landing sequences and schedules for arrival traffic to minimize delays. DA provides recommended cruise speed and descent clearances to help aircraft meet the schedule set by TMA with minimum fuel consumption. FAST assists terminal area controllers in spacing aircraft accurately on final approach. (For further information on CTAS, see Erzberger and Nedell, 1989; Davis, Erzberger, and Green, 1991; Tobias, Volcker, and Erzberger, 1989; ATC Field Systems Office, 1992).

CTAS development has involved thousands of hours of laboratory simulation with controllers to refine and extend algorithms and to enhance the user interface. In order to bring the system functionality to a level of operational stability and to provide information to Air Traffic and System Development Organizations on a possible national deployment decision, further development, validation and verification will be conducted in the field at four ATC facilities. TMA is the first CTAS component to undergo this field development and testing process and will be the focus of discussion for the remainder of this paper.

TMA has been developed for use by the traffic manager at Traffic Management Units (TMUs) within Air Route Traffic Control Centers (ARTCCs). The traffic manager's duties differ from a controller's duties in that traffic managers do not control aircraft directly. Instead they monitor the demand of arrival traffic into the center, coordinating with terminal personnel, area supervisors, and adjacent facilities, making decisions to balance the flow of traffic so that demand does not exceed capacity in the center and terminal areas. TMA is designed to assist the traffic manager by assigning the most efficient landing order and optimal landing times to all arrival aircraft.

TMA is also designed to assist in the re-routing of traffic in response to a runway reconfiguration or weather disturbance, or to balance the traffic load across arrival sectors. The traffic manager can override TMA's automatically-generated schedule at any time by resequencing aircraft, inserting slots for additional aircraft, or changing airport acceptance rates. Aircraft data tags are displayed on configurable moving timelines and are color coded to portray landing schedule and sequence status information. A traffic load display provides a graphical representation of various traffic load characteristics, and several configuration panels are available for modifying timeline displays and setting scheduling parameters. The workstation consists of a SUN4 Sparc workstation with keyboard and mouse input devices.

TMA presents the traffic management coordinator with new functionality, new display representations, and a new workstation. Evaluation of such new system components requires consideration of technical usability, task suitability, and user acceptability. Techniques for disclosing evaluation issues in these areas are discussed next.

Technical Usability

Technical usability refers to the perceptual and physical characteristics of the human-system interface, and includes general issues regarding the ability of users to read, detect, access, and manipulate information. A tremendous amount of research in human factors engineering and human-computer interaction has contributed to the development of principles and guidelines for designing and evaluating human-system interfaces (see for example, Van Cott and Kincaid, 1974; Smith and Mosier, 1986; Shneiderman, 1987; Department of Defense, 1989). These principles and guidelines form the basis for defining technical usability issues.

The identification of technical usability issues consists of three parts. The first involves identifying the human-system interface functions that characterize the system; for example, data entry, dialog type, and data display (e.g., Smith and Mosier, 1986). For TMA these are data display, dialog, and user guidance. Next, design features for each general function are identified. For TMA, some design features of the data display are color coding, timeline scales, abbreviations and labels. These first two parts for defining technical usability issues are necessary for narrowing the selection of relevant interface principles from all possible principles and guidelines and ensuring systematic coverage of all design features.

The third part involves defining technical usability issues. Here, general research principles on perception and information processing (Boff and Lincoln, 1988; Wickens, 1992) and guidelines for human-computer interaction and workstation configuration (Van Cott and Kincaid, 1974; Smith and Mosier, 1986; Shneiderman, 1987) are selected and tailored for specific design features of the system. For issues to be relevant they must reflect the constraints of the physical work environment into which the system will be integrated. For ATC traffic management operations, such constraints include low levels of lighting, physically separated sources of information, a mix of hard copy and computer-generated display media, and access to information from seated and standing positions. Definition of issues must acknowledge these constraints.

Some examples of technical usability issues for the TMA include:

- Are standard meanings used for aircraft size symbols?
- Do colors represent only one category of aircraft scheduling status?
- Can colors be discriminated under low lighting levels?
- Are labels displayed consistently across displays?
- Are abbreviations commonly recognized by the traffic manager?
- Can aircraft identification tags be read easily from operational display viewing distances?

Technical usability issues focus exclusively on the surface characteristics of display and input device interfaces. Addressing human-system interface issues is essential for effective system performance. No matter how elegant the algorithms, a poor user interface will contribute to degraded system performance and negative impressions of the system (Smith and Mosier, 1986). However, to ensure that the system supports the problem solving requirements of the domain, domain suitability must be considered.

Domain Suitability

As intelligent automation and new technology are gradually added to ATC, the controller and computer will become partners in traffic management decisions. For example, one type of partnership might involve the computer generating aircraft separation advisories and the controller evaluating and issuing the advisories to aircraft. Domain suitability refers to the effectiveness of such decision-aiding algorithms and display representations in supporting the requirements of domain tasks. In contrast to technical usability, which is driven by issues of technology utilization, domain suitability requires an understanding of the "cognitive problems to be solved and challenges to be met " (Woods and Hollnagel, 1987; p. 257; see also Rasmussen, 1986; Rasmussen and Goodstein, 1988). Rasmussen and Woods and their colleagues have argued extensively for a problem-driven approach to designing and evaluating decision support systems that will effectively support problem solving in large-scale, complex systems like ATC.

The fundamental basis for understanding the types of cognitive demands that can arise is a description of the domain in terms of domain goals to be achieved, the relationships between these goals, and the means for achieving goals (Rasmussen, 1985; 1986; Woods and Hollnagel, 1987; Rasmussen and Goodstein, 1988). This sort of system description, in terms of a goal-means decomposition, is particularly useful for system evaluation: it guides the description of the cognitive situations that the design must support and it guards against narrowly focusing on problem-solving demands in only one aspect of the work domain. This approach is illustrated for TMA.

A partial goal-means decomposition of the FAA Traffic Management System is shown in Figure 2. This is the domain for which TMA is designed. Sources of information for the goal decomposition were official FAA operational orders for the Traffic Management System, observation of traffic management coordinator activities at the ARTCC and terminal facilities, and discussion with Traffic Management Supervisors.

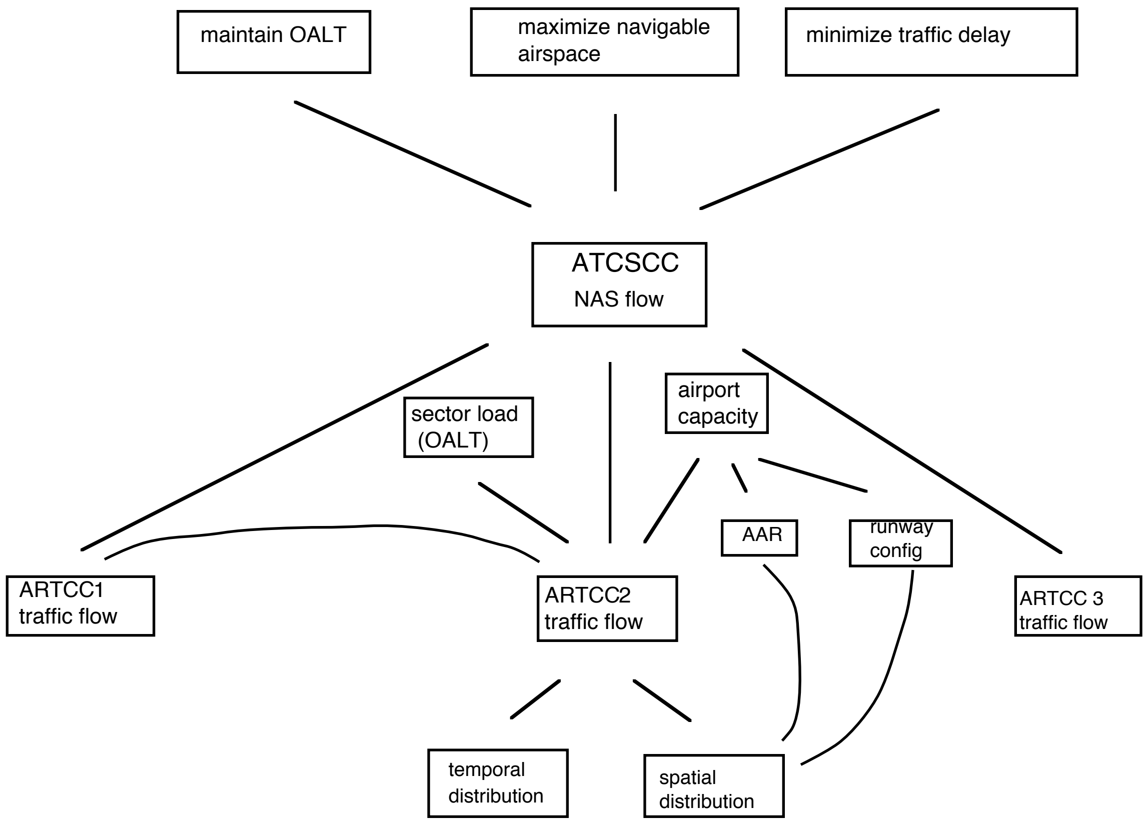

Figure 2. Partial goal decomposition of the FAA Traffic Management System

Causal relationships between goals and functions are indicated by straight lines, while relationships between functions are depicted by curved lines. Ultimate goals of the Traffic Management System are located at the top of Figure 2 --namely, maintain Operationally Acceptable Levels of Traffic (OALT), maximize navigable airspace, and minimize traffic delay. Travelling down the causal links from goals to functions indicate functions for achieving goals. For example, ultimate goals of the Traffic Management System are achieved by managing the flow of traffic in the National Airspace System (NAS). In turn, requirements for NAS flow management are satisfied by adjustments to the traffic flow at ARTCCs across the United States. Appropriate arrival flow at ARTCCs is achieved by adjusting the temporal and spatial distribution of traffic. Moving upwards from functions to goals provides reasons for conducting functions. For example, ARTCC 2 arrival flow is adjusted to meet the goals of sector load, NAS flow requirements and airport capacity. Thus goals can operate as functions and vice versa depending on the direction of travel through the causal network. Note that the Traffic Management System is described in terms that are independent of a particular technology or how the job is done. This type of description is important for understanding the kinds of "cognitive situations" to be confronted by traffic management coordinators (cf Rasmussen, 1986; Woods and Hollnagel, 1987; Rasmussen and Goodstein, 1988).

Various kinds of relationships between goals and functions are indicated by the network of linkages in Figure 2. A single function can satisfy more than one goal; for example, balancing arrival flow keeps the sector load within limits and ensures airport capacity is not exceeded. Functions for achieving one goal can constrain functions for achieving another goal; for instance, airport acceptance rate (AAR) and runway configuration constrain the spatial distribution of aircraft in the arrival sectors. Different functions for achieving the same goal can impose constraints on each other; for example restrictions to the flow in one ARTCC can affect the flow of traffic in another. Understanding these relationships is necessary for designing and verifying effective display representations and appropriate system functionality for helping the traffic management coordinator cope with the complexity inherent in ATC (Rasmussen and Lind, 1981; Rasmussen, 1986; Woods, 1988).

Using the framework provided by the goal decomposition, the next step is to elucidate the cognitive demands of the domain. Two types of cognitive demands are 1) requirements for gathering evidence about the state of system, and 2) demands posed by reasoning and problem solving situations (Woods and Hollnagel, 1987). For traffic management systems, evidence-gathering requirements for determining the state of the system fall into several areas, such as:

- What are characteristics of the traffic flow (mix of traffic, direction of arrival)?
- What is the runway configuration and terminal AAR ?
- What is the sector load? Has Operationally Acceptable Level of Traffic (OALT) been exceeded?
- What ARTCC national directives are in effect?
- Is ARTCC1 imposing any restrictions on ARTCC 2 ?

Evidence regarding the state of the system, together with specifying the relationships depicted in the the goal-means framework, suggests various problem-solving situations. Examples of situations for the traffic management coordinator are:

- Given the characteristics of the arrival flow and AAR, should the temporal distribution of traffic be adjusted (i.e., delay aircraft) ? If so, when?

- Given sector load and airport runway configuration, should aircraft be re-routed to another gate?

- Can airport terminal capacity be increased to accommodate the volume of traffic in arrival sectors?

Based on the types of evidence to be gathered and problem-solving situations, domain suitability issues for evaluation of the TMA can be specified. For example:

- Can the traffic management coordinator determine characteristics of the arrival flow from the TMA scheduling representations?

- Does the TMA representation of arrival flow and traffic load characteristics support judgments for adjusting the temporal distribution of traffic ?

- Can the traffic management coordinator determine the distribution of traffic load across arrival gates, as well as the relationship between spatial distribution of traffic and runway configuration from the TMA display representation?

- Does the TMA support decisions for spatial distribution of traffic that minimizes the impact on sector load and traffic delays?

Note that domain suitability issues are described in terms of the cognitive abilities of the traffic management coordinator (e.g., can the traffic management coordinator make judgements? decisions?), the display representations or functionality (e.g., TMA), and causal relationships of the work domain (e.g., temporal distribution as a function of arrival flow and sector load). These three factors and relationships between them define the complexity of traffic management problem-solving situations (cf Woods, 1988). Focusing on only one or two of these factors while verifying and validating a system raises the risk of collecting data that will not provide insight into system suitability for supporting "cognitive situations" (cf Woods, 1988). When complexity is the essence of a system, it must be embraced in its entirety for effective design and evaluation (Brooks, 1987).

User Acceptability

User acceptability of a new system upgrade is a key factor for determining the extent to which the upgrade will actually be used (Rasmussen and Goodstein, 1988). A general assumption is that if information about system state is easy to access and manipulate (technical usability), and if the system supports the right kinds of problem solving situations confronted by the user (domain suitability), then user acceptability will be high. Clearly, user acceptance is influenced by these two factors. However, for ATC, there is a growing awareness that "incidental consequences" of information technology and automation-- for example, effects on job satisfaction, self-esteem, and professional standing among colleagues, may also influence user acceptability of the system and ultimately system safety and efficiency (Hopkin, 1980; 1992). Validation and verification of ATC systems must acknowledge these potential "incidental consequences" of automation.

New system upgrades can affect sources of job satisfaction and opportunities for recognizing individual merit in two ways:

1. What was satisfying about the job in the current system may be disrupted by the new system upgrade.

2. New situations may emerge (as a consequence of integrating the new system upgrade into the existing system) that make the job less satisfying and preclude opportunities for individual merit.

User acceptability issues for the first case can be identified by determining sources of job satisfaction in the current system and then predicting possible impacts on these sources by the new system upgrade. The approach described next focuses on identifying such issues. Determining issues for the second case is extremely difficult since emergent properties associated with the integrated system upgrade usually elude identification until the system is actually deployed in the field.

Context is critical for understanding the impact of new system upgrades on sources of job satisfaction, individual merit, and opportunities for evaluation by colleagues. What is satisfying and motivating about a job is as much a factor of the individual as it is the nature of the tasks and work domain. Ethnographic techniques for understanding the work environment are thus instructive for capturing valid descriptions of sources of job satisfaction. Such techniques are geared to the study of complex social settings to understand what aspects of activities are important and relevant to individuals. In general, ethnographic techniques have been recognized as essential to understanding, designing, and evaluating complex systems (e.g., Suchman, 1987; Whiteside, Bennet, and Holtzblatt, 1988; Suchman and Triggs, 1991; Hutchins, 1992).

Efforts to understand the user's work environment require a careful balance between the different frames of reference for observation. A description of the situation should be obtained "as the native sees it, " looking from the inside out, as well as from the observer's perspective looking from the outside in (Sanjek, 1990). Video records of situations for later analysis ("inside out" observation) as well as contextual interviews (Whiteside, Bennet, and Holtzblatt, 1990) ("outside in" observation) are helpful in this regard. In addition, different shifts should be sampled as well as the variety of individuals who will ultimately use the system upgrade--for example, supervisors, area managers, and traffic management coordinators (Johnson and Johnson, 1990).

From observations of the Traffic Management Unit to date, what is highly satisfying about the job of the traffic management coordinator is creating a plan to manage a disruption to arrival flow, negotiating with facilities to change various parameters to modify the flow, and keeping different parties happy by striking an equitable balance of restrictions across facilities. Having a plan "work" is a great source of pride and satisfaction as well as an opportunity to reveal individual abilities to colleagues. This was demonstrated particularly well during a weather disturbance.

A line of thunderstorms had moved into the arrival area from the south-west and was heading toward the airport. Two of the four arrival gates and the south departure gate had to be closed. Active discussion of a plan for handling the disturbance ensued between a junior and senior traffic manager and supervisor. Several options were actively considered: rerouting arrival aircraft and lowering their altitude to skirt the weather (this option was suggested by the junior traffic manager who had been actively monitoring the situation display and activity in the arrival sectors), placing a restriction on aircraft that were still outside the center's airspace, and re-routing a portion of the arrival flow to the north-west arrival gate. The option suggested by the junior traffic manager was selected and later commended by the supervisor as it eliminated the need for restricting aircraft outside of the center's airspace. During the planning activities, the supervisor had been negotiating with Terminal personnel to raise the airport acceptance rate slightly to allow aircraft in the arrival sectors to come out of holding, thereby relieving congestion in these sectors. After much discussion and analysis of the traffic situation with the supervisor, the Terminal agreed to raise the rate. When the supervisor reported the new rate to the other traffic managers they looked at each other, obviously impressed, and one exclaimed "How did you do that? " He went on to say that he could never get a rate change. The supervisor said, with feigned humility, " What can I say?" and then slapped hands with the traffic manager in a victory fashion.

Understanding sources of job satisfaction in the current system is necessary for assessing potential user acceptance of the new system. Some examples of user acceptability issues for the TMA are listed below. Issues are presented from the perspective that those aspects of the job that are satisfying and provide opportunity for individual merit should be supported or enhanced by the new system upgrade. This perspective is in keeping with the FAA policy for gradually evolving of system enhancement with minimal disruption to operational personnel (Hunt and Zellweger, 1987) and also in general with principles of human-centered automation (Billings, 1992) and system design (Norman and Draper, 1986).

- Does the TMA support planning for handling disruptions to traffic flow? What level of planning is required by the traffic management coordinator when using the TMA?

- Does the TMA support effective negotiations with external facilities regarding modifications to the traffic flow by making apparent appropriate parameters for flow modification?

- Does the TMA facilitate TMC decisions regarding equitable restrictions to traffic flow across facilities?

The context of the user's job must be acknowledged for deriving meaningful user-acceptance issues. The current proposed effort by the FAA TATCA Program Office for conducting the final stages of developing and evaluating of a system upgrade in the field recognizes the importance of context. One of the many benefits of this approach is the considerable opportunity it will provide for grasping issues associated with the impact of system upgrades on job satisfaction.

Conclusion

Before criteria and measures can be specified for verifying and validating ATC systems, an explicit definition is required of the issues associated with technology upgrades. Quite simply, we need to know what aspects of the system should be measured before we measure the system. This claim may be stating the obvious, but achieving it is perhaps one of the most important and challenging steps in evaluating ATC systems. Performance criteria for complex ATC systems are not obvious. We lack detailed knowledge of the kinds of problem-solving situations confronting controllers, of the social aspects of the work context, of job performance by individual controllers and controller teams in current and future ATC environments, and of the impact and consequences of automation on controller management of traffic. Thus, to compensate, but not by-pass this knowledge gap, considerable effort must be devoted to elucidating human-centered issues associated with ATC system upgrades.

A cookbook approach that describes how to evaluate complex systems will not guarantee that the appropriate data are collected. Instead, a process is required for mediating the right kinds of evaluation questions. The tripartite framework proposed in this paper is offered as one such approach and is generalizable beyond ATC. Technical usability, domain suitability, and user acceptability provide multiple perspectives of the user's experience in a complex system. A system may be technically usable but not suitable for the domain, and even if it is both of these things, it may not be readily accepted by the user. Thus, issues for all three of these human-center system categories must be considered.

Methods and techniques were suggested for identifying human-centered issues in each of the three categories of the framework. The methods described are not exhaustive of all possible methods, and others may be appropriate. Whatever the method or technique, it must generate issues at an intermediate level of describing of the system. Ultimate system goals such as safe, expeditious, and orderly flow of traffic are too general to be assessed directly. Instead, issues for evaluation must be couched at a lower level, in terms of the relationships between characteristics of the user, features of the system upgrade, and aspects of the domain environment (Rasmussen, 1986; Woods, 1988). Examples of issues that were provided for the TMA illustrated this level of description. In addition, it is important that methods for disclosing issues must be contextually based; that is, they are grounded in an understanding of the physical environment, domain, and work activities. Issues that are detached from the context of the system will most likely result in data that are irrelevant for validating and verifying complex systems.

In discussing present and future trends in human factors, Christensen stated that "...the criterion problem will never be completely resolved" (Christensen, 1958; p. 3). However, approaches that mediate the process of defining appropriate issues for evaluation hold promise for coping with the criterion problem. Effort must be directed at defining meaningful human-centered system issues prior to evaluating complex systems. Such issues are essential for identifying criteria and measures that will help guide the collection of data for supporting informed decisions on ultimate system safety and efficiency.

Acknowledgements

The author acknowledges the interest and support of Peter Challan, ARD-40, FAA TATCA Program Manager, and Rick Page, ACD-340, FAA Technical Center TATCA Senior Technical Program Manager. C. Halverson, E. Murphy, and L. Tobias are also acknowledged for their insightful comments on earlier versions of this paper.

References

ATC Systems Field Office (1992). Traffic Management Advisor (TMA) Reference Manual. Moffett Field, CA: NASA-Ames Research Center.

Billings, C.E. (1992). Human-Centered Aircraft Automation: A Concept and Guidelines. NASA Technical Memorandum 103885. Moffett Field, CA: NASA-Ames Research Center.

Boff, K.R. and Lincoln, J.E. (Eds.) (1988). Engineering Data Compendium. Wright Patterson Air Force Base, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.

Brooks, F.P., Jr. (1987). No silver bullet: Essence and accidents of software engineering. Computer, 20, 10-19.

Christensen, J.M. (1958). Trends in human factors. Human Factors, 1 (1), 2-7.

Davis, T.J., Erzberger, H., and Green, S.M. (1991). Design and evaluation of air traffic control final approach spacing tool. NASA Technical Memorandum 10287. Moffett Field, CA: NASA-Ames Research Center.

Department of Defence (1989). Military Standard: Human Engineering Design Criteria for Military Systems, Equipment, and Facilities. MIL-STD-1472D. Washington, D.C.: Department of Defence.

Federal Aviation Administration. (1989). FAA NAS Test and Evaluation Program. ORDER No. 1810.4A. Washington, D.C.: Federal Aviation Administration.

Federal Aviation Administration. (1990). The National Plan for Aviation Human Factors (Draft). Washington, D.C.: Federal Aviation Administration.

Erzberger, H. and Nedell, W. (1989). Design of Automated Systems for Management of Arrival Traffic. NASA Technical Memorandum 102201. Moffett Field, CA: NASA-Ames Research Center.

Fitts, P.M. (Ed.). (1951). Human Engineering for an Effective Air-Navigation and Traffic Control System. Washington, D.C.: National Research Council.

Hollnagel, E. and Woods, D.D. (1983). Cognitive systems engineering: new wine in new bottles. International Journal of Man-Machine Systems, 18, 583-600.

Hopkin, V.D. (1980). The measurement of the air traffic controller. Human Factors, 22 (5), 547-560.

Hopkin, V.D. (1991). Automated flight strip usage: Lessons from the functions of paper strips. In Challenges in Aviation Human Factors: The National Plan, Book of Abstracts. Washington, D.C.: AIAA. (pp.62-64).

Hopkin, V.D. (1992). Human factors issues in air traffic control Human Factors Society Bulletin, 35 (6). Santa Monica, CA: Human Factors Society.

Hunt, V.R. and Zellweger, A. (1987). The FAA's advanced automation system: Strategies for future air traffic control systems, Computer, 20, 19-32.

Hutchins, E. (1991). How a Cockpit Remembers its Speed. Technical Report, Distributed Cognition Laboratory, University of California at San Diego.

Johnson, A. and Johnson, O.R. (1990). Quality and Quantity: On the measurement potential of ethnographic fieldnotes. In R. Sanjek (Ed.), Fieldnotes: The Making of Anthropology. Ithaca, NY: Cornell University Press. (pp. 161-186).

Kloster, G.V. and Zellweger, A. (1987). Engineering the man-machine interface for Air Traffic Control. Computer, 20, 47-62.

Meister, D. (1985). Behavioral Analysis and Measurement Methods. New York, NY: John Wiley and Sons

Norman, D.A. and Draper, S.W. (1986). User-Centered System Design: New Perspectives on Human Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Parsons, M. (1972). Man Machine System Experiments. Baltimore, MD: The Johns Hopkins Press.
Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. IEEE Transaction on Systems, Man, and Cybernetics, 15, 234-243.
Rasmussen, J. (1986). Information processing and Human-Machine Interaction: An Approach to Cognitive Engineering. Amsterdam: North-Holland.
Rasmussen, J. and Goodstein, L.P. (1988). Information technology and work. In M. Helander (Ed.), Handbook of Human-Computer Interaction. Elsevier Science Publishers BV (North Holland) : New York. (pp. 175-201)
Rasmussen, J. and Lind, M. (1981). Coping with complexity. In H.G. Stassen (Ed.), First Annual European Conference on Human Decision-Making and Manual Control. New York: Plenum Press. (pp. 70-91).
Sanjek, R. (1991). The secret life of fieldnotes. In R. Sanjek (Ed.), Fieldnotes: The Making of Anthropology. Ithaca, NY: Cornell University Press. (pp. 187-270).
Shneiderman, B. (1987). Designing for the User Interface. Reading, MA: Addison-Wesley.
Smith, S.L. and Mosier, J.N. (1986). Guidelines for designing user interface software. (Technical Report NTIS No. A177 198). Hanscom Air Force Base, MA: USAF Electronic Systems Division.
Suchman, L.A. (1987). Plans, and Situated Actions: The Problem of Human Machine Communication Cambridge, MA: Cambridge University Press.
Suchman, L.A. and Trigg, R.H. (1991). Understanding practice: Video as a medium in reflection and design. In J. Greenbaum and M.Kyng (Eds.), Design at Work. Hillsdale, NJ: Lawrence Erlbaum Associates. (pp. 65-89).
Tobias, L., Volckers, U., and Erzberger, H. (1989). Controller evaluations of the descent advisor automation aid. In Proceedings of the AIAA Guidance, Navigation , and Control Conference. Washington, D.C.: AIAA. (pp. 1609-1618).
Van Cott, H.P. and Kincaid, R.G. (1972). Human Engineering Guide to Equipment Design. Washington, D.C.: U.S. Government Printing Office.
Whiteside, J., Bennet, J., and Holtzblatt, K. (1988). Usability engineering: Our experience and evolution. In M. Helander (Ed.), Handbook of Human-Computer Interaction. Elsevier Science Publishers BV (North Holland): New York. (pp. 791-817)
Whitfield, D., and Stammers, R.B. (1978). The air traffic controller. In W.T. Singleton (Ed.), The Study of Real Skills: Volume 1 The Analysis of Practical Skills. Lancaster, U.K.: MTP Press. (pp. 209-235).
Wickens, C.D. (1992). Engineering Psychology and Human Performance, Second Edition. New York, NY: Harper Collins Publishers.
Woods, D.D. and Hollnagel, E. (1987). Mapping cognitive demands in complex problem-solving worlds. International Journal of Man-Machine Systems, 26, 257-275.
Woods, D.D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L.P. Goodstein, H.B. Andersen and S.E. Olsen (Eds.), Tasks, Errors and Mental Models. London: Taylor Francis. (pp. 128-148).