

NASA Contractor Report 177617

Cognitive Engineering in Aerospace Application: Pilot Interaction with Cockpit Automation

Nadine B. Sarter and David D. Woods

Cognitive Systems Engineering Laboratory
Department of Industrial and Systems Engineering
Ohio State University
Columbus, OH 43210

Prepared for
Ames Research Center
CONTRACT NCC2-592
August 1993



National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

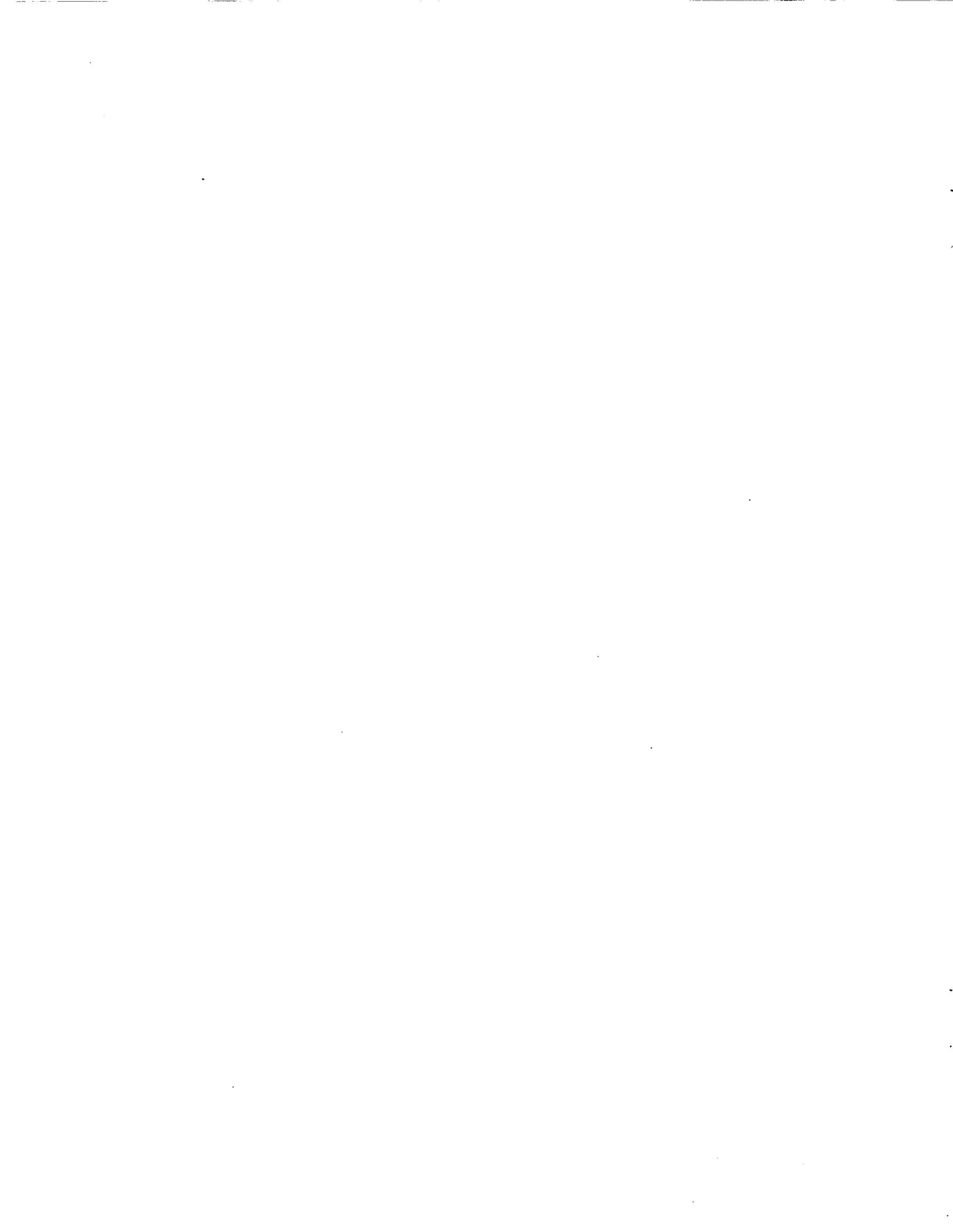


Table of Contents

	Page
Introduction	1
Cognitive Systems in Context	2
Pilot Interaction with Cockpit Automation: Operational Experiences with the Flight Management System	10
Pilot Interaction with Cockpit Automation II: An Experimental Study of Pilots' Model and Awareness of the Flight Management System (FMS)	29
"How in the world did I ever get into that mode?" Mode Error and Awareness in Supervisory Control	51
Acknowledgment	64
Glossary	65
Project Publications	66

PRECEDING PAGE BLANK NOT FILMED

Introduction

This Contractor Report starts out with a paper providing an overview of the Cognitive Systems Engineering approach. Following this introduction, two papers present the results of a series of complementary studies of pilot interaction with one of the core systems of cockpit automation – the Flight Management System (FMS)¹.

These studies apply concepts and techniques from Cognitive Systems Engineering to problems of human interaction with advanced automated systems in aerospace applications. Thus, the results have implications for the aviation community. They expand on Wiener's concept of 'clumsy automation' and may therefore be of great interest to designers of flight deck automation and designers of training programs for glass cockpit pilots.

At a more general level, exploring issues surrounding human-automation interaction in the context of advanced flight decks provides new insights that add to and extend the research base relevant to Cognitive Systems Engineering. They speak to issues such as human-computer cooperation in event-driven domains and techniques for studying cognitive systems in the field.

The fourth paper in this report addresses one of the challenges for design and training in aviation as well as in a variety of other domains, the problem of mode awareness and mode error. The need for a better understanding of the processes underlying these phenomena and for a revision of the concept in the face of increasingly automated environments is emphasized.

The studies reported here are one part of a larger program of research designed to advance the foundations of Cognitive Systems Engineering in order to bring new concepts and techniques to bear on questions surrounding human-centered automation in aerospace applications.

¹Other papers on related research about human interaction with advanced technology produced by the researchers in the Cognitive Systems Engineering Laboratory can be obtained by contacting: David Woods, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH 43210; Fax: 614-292-7852; Email: woods@csel.eng.ohio-state.edu

Cognitive Systems in Context

- Joint Cognitive Systems and Research on Human-Machine Systems -

David D. Woods

Cognitive Systems Engineering Laboratory
Department of Industrial and Systems Engineering
The Ohio State University

NEW INFORMATION TECHNOLOGY IS A DOUBLE-EDGED SWORD

Many organizations have experienced significant difficulties in turning AI research and other new developments in computational technology into systems that actually improve performance in the target field of practice (e.g., space, flightdecks, air traffic control, nuclear power plant control rooms, communication network management, ground satellite control stations). Flexibility and customizability are the hallmarks of today's information technology as compared to previous media for data handling and display. First, ubiquitous computerization has tremendously advanced our ability to collect, transmit and transform data. In all areas of human endeavor, we are bombarded with computer processed data, especially when anomalies occur. Second, user interface technology has allowed us to concentrate this expanding field of data into one physical platform, typically a single VDU, by providing the capability to manage multiple windows and the capability to generate tremendous networks of computer displays as a kind of virtual perceptual field viewable through the narrow aperture of the VDU. Third, heuristic and algorithmic technologies expand the range of subtasks and cognitive activities that can be automated. These intelligent space machines create joint cognitive systems that distribute cognitive work across multiple agents (Woods, 1986; Roth, Bennett and Woods, 1987; Hutchins, 1990).

But these and other vectors of technological change can create new burdens and complexities for the human practitioners responsible for achieving goals within some field of activity (Woods, Cook and Sarter, 1992; Woods, 1993). Our ability to digest and interpret data has failed to keep pace with our abilities to generate and manipulate greater and greater amounts of data. Thus, one result of the vectors of technology change is the data overload syndrome. Practitioners get lost in large networks of computer based display frames; they experience keyhole effects in trying to monitor dynamic systems through the narrow aperture of even a windowed computer screen; they fail to focus on the critically relevant data in a specific context out of the massive field of available data (Woods, in press). It has turned out that using new computational possibilities to create effective human-machine ensembles, what we will refer to as joint or distributed cognitive systems, is a substantive issue at the intersection of cognitive psychology, software engineering, social psychology and artificial intelligence (Hollnagel and Woods, 1983; Woods, 1986).

COGNITIVE SYSTEMS ENGINEERING

Cognitive Systems Engineering is the emerging discipline which directly confronts the question of how to use the huge space of possibilities provided by today's computational power (Hollnagel and Woods, 1983; Woods and Roth, 1988). This means that cognitive engineering research focuses on how people solve problems with tools. This is in contrast to the typical model for laboratory research which continues to examine human problem solving in the absence of even simple tools, e.g., pencil and paper. The tools in question include different kinds of external representations of the domain via control boards, display panels, and computer displays/controls and different kinds of support systems such as AI diagnostic systems, alerting and alarm systems, and new forms of information.

Cognitive Work in the Wild

Technology change is one departure point for seeing the role of Cognitive Systems Engineering. Another departure point is to think about cognition in the wild, particularly fields of activity that involve advanced technology. If we look at flightdecks of commercial jet airliners or control centers that manage space missions, or surgical operating rooms or control rooms that manage chemical or energy processes or control centers that monitor telecommunication networks or many other fields of human activity, what do we see?

First, we do not see cognitive activity isolated in a single individual, but rather cognitive activity goes on distributed across multiple agents (Hutchins, in press). Second, we do not see cognitive activity separated in a thoughtful individual, but rather as a part of a stream of activity (Klein et al., 1992). Third, we see these sets of active agents embedded in a larger group, professional, organizational, institutional context which constrains their activities, sets up rewards and punishments, defines not altogether consistent goals, and provides resources. Fourth, we see phases of activity with transitions and evolutions. Cognitive and physical activity ebbs and flows, with periods of lower activity and more self paced tasks interspersed with busy, high tempo, externally paced operations where task performance is more critical. Higher tempo situations create greater need for cognitive work and at the same time often create greater constraints on cognitive activity (e.g., time pressure, uncertainty, exceptional circumstances, failures and their associated hazards). We see that there are consequences at stake for the individuals and the groups and organizations involved in the field of activity or affected by that field of activity – economic, personal, safety.

Fifth, even a causal glance at these domains reveals that tools of all types are everywhere; almost all activity is aided by something or someone beyond the unit of the individual cognitive agent. More in depth observation reveals that the technology is often not well adapted to the needs of the practitioner – that much of the technology is clumsy in that it makes new demands on the practitioner, demands that tend to congregate at the higher tempo or higher criticality periods (Woods, 1993). Close observation reveals that people and systems of people (operators, designers, regulators, etc.) are adapting the tools and adapting their activities continuously to respond to indications of trouble or to meet new demands. Furthermore, new machines are not used as the designers intended, but are shaped by practitioners to the contingencies of the field of activity in a locally pragmatic way (Woods et al., 1992).

Cognitive Systems

The reverberations of technology change and observations of cognitive work in the wild lead us to the basic point of departure for a Cognitive Systems Engineering. This is the idea suggested by Hollnagel and Woods (1983) and Hutchins (1991) among others that one can look at operational systems – the individual people, the organization both formal and informal, the high technology artifacts (AI, automation, intelligent tutoring systems, computer-based visualization) and the low technology artifacts (displays, alarms, procedures, paper notes, training programs) intended to support human practitioners – that one can look at all of these things as a single cognitive system.

Operational systems can be thought of as joint and distributed human-machine cognitive systems in that:

1. one can describe and study and design these systems in terms of cognitive concepts such as information flow, knowledge activation, control of attention, etc.,
2. cognitive systems are distributed over multiple agents both multiple people and mixtures of people and apparently animate, agent-like machines,
3. external artifacts function as cognitive tools that modify the activities of agents within the cognitive system,

There is a reciprocal relationship or mutual shaping between properties of external artifacts and representations of aspects of the field of activity and the cognitive activities distributed within the cognitive system. Properties of these artifacts and representations shape practitioner cognitive strategies and in turn these artifacts are shaped by practitioners to function as tools within the field of activity.

4. cognitive systems adapt to the demands of the field of practice.

Hence, the cognitive systems perspective can be summarized by the triad – cognition in context, cooperation, and tools.

COGNITIVE SYSTEMS AND CONTEXT

In Cognitive Systems Engineering we are concerned with cognitive work within complex fields of practice – highly coupled domains, event-driven dynamic situations, the possibility of multiple interacting faults. This means that a cognitive systems approach is context bound – the study of cognition is intimately linked to the study of the situation in which it occurs. To do this requires studies of specific meaningful tasks as carried out by actual practitioners. However, to say that the study of human-machine systems can and should be context bound is not simply to call for more applied studies in particular domains. It is, ..., the fundamental principle of cognition that the universal can be perceived only in the particular, while the particular can be thought of only in reference to the universal" (Cassirer, 1953, p. 86). As Hutchins puts it, "There are powerful regularities to be described at a level of analysis that transcends the details of the specific domain". It is not possible to discover these regularities without understanding the details of the domain, but the regularities are not about the domain specific details, they are

about the nature of human cognition in human activity."² This reveals the proper complementarity between so called basic and applied work where the experimenter functions as designer and the designer as experimenter (Woods, 1992). 'New technology is a kind of experimental investigation into fields of ongoing activity. If we truly understand cognitive systems, then we must be able to develop designs that enhance the performance of operational systems; if we are to enhance the performance of operational systems, we need conceptual looking glasses that enable us to see past the unending variety of technology and particular domains (Woods and Sarter, in press).

The experimenter as designer? Cognitive tools are ubiquitous; technology change implicitly changes cognitive systems by changes in agent-like machines and by changes in artifacts that constrain cognitive work. This means new technology is a kind of experimental manipulation that can be exploited to help understand human cognition as expressed in meaningful fields of practice.

The designer as experimenter? The possibilities of technology seem to afford designers great degrees of freedom. The possibilities seem less constrained by questions of feasibility and more by concepts about how to use the possibilities skillfully to meet operational and other goals. In other words, in order for designs to be developed in a problem-driven manner, as opposed to the more typical technology-driven fashion, in order for designs to provide real and not illusory benefits for real operational systems, the designer must adopt the attitude of an experimenter trying to understand and model of the dynamics of cognitive systems.

To conceive of the experimenter as designer and designer as experimenter requires a drastic shift in the normal attitude of researchers and developers towards the subjects of their work – people and technology. Instead of separate and independent topics they are intimately interconnected as parts of a larger and more useful system boundary – a joint cognitive system. To conceive of the experimenter as designer and designer as experimenter shifts the relationship between 'basic' and 'applied' research. The above concepts mean that these activities are complimentary where growing the research base and developing effective applications are mutually inter-dependent (Woods, 1992).

Cognitive Systems and Complexity

In Cognitive Systems Engineering concepts about cognition that were developed in simpler task environments may be used to begin to chart more complex fields of human activity. But the target is cognition in context which means a much greater emphasis on research techniques honed for field or relatively high fidelity simulation settings. In more laboratory-like studies, tasks must be modeled after naturally occurring ones; subjects must receive extensive practice and possess significant domain knowledge; interface and support systems must be representative of actual or proposed systems. There is an assumption that research results derived from studies of simple situations will transfer directly to more complex situations. However, simple situations may leave out parts of the problem solving process or underestimate their importance. Furthermore, some aspects of problem solving may only emerge when more complex situations are directly examined. Cognitive Systems Engineering directly studies cognition in context.

²Hutchins, 1992, personal communication

The ultimate target of Cognitive Systems Engineering research is the domain practitioner. From a practical point of view, what tools/resources can help him/her to be a better performer? Usually this person will be experienced and domain knowledgeable. The usual distinction between novice and expert is over simplistic. This means that Cognitive Systems Engineering must use a different subject population than the typical subject of psychology experiments—either experienced, domain knowledgeable practitioners, people who are similar to this group (e.g., engineering graduate students who receive extensive practice at a highly realistic task environment), or people who contrast practitioners on some important dimension (e.g., engineering staff may possess more detailed theoretical knowledge about a system than the operational staff).³

Artifacts and Tools

Cognitive Systems Engineering is sine qua non for human cognition with tools. Again, tools are ubiquitous in real cognitive systems, and these tools influence the cognitive activities of practitioners within a cognitive system. For example, cognitive science has shown repeatedly that the representation of the problem provided to a problem solver can affect his, her or its task performance (cf., Woods, in press for the implications of this for designing cognitive systems). Thus, understanding cognitive systems means understanding how properties of artifacts and representations influence and transform cognitive activities (e.g., Hutchins, 1991).

Technology change is one of the driving forces for a cognitive systems approach. New technology introduced for putative benefits in fact often introduces new demands and complexities into already highly demanding fields of practice. Practitioners develop and use a variety of strategies to cope with these new complexities. In other words, practitioners adapt information technology provided for them to the immediate tasks at hand in a locally pragmatic way, usually in ways not anticipated by the designers of the information technology (Roth et al., 1987; Flores et al., 1988; Cook et al., 1990; Hutchins, 1990). While characteristics of artifacts shape practitioner strategies, it is important to remember that tools are shaped by their users (Woods et al., 1992). One of the goals of a cognitive systems approach is to understand how practitioners, individually, as a group, and as operational organizations, shape artifacts to function as cognitive tools.

The above means that research in this area must involve the availability of significant tools that differ in relation to their impact on the practitioner's information processing resources and strategies (Woods and Sarter, in press). The representations of the problem domain available to the practitioner can degrade or support information processing tasks and strategies related to task performance. Thus, there is an increasingly pressing need "... with developing a theoretical base for creating meaningful artifacts and for understanding their use and effects" (Winograd, 1987, p.10), in other words, to develop a theoretical base for representation design in the computer medium (Woods, in press). Thus, we must confront the difficult problem of the interaction and inter-dependence of investigations and design. Effective design depends on the results of investigations of the cognitive activities in the field of practice. But how does one utilize this information in design (a) when the new artifacts transform the cognitive system in question potentially in radical ways and (b) while recognizing that design is an open-ended

³Note the use of the word participant rather than the traditional term of subject in recognition of the need to be relevant to practitioners functioning in their field of activity.

process of discovering a space of possibilities that balance multiple interacting or competing constraints on what would be useful?

STUDYING COGNITIVE SYSTEMS IN THE WILD

Research on problem solving in more complex situations, where significant tools are available to support the human and where experienced domain knowledgeable people are the appropriate study participants, requires a shift in research methodology from typical laboratory studies. Natural history techniques (critical incidents, corpus building, direct observation 'in situ,' elicitation of practitioner descriptions) are an important element for identifying and abstracting the phenomena of joint cognitive systems. Cognitive simulation is a tool that can help model data about these phenomena and to stimulate more directed studies (Woods and Roth, in press). 'Field experiments' investigate cognitive systems in scaled worlds shaped by (but not created by) the investigators, where the situation where behavior is under investigation is seen as a particular instance of a larger class of related situations over which the results can be generalized (Woods, 1992).

These techniques do not mean that rigor or control or generalizability always must be sacrificed to the goal of relevance. Rather it assumes that there are research approaches that can achieve both generalizability and relevance. If we abandon methodological imperialism and adopt a wide set of research tools, then we can identify recurring phenomena involving joint cognitive systems and converge on an understanding of the factors involved in these phenomena including how to control the phenomenon (i.e., how properties of artifacts, automation or support systems exacerbate it and other types eliminate or reduce it).

A common complaint in cognitive psychology is that the research is driven too much by experimental paradigms or procedures (e.g., dual task studies, the Sternberg paradigm). A parallel common complaint in human-computer interaction is that research is driven too much by the current technology (the current example in a long history is the question of tiled versus overlapping windows). Similarly, research in developing machine cognitive systems (automation, intelligent systems) has been framed in terms of the technologies from which the systems are built. For example, knowledge engineering techniques are usually described in terms of the syntax of computational mechanisms, i.e., the language of implementation is used as a cognitive language. As a result, questions about what would be effective assistance are displaced either to whomever selects the computational mechanisms or to the domain expert who enters knowledge. A cognitive systems approach directly attacks the question of what would be useful as a necessary complement to technology driven approaches.

If the context bound, cognitive systems approach is to succeed, it must meet several challenges that go beyond past efforts in work analysis (De Keyser, 1990) or ethnography: (a) methods and results for building models of how joint and distributed cognitive systems function in fields of practice, (b) methods and theory building to generate generalizable results from context bound studies (producing distilled results transportable across scenarios, participants and domains rather than just diluted motherhood generalizations), (c) methods and theories to stimulate critical growth of knowledge across specific context bound studies, (d) establish the complementary link between growing the research base and using it in the process of design of cognitive artifacts to enhance the performance of operational systems. If we can achieve this complementarity between what is usually seen as separate domains of basic and applied work, then we can 'ascend to the particular.'

REFERENCES

- Cassirer, E. (1953). *The Philosophy of Symbolic Forms, Vol. 1: Language*. New Haven CT: Yale University Press. (Translated by R. Manheim, original work published 1923).
- Cook, R.I., Woods, D.D. and Howie, M.B. (1990). The natural history of introducing new information technology into a dynamic high-risk environment. In *Proceedings of the Human Factors Society 35th Annual Meeting*. Orlando, FL.
- De Keyser, V. (1990). Why field studies? In M. Helander (Ed.), *Human Factors in Design for Manufacturability and Process Planning*. International Ergonomics Association.
- Flores, F., Graves, M., Hartfield, B. and Winograd, T. (1988). Computer systems and the design of organizational interaction. *ACM Transactions on Office Information Systems*, 6:153-172.
- Hollnagel, E. and Woods, D.D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, 18:583-600.
- Hutchins, E. (1990) The Technology of Team Navigation. In J. Galegher and R. Kraut, and C. Egido (Eds.), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hutchins, E. (1991) *How a Cockpit Remembers Its Speed*. (Technical Report). Distributed Cognition Laboratory, San Diego: University of California.
- Hutchins, E. (in press). *Distributed Cognition*.
- Klein, G., Orasanu, J. and Calderwood, R. (Eds.).(1992). *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex.
- Roth, E.M., Bennett, K. and Woods, D.D. (1987). Human interaction with an intelligent machine. *International Journal of Man-Machine Studies*, 26, 257-275.
- Winograd, T. (1987) *Three responses to situation theory*. Technical Report CSLI-87-106, Center for the Study of Language and Information, Stanford University.
- Woods, D.D. and Roth, E.M. (1988). Cognitive Systems Engineering. In M. Helander, (Ed.) *Handbook of Human-Computer Interaction*, New York, NY: North-Holland.
- Woods, D.D. and Roth, E.M. (in press). Symbolic AI-Based Computer Simulations as a Tool for Investigating the Dynamics of Joint Cognitive Systems. In P.C. Cacciabue, B. Pavard and J-M. Hoc (Eds.) *Simulation of Cognition in Dynamic Environments*. Academic Press.
- Woods, D.D. and Sarter, N.B. (in press). Evaluating the Impact of New Technology on Human-Machine Cooperation. In J. Wise, V. D. Hopkin, and P. Stager (Eds.) *Verification and Validation of Complex and Integrated Human-Machine Systems*, Springer-Verlag: New York.

- Woods, D.D., Cook, R.I. and Sarter, N.B. (1992). *Clumsy Automation, Practitioner's Adaptive Response and Human Error*. Cognitive Systems Engineering Laboratory Report, The Ohio State University, Columbus OH.
- Woods, D.D. (1986). Cognitive Technologies: The design of joint human-machine cognitive systems. *AI Magazine*, 6:86-92.
- Woods, D.D. (1992). Process Tracing Methods for the Study of Cognition Outside of the Experimental Psychology Laboratory. In G. A. Klein, J. Orasanu and R. Calderwood (Eds.). *Decision Making in Action: Models and Methods*, New Jersey: Ablex.
- Woods, D.D. (1993). The price of flexibility. In W. Gray, W. Hefley and D. Murray, (Eds.) In *Proceedings of the 1993 International Workshop on Intelligent User Interfaces*, New York, ACM Press.
- Woods, D.D. (in press). Towards a Theoretical Base for Representation Design in the Computer Medium: Ecological Perception and Aiding Human Cognition. In J. Flach, P. Hancock, J. Caird, and K. Vicente (Eds.) *The Ecology of Human-Machine Systems*.

Pilot Interaction With Cockpit Automation: Operational Experiences with the Flight Management System¹

Nadine B. Sarter and David D. Woods

Cognitive Systems Engineering Laboratory
Department of Industrial and Systems Engineering
The Ohio State University

ABSTRACT

Because of recent incidents involving glass-cockpit aircraft, there is growing concern with cockpit automation and its potential effects on pilot performance. However, little is known about the nature and causes of problems that arise in pilot-automation interaction. In this paper, we report the results of two studies that provide converging, complementary data on pilots' difficulties with understanding and operating one of the core systems of cockpit automation, the Flight Management System (FMS). A survey asking pilots to describe specific incidents with the FMS and observations of pilots undergoing transition training to a glass cockpit aircraft served as vehicles to gather a corpus on the nature and variety of FMS-related problems. The results of both studies indicate that pilots become proficient in standard FMS operations through ground training and subsequent line experience. But even with considerable line experience, they still have difficulties tracking FMS status and behavior in certain flight contexts, and they show gaps in their understanding of the functional structure of the system. The results suggest that design-related factors such as opaque interfaces contribute to these difficulties which can affect pilots' situation awareness. The results of this research are relevant for both the design of cockpit automation and the development of training curricula specifically tailored to the needs of glass cockpits.

INTRODUCTION

There is growing concern with the potential effects of increasing levels of cockpit automation on pilots' performance. These effects seem to be related to the fact that automation changes the nature of the pilot's role on the flight deck. Pilots become system managers who are monitoring systems and who intervene only when changes are necessary or unanticipated situations occur (Billings, 1991). Instead of handflying the airplane, pilots act indirectly through instructions to the automation in order to control the aircraft. This may remove the pilot from the loop decreasing system awareness, especially if feedback on automation status and behavior is limited.

Despite the growing interest in pilot-automation interaction, only limited empirical data are available about the nature of problems that occur with the current generation of automated cockpit systems. Pilot reports to the Aviation Safety Reporting System

¹Paper appeared in *The International Journal of Aviation Psychology*, 2(4), pp. 303-321, 1992.

(ASRS) have been analyzed but these data are limited to a subset of incidents that were severe enough to threaten safety (e.g. Eldredge et al., 1991). Some analyses of selected incidents have been conducted (e.g. Norman, 1990) in the context of larger theoretical treatments of human-automation interaction. Questionnaire techniques have been used to obtain ratings about glass cockpit pilots' attitudes and opinions concerning current cockpit automation (e.g. Wiener, 1989; Lyall, personal communication²; James et al., 1991; "Automation Comment", 1991³). While these rating data provide interesting suggestions, they do not reveal the dynamics of pilot-automation interaction or specific areas of difficulty. Also, these pilot opinion data would be more informative if they were complemented by observational data.

We utilized two complementary approaches to obtain data about pilot-interaction with one of the core systems of cockpit automation, the Flight Management System (FMS). In one study, we asked pilots to describe in detail problems or incidents that they had experienced with the FMS, especially ones where FMS behavior surprised them. The corpus of incidents generated by this self-report technique was analyzed to identify the nature of pilot difficulties and the flight contexts in which they occurred.

In a second study, we observed crews transitioning to a glass cockpit aircraft during a number of line-oriented simulation (LOS) on a fixed-base simulator. Pilot-FMS interaction and crew-instructor communications during and after the simulated scenarios were analyzed to identify difficulties in pilot-FMS interaction. The two studies are complementary because they use different "corpus gathering" techniques, and because they sample both experienced glass cockpit pilots and experienced pilots in transition to a glass cockpit aircraft.

The results of this research add to a better understanding of the effects of flight deck automation on pilot performance. This data may be helpful in the design of future flight decks by pointing at specific sources of difficulty such as poor feedback on automated system status and behavior. The results can also be used to refine and expand training programs for glass cockpit aircraft. They provide information on specific FMS modes, functions, and flight situations where pilot-FMS interaction is most troublesome.

INTRODUCTION TO THE FLIGHT MANAGEMENT SYSTEM

The following section provides a brief, simplified overview of the Flight Management System (FMS). The FMS supports pilots in a variety of tasks such as flight planning, navigation, performance management, and flight progress monitoring. One of its major functions, and the function of primary interest in the context of the reported studies, is automatic flight path control.

The major FMS controls in the cockpit are the Mode Control Panel (MCP) and the multifunction keyboards of two Control Display Units (one for each pilot). FMS-related cockpit displays are the CDU multifunction display, two Attitude Director Indicators (ADI), and two Horizontal Situation Indicators (HSI). Figure 1 illustrates the typical

² Beth Lyall, America West Airlines, 1991.

³ Automation comment. (1991, February). Feedback, No.23 pp. 2-5.

location of these different FMS components within a generalized glass cockpit.

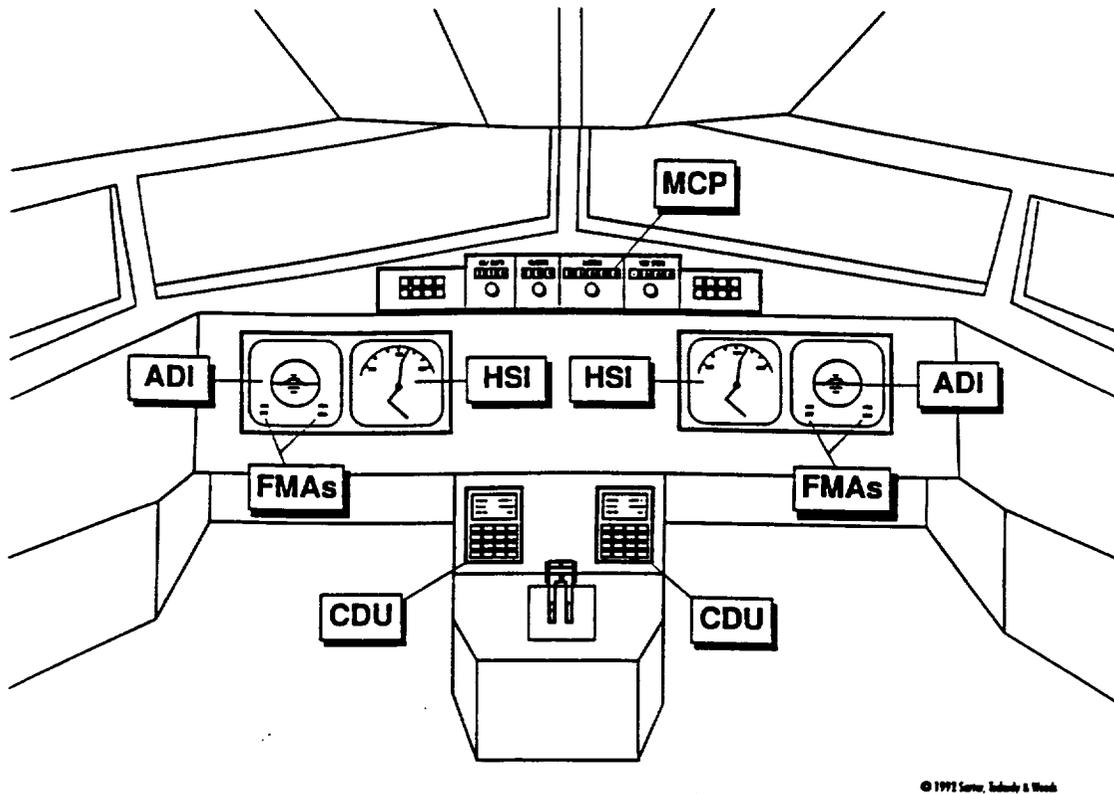


Figure 1. Flight deck controls and displays related to pilot-FMS interaction within a generalized glass cockpit

The Control Display Units (CDU) consist of a multifunction control unit (keyboard) and data display. The keyboard is used by pilots to enter data that define a flight path and to access flight-related data available on various pages within the CDU page architecture. The pilot-entered flight path, continuously updated to reflect the current flight status, is presented on the map display of the Horizontal Situation Indicator (HSI). This allows pilots to monitor progress along the path. In the HSI Plan Mode, the pilot can visually check modifications to the active flight plan.

The Mode Control Panel is used to activate different automatic flight modes (e.g. VNAV, LNAV, HDG SEL, LVL CHG). The pilot can also use knobs on the MCP to dial in targets for individual flight parameters (airspeed, heading, altitude, and vertical speed) which are tracked by the system if a corresponding automatic flight mode is activated. To find out which FMS modes are currently active, the pilot can monitor the Flight Mode Annunciations on the Attitude Director Indicator (ADI). These provide data on the active (or armed) pitch and roll modes and on the status of the autopilot(s). They also indicate the status and mode of the autothrottles which can be set to manual or automatic mode for speed and altitude control. The various FMS interfaces and autoflight functions provide the pilot with a high degree of flexibility in terms of

selecting and combining levels of automation to respond to different situations and requirements.

It is important to remember that there are various modes of automatic flight control that range between the extremes of automatic and manual. The highest level of automatic control occurs in the VNAV (Vertical Navigation) and LNAV (Lateral Navigation) mode. In these modes of control, the pilots enter (or, in their words, "program") a sequence of targets that define an intended flight path into the CDU, and then activate the automatics by selecting VNAV (Vertical Navigation) and/or LNAV (Lateral Navigation) through controls on the Mode Control Panel (MCP). The Flight Management Computer (FMC) automatically controls the aircraft to follow the desired flight path. At this strategic level of automation, the FMS pursues a sequence of target values without the need for further intervention by the pilot. This is particularly helpful in situations that allow for long-term planning with a low likelihood of deviations from the plan (e.g. cruise phase of flight).

When the pilot needs to quickly intervene and change flight parameters (e.g. in terminal areas), other lower levels of automation are available. The pilot can enter target values for different flight path parameters (i.e. airspeed, heading, altitude, vertical speed) on the Mode Control Panel (MCP). He then activates one of the corresponding modes (e.g. Heading Select or Level Change), and the target will be captured and maintained automatically until target or mode of control are actively changed by the pilot.

An important characteristic of automatic flight path control is the high degree of dynamism. Transitions between modes of control occur in response to pilot input and to changes in flight status. Automatic mode changes can occur automatically when a target value is reached (e.g. when leveling off at a target altitude) or based on protection limits (i.e. to prevent or correct pilot input that puts the aircraft into an unsafe configuration).

Both the flexibility of the FMS and the dynamism of flight path control impose cognitive demands on the pilot. He has to decide which level and mode of automatic control to use in a given set of circumstances, and he also has to track the status and behavior of the automation. This latter task requires that he attends to and integrates data from a variety of indications in the cockpit such as the Flight Mode Annunciations on the Attitude Director Indicator, the visualization of the programmed route of flight on the Horizontal Situation Indicator, or the display of target values on the Mode Control Panel.

RESEARCH ACTIVITIES

Study 1: Pilot Reports of FMS-Related Surprises

Background and Methods

The pilot report corpus was generated through a questionnaire, distributed to experienced airline pilots flying the B-737-300. This survey expands on results from a portion of a study by Wiener (1989) who asked B-757 pilots to rate statements concerning their attitude towards cockpit automation. Two of the statements were specifically related to FMS operations (see Figure 2).

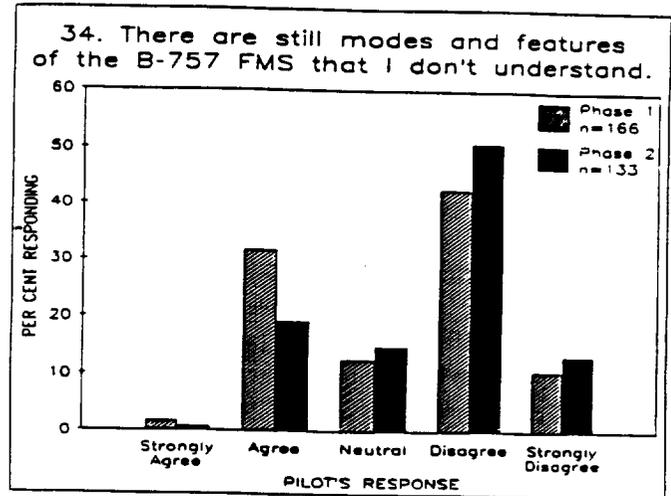
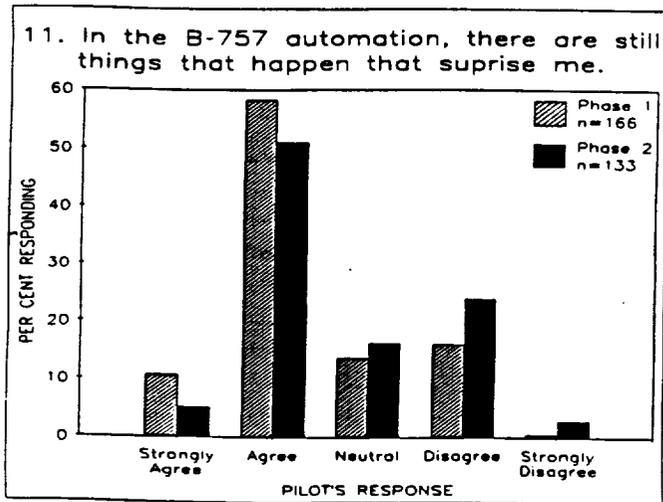


Figure 2. Results of a pilot survey concerning "glass cockpit"-related issues (adopted from Wiener, 1989, p. 28 and 58). (Phase 1 = data collected in 1986; Phase 2 = data collected in 1987; volunteer 757 pilots served as subjects).

Interestingly, the responses show that a rather large number of pilots with more than one year of experience on the B-757 agree that they are still being surprised by the automation (~ 55% of the pilots) or that they do not understand all of the FMS modes and features (~ 20% of the pilots). Given the implications and potential consequences of these rating data, it seems important to examine in detail what are the nature and circumstances of these surprises and gaps in pilots' mental models. We followed up Wiener's results by asking B-737-300 pilots to rate their agreement/disagreement with the above two statements on a five point scale (strongly agree, agree, neutral, disagree, strongly disagree). But more critically, we asked them to describe in detail as many instances as possible of surprises they had actually experienced and modes they did not understand. The survey vehicle thus served as a "corpus gathering" technique, that is, a means for "the identification and description of naturally occurring phenomena" (Reason, 1990). In other words, it captured some of the variety of real-world difficulties and recurrent patterns of pilot-automation interaction.

Results

Table 1 summarizes the background data on the pilots who responded to the survey. The survey was distributed to 887 B-737-300 line pilots from one airline company; responses were received from 135 pilots.

Table 1. Background and flight experience of pilots responding to the survey.

Age (n=134):	42.6 (8.5) y/o	[mean (std. dev.)]
Flight Time on the B-737-300 (n=132):	944 (613) hrs	[mean (std. dev.)]
Seat on the B-737-300 (n=134):	Captain	75 pilots
	F/O	59 pilots
Total Flying Time (n=134):	7714 (4978) hrs	[mean (std. dev.)]
Previous Commercial Jet Aircraft Flown (n=124):		
	B 727	38
	B 737-200	19
	DC - 8	16
	B 747	2
	DC - 9	1
	B 757/767	18
	B 747-400	1
	DC - 10	29

The pilots' ratings of the two statements on cockpit automation basically replicate Wiener's (1989) results.

Table 2. Percentages of pilots' responses to the first statement "In the B-737-300 automation, there are still things that happen that surprise me."

RATING	ALL PILOTS (N=135)	< 1,200 HRS OF LINE EXPERIENCE (N=97)	>= 1,200 HRS OF LINE EXPERIENCE (N=37)	NO PREVIOUS GLASS COCKPIT EXPERIENCE (N=104)	PREVIOUS GLASS COCKPIT EXPERIENCE (N=19)
STRONGLY AGREE	18	22	5	72	42
AGREE	49	57	33		
NEUTRAL	7	5	11	7	5
DISAGREE	22	14	43	21	53
STRONGLY DISAGREE	4	2	8		

Table 3. Percentages of pilots' responses to the second statement "There are still modes and features of the B-737-300 FMS that I don't understand."

RATING	ALL PILOTS (N=135)	< 1,200 HRS OF LINE EXPERIENCE (N=97)	>= 1,200 HRS OF LINE EXPERIENCE (N=37)	NO PREVIOUS GLASS COCKPIT EXPERIENCE (N=104)	PREVIOUS GLASS COCKPIT EXPERIENCE (N=19)
STRONGLY AGREE	12	15	0	54	11
AGREE	33	36	25		
NEUTRAL	16	21	8	13	21
DISAGREE	25	20	39	33	68
STRONGLY DISAGREE	14	8	28		

More important for the purpose of developing countermeasures to any existing problems related to pilot-automation interaction are pilots' descriptions of specific instances of FMS surprises and modes/features that they had difficulties with.

Corpus of Pilot Reported FMS Surprises and Problematic FMS Modes/Features

Pilots were asked to describe instances where FMS behavior surprised them and to indicate modes/features of FMS operation that they did not understand. There were no sharp boundaries between the incidents elicited by the two questions. Pilot reports are categorized according to their underlying theme. The major categories refer to a) Vertical Navigation (VNAV) modes, b) data entry, c) uncommanded mode transitions, d) infrequently used modes and features of the FMS, e) surprising flight director commands, f) monitoring of active target values, g) the availability of multiple methods for achieving a goal, h) the lack of data propagation within the control display unit (CDU), and i) the effects of partial system failures. For each category in the corpus we provide a short description of the kinds of problems reported, the number of pilot reports on problems in this category, and, for some categories, an abbreviated example. Surprises or unclear features of the FMS that were only reported by one pilot are not included in the corpus.

A) VNAV-related problems

The largest number of reported problems refer to Vertical Navigation Modes (VNAV). These are subdivided into the four categories "VNAV logic and calculations", "Switching between VNAV and MCP descent modes", the "VNAV Speed Descent mode in general", and the "Disengagement of the APPROACH mode".

VNAV logic and calculations

38 Reports

Pilots indicate that the algorithms underlying the calculation of a VNAV path are not transparent to them. They can not visualize the intended path, and therefore they are sometimes unable to anticipate or understand VNAV activities initiated to maintain target parameters (25 reports). VNAV control actions are often described as being surprisingly abrupt (4 reports). Several pilots report that they have been surprised by VNAV when it failed to start the descent upon reaching the top-of-descent point (TOD) (9 reports).

Abbreviated example:

VNAV was used for a path descent. Although the displayed TOD was reached, and autothrottles (A/Ts), autopilot (A/P) and VNAV were engaged, the aircraft did not start to go down. The pilots finally figured out that this happened because they had not changed their initial cruise altitude entry in the CDU after ATC told them to level off at FL 290 instead of the originally planned FL 310. Meanwhile, the TOD point had been passed. The airspeed had rolled back from 280 kts to 190 kts. When the descent was initiated by the pilots, the FMC (Flight Management Computer) used an excessive rate of descent (6,000 fpm) to get down to the path. This caused an ATC alert, and the actual airspeed increased to the maximum limit speed.

Switching between VNAV and MCP descent modes

11 Reports

These examples refer to situations where pilots had a descent properly programmed and both VNAV (Vertical Navigation mode) and LNAV (Lateral Navigation mode) engaged when ATC asked them for an unanticipated level-off or change in heading. They report uncertainty as to whether or not the reengagement of VNAV after compliance with the clearance by means of MCP interventions would bring them "back on track". They have problems with keeping track of active target values related to different FMS subsystems under such circumstances.

VNAV Speed Descent Mode in general

8 Reports

Pilots indicate that they do not understand how the VNAV Speed Descent works in terms of its targets, protections, and its operational logic.

Disengagement of the Approach (APPR) mode

6 reports

Some pilots report that they were not able to disengage the APPR mode when required to do so. This problem is especially important as it occurs at a fairly low altitude, under time pressure and sometimes in congested traffic areas.

Example:

During the final descent, the pilots were unable to deselect the APPR mode after localizer and glideslope capture when ATC suddenly requested that the aircraft maintain the current altitude and initiate a 90 degree left turn for spacing. They tried to select ALT HOLD (Altitude Hold mode) and HDG SEL (Heading Select mode) on the MCP to disengage the APPR mode and comply with the clearance but neither mode would engage and replace the APPR mode. They finally turned all autoflight systems off.

B) Data Entry

54 reports

There was a large number of reports related to the rejection of attempted input into the CDU due to different software versions running on the FMS. While the survey was underway, three slightly different FMS software versions were in use. According to the reports, this resulted most frequently in unsuccessful attempts to enter a new crossing restriction during the approach because the required data entry format and procedure is not the same for the three software versions. Pilots also commented that the "Invalid Entry" message they received in these cases did not help them find the correct input format. These data entry problems frequently occurred when the pilots were working under time pressure, and in some cases they contributed to altitude violations.

C) Uncommanded Mode Transitions

28 reports

Pilots report that they are surprised by "uncommanded" mode transitions which occur upon reaching a target state or for protection purposes. Most often, the reports refer to the automatic reversion from Vertical Speed mode (V/S) to Level Change mode (LVL CHG) which occurs if the airspeed deviates from the target range due to an excessive rate of climb or descent. One potential consequence of this automatic mode transition is that the vertical rate changes dramatically without any intervention by the crew. Pilots' reports seem to indicate that such uncommanded changes are difficult to track given current cockpit displays and indications.

D) Infrequently used features/modes

14 reports

Pilots report that they do not understand modes and features of the FMS that they rarely use (e.g. the Required Time of Arrival (RTA) feature). However, they also comment that they do not think of these as critically important features.

E) Flight Director (FD) Bars

11 reports

Pilots describe cases where the FD bars commanded pitch attitudes which seemed to be inadequate or unnecessarily abrupt. Some pilots report that, as a result, they lose confidence in the FD bars.

F) Active Target Values

10 reports

In some situations, it seems to be difficult for pilots to keep track of what are the currently active target values. The pilot reports indicate that one of the major sources of this problem is the interaction between the values selected on the MCP and those selected within the CDU. Pilots also commented that, while the MCP targets can immediately be seen on the MCP, the FMS targets are sometimes "hidden" in the CDU page architecture.

Example:

As a protective measure, VNAV climbs and descents are constrained by the selected MCP altitude. For example, in order for a preprogrammed FMC descent to begin upon reaching the TOD point, a lower than CRZ altitude has to be selected on the MCP. If

pilots forget to do so, the aircraft will maintain cruise altitude beyond the TOD and the airspeed will slow down. Some pilots report that they have been surprised by this aircraft behavior because they did not realize that in this case the MCP target overrides the CDU target.

G) Multiple Methods

10 reports

Some pilots mention that, for certain tasks, there seems to be an overwhelming number of possible methods to do the job. Their reports indicate that there is a cognitive load associated with learning and deciding on which method to use for a particular task in a particular flight context. The reports point to the tradeoff between providing pilots with flexibility and imposing additional cognitive load on them.

H) Lack of data propagation

9 reports

Pilots report that they are sometimes surprised by the effects of interactions between target values entered on different but interrelated CDU pages. They suggest that certain data should propagate automatically to functionally interrelated CDU pages.

Example:

A frequently described example is the case where, during the cruise phase of flight, the airspeed entry on the CRZ page of the CDU is changed but the new data do not propagate to the DES page. If the new cruise speed is lower than the originally programmed descent speed pilots are surprised upon reaching the TOD point, when the aircraft starts to accelerate rather than decelerate.

I) The effects of partial system failures

3 reports

These pilots report that they are unsure of the consequences of partial FMS failures. After such failures, they can not tell which subsystems are still active, which subsystems are available, or how the failure may interact with the active flight control mode. These reports implicate potential problems with both pilots' mental model of the FMS structure and with the indications of FMS status and behavior.

The corpus of reported difficulties in pilot-FMS interaction was generated through one technique that sampled a small part of the relevant user population. In order to converge on a more comprehensive and meaningful assessment of existing difficulties in pilot-automation interaction we conducted a complementary study where we observed crews in transition to highly automated aircraft during a number of simulated flight scenarios.

Study 2: Observation of Crews in Transition Training

Background and Methods

To complement the data gathered through pilot reports, we also observed the behavior of experienced pilots who were in the process of transitioning to the B-737-300 aircraft. This transition training involves classroom, computer-based training (CBT), LOS (line-

oriented simulation) sessions on a fixed-base trainer, and LOFT (line-oriented flight training) sessions on full-mission simulators. At the end of training, pilots take a 4-hr simulator check-ride in which they have to demonstrate that they are proficient in the following autoflight systems operations: Active Data Base Check, FMS and Performance Initialization, Flight Plan Entry, Direct To/Intercept Leg To, Holding Pattern, Installing an Approach, Closing a Route Discontinuity, and MCP (Mode Control Panel) Speed Interventions.

We observed 10 pilot crews during fifteen LOS sessions with 6 different scenarios during their transition training on a fixed-base B-737-300 trainer (see Table 4 for a breakdown of crews by scenario).

Table 4. Observed training sessions on the FMS-part task trainer.

LOS (Line-Oriented Simulation) Scenario	No. of Observations	Crew
LOS 2	1	A
LOS 3	2	B C
LOS 4	3	D E F
LOS 5	1	F
LOS 6	4	F G H I
LOS 7	4	B D E

Crews A, C, G, H, I, and K were observed only during one of the seven LOS sessions. The other four crews were observed more than once. For example, crew B was observed early in their training (session 3) and again during their last LOS session. The advantage of multiple observations is that the progress of these crews could be examined.

The transition training is carried out using a fixed-base simulator which allows for all flight operations except hand-flying the aircraft below 1,000 ft AGL. It is equipped with all relevant cockpit instruments and displays including a Flight Management System with its associated electronic flight displays and control interfaces described earlier (Figure 1).

Each of the observed LOS sessions requires 3 hours to complete. As in line operations, one of the pilots is assigned the role of pilot-flying, the other carries out the tasks of the pilot-not-flying. From time to time, the simulation is interrupted by the instructor to ask questions or to discuss the flight situation with the pilots.

The simulation scenarios consist of a complete flight, including cockpit setup, takeoff and landing, and they are designed to cover predefined sets of objectives emphasizing

FMS operations. Abnormal and increasingly difficult situations such as system failures are introduced at the later stages of training.

Throughout each LOS session, an observer was present (the first author) who was knowledgeable about both the scenarios and the FMS procedures and activities required to handle each scenario.

The observer collected two types of data. First, she encoded crew-FMS interactions - the methods used to carry out given tasks and errors or difficulties that occurred. A second source of data was the discussion between the instructor and the crew which occurred during the scenario and after the scenario was completed. These instructor-crew communications help to reveal gaps in FMS-related knowledge and misconceptions in the pilots' model of the system. For example, the discussions indicated whether the pilots were capable of explaining their interaction with the FMS, or whether they simply used "recipes" to operate the system without fully understanding how their input lead to the desired outcome.

The LOS scenarios on the fixed-based simulation facility provide a meaningful window on pilot-automation interaction. They allow for the collection of verbal reports as frequent and extended interruptions naturally occur to answer pilots' questions and to comment on their performance (Woods, 1992). Such interventions are not desirable in the context of real-time full-mission simulation training.

Results

Table 5 contains the flight background of the ten crews. Note that 6 of the 10 observed crews were "mixed crews" in the sense that one of the pilots had previous "glass-cockpit" experience while the other one came from a "non glass-cockpit."

Table 5. Previous aircraft flown by pilots observed during transition training.

Crew	Captain	First Officer
A	B 767	B 727
B	B 767	B 767
C	DC - 10	B 727
D	B 767	DC - 8
E	B 767	B 727
F	B 767	B 727
G	B 767	B 727
H	B 767	B 727
I	DC - 8	B 727
K	DC - 10	DC -8

The observations indicated that pilots can become proficient in basic FMS operations in a fairly short amount of time. Difficulties with these basics were observed only, with

very few exceptions, during the first three training sessions (LOS 2, 3, 4). The few difficulties observed concerned basics such as entering data in the correct format, finding relevant data in the CDU pages, or carrying out tasks such as FMS Initialization. During these first 3 sessions, it was, in some cases, difficult for pilots to keep track of who is in charge and what are the currently active target values. Difficulties arose in managing the Horizontal Situation Indicator (HSI), i.e., selecting ranges and modes of presentation.

During the last three training sessions (LOS 5, 6, 7), pilot errors and questions focused on gaps in their understanding of the underlying functional structure of the FMS. Table 6 provides an overview of the most frequently encountered problems and questions.

Table 6. Most frequently observed problems during transition training to the B-737-300.

<ul style="list-style-type: none">- Availability/Disengagement of Modes For example: Knowledge of LNAV capture criteria or ways to disengage the APPR mode- Keeping track of Uncommanded Mode Transitions For example: Awareness of automatic transition to ALT HOLD mode upon level-off, and subsequent requirement to re-engage a Climb-Mode for changing altitude- VNAV Targets and Logic For example: Visualization of FMS-calculated vertical profile- Multiple Methods For example: Pilots often indicated that they were not sure whether there were other ways of achieving a goal or how to choose among multiple methods

Frequently, pilots were able to describe FMS behavior during standard operations. For example, a pilot could describe the states and activities of the autothrottles (A/Ts) during takeoff and climb. But the same pilot would have difficulties applying this knowledge to a specific and more complicated operational situation, e.g. an aborted takeoff. This is often referred to as the problem of inert knowledge (Glaser, 1984).

In summary, the physical appearance and the "recipes" for carrying out standard tasks could be learned in a fairly short amount of time. However, even during the last training sessions, many of the observed pilots show gaps in their understanding of the overall functional structure of the system as indicated by their problems in dealing with complex or novel tasks and situations. The above problems were most often seen with pilots who transitioned from a "non-glass cockpit" aircraft. While their "glass-experienced" colleagues "only" had to get used to minor differences between their previous aircraft and the B-737-300, these pilots had to learn a whole new cockpit concept.

As a result, it appears from our observations that there are disadvantages to "mixed" training crews, i.e. crews where only one of the two pilots has previous glass-cockpit experience. This pilot who could focus on deeper issues about how to manage the automation capabilities in diverse contexts often was forced to wait while the other crew member explored more basic concepts and flight situations. In turn, the pilot with no previous experience on a highly automated aircraft sometimes did not ask all of his questions because he felt that he was slowing down the training process.

Overall, we observed the same kinds of difficulties during the late stages of transition training as were reported by line pilots in the survey study. The two studies used complementary data collection techniques (pilot reports and behavioral data through observation of training simulations) and sampled different levels of experience with glass cockpit aircraft. The combined results create a corpus of specific flight situations and FMS behaviors where difficulties arise in pilot-automation interaction. At this level, the results may be useful to system designers interested in incremental improvements of the current system and its pilot interfaces. Similarly, the results may be useful to those responsible for training pilots to work with current cockpit automation by highlighting particular modes of the FMS and particular flight situations where pilots have difficulty tracking and anticipating FMS behavior. However, one may also interpret the specific reported and observed difficulties in a larger perspective – what do these results tell us about the factors that are important for effective human-automation cooperation?

DISCUSSION

Breakdowns in Pilot-Automation Interaction

The corpus of reported and observed difficulties provides a picture of the kinds of complexities that can arise in pilot-automation interaction and the kinds of task contexts where these complexities can affect performance. Knowledge of these mechanisms is essential to be able to better design the interface between pilots and automation from the point of view of a cooperative human-machine cognitive system (Woods, 1986; Hutchins, 1991). This knowledge also indicates how training programs may need to change in fundamental ways to accommodate the changes in the human's role in highly automated aircraft.

The corpus of reported and observed difficulties indicates that pilots can lose situation awareness (Sarter and Woods, 1991) with respect to FMS status and behavior. Wiener (1989) has summarized his results on cockpit automation in the phrase, "the three most commonly asked questions in glass cockpits are: 'What is it doing?' 'Why did it do that?' 'What will it do next?'" The corpus reveals specific pilot-FMS interaction difficulties that can be grouped under these three questions. For example, difficulties in tracking active target values and FMS behavior in some modes can contribute to losing track of "what the automation is doing". Uncommanded mode transitions can create situations where the crew can be surprised – "why did it do that?" One common factor contributing to such an incomplete or faulty assessment of system status and behavior seems to be weak feedback from the FMS displays and interfaces (Norman, 1990). Another common factor implicated in many of the problems noted in the corpus is incomplete or buggy mental models of how various modes of the FMS work and especially how they interact with each other in different flight contexts. If the pilot has difficulty monitoring and

understanding automatic system behavior, it will also be difficult for him to project or anticipate future states – “what will it do next”.

The problem of weak feedback on system status and behavior is a common deficiency in human-computer interfaces (Norman, 1990). While all of the necessary data on FMS status may be available somewhere in the cockpit displays and the CDU page architecture (see Figure 1), finding, integrating, and interpreting all of the relevant data to build an assessment of current and future FMS behavior can be a demanding cognitive task, especially given the time demands of actual flight operations (Woods, 1991). Many examples of inadequate feedback occurred in the corpus including difficulties integrating data on FMS status distributed over different cockpit displays or CDU pages, difficulties anticipating uncommanded mode changes, difficulties assessing the implications of changes to the instructions given to the FMS (e.g., enroute changes in cruise speed may interact with pre-programmed values for the descent phase on a different CDU page), difficulties visualizing the descent profile programmed in VNAV. Weak feedback can increase cognitive workload in several ways: by increasing demands on pilots to remember information and by increasing the need to rely on mental models of FMS structure and function to assess or project FMS behavior.

Another factor that seemed to contribute to difficulties noted in the corpus is incomplete or buggy mental models of how various modes of the FMS work and how they interact with each other in different flight contexts. The various FMS interfaces and autoflight functions provide the pilot with a high degree of flexibility in terms of selecting and combining levels of automation to respond to different situations and requirements. However, this flexibility creates new sources of cognitive workload for the pilot. One issue is simply that there are a large variety of ways that the automation can be used and that having a detailed and complete understanding of how these various automation modes work in detail is a demanding new knowledge requirement for the glass cockpit pilot. The corpus results indicate that there are infrequently used modes (e.g., VNAV speed descent is rarely used in US airspace) that pilots do not understand completely. Second, the flexibility of the automation requires that pilots understand how different modes interact and the consequences of transitions between modes in various flight contexts. Third, the pilot needs to develop knowledge and strategies for how to use the flexibility of the automation in different flight circumstances. The corpus results indicate that pilots tend to adopt and stick with a small repertoire of strategies because their knowledge about the advantages and disadvantages of the various options for different flight contexts is incomplete.

Both the self-reports and the training observations indicate that pilots do not perceive the FMS as one large integrated system consisting of a variety of closely related, interacting subsystems such as the MCP or the CDU. They rather tend to refer to the MCP as a separate system to which they “escape” in case things become too complex or time pressure is too high while working with the CDU. From an engineering perspective, the FMS works in an integrated way. But this property is not sufficiently emphasized in training, and it is not clearly represented in the image the system presents to the pilot through the various displays of FMS status. Our data show that pilots think of and operationally use the MCP and CDU as, at least, two different systems.

Discussions of issues on pilot-automation interaction often focus on the transition from automated to manual control of the aircraft. Our data show that the problematic issue is in fact different. It is important to remember that there are various modes of automatic flight control that range between the extremes of automatic and manual. The FMS provides the pilot with the opportunity to select among and combine a wide variety of modes which results in different levels of automation. The process of gradually moving up and down between these levels of automation is where difficulties in managing the system occur frequently due to problems with keeping track of the states and target values of the various modes. This problem is aggravated by the fact that these transitions are most likely to occur during busy climb and descent phases of flight.

New technology that creates or exacerbates bottlenecks during busy, high tempo, high criticality, event-driven operations, while its benefits tend to occur during routine, low workload situations has been termed "clumsy automation" (Wiener, 1988; 1989). Clumsy automation or the clumsy use of technology is a form of poor coordination between the human and the machine. The concept is based on the fact that in complex systems human activity ebbs and flows, with periods of lower activity and more self paced tasks interspersed with high tempo, externally paced operations where task performance is more critical (Rochlin, et al., 1987). An important design feature for well integrated cooperative work between the automation and the human is how the automation supports high workload periods or more difficult tasks. As a result, the effects of factors such as weak system feedback and incomplete mental models of the functional structure of the FMS may only be visible during more difficult or unusual situations (Roth et al., 1987).

The corpus of reported and observed difficulties show that, while pilots can make the FMS work (e.g., by using familiar modes or by switching to less automated modes), they are not always capable of explaining why their input resulted in the desired outcome. In addition, they do not fully exploit the range of capabilities of the system. In case of unusual or novel situations, it may be essential, however, to have a thorough understanding of the functional structure of the FMS and to be able to use this knowledge in an operationally effective way.

While some of this knowledge about how to manage the FMS capabilities is acquired during training, initial operating experience, and line operations, our data from experienced pilots show that there may not be enough time to explore all system options or to figure out the reasons underlying surprises or unclear modes. Furthermore, since the pilots can work around areas in which their knowledge may be buggy or which occur infrequently, incentives for deepening their understanding of the FMS may diminish with time. People are not always accurate in their judgments about how much they know (the degree of calibration) and can overestimate how much they know (an overconfidence bias), especially when the device in question has an opaque interface that provides weak feedback about actual status and behavior.

The concept of clumsy automation reminds us that cognitive work in the automated cockpit is inherently cooperative – between the human crew members and between the pilots and the automation (Woods, 1986; Hutchins, 1991). Therefore, the fact that pilots tend to adopt a limited repertoire of strategies for using the capabilities of the automation creates a potential coordination problem. When two pilots with different preferences have to coordinate their activities and crosscheck inputs without fully

understanding the strategies preferred or used by their colleague, they may have problems to maintain situation awareness. Cooperation and coordination is also necessary between the crew and the automation. Thus, for glass cockpit aircraft, cockpit resource management training should be concerned with pilot-automation as well as pilot-pilot coordination and communication.

An additional factor complicating pilot-automation cooperation is the difficult problem of software configuration control. One can assume that software is not a static entity but changes and evolves throughout the life of the system. Our results show that there are operational consequences and design implications which should be taken into account in managing software changes and that version control problems can be a source of difficulties for the crew.

Implications for Design and Training

The corpus of reported and observed difficulties in pilot-automation interaction suggests approaches to improving coordination and cooperation in current and future systems. First, better feedback on FMS status and behavior can support pilots in maintaining situation awareness in high tempo, high workload or unusual flight contexts. One part of this may be to explore new concepts that help pilots integrate diverse data into a coherent, operationally relevant picture of FMS status and behavior, including past behavior, current activities and setup, and future implications (e.g., Woods, 1991). In addition, the pilot-FMS interfaces can be modified to support data access and interface management tasks. Second, training programs and design efforts can address new ways to support pilots in forming and refining their mental model of the functional structure of the FMS.

The current training programs for pilots in transition to glass cockpits can provide pilots with the basic knowledge required to "make the FMS work", especially in standard situations. However, the data in the corpus show that this training may not be sufficient to prepare pilots for dealing with all operationally significant FMS procedures and information for coping with non-standard situations. It may prove important to revise our conception of training experienced pilots for transition to glass cockpit aircraft where the initial training is one part of a longer, continuing learning process with respect to how cockpit automation functions and how it can be utilized as a resource in a wide range of operational circumstances. Training opportunities for pilots flying glass cockpits in line operations may need to be expanded to establish ongoing progressive training through additional information about FMS features that are used less frequently or that can not be tried out in line operations for safety reasons, through opportunities to test and to extend their skills in managing the automation especially in more difficult or unusual flight contexts, and through opportunities to follow up and learn from surprises that they or their fellow pilots have experienced.

The FMS training that we observed emphasized a bottom-up approach oriented towards proficiency in specific tasks by providing "recipes" for system operation. The result that most of the difficulties in the corpus involved non-standard situations and complex interactions of FMS subsystems seems to suggest that a top-down approach would be desirable as an addition or complement. If pilots were provided with an overall mental representation of the functional structure of the FMS, they would be better able to

manage and utilize the automated systems in unusual or novel situations. Given that their role has shifted towards the detection of deviations from the expected and towards troubleshooting and managing such situations, this capability seems to be very important for pilots in highly automated aircraft.

In summary, the corpus of observed and reported difficulties in pilot-automation interaction suggests the need for the following improvements in the design and training of the FMS to help pilots exploit the full range of capabilities provided by flight deck automation:

- system states and transitions, goals, and options need to be clearly and coherently indicated to the pilot;
- the user needs to be supported in forming an accurate mental model of the device functionality which is critical for coping with more difficult and unusual flight situations;
- the display and interaction capabilities that mediate pilot-FMS communication need to be tailored to high demand situations and circumstances.

REFERENCES

Billings, C.E. (1991). *Human-centered aircraft automation philosophy* (Technical Memorandum 103885). Moffett Field, CA: NASA-Ames Research Center.

Eldredge, D., Dodd, R.S., and Mangold, S.J. (1991). *A review and discussion of Flight Management System incidents reported to the Aviation Safety Reporting System*. Battelle Report, prepared for the Department of Transportation, Volpe National Transportation Systems Center.

Glaser, R. (1984). Education and thinking: the role of knowledge. *American Psychologist*, 39, 93-104.

Hutchins, E. (1991). *How a cockpit remembers its speed*. (Technical Report). San Diego: University of California, Distributed Cognition Laboratory.

James, M., McClumpha, A., Green, R., Wilson, P., and Belyavin, A. (1991). *Pilot attitudes to flight deck automation*. Paper presented at the One Day Conference "Human Factors on Advanced Flight Decks", The Royal Aeronautical Society, London.

Norman, D.A. (1990). The 'problem' with automation: Inappropriate feedback and interaction, not "over-automation". *Philosophical Transactions of the Royal Society of London*, B 327.

Reason, J.T. (1990). *Human error*. Cambridge: Cambridge University Press.

Rochlin, G.I., La Porte, T.R., and Roberts, K.H. (1987). The self-designing high-reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review*, Autumn, 77-91.

- Roth, E.M., Bennett, K., and Woods, D.D. (1987). Human interaction with an intelligent machine. *International Journal of Man-Machine Studies*, 26, 257-275.
- Sarter, N.B. and Woods, D.D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
- Wiener, E.L. (1988). Error exposure and control in high technology aircraft. *Proceedings of the International Conference on Human Machine Interaction and Artificial Intelligence in Aeronautics and Space*, (pp.241-248). Toulouse-Blagnac, France: CERT.
- Wiener, E.L. (1989). *Human factors of advanced technology ("glass cockpit") transport aircraft* (NASA Contractor Report No. 177528). Moffett Field, CA: NASA-Ames Research Center.
- Woods, D.D. (1986). Paradigms for intelligent decision support. In E. Hollnagel, G. Mancini, & D.D. Woods (Eds), *Intelligent decision support in process environments* (pp. 153-173). New York: Springer Verlag.
- Woods, D.D. (1991). The cognitive engineering of problem representation. In G.R.S. Weir & J.L. Alty (Eds). *Human-computer interaction and complex systems* (pp. 169-188). London: Academic Press.
- Woods, D.D. (1992). Process tracing methods for the study of cognition outside of the experimental psychology laboratory. In G.A. Klein, J. Orasanu & R. Calderwood (Eds). *Decision making in action: models and methods*. New Jersey: Ablex.

Pilot Interaction with Cockpit Automation II: An Experimental Study of Pilots' Model and Awareness of the Flight Management System (FMS)¹

Nadine B. Sarter and David D. Woods

Cognitive Systems Engineering Laboratory
Department of Industrial and Systems Engineering
The Ohio State University

ABSTRACT

Technological developments have made it possible to automate more and more functions on the commercial aviation flight deck and in other dynamic high consequence domains. This increase in the degrees of freedom in design has shifted questions away from narrow technological feasibility. Many concerned groups from designers to operators to regulators and researchers have begun to ask questions about how should we use the possibilities afforded by technology skillfully to support and expand human performance. In this paper we report on an experimental study that addresses these questions by examining pilot interaction with the current generation of flight deck automation. Previous results on pilot-automation interaction derived from pilot surveys, incidents reports and training observations that have produced a corpus of features and contexts where human-machine coordination is likely to break down (e.g., automation surprises). We used these data to design a simulated flight scenario that contained a variety of probes designed to reveal pilots' mental models of one major component of flight deck automation, the Flight Management System (FMS). The events within the scenario also were designed to probe pilot's ability to apply their knowledge and understanding in specific flight contexts and to examine their ability to track the status and behavior of the automated system (mode awareness). While pilots were able to "make the system work" in standard situations, the results reveal a variety of latent problems in pilot-FMS interaction that can affect pilot performance in non-normal time critical situations.

INTRODUCTION

The introduction of advanced technology to modern flight decks has succeeded in terms of increasing the precision and efficiency of flight operations. However, recent accidents and incidents involving glass cockpit aircraft have suggested that the current generation of cockpit automation may have created new operational burdens and new kinds of failure modes in the overall human-machine system (Billings, 1991). Only a limited empirical data base is available concerning the nature and circumstances of existing problems in pilot-automation interaction (Wiener, 1989; Eldredge, Dodd and Mangold, 1991; James et al., 1991). These data consist primarily of either subjective data obtained from questionnaires and interviews or of in-flight observations of pilot interaction with one of the core systems of cockpit automation, the Flight Management System (FMS). The resulting data about pilots' attitude towards the system and the anecdotal reports of problems indicate that there is a need for further research that

¹Submitted for publication

systematically analyzes the nature of and the reasons for FMS-related problems. These results will be critical in order to develop countermeasures and to improve pilot-automation interaction.

With this goal in mind, we studied pilot-FMS interaction through three different methodological approaches that allowed us to systematically collect converging data to describe existing problems and to understand why they exist. In the first report on our work (Sarter and Woods, in press), two exploratory research activities were described. A survey of pilots' self-reports of their operational experiences with the FMS and observations of transition training from a conventional to a "glass cockpit" aircraft were used to gather a corpus of problems with the operation of the FMS. This corpus consisted of detailed incident descriptions from which major underlying problem categories were extracted.

These categories provided the basis for the design of a scenario for an experimental study of pilots' mental model and their awareness of the FMS. In this study we confronted twenty experienced pilots with situations and tasks that are instances of the previously identified FMS-related problem categories. The pilots flew the scenario on a part task training simulator that had been developed to teach FMS operations. As a result, it was possible to test the completeness and accuracy of their FMS-related knowledge as well as their ability to apply this knowledge in specific situations.

INTRODUCTION TO THE FLIGHT MANAGEMENT SYSTEM

The following section provides a brief, simplified overview of the Flight Management System (FMS). The FMS supports pilots in a variety of tasks such as flight planning, navigation, performance management, and flight progress monitoring. One of its major functions, and the function of primary interest in the context of the reported studies, is automatic flight path control.

The major FMS controls in the cockpit are the Mode Control Panel (MCP) and the multifunction keyboards of two Control Display Units (one for each pilot). FMS-related cockpit displays are the CDU multifunction display, two Attitude Director Indicators (ADI), and two Horizontal Situation Indicators (HSI). Figure 1 illustrates the typical location of these different FMS components within a generalized glass cockpit.

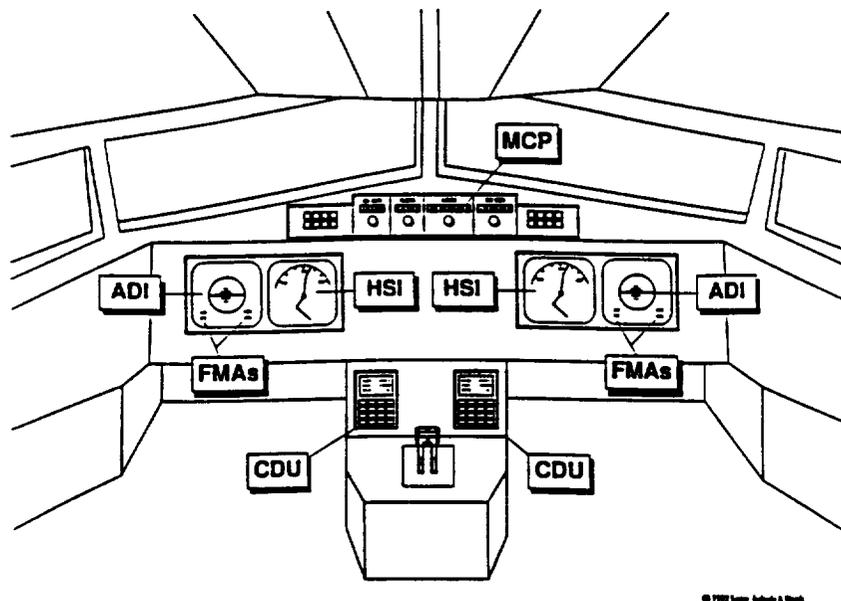


Figure 1. Flight deck controls and displays related to pilot-FMS interaction within a generalized glass cockpit

The Control Display Units (CDU) consist of a multifunction control unit (keyboard) and data display. The keyboard is used by pilots to enter data that define a flight path and to access flight-related data available on various pages within the CDU page architecture. The pilot-entered flight path, continuously updated to reflect the current flight status, is presented on the map display of the Horizontal Situation Indicator (HSI). This allows pilots to monitor progress along the path. In the HSI Plan Mode, the pilot can visually check modifications to the active flight plan.

The Mode Control Panel is used to activate different automatic flight modes (e.g. VNAV, LNAV, HDG SEL, LVL CHG). The pilot can also use knobs on the MCP to dial in targets for individual flight parameters (airspeed, heading, altitude, and vertical speed) which are tracked by the system if a corresponding automatic flight mode is activated. To find out which FMS modes are currently active, the pilot can monitor the Flight Mode Annunciations on the Attitude Director Indicator (ADI). These provide data on the active (or armed) pitch and roll modes and on the status of the autopilot(s). They also indicate the status and mode of the autothrottles which can be set to manual or automatic mode for speed and altitude control. The various FMS interfaces and autoflight functions provide the pilot with a high degree of flexibility in terms of selecting and combining levels of automation to respond to different situations and requirements.

It is important to remember that there are various modes of automatic flight control that range between the extremes of automatic and manual. The highest level of automatic control occurs in the VNAV (Vertical Navigation) and LNAV (Lateral Navigation) modes. In these modes of control, the pilots enter (or, in their words, "program") a sequence of targets that define an intended flight path into the CDU, and then activate the automatics by selecting VNAV (Vertical Navigation) and/or LNAV (Lateral Navigation) through controls on the Mode Control Panel (MCP). The Flight Management Computer (FMC) automatically controls the aircraft to follow the desired flight path. At this strategic level of automation, the FMS pursues a sequence of target values without the need for further intervention by the pilot. This is particularly helpful in situations that allow for long-term planning with a low likelihood of deviations from the plan (e.g. cruise phase of flight).

When the pilot needs to quickly intervene and change flight parameters (e.g. in terminal areas), other lower levels of automation are available. The pilot can enter target values for different flight path parameters (i.e. airspeed, heading, altitude, vertical speed) on the Mode Control Panel (MCP). He then activates one of the corresponding modes (e.g. Heading Select or Level Change), and the target will be captured and maintained automatically until target or mode of control are actively changed by the pilot.

An important characteristic of automatic flight path control is the high degree of dynamism. Transitions between modes of control occur in response to pilot input and to changes in flight status. Automatic mode changes can occur automatically when a target value is reached (e.g. when leveling off at a target altitude) or based on protection limits (i.e. to prevent or correct pilot input that puts the aircraft into an unsafe configuration).

Both the flexibility of the FMS and the dynamism of flight path control impose cognitive demands on the pilot. He has to decide which level and mode of automatic control to use in a given set of circumstances, and he also has to track the status and behavior of the automation. This latter task requires that he attends to and integrates data from a variety of indications in the cockpit such as the Flight Mode Annunciations on the Attitude Director Indicator, the visualization of the programmed route of flight on the Horizontal Situation Indicator, or the display of target values on the Mode Control Panel.

METHODOLOGY

General Approach

The study was designed based on a phenomenon-driven ethnographic approach to studying cognitive systems in high-tempo event-driven worlds (Woods, in press). First, we had to identify an experimental environment for studying pilot-FMS interaction. It seemed important to account for the numerous concurrent tasks that have to be carried out by the pilot in the real operational environment and that may affect his FMS-related performance. Also, the impact of the high-tempo nature of flight had to be captured to arrive at valid results. Therefore, a strict laboratory study with a restricted set of tools and environmental fidelity was rejected. The other extreme on the scale of possible approaches, i.e. a high-fidelity full-mission simulation study, was not selected because some of its inherent capabilities (e.g. aircraft motion, outside view) were not essential for the purpose of this study and because there were high costs associated with obtaining access to such facilities. As a result, we chose an environment that allows for both realistic tools and tasks as well as for a fairly high level of fidelity - a part-task training simulator for FMS operations.

The next important step in conceptualizing the study was designing the scenario based on predefined phenomena of interest (Woods and Sarter, in press). This is much more than simply making the scenario as realistic as possible. A realistic setting only provides the background on which the scenario needs to be staged. In this study, the problem categories identified by our survey and the training observations represented the phenomena of interest. The scenario design process involved identification of specific tasks and events linked together in a coherent scenario that would probe these phenomena. This approach enables the experimenter to trigger behavior of interest rather than hope for it to happen accidentally. While this approach may underlie a large number of simulation studies, it is often not explicitly laid out for the reader of a research report. In contrast, this paper will provide a detailed description of the match between phenomena of interest and events within the simulated scenario.

The data collection included both verbal and behavioral reports. An observer knowledgeable about FMS operations and about the test scenario kept track of pilots' interaction with the FMS on-line by means of a data-collection sheet that laid out the possible trajectories of the scenario and pilot behavior. In addition, pilots were asked to describe their reactions to hypothetical events which could not actually be simulated due to time restrictions and about FMS-related knowledge in general. These questions were asked in low workload phases of the flight without interrupting the simulation. This allowed us to probe pilots' knowledge within the actual operational environment rather than questioning them out of context where their task would be more related to the retrieval of information than to its application. A few questions were asked before or after completion of the flight as they related to more general topics or to preflight activities.

The data were collected throughout the experiment rather than extracted from video- and audio-recordings of the simulation runs after the fact. Such recordings can be helpful for exploratory studies or in cases where a knowledgeable observer is not available. But even though the retrospective analysis of videotapes may sometimes reveal unexpected or previously unattended but interesting behavior, there are disadvantages as well (e.g., investigators who are overwhelmed by the amount of data and unsure of how to abstract broader results from all the details). In this case getting actual line pilots to volunteer to participate in the study virtually ruled out the use of videotape (e.g., getting practitioners and their representatives, unions, to agree is prohibitively difficult). In addition, videotape is no substitute for careful and detailed identification of what one is looking for based on the mapping between phenomena of interest and the specific scenario; and videotape is no substitute for careful and detailed identification of what one might expect as canonical behavior based on knowledge of the field of practice (Woods, in press; Woods and Sarter, in press).

Experimental Scenario

The experimental scenario for this study was designed to address predefined phenomena of interest. These phenomena had been identified by the corpus gathering activities (pilot survey and training observations) preceding the study (see Sarter and Woods, in press). The issues were related to (a) pilots' proficiency in standard tasks, (b) pilots' mental model of the functional structure of the FMS and (c) their awareness of system state and behavior (mode awareness). In cooperation with a flight instructor, we identified tasks and events that would best serve to probe these phenomena. The basic flight context consisted of a flight from Los Angeles to San Francisco which took approximately 60 minutes to complete².

The following paragraphs provide an overview of the mapping between phenomena of interest and specific tasks and events within the scenario. Figure 2 illustrates the flight route and the timing of the tasks and events throughout the scenario. In order to better understand the following description of the scenario, it may be helpful for the reader who is not familiar with glass-cockpit technology to take a look at the short introduction to the FMS that was provided in the first part of this paper.

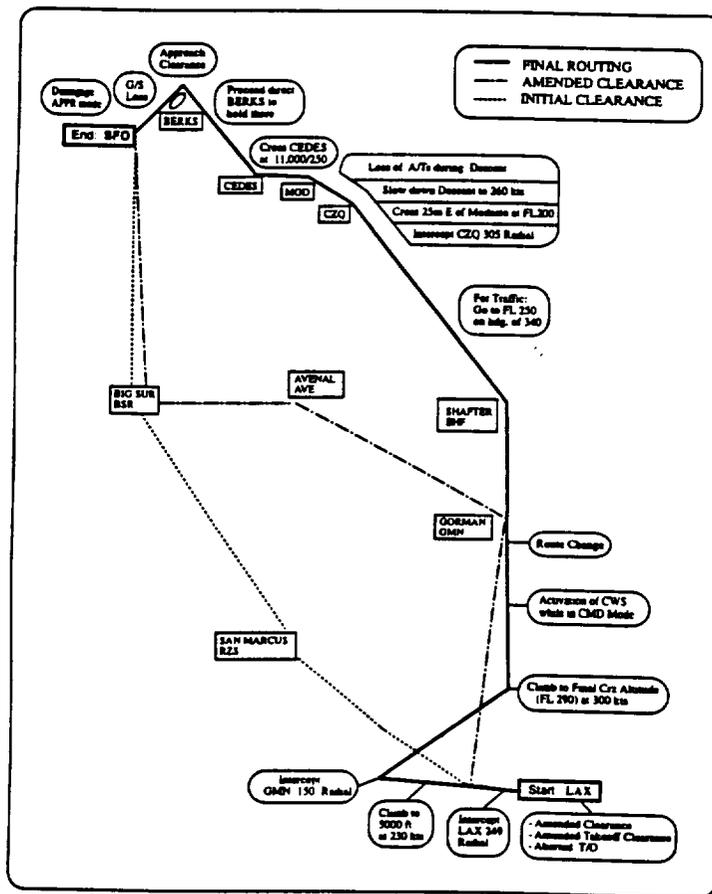


Figure 2. The Timing of Scenario Tasks and Events along the Flight Route

²The actual flight time would be longer but temporary increases in the simulated aircraft speed were used during quiescent phases of flight to reduce time on the simulator.

Proficiency in Standard Tasks

Table 1 lists the standard tasks that pilots had to carry out in the course of the scenario. The previous study (Sarter and Woods, in press) showed that pilots' proficiency at standard tasks did not seem to be a major source of difficulties. It was included as part of this scenario to provide additional converging evidence based on a scenario evolving in real time and involving pilots with line experience in glass cockpits to confirm the previous results³.

Table 1. Scenario Probes of Pilots' Proficiency in Standard FMS-Related Tasks

- | |
|---|
| <ul style="list-style-type: none">- Route Changes- Intercepting a Radial- Going direct to a waypoint- Building and Executing a Hold- Installing an ILS Approach- Entering Crossing Restrictions- Unplanned Level-Off- Extending the Final Approach Fix |
|---|

Pilots' Knowledge Of The Functional Structure Of The System

The second phenomenon of interest is the pilots' knowledge of the functional structure of the FMS. By functional structure of the FMS we are referring to their knowledge about how the FMS behaves in different flight situations rather than their ability to simply recite facts about the FMS. For example, do they understand the sequence of mode changes, their associated indications and the corresponding aircraft behavior throughout the takeoff roll.

To probe this phenomenon of interest, we built into the scenario a variety of tasks and situations that permit inferences about pilots' knowledge of the system and their ability to apply this knowledge in actual task contexts. Knowledge of overall FMS functionality was subdivided into six subtopics, and specific probes were developed for each subtopic (Table 2 summarizes the scenario probes).

A. Knowledge of the CDU Page Architecture

The page architecture of the FMS control display unit contains a huge amount of data that may be relevant at some point during the flight. Since only a very limited set of data can be presented on the CDU screen at any given time, pilots need to be able to navigate through the "hidden" data space. To find out about problems related to this task, pilots were asked to locate information on CDU pages on the following topics:

- Single engine capabilities
- Wind data for fixes of flight
- Available fuel
- Localizer Frequency and Front Course for a Runway

³The group of standard tasks presented in this experiment did not include the FMCS/Performance Initialization as we have already seen during the training observations that these tasks did not challenge the pilots. Also, we wanted to focus on tasks that have to be performed in the dynamic airborne portion of flight rather than on ground tasks that are not as much affected by time pressure or concurrent tasks.

We also asked pilots about their expectations concerning data propagation throughout the CDU page architecture. After pilots had entered speed and altitude target values on the Cruise Page to comply with an amended clearance by ATC, we asked whether they expected these data to propagate to the CDU Descent Page to become the targets for their descent.

These probes were supposed to test pilots' knowledge of the page architecture of the CDU as well as their ability to use the CDU interface to call up information/pages.

B. Mode Availability - Mode Disengagement

After being vectored off-course by ATC, pilots were asked to recapture the preprogrammed route. This task was introduced to find out whether pilots were aware of the criteria that have to be met in order for the LNAV mode to capture the original flight path.

When being cleared by ATC for the ILS approach, pilots were asked to properly set up the FMS to be able to use the automatic APPROACH mode. They had to remember that a lower MCP altitude had to be selected before engaging the APPROACH mode. Without this first step, the APPROACH mode engagement would not result in the desired start of descent; rather, the FMS would control the aircraft to maintain the MCP target altitude.

After Localizer and Glideslope capture on final descent, pilots were asked to describe how they would disengage the APPROACH mode if ATC told them to change heading and altitude for traffic.

The above probes allow us to determine whether pilots are familiar with the general prerequisites and procedures for engaging or disengaging a mode and whether they can apply this knowledge to a specific flight context.

C. FMC Logic

After takeoff from Los Angeles, pilots were asked to intercept the LAX 248 degree radial outbound. In order to successfully perform this task using LNAV, the pilot had to understand that the FMS logic is to always fly towards, not away from a waypoint. As his original flight plan did not include any waypoint on the radial, he first had to create a fictitious fix somewhere on the radial that the FMS could fly to.

After completion of the flight, we asked pilots about functional characteristics of the VNAV Path Descent in comparison to the VNAV Speed Descent. The questions referred to the way either one of these types of descent is initialized, what control mode the system uses to maintain target speed in either mode, and what is the lowest altitude that the system automatically descends to.

D. Effects of partial system failures

During a descent, pilots were asked about the expected consequences of losing the autothrottles in that situation. Would the aircraft still level off at target altitude, and what would be the consequences in terms of airspeed? How would they intervene in that case?

After capturing the glideslope, the glideslope was failed due to a signal loss. This allowed us to test whether pilots would realize what happened, whether they would understand the implications of losing the glideslope, and how they would react to the failure. In addition, they were asked about the differences between a glideslope failure above versus below 1,500 ft AGL.

If the glideslope signal is lost above 1,500 ft, the G/S indicator and the F/D bars disappear from the ADI, and the aircraft continues its descent at the current rate of descent. A flag indicating unreliable

glide slope input appears only on the standby attitude indicator. Glideslope loss below 1,500 ft (where automatic system tests are conducted) results in both autopilots disengaging and in changes in the mode indications (FLARE armed is not annunciated).

E. Protections

While climbing to 5,000 ft with VNAV engaged, pilots were asked what other modes they could use for the climb. With respect to one of the possibilities, the V/S mode, they also were asked what happens if an excessive rate of climb is used (i.e. the FMS automatically reverts to the LVL CHG mode to maintain a safe airspeed).

F. Various Options for Carrying Out a Task

Pilots were asked to comply with ATC clearances by using the FMS the same way as in real line operations. Once they had decided to use a certain mode for a given task, they were asked about other possible ways of achieving the same goal. This provided us with information on their knowledge about options provided by the FMS as well as about their criteria for selecting modes under different circumstances.

Table 2. Scenario Probes of Pilots' Knowledge of the Functional Structure of the FMS

- | |
|---|
| <ul style="list-style-type: none">- Locating data in the CDU page architecture- Tracking data propagation within the CDU- Applying knowledge about mode capture criteria- Disengaging the automatic APPR mode after capturing localizer and glideslope- Intercepting a radial outbound- Questions concerning VNAV Speed versus VNAV Pathdescent- Loss of autothrottles during a descent- Loss of G/S signal / G/S failure- Predicting effects of excessive rate of climb in V/S mode- Describing the different possible ways of doing a task |
|---|

Mode Awareness

Table 3 summarizes the probes built into the scenario for testing pilots' mode awareness. They help determine whether pilots know who/which system is in charge of controlling the aircraft, what the active target values are, and whether they can anticipate the status and behavior of the FMS.

"Who is in charge?"

Immediately before takeoff, pilots were asked how they would abort the takeoff if necessary at approximately 40 kts with the autothrottles turned on. In order to adequately cope with the situation, pilots have to understand what regime the autothrottles follow during takeoff. Until reaching 64 kts, the autothrottles will automatically go to N1. At and above 64 kts, pilots can manually override the autothrottles. Thus, if aborting the takeoff before 64 kts, the autothrottles have to be disengaged to prevent them from advancing again to reach N1.

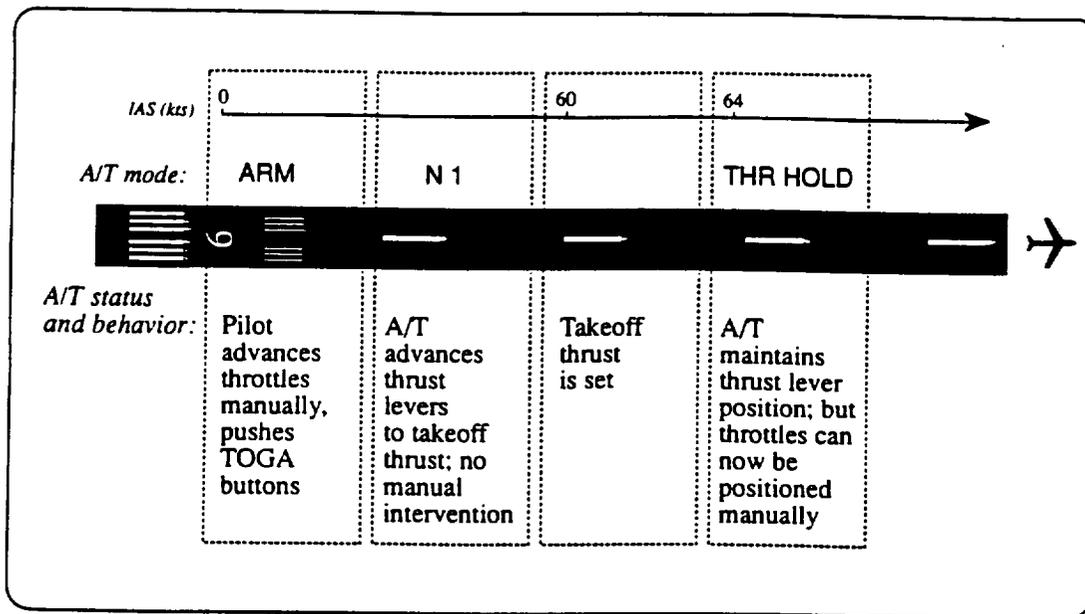


Figure 3. Autothrottle Status, Behavior, and Indications throughout the Takeoff Roll

Pilots' awareness of active mode settings was also probed by checking whether they (re)activated a corresponding mode after modifying target data in order to make the system work on reaching a new target state.

"What are the active target values?"

Several probes were used to find out about pilots' awareness of the current FMS target values. Shortly before takeoff, they were given an amended takeoff clearance involving a tailwind component. This requires that they remember to change their N1 setting from reduced to full takeoff thrust.

During the cockpit setup, a pointer to the pilot-calculated N1 target value can be manually positioned on the forward engine display for reference purposes. However, if the autothrottles are active during takeoff, as in this scenario, they use the FMS-calculated N1 target which is shown on the CDU Takeoff Reference page. To probe pilots' awareness of the relevant N1 value, the instructor manually positioned the N1 pointer on the engine display to a different value than the one indicated on the FMS-CDU. Pilots were asked which of the two values would be the target for the autothrottles during takeoff.

During an intermediate climb, the pilot-not-flying activated the CONTROL WHEEL STEERING (CWS) pitch mode by pulling on the yoke, thus overriding the active LVL CHG pitch. The CWS pitch mode maintains the vertical rate that corresponds to the pilot-induced yoke position. The pilot-flying had to determine whether the aircraft would still level off at the target altitude that had been preselected on the Mode Control Panel for the LVL CHG mode.

Anticipation of system status and behavior

Whenever transitions in aircraft behavior were imminent (e.g. level-off at a target altitude), the participants were asked what flight mode annunciations they expected to see on the ADI throughout the transition.

Table 3. Scenario Probes of Pilots' Mode Awareness

- Aborted Takeoff below 64 kts
- Frequent changes in clearances involving mode transitions
- Tailwind in takeoff clearance
- Incorrect N1 manual setting
- Activation of CWS during climb
- Ask for predictions of ADI mode indications

Study Participants

The participants in this study were 20 airline pilots who responded to postings or who were approached by the airline's training department. Participation was voluntary and pilots were paid a nominal compensation for their cooperation. The participating pilots either had a considerable amount of line experience on the B-737-300 (n=14), or they were about to finish their fixed-base transition training to the B-737-300 (n=6). Table 4 describes their biographical data and flight background.

Table 4. Biographical Data and Flight Background of the Participating Pilots

	Experienced Pilots (n=14)	Transitioning Pilots (n=6)
Age [mean(std. dev.)]	41.1 (10.1) years	41.0 (4.0) years
Total Flying Time [mean(std. dev.)]	8,471 (4,539) hours	5,183 (1,273) hours
"Glass Cockpit" Experience [mean(std. dev.)]	1,011 (582) hours	0 hours

Procedure

Pilots were asked to fly individually a 60 minute scenario on a fixed-base B-737-300 part-task trainer. This simulator works based on an actual aircraft data base. It is equipped with all relevant cockpit instruments, and it allows for any operation except hand-flying the aircraft below 1,000 ft AGL.

Upon arrival at the simulator, pilots were provided with the necessary paperwork (e.g. charts, approach plates, weather, weight manifest) as well as the LAX-ATIS and their clearance (see Appendix A). The

participants were asked to take their seat in the cockpit, and to act as Pilot-Flying (PF) during this flight. They were given time to familiarize themselves with the cockpit set-up and the intended flight. The instructor told them that weather was not a consideration, no NOTAMs existed for the flight, and all appropriate checklists would be completed during the flight.

The instructor took care of the cockpit set-up for the participant. He occupied the empty seat and acted as Pilot-Not-Flying (PNF) and ATC throughout the flight. An observer was seated behind both pilots to collect behavioral and verbal data throughout the test run and to introduce scenario events through manipulation of the simulator (e.g., introduction of failures).

With respect to FMS-related tasks, each pilot was given the following instructions:

- All FMS-related work has to be done by the PF (the participant) after activation of the autopilot at 1,000 ft AGL.
- Altitude changes on the MCP will be taken care of by the PNF (the instructor) as in actual line operations and the PF can command the PNF to carry out specified MCP manipulations for him.
- All tasks should be carried out by the participant the same way as in real line operations.
- Don't be in a hurry on the CDU or MCP! We want to keep track of what you are doing. Speed is not important for our purposes.

At various points during the scenario, pilots were asked to perform or describe FMS-related tasks, or they were asked questions concerning their FMS-related knowledge. After completion of the flight, additional questions were asked concerning FMS logic and operations, and the pilots were given the chance to ask the instructor about tasks and events that occurred during the test run.

RESULTS

The data were first analyzed across all of the participants to identify tasks and events that posed problems to the majority of pilots. Subsequently, pilots' behavior and misconceptions with respect to these probes were looked at in greater detail. A dedicated section will deal with any significant differences between the performance of pilots with glass cockpit line-experience versus those without glass cockpit experience. For some tasks that allow pilots to choose among several different approaches, the preferred strategies for the two pilot groups will be presented. Finally, problems related to mode activation which occurred across different tasks are examined more closely.

Problematic tasks/events

Less than six pilots (30 %) had any difficulties carrying out the routine tasks of changing a route (i.e. creating/entering new waypoints/airways), intercepting a radial, building/executing a holding pattern, installing an ILS approach, and entering crossing restrictions for waypoints along the route.

On the contrary, more than 14 pilots (70 %) showed deficiencies in performing the following tasks:

- Aborting a takeoff at 40 kts with A/Ts on
- Anticipating ADI mode indications throughout TO roll
- Anticipating when GA mode becomes armed throughout landing
- Disengaging APPR mode after LOC and G/S capture
- Explaining speed management for VNAV Path vs. VNAV Speed Descent
- Defining end of descent point for VNAV Path vs. VNAV Speed Descent
- Describing consequences of G/S loss above/below 1,500 ft.

The first three of these tasks are related to mode awareness either in the context of dealing with an FMS-related failure or in the sense of anticipating system status and behavior. The last four tasks point out deficiencies in pilots' knowledge of the functional structure of the system. The results revealed in detail the kinds of problems that can arise in pilot-automation interaction and the misconceptions that pilots can have about the FMS.

A. Aborted Takeoff

Immediately before receiving their takeoff clearance, pilots were asked what procedure they would use to abort the takeoff at 40 kts. Although it was emphasized that the takeoff had to be aborted at 40 kts (i.e. before THR HOLD is reached at 64 kts when the pilot can manually position the throttles), 16 pilots (80 %) described the procedure as follows: "Throttles back, reversers, and manual brakes". They did not mention that the autothrottles would have to be disconnected to prevent the throttles from coming back up again after manual intervention. When explicitly asked whether they would also disconnect the autothrottles, 3 participants (15 %) realized that they had missed that item. Two pilots (10%) were not sure about this question and suggested that they would hold the A/Ts back manually, "just in case"⁴.

Only 4 pilots (20 %) responded immediately by disconnecting the autothrottles to abort the takeoff. They were asked why this action is necessary, and all but one pilot properly described the reason. This one pilot explained that he would disconnect the autothrottles because he thought that this was standard procedure, but he indicated that he was not aware of the consequences of failing to carry out this step.

B. Anticipation of ADI indications during takeoff

Pilots were asked for their expectations concerning ADI mode indications throughout the takeoff roll as these indications are supposed to help monitor whether the system is working properly and as expected.

The relevant indications that appear in the lower left corner of the ADI are N1, i.e., the autothrottles are in charge and will go to takeoff thrust, and THR HOLD, i.e., the aircraft has reached 64 kts, the autothrottles will go to takeoff thrust but they can now be overridden manually by the pilot. Five of the pilots (25 %) expected to see both these indications. Twelve subjects (60 %) only mentioned either THR HOLD (15 % of the pilots) or N1 (45 % of the pilots) as an indication during takeoff. And another three pilots (15 %) could not predict any of the mode indications.

C. GA mode arm

The GA mode becomes available when descending below 2,000 ft radio altitude with the autothrottles armed. Out of 20 pilots, only 5 recalled the altitude at which this occurs. Eight pilots (40%) knew that the availability of the mode depends upon reaching a certain altitude but they did not remember the actual height. Another 4 pilots (20%) replied that they had no idea when the mode becomes available, and the remaining 3 pilots (15%) assumed that the GA mode is available upon glideslope capture.

D. Disengagement of the APPR mode after LOC and G/S capture

When asked to disengage the APPR mode after localizer and glide-slope had been captured, only 3 pilots (15 %) could recall the three ways of accomplishing this: either pushing the TOGA buttons on the

⁴In the debriefing, these pilots argued that they could still hold the throttles back manually to prevent them from advancing without disengaging them. But it is not clear that they would do so in the actual situation because, without understanding the FMS behavior, it seems unlikely that they would anticipate the need for manual intervention.

throttles, turning both FDs and the A/P off, or retuning the VHF radio¹. Seven pilots (35 %) did not know of any procedure for disengaging the APPR mode. Three participants (15 %) were familiar with two of the three different options.

The solutions suggested by the remaining seven pilots (35 %) included at least one possible approach, but also at least one approach that would not result in the disengagement of the APPR mode:

- 6 pilots (30%) thought that they could disengage the APPR mode by pushing the APPR key again,
- 5 pilots (25%) expected that engaging another pitch mode such as V/S or ALT HOLD would get them out of the APPR mode,
- 5 pilots (25%) thought that they would have to disengage either the A/P or the FDs, but not both,
- 4 pilots (20%) assumed that choosing another roll mode would solve the problem (e.g. HDG SEL or VORLOC).

E. Speed Management and End of Descent Point -- VNAV Path vs. VNAV Speed

Knowledge of the control modes (pitch and power) used to maintain a target airspeed during a descent is important for pilots to be able to monitor and anticipate aircraft behavior. It allows them to recognize unexpected activities or the lack of timely aircraft response. Nine out of 20 pilots knew how the FMS maintains target speed during a VNAV Path descent. Eight pilots (40%) were aware of the speed control mode during a VNAV Speed descent. With respect to the end-of-descent point of a Path vs. Speed descent, the results were similar. Twelve pilots (60 %) were aware of the end of descent during a VNAV Path descent, 9 pilots (45 %) knew at what point the VNAV Speed descent would end.

E. The consequences of a G/S failure above/below 1,500 ft

After G/S capture, a G/S signal loss was simulated at approximately 3,000 ft. Upon realizing the problem, pilots were asked about the consequences of this event. Fifty four per cent of the pilots could provide the correct answer. When asked whether a G/S failure at a lower altitude (below 1,500 ft) would have different effects, only 15 % of the pilots knew the answer. Twenty-three percent of the participants did not know the answer to either question.

Although detection time was not measured for this failure, it was observed that it took some pilots a rather long time (in some cases, several minutes) to even realize the problem although they were looking directly at the ADI (with the G/S indications and FD bars disappearing) during this phase of flight.

Differences Between Line-Experienced Versus Inexperienced Pilots

Major differences in performance between line-experienced versus transitioning pilots were seen only with respect to three of the tasks within the scenario.

When asked to intercept the LAX 248 degree radial, all 6 inexperienced pilots had difficulties carrying out the task using LNAV compared to only 50% of the 14 experienced pilots. None of the inexperienced pilots realized the need for building a fictitious waypoint on the radial. When asked about the consequences of using an excessive vertical rate of climb in the V/S mode, again 100 % of the inexperienced six subjects could not provide the correct answer compared to only 5 of the participants with line experience (36%). And finally, 83% of the six pilots without line experience could not describe how to program an intermediate descent on the VNAV CRZ page for avoiding traffic whereas none of the 14 experienced pilots had any problem with this task.

Preferred Strategies of FMS Usage

In addition to probes that only allowed for one correct answer or reaction, some situations were built into the scenario that required that pilots choose among different options to carry out the task. We asked pilots to use the automation as they would in real line operations. This provided us with behavioral data on their primary choice of modes for a given task under specified circumstances. Subsequently, we asked them about other possible strategies for achieving the same objective.

A. Intercepting a Radial outbound without a waypoint at a low altitude

There are two possible methods for accomplishing this task. Pilots can use the VOR/Localizer mode (VORLOC) which involves MCP manipulations, or they can use LNAV which requires working with the CDU. As Figure 4 illustrates, most of the pilots with glass cockpit experience preferred to use VORLOC (93 %), while the pilots in transition to glass cockpits preferred to use LNAV (83 %)⁵. While it is possible to use LNAV for this task, after one creates a fictitious fix outbound, MCP-VORLOC is the faster and easier method to do the job at low altitudes. It requires less pilot input and no heads-down time as compared to creating a fix using the CDU.

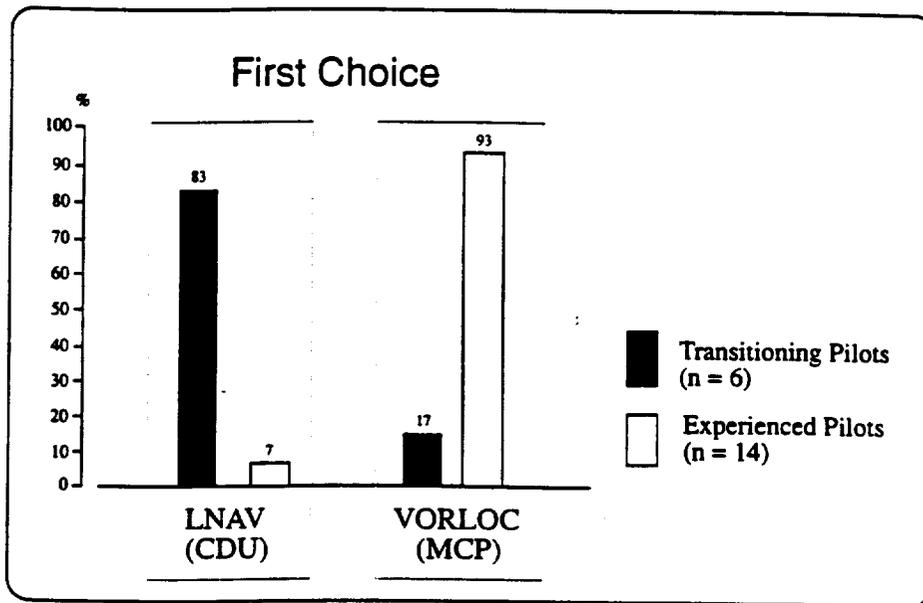


Figure 4. Preferred Mode and Level of Automation for Intercepting a Radial Outbound For Experienced versus Transitioning Pilots

B. Speed-Restricted Climb to 5,000 ft

Again there are two options available to pilots – using the LVL CHG mode via MCP manipulations or modifying data on the CDU CLB page and activating the VNAV mode. In this case, all of the pilots in transition and 79 % of the experienced glass cockpit pilots preferred the MCP-LVL CHG mode. Again, using the MCP minimizes heads down time, which is important as the aircraft is still at a very low altitude during this task.

⁵Some of the pilots in transition (16%) could not think of any second method at all.

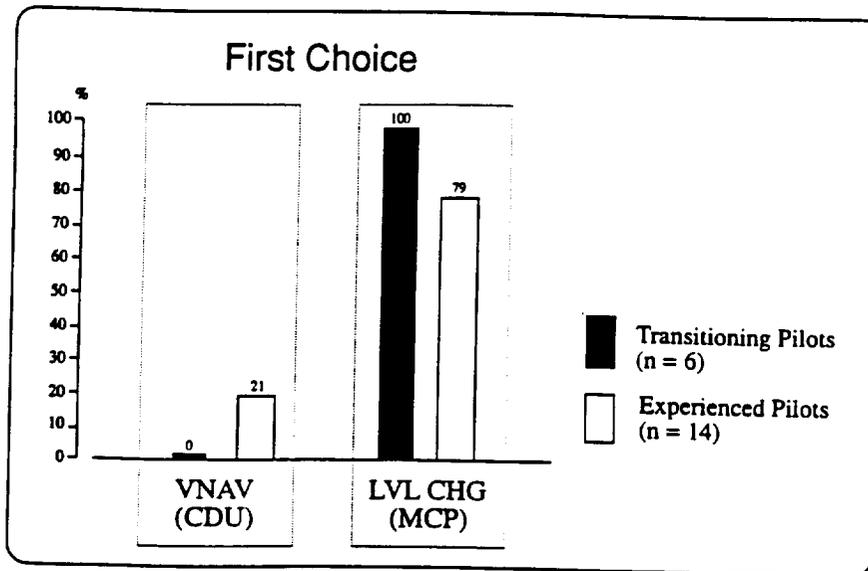


Figure 5. Preferred Mode and Level of Automation for a Speed-Restricted Climb at Low Altitude for Experienced and Transitioning Pilots

C. Unplanned Descent for Traffic at FL 290

In this situation, the pilots could either choose the LVL CHG mode on the MCP or they could "program" the descent on the CDU CRZ page and then activate VNAV. As Figure 6 shows, the majority of line-experienced pilots chose to descend using VNAV (79 %) while most of the less experienced pilots preferred to use the LVL CHG (MCP) mode (83 %). When asked why they preferred VNAV, the experienced pilots explained that, since they were at FL 290, they felt they had enough time to program the CDU. They also said that they would prefer to modify the VNAV data right away rather than switch between VNAV and another descent mode at a lower level of automation which makes it more difficult for them to keep track of active modes and targets.

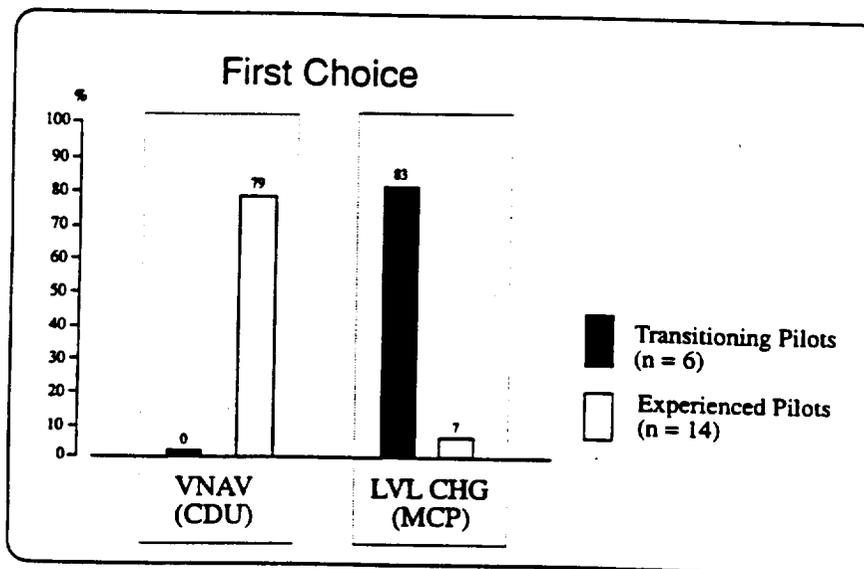


Figure 6. Preferred Mode and Level of Automation for an Unplanned Descent for Traffic at High Altitude for Experienced and Transitioning Pilots

Problems of Mode Activation

Another interesting result refers to failures to engage or re-engage a mode after entering (new) target values into either the MCP or the CDU. This omission occurred at least once during the scenario for 5 of the 6 transitioning pilots (the total number of omissions for this group was 9). Only two of the 14 experienced pilots forgot to engage an appropriate mode, and this occurred only once for each of them. The problem occurred four times in regard to the LNAV mode, six times with respect to the VNAV mode and once concerning the LVL CHG mode.

In seven of the failures to engage a mode, all required entries into the CDU or MCP were made, but no mode was activated. In the remaining four instances, the pilot would first use an MCP mode (e.g., HDG SEL) to get the system started towards the target, then he would enter the new target data into the CDU, but ultimately he would forget to switch from the MCP mode to VNAV or LNAV which use the entered CDU values as targets. The fact that in the majority of cases pilots forgot to engage VNAV or LNAV (rather than an MCP mode) after entering new target data may be related to the spatial separation between the data entry unit (CDU) and the VNAV- and LNAV-buttons which are located on the MCP.

Another problem related to mode engagement was the attempt to activate a mode without the prerequisites for this activation being met: Three (50%) of the transitioning and one of the 14 experienced pilots tried to engage VORLOC without being in the manual radio mode as required. Three (50%) of the transitioning and 5 of the 14 experienced pilots engaged the APPR mode without lowering the MCP altitude first, and they were surprised to find that the aircraft did not start the descent.

DISCUSSION

This study verifies and expands on the results obtained from the previous corpus gathering studies of pilot-automation interaction (Sarter and Woods, in press). It confirms that most of the difficulties in pilot-automation interaction are related (a) to a lack of mode awareness and (b) to gaps in pilots' mental models of the functional structure of the automation. These kinds of problems seem to occur primarily in the context of non-normal time critical situations such as an aborted takeoff. Problems related to such situations may be under-reported in surveys because these situations rarely occur in line operations. In this study, however, every participant was forced to cope with non-normal events in the scenario. In this way, latent problems in pilot-FMS interaction could be revealed.

For the majority of pilots, it was difficult or impossible to manage the cockpit automation in three non-normal situations in the scenario – an aborted takeoff, the need to disengage an automatic approach mode for collision avoidance, and the loss of the glideslope during final descent. In the case of the aborted takeoff, 65% of all participants did not understand how the autothrottle controls the aircraft throughout the takeoff. Fifteen per cent of the pilots knew about the ongoing mode activities and transitions, but they were not capable of applying this knowledge to the situation at hand. In terms of behavior, this resulted in only 4 pilots responding correctly, and one of them did not seem to completely understand the basis for this action. Another non-normal time-critical event in the scenario was the request to disengage the APPROACH mode after localizer and glideslope capture. While most of the pilots knew about at least one way of complying with this request, 14 pilots also suggested at least one ineffective approach. If, in a real world case, ATC told the pilot to immediately change heading and/or altitude to avoid a collision, there would be no time for failed attempts to disengage the mode. The pilot would have to respond immediately. This problem is related to the need for an interface design that indicates available options to help the pilot intervene quickly and directly when necessary. In the case of the third non-normal event in the scenario, the loss of the G/S during final descent, it was observed that it took many pilots fairly long to even realize that an anomaly had occurred, even though delay times could not be measured precisely. Although they were looking directly at the ADI display at this stage of

the simulated flight, it took some pilots several minutes to realize that the G/S indication and the FD bars had disappeared. This problem illustrates that cueing by absence may not be a good technique for indicating the presence of an anomaly. Not only was anomaly detection relatively slow, about one half of the participants were not aware of the consequences of a loss of the G/S. The scenario contained a variety of other probes of the pilots' ability to be "ahead of the FMS", i.e., the ability to anticipate future system behavior which can change not only in response to current pilot input but also as a result of changes in the environment, previous pilot input, or for protection purposes. For example, only one out of 20 participants could predict the entire sequence of expected mode indications for the takeoff roll. Similarly, only five of the participants knew when to expect the indication that the Go-Around mode is now available.

The underlying reason for the observed problems seems to be a lack of mode awareness. In the context of simpler devices and environments, mode awareness usually refers to the adequate assessment of the currently active mode status. But our results show that in the context of the highly dynamic and complex cockpit environment, other aspects of mode awareness are more critical. In these systems, the pilots' role has changed from active manipulator of the aircraft to supervisor of the automated systems. To fulfill this role, pilots have to a) have a thorough understanding of what a mode means in terms of system behavior and b) have to be "ahead of the FMS", i.e. they have to be able to anticipate future system behavior which can change not only in response to his own input but also as a result of changes in the environment or for protection purposes (see Reason, 1990).

Operational Costs of Technology Centered Automation

New automation is developed because of some payback (precision, more data, reduced staffing, etc.) for some beneficiary (the individual practitioner, the organization, the industry, society). But often overlooked is the fact that new automated devices also create new demands for the individual and groups of practitioners responsible for operating and managing these systems. The new demands can include new or changed tasks (setup, operating sequences, etc.), and new cognitive demands are created as well. There are new knowledge requirements (e.g., how the automation functions), new communication tasks (e.g., instructing the automation in a particular case), new data management tasks (e.g., finding the relevant page within the CDU page architecture), new attentional demands (tracking the state of the automation), and new forms of error or failure (e.g., mode error). This study reveals some of these kinds of costs that occur in the context of the current generation of cockpit automation – costs that could be minimized or eliminated through skillful design of human-centered automation (Billings, 1991).

Mode Error and Mode Awareness

Two of the cost centers associated with changes in automation are the possibility of new forms of error or failure and the possibility of creating new cognitive demands for practitioners. Interlinked examples of these effects of automation for the glass cockpit case seem to be mode error and mode awareness.

Devices that allow some thing to be done one way in one mode and another way in another mode create the possibility of mode errors where one executes an intention in a way appropriate to one mode when the device is actually in another mode (Norman, 1988). Automated systems like those in the glass cockpit cannot be characterized by a single mode setting. There are a number of subsystems each involving a number of possible mode settings. This increase in the power and flexibility of automated resources creates a form of operational complexity that increases the potential for mode errors.

But advanced automation like the FMS extends the kinds of mode related problems that can occur because system status and behavior can change independent of immediate and direct pilot commands due to situation factors or protection limits (Sarter and Woods, 1992). This means that a new cognitive

demand is created: the need to maintain awareness of externally induced mode transitions. As the pilot's role has changed from active manipulator of the aircraft to supervisor of automated systems, effective situation awareness requires pilots to stay 'ahead of the FMS,' i.e., he or she has to be able to anticipate future system behavior or detect system failures (Sarter and Woods, 1991). However, in this study, only five out of 20 participants could predict the operationally most significant mode indications (N1 and THR HOLD) for the takeoff roll and only 5 of the participants knew when to expect the indication that the Go-Around mode is available.

One way to interpret the results of this study and the complementary results of Sarter and Woods (in press) is that many of the observed problems result from a lack of mode awareness – the pilots lost track of system targets and missed mode changes that occurred independent of immediate pilot commands. Maintaining mode awareness requires that pilot attend to and integrate data from a variety of indications in the cockpit such as the Flight Mode Annunciations on the Attitude Director Indicator, the visualization of the programmed route of flight on the Horizontal Situation Indicator, or the display of target values on the Mode Control Panel. Breakdowns in mode awareness may be due to characteristics of these indications given the nature of the cognitive demands of high tempo phases of flight or non-normal flight situations. Another contributor to these attentional breakdowns may be limits and gaps in the pilots mental models of the automated resources.

New Knowledge Requirements

The transition to glass cockpit aircraft requires pilots to learn a great deal about the FMS and other flight deck automated subsystems. As the results of this study show and given the results of the previous corpus building studies, there are a number of areas where pilots have gaps in their understanding of the functional structure of the FMS. By forcing pilots to deal with various non-normal situations, gaps or errors in their understanding of how the automation works in various situations were revealed. Again, the results indicated that pilots do not have an accurate model of how VNAV descent modes work and that the displays do not help them in tracking either the targets or the control modes used by VNAV Path and VNAV Speed descents. Overall, this study confirms previous results (Sarter and Woods, in press) and shows that these problems can occur even with pilots who have relatively extensive glass cockpit experience.

Note the interaction between two factors. Breakdowns in mode awareness can be due in part to a lack of effective feedback on the state of the automation and in part due to buggy mental models of the automation. The lack of feedback on the state of the automation can in turn limit pilots' ability to learn from experience and correct or elaborate their mental models of system function over time. It also limits their ability to learn to perceive the state of the automation from the available indications. A third factor further complicates the difficulty. Many of flight situations that stress these problems occur relatively rarely in line operations. This combination has broad repercussions for training pilots to manage highly automated aircraft. First, training must go beyond simply providing pilots with facts about the FMS. The results showed that sometimes pilots possessed knowledge in the sense of being able to recite the facts, but that they were unable to apply the knowledge successfully in an actual flight context. This is called the problem of inert knowledge. Training must conditionalize knowledge to the contexts where it is utilized. Second, pilots need to learn not simply how the automated system works, but also how to work the system. This will require scenarios and instruction designed around managing the transitions between different modes of automation. Third, since pilots do learn a subset of methods to be able to make the system work under routine conditions, situations that challenge their current understanding may arise relatively infrequently (or go unnoticed as such in part due to lack of feedback about the state and behavior of the FMS). This means that an ongoing learning programs will need to be devised that help even experienced glass cockpit pilots discover and correct subtle bugs in their mental models of the FMS or to elaborate their understanding of how the automation works in particular situations in a risk-free environment.

Knowledge Miscalibration

The results indicate that pilots have gaps in their understanding of the functional structure of the FMS. Furthermore, there are some indications in the data that pilots are miscalibrated with respect to their understanding of the FMS, that is, the pilots may not be aware of the gaps in their mental models. An expert is well calibrated if they are aware of the areas and circumstances where they have correct knowledge and the areas where their knowledge is incomplete or limited. If the expert is overconfident and believes that they understand areas where in fact their knowledge is incomplete or limited, then that person is said to be miscalibrated (e.g., Wagenaar and Keren, 1986). Note that degree of calibration is not necessarily correlated with expertise.

When we compare pilot responses to questions like, "how much do you agree or disagree with the statement: 'there are modes and features of the FMS that I still don't understand' " (Wiener, 1989; Sarter and Woods, in press) to the behavioral data in this study, there is some indication that glass cockpit pilots are overconfident and miscalibrated about how well they understand the FMS. When forced to cope with flight situations that challenge their ability to monitor and manage cockpit automation by the design of the scenario, the number and severity of pilots' problems was higher than would be expected from previous survey data, in particular for pilots with line experience in glass cockpits. Some of the participants in this study made comments in the post-scenario debriefings such as: "I never knew that I did not know this. I just never thought about this situation." Similar results have been obtained in studies of physician interaction with computer based automated devices in the surgical operating room (Cook et al., 1991; Moll van Charante et al., 1992)

There are several factors that could contribute to the observed miscalibration. First, areas of incomplete or buggy knowledge can remain hidden from pilots because pilots have the capability to work around these areas by sticking with a few well practiced and well understood methods. In addition, flight situations that force pilots into areas where their knowledge is limited and miscalibrated may arise infrequently. Second, studies of calibration have indicated that the availability of feedback, the form of feedback and the attentional demands of processing feedback can effect knowledge calibration (e.g., Wagenaar and Keren, 1986). Problems with ineffective feedback on the state and behavior of the FMS that were observed in this study and reported in previous studies of pilot interaction with cockpit automation (e.g., Norman, 1990) could be a factor that contributes to poor calibration of pilots, i.e., a lack of awareness of the gaps in their mental models of the FMS. The relationship between poor feedback and miscalibrated practitioners was also found in studies of physician-automation interaction (e.g., Cook et al., 1991). Knowledge miscalibration in pilots, if it is widespread, is one factor that could lead to under-reporting of problems with cockpit automation in survey studies.

How to Manage Automated Resources

Cockpit automation provides a large number of functions and options for carrying out a given flight task under different circumstances. For example, the FMS provides at least five different mechanisms at different levels of automation for changing altitude. This flexibility is normally construed as a benefit that allows the pilot to select the mode or option best suited to a particular flight situation (e.g., time and speed constraints). However, this flexibility creates new demands as well. Pilots must learn and know about the functions of the different modes, how to coordinate which mode to use when, how to switch from one mode to another smoothly. In other words, the pilots must know how the automated system works and he or she must develop skill at how to work the system. To meet the latter criterion, a pilot must:

- learn about all of the available options,
- learn and remember how to deploy them across a variety of operational circumstances, especially rarely occurring but more difficult or more critical ones,

- learn and remember the interface manipulations required to invoke the different modes or features,
- learn and remember how to interpret or where to find the various indications about which option is active or armed and what are its associated target values.

The results of this study indicate that pilots become proficient and maintain their proficiency on only a subset of the modes and options provided by the FMS. Further evidence for this phenomenon was provided by previous FMS-related studies (Sarter and Woods, in press) and by studies of human-machine interaction in other domains (e.g., Rosson, 1983; Cook et al., 1990) where users hardly ever use more than a small subset of the options provided. This is, in part, a consequence of the increased costs involved in learning extra functions, but it also allows practitioners to protect themselves from having to make difficult decisions due to an increased number of alternatives. In the case of the FMS, pilots try to manage the system within a set of stereotypical responses or techniques. In this study, we were able to compare the tactics selected by pilots with line experience in glass cockpits versus pilots without previous glass cockpit experience. The results indicate that, over time, pilots learn to select among the various options depending on situation factors (e.g., altitude, time constraints) and on expectations (e.g., the likelihood of deviation from plan). But pilots who just had finished their transition training were much less sensitive to these contextual factors. They tended to always use the highest level of automation independent of context.

Note that, in higher tempo phases of flight, more experienced pilots chose to use intermediate levels of automation which use the MCP as the interface over higher levels of automation that require CDU interaction. The MCP based modes generally require less interaction, less head down time, less diversion of attention to the interface itself (e.g., remembering the necessary interface manipulations). In addition, the modes of automation accessed through the MCP as an interface tend to respond only to direct pilot input (e.g., the pilot enters a target value, activates a mode of control, the automation then responds by capturing and maintaining that target value until another pilot command is received) and do not initiate a sequence of automated system activities. This may explain previous results that pilots see the MCP and the CDU as separate systems (Sarter and Woods, in press) despite the fact that from an engineering point of view both are part of an integrated FMS. Operationally, interacting with the MCP modes has a different character than 'programming' the CDU. This means that general questions about pilots' attitudes towards cockpit automation in general are ambiguous, and pilots may vary from each other and from the investigator in their interpretation of what aspects of cockpit automation the question refers to.

SUMMARY

The results of this and previous studies of pilot interaction with cockpit automation in commercial aviation yield consistent results across diverse methods. While pilots seem to be able to "make the system work" in standard situations, one of the most important results of this study is the discovery of latent problems with pilot-FMS interaction that can affect even experienced pilots' performance in non-normal time critical situations. The severity and importance of these problems is underestimated due to several interacting factors:

- there are gaps in pilots' understanding of the functional structure of the automation,
- the opaque interface between pilots and automation makes it difficult for pilots to track the state and activity of the automation,
- pilots may not be aware of the gaps in their knowledge about FMS function,
- pilots can 'escape' from the CDU to the MCP whenever a situation gets too complicated or time pressure is too high, and
- the flight situations where these problems help produce unmistakable performance difficulties may occur infrequently in line observations.

The data in this study, in conjunction with the data from previous studies (e.g., Wiener, 1989; Norman, 1990; Sarter and Woods, in press), point out some of the costs of the 'clumsy' use of technological possibilities from an operational point of view. These costs should provide input to designers trying to develop human-centered automation and to trainers trying to develop new instructional programs for developing, maintaining and testing pilot proficiency in managing automated resources. However, it is important to remember that the problems in pilot interaction with cockpit automation are not inherent in the technology itself, but rather these problems result from limitations in how the automation and the human pilots are integrated together as a joint, distributed cognitive system (Hutchins, 1991; Woods, 1993).

ACKNOWLEDGEMENTS

This work was supported under a Cooperative Agreement (NCC 2-592) with the Aerospace Human Factors Research Division of the NASA-Ames Research Center (Technical Monitor: Dr. Everett Palmer).

The authors would like to very much thank all of the pilots who participated in the study and shared their experience with us. We are also very grateful for the support and patience of a large number of people at the collaborating airline who made it possible to carry out this line of research.

REFERENCES

- Billings, C.E. (1991). *Human-centered aircraft automation philosophy* (Technical Memorandum 103885). Moffett Field, CA: NASA-Ames Research Center.
- Cook, R.I., Woods, D.D., and Howie, M.B. (1990). The natural history of introducing new information technology into a dynamic high-risk environment. *Proceedings of the Human Factors Society, 35th Annual Meeting*. Orlando, FL.
- Cook, R.I., Potter, S.S., Woods, D.D. and McDonald, J.M. (1991). Evaluating the human engineering of microprocessor-controlled operating room devices. *Journal of Clinical Monitoring*, 7, 217-226.
- Eldredge, D., Dodd, R.S., and Mangold, S.J. (1991). *A review and discussion of Flight Management System incidents reported to the Aviation Safety Reporting System*. Battelle Report, prepared for the Department of Transportation, Volpe National Transportation Systems Center.
- Hutchins, E. (1991). *How a cockpit remembers its speed*. (Technical Report). San Diego: University of California, Distributed Cognition Laboratory.
- James, M., McClumpha, A., Green, R., Wilson, P., and Belyavin, A. (1991). *Pilot attitudes to flight deck automation*. Paper presented at the One Day Conference "Human Factors on Advanced Flight Decks", The Royal Aeronautical Society, London.
- Moll van Charante, E., Cook, R.I., Woods, D.D., Yue, L., and Howie, M.B. (1992). Human-computer interaction in context: Physician interaction with automated intravenous controllers in the heart room. *Proceedings of the Fifth IFAC/IFIP/IFOE/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. The Hague, Netherlands.
- Norman, D.A. (1988). *The psychology of everyday things*. New York: Basic Books.

- Norman, D.A. (1990). *The 'problem' with automation: Inappropriate feedback and interaction, not "over-automation"*. Philosophical Transactions of the Royal Society of London, B 327.
- Rosson, M.B. (1983). Patterns of experience in text editing. In *CHI83: Human Factors in Computing Systems*, pp. 171-175. New York: ACM.
- Sarter, N.B. and Woods, D.D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), p.45-57.
- Sarter, N.B. and Woods, D.D. (1992). Mode error in supervisory control of automated systems. *Proceedings of the Human Factors Society 36th Annual Meeting*, (pp. 26-29). Atlanta, GA.
- Sarter, N.B. and Woods, D.D. (in press). Pilot interaction with cockpit automation I: Operational experiences with the Flight Management System. *International Journal of Aviation Psychology*.
- Wagenaar, W.A. and Keren, G.B. (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Mancini and D.D. Woods (eds.) *Intelligent Decision Support in Process Environments*. Springer-Verlag, New York.
- Wiener, E.L. (1989). *Human factors of advanced technology ("glass cockpit") transport aircraft* (NASA Contractor Report No. 177528). Moffett Field, CA: NASA-Ames Research Center.
- Woods, D.D. (1993) Process tracing methods for the study of cognition outside of the experimental psychology laboratory. In G.A. Klein, J. Orasanu and R. Calderwood (eds.), *Decision making in action: models and methods*. Ablex: New Jersey.
- Woods, D.D. (1993) The price of flexibility. In W. Hefley and D. Murray (eds.) *Proceedings of the ACM International Workshop on Intelligent User Interfaces*.
- Woods, D.D. and Sarter, N.B. (in press) Evaluating the Impact of New Technology on Human-Machine Cooperation. In J. Wise, V.D. Hopkin, and P. Stager (eds.) *Verification and Validation of Complex and Integrated Human-Machine Systems*. Springer-Verlag, New York.

"How In the world did I ever get into that mode?" Mode Error and Awareness in Supervisory Control¹

Nadine B. Sarter and David D. Woods
Cognitive Systems Engineering Laboratory
The Ohio State University
Columbus, OH

ABSTRACT

New technology is flexible in the sense that it provides practitioners with a large number of functions and options for carrying out a given task under different circumstances. But this flexibility also has a price: It is the job of the human supervisor to select the mode best suited to a particular situation. The practitioner must know more and has new monitoring and attentional demands to track which mode the automation is in and what it is doing to manage the underlying processes. When designers proliferate modes, these new cognitive demands contribute to new mode-related error forms and failure paths. While mode error has been discussed in human-computer interaction for some time, the increased capabilities of new automated systems appear to have created new types of mode-related problems, in particular, a decrease in mode awareness. We explore these new aspects based on results from our own and related studies of human-automation interaction. In particular, we draw on empirical data from a series of studies of pilot-automation interaction in commercial glass cockpit aircraft to illustrate the nature, circumstances, and potential consequences of mode awareness problems in supervisory control of automated resources. The result is an expanded view of mode error that takes into account the new demands imposed by more automated systems. We also argue that in order to be able to develop countermeasures to the new varieties of mode related problems, a better understanding of the cognitive processes underlying mode awareness and situation awareness, in general, is needed.

INTRODUCTION

New technology is increasing the potential for automated resources to support human supervisory controllers. The technology's inherent flexibility allows designers to add a wide range of capabilities in the name of providing the practitioner with a set of tools that can be used to optimize system performance across a wide range of circumstances. However, the same flexibility tends to create and proliferate various modes of operation. This proliferation of modes that can so easily accompany new levels of automation in complex systems also creates new cognitive demands on practitioners (Woods, 1993). Practitioners must know more – both, about how the system works in each different mode and about how to manage the new set of options in different operational contexts (Sarter and Woods, 1992; 1993). New attentional demands are created as the practitioner must keep track of which mode the device is in, in order to select the correct inputs when communicating with the automation and in order to track what the automation is doing now, why it is doing it, and what it will do next. For example, an automated cockpit system such as the Flight Management System (FMS) is flexible in the sense that it provides pilots with a large number of functions and options for carrying out a given flight task under different circumstances. There are at least five different methods at different

¹Submitted for Publication

levels of automation that the pilots could invoke to change altitude. This flexibility is usually portrayed as a benefit that allows the pilot to select the mode best suited to a particular flight situation. But this flexibility also has a price: the pilots must know about the functions of the different modes, how to coordinate which mode to use when, how to 'bumplessly' switch from one mode to another, how each mode is set up to fly the aircraft, and he has to keep track of which mode is active.. These new cognitive demands can easily congregate at high tempo and high criticality periods of device use thereby adding new workload at precisely those time periods where practitioners are most in need of effective support systems. Clumsy use of technological possibilities, such as the proliferation of modes, creates the potential for new forms of human-machine system failure and new paths towards serious critical incidents, e.g., the air crashes at Bangalore (e.g., Lenorovitz, 1990) and Strasbourg (Monnier, 1992).

In a variety of studies, we have investigated human-automation interaction in the context of commercial 'glass' cockpits (Sarter and Woods, 1992; 1993) and in the context of anesthetic management under surgery (Cook et al., 1990; Moll van Charante et al., 1992). Based on these and other studies, we think that the classic concept of mode error is inadequate to describe the problems in human interaction with today's automated resources. In this paper, we extend the concept of mode error to take into account the problems caused by new automation capabilities – mode awareness.

MODE AWARENESS

Multiple modes in devices can create the potential for mode errors. The concept of mode error has been established as one kind of problem that can occur in human interaction with computerized devices (Lewis and Norman, 1986) and as a basic kind of erroneous action in psychological taxonomies of error forms (Norman, 1981). Norman (1988) summarizes the source of mode error quite simply by suggesting that if one wishes to create or increase the possibilities for erroneous actions one way is to "... change the rules. Let something be done one way in one mode and another way in another mode." When this is the case, a human user can commit an erroneous action by executing an intention in the way appropriate to one mode of the device when the device is actually in another mode. Note that mode error is inherently a human-machine system breakdown in that it requires that the users lose track of which mode the device is in (or confuse which methods or actions are appropriate to which mode) and it requires a machine where the same actions and indications mean different things in different modes of operation. Several studies have shown that human-computer interface design and evaluation should identify computerized devices which have a high potential for mode errors (e.g., Lewis and Norman, 1986; Cook et al., 1991), and several design techniques have been proposed to reduced the chances for mode errors (Monk, 1986; Sellen et al., 1992).

The original work on mode error was done primarily in reference to relatively simple computerized devices, such as word processors. The erroneous actions in question were acts of commission in carrying out self paced tasks with devices that only reacted to user inputs and commands. Increases in the complexity and autonomy of automated systems for event-driven, dynamic task environments such as commercial aviation flightdecks and anesthetic management under surgery have resulted in a proliferation of system and interface "modes." Human supervisory control of automated resources in event-driven task domains is a quite different type of task environment as compared to the applications in the original research on mode error. Automation is often introduced as a resource for the human supervisor providing him with a large number of modes of operation for carrying out tasks under different circumstances. The human's role is to select the mode best suited to a particular situation, but to accomplish this he or she must know more and must meet new monitoring and attentional

demands to track which mode the automation is in and what it is doing to manage the underlying process. These cognitive demands can be particularly challenging in the context of highly automated systems which can change modes on their own based on environmental inputs or for protection purposes, independent of direct and immediate instructions from the human supervisor. This capability of highly automated systems drives the demand for mode awareness, that is, the ability of a supervisor to track and to anticipate the behavior of automated systems.

What is involved in maintaining mode awareness is determined to a large extent by the design and capabilities of the automated resources and especially the interface between the automation and the people in the system. Therefore, how have changes in automation and in the interface between person and automated resources impacted mode awareness? How has the human's role and tasks changed and how can they be supported?

THE COMPLEXITY OF MODES IN AUTOMATED SYSTEMS AND THE CHALLENGE TO MODE AWARENESS

Early automated systems were characterized by a fairly small number of modes. In most cases, these modes provided the passive background on which the operator would act by entering target data and by requesting system operations. Another characteristic of these early systems was that they would only have one overall mode setting for each function to be performed. Consequently, mode annunciations (indications of the currently active mode and of transitions between modes) could be dedicated to one spatial location on the display. Finally, consequences of a breakdown in mode awareness tended to be fairly small, in part because of the short time-constant feedback loops involved in these systems. The operator seemed to be able to detect and recover from erroneous actions relatively quickly.

The flexibility of more advanced technology allows automation designers to develop much more complicated mode-rich systems. Modes proliferated by providing multiple levels of automation and by providing more than one mode option for many individual functions. The result is numerous mode indications distributed over multiple displays each containing just that portion of the mode status data corresponding to a particular system or subsystem. Furthermore, the designs allow for interactions across the various modes. The increased capability of the automated resources themselves creates increased delays between user input and feedback about system behavior. This increase to longer time-constant feedback loops increases the difficulty of error or failure detection and recovery and challenges the human's ability to maintain awareness of the active modes, the armed modes, the contingent interactions between environmental status and mode behavior, and the contingent interactions across modes.

A very important trend relates to the input sources that can evoke changes in system status and behavior. Early systems would change their mode status and behavior only in response to operator input. Advanced technology, on the other hand, responds to operator input as well as situational and system factors. In the case of the Flight Management System in highly automated cockpits, for example, a mode transition can occur as an immediate consequence of operator input. But it can also happen when a preprogrammed intermediate target (e.g., a target altitude) is reached or when the system changes its mode in order to prevent the pilot from putting the aircraft into an unsafe configuration.

Even the aspect of operator input has itself become more complicated as the complexity of the system of automated resources has increased. Incidents and accidents have shown that there is

an increased risk of inadvertent activation of modes by the operator. A mode can not only be activated through deliberate explicit selection of the mode by pushing the corresponding button. In addition, pushing a button can result in the activation of other different modes depending on the system status at the time of manipulation. The resulting system behavior can be disastrous but may be missed by the operator if adequate feedback is not provided to support mode awareness.

An example of such an inadvertent mode activation contributed to a major recent accident in the aviation domain (the Bangalore crash; e.g., Lenorovitz, 1990). In that case, the pilot put the aircraft into a mode called OPEN DESCENT without realizing it. This resulted in the aircraft speed being controlled by pitch rather than thrust, i.e., throttles went to idle. In that mode, the automation ignores any preprogrammed altitude constraints. To maintain the pilot-selected target speed without power, the automation had to use an excessive rate of descent which ultimately led to the crash of the aircraft short of the runway. How could this happen?

There are at least three different ways of activating the OPEN DESCENT mode. First, it can be selected by pulling the ALTITUDE knob after selecting a lower altitude. Second, it can be activated by pulling the SPEED knob provided the aircraft is in the so-called EXPEDITE mode at that point in time. And third, the OPEN DESCENT becomes active when selecting a lower altitude while in the ALTITUDE ACQUISITION phase. This latter indirect option contributed to the above accident. The pilot must not have been aware of the fact that the aircraft was within 200 feet of the previously entered target altitude (which puts the system into the ALTITUDE ACQUISITION mode). Consequently, he may not have expected the selection of a lower altitude at that point in time to result in a mode transition. As he did not expect any mode change, he may not have closely monitored his mode annunciations and thus missed the transition. It was not until 10 seconds before impact that the crew discovered what had happened; too late for them to recover with the engines at idle.

Display modes are another factor aggravating the problem of mode awareness. In some devices, the current mode configuration does not only determine what control functions become activated by a given input; rather, these devices also interpret user-entered target values differently depending on the active display mode. In the following example, it is easy to see how this may result in unintended system behavior. In a current glass cockpit aircraft, pilots enter a desired vertical speed or a desired flight path angle via the same display. The interpretation of the entered value depends on the active display mode. But although the verbal expressions for different targets differ considerably (for example, a vertical speed of two thousand five hundred feet versus a flight path angle of two-point-five degrees), these two targets on the display look almost the same. The pilot has to verify the mode indication for this display instead of the display format supporting an intuitive, mentally economical apprehension of the active mode. In this case, the problem is further aggravated by the fact that the pilot is increasingly removed from the actual ongoing process as previously available cues about system behavior such as moving throttles or noise may have been reduced or removed in the design process .

The behavior and capabilities of the machine agent in human-machine systems have changed considerably. In simpler devices, each system activity was dependent upon operator input; consequently, in order for a lack of mode awareness to become operationally significant, the operator had to act to evoke undesired system behavior. In more automated systems, the level of animacy of machine agents has dramatically increased. Once activated, systems are capable of carrying out long sequences of tasks autonomously. For example, advanced Flight Management Systems can be programmed to automatically control the aircraft from takeoff through landing. Inadvertent mode settings and selections may not produce visible consequences for a long time complicating the process of error or failure detection. This creates the possibility of errors of

omission (i.e., failure to intervene) in addition to errors of commission as a consequence of a lack of mode awareness.

Another complicating factor that makes it difficult to maintain awareness of the active mode configuration is the fact that many systems allow for simultaneous use by multiple practitioners rather than input by just one individual user. Tracking system status and behavior becomes more difficult if it is possible for other users to interact with the system without the need for consent by all of the practitioners involved. This problem is most obvious when two experienced operators have developed different strategies of system use. When they have to cooperate, it can be particularly difficult for them to maintain awareness of the history of interaction with the system which may determine the effect of the next system input.

All of the factors mentioned above challenge a human supervisor's ability to maintain mode awareness in highly automated systems. The results of a number of studies of human-automation interaction in a variety of domains have indicated that problems in mode awareness are often a consequence of technology centered automation (e.g., Sarter and Woods, 1992 and 1993; Cook et al., 1991; Moll van Charante et al., 1992). In the following section, we will examine in more detail results of a series of studies on pilot-automation interaction that illustrates the trends in mode awareness problems in the context of the mode-rich cockpit environment.

SOME EMPIRICAL RESULTS ON MODE AWARENESS IN PILOT-AUTOMATION INTERACTION

The role of pilots in modern glass cockpit aircraft has shifted from direct control of the aircraft to supervisory control of automated machine agents. One of the core automation systems in these cockpits is the Flight Management System (FMS) which can be programmed by the pilot to automatically follow a desired flight path and profile from takeoff through landing. To maintain awareness of the status and behavior of the various modes of operation within the FMS, pilots have to gather and integrate a variety of data from numerous different displays in the cockpit. In addition to monitoring these nominal indications of system targets and status, pilots need to be able to interpret the indications to extract what is implied about current and future system and aircraft behavior. In other words, the automation is becoming more of a dynamic process in itself, where the indications are a kind of 'raw' data which require an act of interpretation in order to become information about current or future states. Interpreting the raw indications requires the human supervisor to have an adequate mental model of the various automated modes, their inter-relationships, and knowledge of how to use these as resources in various contexts.

Given the fairly low rate of change in aircraft behavior throughout large parts of the flight, the pilot does not have to continuously monitor the mode annunciators. Rather, he has to be able to predict the occurrence of transitions in system behavior to attend to the right indications at the right time. During busy phases of flight (e.g., final approach), numerous changes in system status and behavior can occur in a very short period of time. During this high tempo phase of flight, with a large number of concurrent tasks, the crew now has another set of cognitive tasks to perform – monitoring and interpreting mode annunciators relative to expected behavior.

In a series of studies of pilot-automation interaction, we had the opportunity to investigate the nature and circumstances of mode-related problems in highly automated glass cockpit aircraft. In one investigation, we built a corpus of FMS-related problems that were encountered in line operations (Sarter and Woods, 1992). For this purpose, descriptions of automation surprises were collected from experienced airline pilots. A second converging activity was to observe

pilots during their transition training from a conventional to a glass cockpit aircraft, i.e. before they had a chance to adapt to the system. Analysis of these corpus data suggested that difficulties in mode awareness and gaps in pilots' understanding of all of the modes of operation and their interactions contributed to automation surprises and related supervisory control difficulties. Based on the results from the corpus gathering studies, a field experiment was carried out to try to examine pilot-automation interaction more closely (Sarter and Woods, 1993). Twenty airline pilots were asked to fly a mission on a part-task flight simulator. The scenario was designed to contain numerous tasks and events that served as probes of pilots' mode awareness and of their mental model of the automation. This phenomenon-driven scenario design permitted on-line data collection on various issues regarding pilot-automation interaction. In addition, we were able to question the pilots about their knowledge and assessments of the status and behavior of the automation during low workload phases of the simulated flight and after completion of the simulation.

These studies provided consistent and converging data to help understand why and under what circumstances pilots encounter problems related to the interaction with cockpit automation. Most of the observed difficulties were related to lack of mode awareness and to gaps in mental models on how the various automated modes work and interact. The problems in coordination between pilot and automation (e.g., automation surprises) occurred primarily in the context of non-normal, time-critical situations, for example, aborted takeoff, disengaging an automatic mode during approach for collision avoidance, and loss of the glideslope signal during final approach. In the case of the aborted takeoff, 65% of the pilots were not aware that the autothrottle system was in charge of thrust control. Consequently, they did not disengage the autothrottles in order to have full manual control of the throttle setting. In the debriefing, 15% of these pilots could describe the active mode settings and the system activities during takeoff. But their knowledge was inert, i.e., they had not been capable of applying this knowledge to the ongoing situation. Overall, only four out of twenty participants responded completely correctly in managing the automation during the aborted takeoff, and one of these four pilots explained that he did so because he was trying to comply with standard procedures, not because he understood what was going on within automation. In the second non-normal situation, the pilots had to quickly comply with an ATC request to disengage the automatic APPROACH mode after localizer and glideslope capture in order to change heading and altitude to avoid a collision. Most of the pilots knew only one of the several methods to disengage the mode, and fourteen pilots also 'knew' at least one inappropriate method which could lead to delayed responses to the ATC request. In the case of the glide slope loss during final approach, about one half of the pilots were not aware of the consequences of this event in terms of FMS behavior. They could not explain the effects in the debriefing, and many of them even had difficulties detecting the occurrence of the problem during the ongoing simulation.

Another interesting result of this study was related to the future-oriented aspect of mode awareness. Pilots sometimes had problems anticipating system behavior and the associated mode annunciations. For example, only five of the participant knew when to expect the indication that the Go-Around mode would be available. Failures to anticipate mode status and transitions, like this one, indicate a lack of mode awareness which degrades the pilot's ability to allocate attention effectively and to detect errors, failures, or miscommunications between pilot and automation prior to explicit flight events – automation surprises. The more experience pilots had with the automation, the more they were capable of applying their knowledge about the advantages and shortcomings of the different modes to manage the automated resources in different contexts. Pilots with less glass cockpit experience tended to utilize a single strategy or mode over a wide range of flight circumstances. One could interpret this as an attempt to cope with the complexity of the automation by ignoring some modes and options, even in situations where the stereotypical strategy was less than ideal relative to other strategies for managing the

automated resources. Finally, there were several instances of pilots who instructed the automation by entering new flight path related targets but who did not activate a mode of the automation to work on acquiring these targets. They were surprised when the aircraft did not respond as expected; they did not realize or understand why their instructions to the automation had not resulted in the desired change. In some sense, this is a good example to show how pilots try to communicate with the system in a way analogous to communication with another human agent. They assume that entering a desired target value is sufficient for the system (as it would be for a human crewmember) to understand that it is supposed to achieve this new target and how it is supposed to do so in detail.

These investigations into one specific field of activity illustrate a trend in human-machine cooperation (Woods, 1993). Technology allows a proliferation of modes of increasing complexity and capability for autonomous activity. These changes create new cognitive demands for human supervisory controllers, demands which tend to congregate at higher tempo epochs where workload demands are highest (cf., also Moll van Charante et al., 1992 for similar results from another field of practice). The complexity of modes challenges the human supervisor's ability to track and anticipate the behavior of the automation – mode awareness. Difficulties in maintaining mode awareness focus on transitions between more quiescent phases or situations where mode behavior is complex or transitions frequent.

SOURCES OF PROBLEMS IN MODE AWARENESS

The data on problems in mode awareness imply that there are two kinds of contributing factors. One is buggy mental models. The other is opaque indications of the status and behavior of the automation. The former derives from a failure of the designers of automation to anticipate the new kinds of knowledge demands their automation creates and a failure to provide mechanisms to help practitioners acquire and maintain this knowledge in ways usable in actual operational contexts. The latter derives from a failure of designers to support the supervisor's increasingly challenging cognitive demand of tracking the state and behavior of the automation as another kind of dynamic process within their scope of responsibility (e.g., Norman, 1990). The indications of the nominal status of the automation are a kind of data; the practitioner must interpret this data to develop and maintain an assessment of the automation as process and the automated process over time. The data on problems in mode awareness strongly imply that this cognitive demand is poorly supported by the kinds of displays on the state of the automation currently provided to practitioners. As Earl Wiener likes to put it: the three most commonly asked questions on the highly automated flightdeck are – what is it doing? why is it doing that? what will it do next? (to which we would like to add a fourth – how in the world did we get into that mode?). The interpretation of data on the automation as process is apparently a cognitively demanding task with these displays rather than a mentally economical one. This is troublesome when this cognitive task is important during high tempo, high workload, high criticality situations.

COPING WITH MODE ERROR AND AIDING MODE AWARENESS

The examination of mode awareness here leads us to several strategic directions for responding to problems in this cognitive task. First, one can say that mode awareness problems are induced by the complexity of the technological system. Technological powers for automation are used clumsily when the cognitive and other kinds of demands on the operational system created by new automation are ignored (Norman, 1990; Woods, 1993; Woods et al., 1993). This is what we mean by technology-centered automation (cf., Billings, 1991 for an extensive discussion of

technology-centered versus human-centered automation). Then one avenue to improve the human-machine system is to reduce the operational complexity induced by how technology is deployed. In the case of mode awareness this can be stated quite clearly – reduce the number and complexity of the modes. However, there may be a variety of pressures, such as marketing demands from a diverse set of customers or the methods for optimizing various parameters like precision or efficiency of resource consumption vary greatly across different operational circumstances, which reduce the designer's ability to counter mode proliferation.

Two other directions for change probably are very tightly coupled in their implementation in a real field of practice, although they can be discussed separately. One is to support the new knowledge demands created by increasingly complex automated resources through new approaches to training human supervisory controllers. This is much more than simply a new list of facts about how the automation works. Instead, it must be focused on knowledge activation in context in order to avoid what we are already seeing – inert knowledge where the user can recite the facts but fails to use the knowledge effectively in real operational contexts. Training to enhance skill at control of attention would also be relevant here (Gopher, 1991).

The new knowledge demands require that more attention be paid to developing and teaching knowledge and strategies for how to work the system of automated resources in varying operational contexts. Finally, the knowledge demands of new levels of automation are strongly conditioned by a major constraint: if the automation is well engineered in a narrow sense, it will work well in a variety of routine situations; the problems of supervisory control will be manifest in situations with complicating factors that go beyond these routines. However, these will be relatively infrequent, at least for the individual practitioner. Thus, meeting the knowledge demands will require investing in maintaining usable knowledge relevant to the more difficult but infrequently occurring situations. In this, as in many other cases of introducing new levels of automation (e.g., Adler, 1986; Bereiter and Miller, 1988), new automation produces new training requirements.

A third direction for change is to develop new forms of aiding mode awareness itself through changes in the interface and displays that reveal what the automation is doing, why it is doing that, and what it will do next. The strategy is to provide better indications of what mode the system is in (to avoid mode errors), how future conditions may produce changes without direct practitioner intervention, and support better detection of and recovery from mode misassessments and mode errors when they do occur. Some attempts to do this have been made by changing the overall format of a display in different modes or changing the cursor shape as the system transitions between modes. However, since the practitioner's visual channels are often heavily loaded in some fields of practice, signaling mode changes through non-visual channels such as aural or kinesthetic feedback may be useful (cf., Monk, 1986 and Sellen et al., 1992 respectively). Another concept is "history" displays of instructions to and of the behavior of automated systems. Such displays would provide a visual trace of past and projected system behavior under the current mode configuration. However, such displays would have to be 'intelligent' in that the future behavior of the automated systems is contingent on future events in the environment.

While there are several suggestions that can be offered on potentially fruitful directions ranging from general strategies on what are effective ways to clearly signal mode status and changes (e.g., use orienting perceptual channels such as auditory or tactile signals; Monk, 1986; Sellen et al., 1992) to particular tips (e.g., display active targets with mode annunciations), it turns out that the human error, cognitive engineering and human-computer interaction communities have barely begun to study the relevant issues to provide the necessary research base to drive or support practical advice to designers. However, developing such aids probably requires that we

advance our understanding of how attention shifts across the perceptual field in dynamic multi-task domains (e.g., Eriksen and Murphy, 1987; Jonides and Yantis, 1988; Theeuwes, 1990). In this kind of field of activity, shifting the focus of attention does not refer to initial adoption of a focus from some neutral waiting state (Kahneman et al., 1973). Instead, one re-oriens attentional focus to a newly relevant event from a previous state where attention was focused on other data channels or on other cognitive activities (such as diagnostic search, response planning, communication to other agents). We need to understand how some practitioners develop a facility with reorienting attention rapidly to new potentially relevant stimuli (Woods, 1992). Thus, investigating how to aid mode awareness and how to provide cognitive tools to avoid or cope with mode related problems is a fruitful avenue for advancing our more basic understanding of more general issues like the cognitive processes such as control of attention, workload management, mental simulation – or more simply the panoply of cognitive processes that go under the generic label of situation awareness.

Another design aiding path that has been proposed to deal with mode related problems is forcing functions. Forcing functions are defined as "something that prevents the behavior from continuing until the problem has been corrected" (Lewis and Norman, 1986). Forcing functions can take a variety of forms: the system can prevent the user from expressing impossible intentions ("gag"), it can react to illegal actions by doing nothing, or it can explore with the user what the user's intention was and then help translate this intention into a legal action ("Self-correct", "Teach me", or "Let's Talk About It"). The problem with such forcing functions is that they require (a) that there is only one legal action/strategy for each intention or (b) that a system is capable of inferring the user's intention to compare it with his input in order to judge the acceptability of the input. Such a system would also have to have access to information on the overall context which can determine whether an action is appropriate. Without these capabilities, it would have to question almost any action – just in case, and run the hazard of over-interrupting at the wrong times.

The last direction is to consider supervisory control of automated resources as a kind of cooperative or distributed multi-agent architecture. One kind of cooperative agent concept would be to support mode awareness as a "management by consent" process which requires that all members of the team, human and machine, need to agree to any input to the system before it is activated. This approach could help a model or trace of all prior system interactions and lead to better prediction of future automated behavior. If automation and team work are supposed to reduce the burden on the practitioner by taking over and sharing tasks, then it seems counterproductive to require that all input is checked and agreed to by every member of the team.

Note that all of the above kinds of recommendation are human-centered in the sense that the costs of clumsy automation are seen in 'human error' and that the avenue for reducing perceived problems in the human element is to recognize that they are symptoms of the complexities produced by the clumsy use of technological possibilities (Woods et al., 1993).

IMPLICATIONS FOR THE CONCEPT AND THE STUDY OF SITUATION AWARENESS

Analyzing the Phenomenon of Situation or Mode Awareness

Our results suggests that the design and the capabilities of advanced automated systems make it more important and at the same time more difficult for their users to maintain awareness of the status and behavior across the different modes of operation of these systems. Despite the fact that vectors of technology change are increasingly challenging mode awareness, little research

has yet been done to better understand the relevant human-machine questions. But without this understanding, it will not be possible to develop effective countermeasures to mode related problems. The same deficit can be observed for the issue of situation awareness in general where a long tradition of research has not brought us much closer to being able to understand and support the phenomenon. What kind of research agenda is needed so that the research base can be expanded and practical countermeasures can be developed before technology change creates mode related problems of such magnitude that individual industries cry out for immediate answers?

First, extended efforts to develop the 'right' definition or a consensus definition of situation (and mode) awareness will probably not be constructive. Rather, the term situation awareness should be viewed as just a label for a variety of cognitive processing activities that are critical to dynamic event-driven and multi-task fields of practice. Control of attention (Gopher, 1991), mental simulation (Klein and Crandall, in press), directed attention (Woods, 1992), mental bookkeeping to track the multiple influences that act on a automated dynamic process and the multiple threads of sub-problems and resulting activity to manage them that go on in dynamic incidents are just a few of the cognitive processes that may pass under the label of situation awareness. Analyzing these cognitive processes and understanding what factors affect these processes should be the focus in the attempt to support situation and mode awareness (cf., Endsley, 1988, Tenney et al., 1992, and Sarter and Woods, 1991 for some initial steps). Second, it appears to us to be futile to try to determine the most important contents of situation awareness because the significance and meaning of any data is dependent upon the context in which they appear.

Measuring Situation or Mode Awareness

Conceptual or theoretical developments about the cognitive processes peculiarly associated with supervisory control of dynamic processes are critical if we are to develop effective measures of mode or situation awareness. Measurement techniques cannot be developed or used in a theoretical vacuum. There are three major categories of measurement, a) subjective ratings, b) explicit performance measures and c) implicit performance measures. The use of subjective measures (e. g., Situation Awareness Rating Technique, SART; Vidulich, 1992), where the operator is expected to rate his own level of awareness, is problematic on a variety of grounds, e.g., confusing process and product; because there is field data that misassessments color that person's whole standing and recall of the incident evolution (for example, video replay of the participant's behavior coupled with replay of the actual state of affairs is often necessary to get participants to recognize their own misassessments). Subjective measures only seem to make sense when combined with other measurement techniques, for example, in order to learn about how well calibrated were the participants in the evolving incident (extending the concept of how well calibrated is a decision maker to control of attentional focus).

An example of an explicit performance measure to assess situation awareness is Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1988). This method requires that subjects, typically pilots, fly a given mission on a flight simulator. At some random point(s) in time, the simulation is halted and the cockpit displays and outside view is blanked. The pilot is then asked a series of questions about the existing situation. The pilot's answers are later compared with what was actually going on in the scenario. The agreement between the two serves as a measure of the pilot's situation awareness. The most important problem associated with this technique is that halting the simulation and prompting the pilot for information concerning particular aspects of his situation is likely to disturb the very phenomena which the investigator wishes to observe. All research methods for the study of attentional and awareness

related processes suffer from this dilemma – the methods of observation disturb and change or eliminate the phenomenon under observation. For example, one of the important cognitive constituents of situation awareness is the ability to activate relevant knowledge during the actual process of handling an evolving incident. Prompting the participant for knowledge concerning particular aspects of a situation is itself a kind of retrieval cue and relevance marker that can change what the participant will call to mind. This will reveal what knowledge the pilot can activate when prompted with investigator cues as to relevance, but it will not shed light on what knowledge the participant would activate on his own or see as relevant in a particular situation.

The third approach, i.e., implicit performance measures, involves the design of experimental scenarios that include tasks and events that probe the subject's situation awareness (e.g., Sarter and Woods, 1993). In order for this technique to work, the probes have to be operationally significant in the sense that they should provide cues to the operator which if perceived by him should lead to an observable change in behavior. The shortcoming of this technique is that it assumes a direct relationship between situation awareness and performance. This problem can be addressed, in part, however, by means of debriefings in which the attempt is made to determine why a certain behavior did or did not occur. The major advantage of the approach is that it allows for a focused on-line collection of data while trying to minimize the disruptive impact of probes on the behavior of the subject any more than is inevitable in any simulated situation (for a more comprehensive critique of techniques for measuring situation awareness see, e.g., Tenney et al., 1992; Sarter and Woods, 1991).

CONCLUSIONS

As technology allows for proliferation of more automated modes of operation, human supervisory control faces new challenges. The technological trends create the need for mode awareness – human supervisory controllers tracking what their machine counterparts are doing, what they will do, and why they are doing it. But there is hysteresis in changing training for supervisory controllers and developing displays and interfaces to support collaboration to catch up with the cognitive demands imposed by clumsy use of technological possibilities. The result is evidence from field experiments, incident sampling and accidents that mode related problems in highly automated systems such as loss of mode awareness can contribute to new error forms and new paths towards disasters.

As a consequence, we need to be concerned with the question of how mode awareness can be supported successfully. In order to support mode awareness, as well as situation awareness in general, we need to better understand the set of cognitive processing activities that are involved in these phenomena. In other words, we need to take a process-oriented rather than a product-oriented approach to the analysis of the phenomena of mode and situation awareness. This approach will enable us to identify the reasons for breakdowns in mode and situation awareness, and it will help point the way towards how to train supervisory controllers and how to design cognitive tools that support the monitoring, assessment and awareness demands on supervisory controllers.

REFERENCES

- Adler, P. (1986). New technologies, new skills. *California Management Review*, 29, 9-28.
- Bereiter, S. and Miller, S.M. (1988). Sources of difficulty in troubleshooting automated manufacturing systems. In W. Karwowski, H. Parsaei and M.R. Wilhelm (Eds.), *Ergonomics of Hybrid Automated Systems*. Amsterdam: Elsevier.
- Billings, C. E. (1991). *Human-centered aircraft automation philosophy*. (NASA Technical Memorandum 103885). Moffett Field, CA: NASA-Ames Research Center.
- Cook, R. I., Potter, S.S., Woods, D.D. and McDonald, J.M. (1991). Evaluating the human engineering of microprocessor-controlled operating room devices. *Journal of Clinical Monitoring*, 7, 217-226.
- Endsley, M.R. (1988, May). *Situation awareness global assessment technique (SAGAT)*. Paper presented at the National Aerospace and Electronic Conference (NAECON), Dayton, OH.
- Eriksen, C.W. and Murphy, T.D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception and Psychophysics*, 42(3), 299-305.
- Gopher, D. (1991). The skill of attention control: Acquisition and execution of attention strategies. In D. Meyer and S. Kornblum (Eds.), *Attention and Performance XIV*, Hillsdale NJ: Erlbaum.
- Jonides, J. and Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, 43, 346-354.
- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs: Prentice-Hall.
- Klein, G. and Crandall, B. (1993). The role of mental simulation in problem solving and decision making. In J. Flach, P. Hancock, J. Caird and K. Vicente (Eds.) *The Ecology of Human-Machine Systems*. Hillsdale NJ: Erlbaum.
- Lenorovitz, J.M. (1990). Indian A320 crash probe data show crew improperly configured the aircraft. *Aviation Week & Space Technology*, 132(6/25/90), 84-85.
- Lewis, C. and Norman, D.A. (1986). Designing for error. In D.A. Norman and S.W. Draper (Eds.), *User-centered system design: New perspectives of human-computer interaction*, (pp. 411-432). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Moll van Charante, E., Cook, R.I., Woods, D.D., Yue, L., and Howie, M.B. (1992). Human-computer interaction in context: Physician interaction with automated intravenous controllers in the heart room. In IFAC/IFIP/IFOEA/IEA (Eds.), *Proceedings of the Fifth Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. The Hague, Netherlands.
- Monk, A. (1986). Mode errors: A user-centered analysis and some preventative measures using key-contingent sound. *International Journal of Man-Machine Studies*, 24, 313-327.
- Monnier, A. (1992). *Rapport preliminaire de la commission d'enquete administrative sur l'accident du Mont Sainte Odile du 20 Janvier 1992*.

- Norman, D.A. (1981). Categorization of action slips. *Psychological Review*, 88(1), 1-15.
- Norman, D.A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.
- Norman, D.A. (1990). The 'problem' of automation: Inappropriate feedback and interaction, not 'over-automation.' *Philosophical Transactions of the Royal Society of London*, B 327.
- Sarter, N.B. and Woods, D.D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
- Sarter, N.B. and Woods, D.D. (1992). Pilot interaction with cockpit automation: Operational Experiences with the Flight Management System. *International Journal of Aviation Psychology*, 2(4), 303-322.
- Sarter, N.B. and Woods, D.D. (1993). *Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the Flight Management System*. Manuscript submitted for publication (Also Cognitive Systems Engineering Laboratory Report CSEL 92-J-11, Columbus, OH: The Ohio State University).
- Sellen, A.J., Kurtenbach, G.P., and Buxton, W.A.S. (1992). The prevention of mode errors through sensory feedback. *Human-Computer Interaction*, 7, 141-164.
- Tenney, Y.J., Jager Adams, M., Pew, R.W., Huggins, A.W.F., and Rogers, W.H. (1992). *A principled approach to the measurement of situation awareness in commercial aviation*. (NASA Contractor Report No. NAS1-18788). Hampton, VA: NASA Langley Research Center.
- Theeuwes, J. (1990). *Perceptual selectivity is task dependent: Evidence from selective search*.
- Vidulich, M.A. (1992). Measuring situation awareness. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 40-41). Santa Monica, CA: Human Factors Society.
- Woods, D.D. (1992). *The Alarm Problem and Directed Attention in Dynamic Fault Management*. Cognitive Systems Engineering Laboratory Report CSEL 92-TR-06, Columbus, OH: The Ohio State University.
- Woods, D.D. (1993). The price of flexibility. In W. Gray, W. Hefley (eds.) *Proceedings of the International Workshop on Intelligent User Interfaces*, New York: ACM.
- Woods, D.D., Cook, R.I. and Sarter, N.B. (1993) *Clumsy Automation, Practitioner Adaptive Response and 'Human Error.'* Cognitive Systems Engineering Laboratory Report CSEL 92-TR-04, Columbus, OH: The Ohio State University.

ACKNOWLEDGMENT

The work described in this report was supported under a Cooperative Agreement (NCC 2-592) with the Aerospace Human Factors Research Division of the NASA Ames Research Center (Technical Monitor: Dr. Everett Palmer). We wish to thank Ev Palmer for his assistance and intellectual support.

The ideas expressed in the section entitled "Cognitive Systems in Context" come from a variety of studies of cognitive systems in the wild over a number of years, in a number of domains, and with a number of collaborators.

We would also like to thank all of the pilots who participated in the study and shared their experience with us. We are very grateful for the support and patience of a large number of people at the collaborating airline who made it possible to carry out this line of research and instructors who generously donated their time, energies and knowledge to make this research possible.

GLOSSARY

ADI	Attitude Director Indicator
AGL	Above Ground Level
ALT HOLD	Altitude Hold Mode
A/P	Autopilot
APPR	Approach Mode
ASRS	Aviation Safety Reporting System
A/P	Autopilot
A/T s	Autothrottles
ATC	Air Traffic Control
ATIS	Automatic Terminal Information Service
CBT	Computer-Based Training
CDU	Control Display Unit
CRZ	Cruise
CWS	Control Wheel Steering
DES	Descent
EFIS	Electronic Flight Information System
FD	Flight Director
FL	Flight Level
FMC	Flight Management Computer
FMS	Flight Management System
GA	Go-Around
G/S	Glideslope
HDG SEL	Heading Select Mode
HSI	Horizontal Situation Indicator
ILS	Instrument Landing System
LNAV	Lateral Navigation Mode
LAX	Los Angeles (Identifier)
LOC	Localizer
LOFT	Line-Oriented Flight Training
LOS	Line-Oriented Simulation
LVL CHG	Level Change Mode
MCP	Mode Control Panel
NOTAM	Notice for Airmen
PF	Pilot-Flying
PNF	Pilot-Not-Flying
RTA	Required Time of Arrival
THR HOLD	Throttle Hold
TO	Takeoff
TOD	Top-of-Descent
VNAV	Vertical Navigation Mode
V/S	Vertical Speed mode

Project Publications

Pilot Interaction with Cockpit Automation:

- [1] N. Sarter and D. D. Woods. Pilot Interaction with Cockpit Automation I: Operational Experiences with the Flight Management System. *International Journal of Aviation Psychology*, 2:303-321, 1992.
- [2] N. Sarter and D. D. Woods. Pilot Interaction with Cockpit Automation II: An Experimental Study of Pilot's Model and Awareness of the Flight Management System. Manuscript submitted for publication.
- [3] N. Sarter and D. D. Woods. "How in the world did I get into that mode?" Mode error and awareness in supervisory control. Manuscript submitted for publication.
- [4] N. Sarter and D. D. Woods. Situation Awareness: A critical but ill-defined phenomenon. *International Journal of Aviation Psychology*, 1(1):43-55, 1991.
- [5] N. Sarter and D. D. Woods. Mode error in supervisory control of automated systems. In *Proceedings of the Human Factors Society, 36th Annual Meeting*, 1992.
- [6] N. Sarter. The Flight Management System: Pilot's Interaction with Cockpit Automation. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1991.
- [7] N. Sarter and D. D. Woods. The Flight Management System: Rumors and Facts. *Sixth International Symposium on Aviation Psychology*, Columbus OH, April, 1991.
- [8] D. D. Woods. The effects of automation on the human's role: Experience from non-aviation industries. In S. Norman and H. Orlady, editors, *Flightdeck Automation: Promises and Realities*, NASA Ames Research Center, Moffett Field CA, 1989.

Methods for Studying and Evaluating Cognitive Systems in the Wild:

- [1] D. D. Woods and N. Sarter. Evaluating the Impact of New Technology on Human-Machine Cooperation. In J. Wise, V. D. Hopkin, and P. Stager, editors, *Verification and Validation of Complex and Integrated Human-Machine Systems*, Springer-Verlag, in press.
- [2] D. D. Woods. Process Tracing Methods for the Study of Cognition Outside of the Experimental Psychology Laboratory. In G. A. Klein, J. Orasanu and R. Calderwood, editors, *Decision Making in Action: Models and Methods*, Ablex, New Jersey, 1993.
- [3] G. A. Klein and D. D. Woods. Decision Making in Action. In G. A. Klein, J. Orasanu and R. Calderwood, editors, *Decision Making in Action: Models and Methods*, Ablex, New Jersey, 1993.
- [4] D. D. Woods and E. M. Roth. Symbolic AI-Based Computer Simulations as a Tool for Investigating the Dynamics of Joint Cognitive Systems. In P. C. Cacciabue, B. Pavard and J-M. Hoc, editors, *Simulation of Cognition in Dynamic Environments*, Academic Press, in press.

Human Error:

- [1] D. D. Woods, R. I. Cook and N. Sarter. Clumsy Automation, Practitioner's Adaptive Response and Human Error. Cognitive Systems Engineering Laboratory Report, The Ohio State University, Columbus OH, 1992.
- [2] D. D. Woods. Modeling and predicting human error. In J. Elkind, S. Card, J. Hochberg, and B. Huey, editors, *Human Performance Models for Computer-Aided Engineering*, Academic Press, New York, 1990, 248-274.
- [3] D. D. Woods and R. I. Cook. Nosocomial automation: Technology-induced complexity and human performance. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1991.
- [4] R. I. Cook, D. D. Woods, E. McColligan, and M. B. Howie. Cognitive consequences of 'clumsy' automation on high workload, high consequence human performance. In *SOAR 90, Space Operations, Applications and Research Symposium*, NASA Johnson Space Center, June 1990.
- [5] D. D. Woods. Risk and human performance: Measuring the potential for disaster. *Reliability Engineering and System Safety*, 29:387-405, 1990.
- [6] D. D. Woods. On taking human performance seriously in risk analysis. *Reliability Engineering and System Safety*, 29:375-381, 1990.

How to Make AI Systems Team Players:

- [1] D. D. Woods. The price of flexibility. In W. Gray, W. Hefley and D. Murray, editors, *Proceedings of the 1993 International Workshop on Intelligent User Interfaces*, New York, ACM Press, 1993.
- [2] D. D. Woods. Cognitive Activities and Aiding Strategies in Dynamic Fault Management. Cognitive Systems Engineering Laboratory Report, The Ohio State University, Columbus OH, October 1992.
- [3] D. D. Woods, S. S. Potter, L. Johannesen and M. Holloway. Human Interaction with Intelligent Systems: Trends, Problems, New Directions. Cognitive Systems Engineering Laboratory Report, February 1991.
- [4] D. D. Woods, L. Johannesen and S. S. Potter. Human Interaction with Intelligent Systems: An Overview and Bibliography. *SIGART Bulletin*, 2:39-50, 1991.
- [5] N. Sarter and D. D. Woods. Human Interaction with Intelligent Systems in Complex Dynamic Environments. In *Proceedings of FAA/NASA Advanced Workshop on "Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance"*, Daytona Beach, FL, June 17-19, 1992.
- [6] D. D. Woods. Human Interaction with Intelligent Systems in Space Applications. *Proceedings of International Symposium Cognitive Interactions: Man and Advanced Technologies*, SITEF 91, Aramihs, Toulouse France, October, 1991.

- [7] S. S. Potter and D. D. Woods. Event-driven timeline displays: Beyond message lists in human-intelligent system interaction. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1991.
- [8] L. Johannesen and D. D. Woods. Human-intelligent system cooperation in space applications: Trends, problems, new directions. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1991.
- [9] S. Potter and D. D. Woods. Guidance for human interface with artificial intelligence systems. In *SOAR 90, Space Operations, Applications and Research Symposium*, NASA Johnson Space Center, June 1990.

Representation Aiding: Enhancing Cognitive Systems through Computer Based Visualizations

- [1] D. D. Woods. Towards a Theoretical Base for Representation Design in the Computer Medium: Ecological Perception and Aiding Human Cognition. In J. Flach, P. Hancock, J. Caird, and K. Vicente, editors, *The Ecology of Human-Machine Systems*, in press.
- [2] D. D. Woods. The Cognitive Engineering of Problem Representations. In G. R. S. Weir and J. L. Alty, editors, *Human-Computer Interaction and Complex Systems*, Academic Press, London, in press.
- [3] D. D. Woods, S. S. Potter and L. Johannesen. Representation Design in Computer-Based Information Displays. *Insight*, Visual Performance Technical Group Newsletter, Human Factors Society, 14(March): 6-10, 1992.
- [4] S. S. Potter, D. D. Woods, T. Hill, R. Boyer and W. Morris. Visualization of dynamic processes: Function-based displays for human-intelligent system interaction. In *IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1992.
- [5] D. D. Woods, L. Johannesen and S. Potter. The sophistry of guidelines: Revisiting recipes for color use in human-computer interface design. In *Proceedings of the Human Factors Society*, 36th Annual Meeting, 1992.
- [6] D. D. Woods. Representation Aiding: A ten year retrospective. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 1991.
- [7] D. D. Woods, E. M. Roth, W. F. Stubler and R. J. Mumaw. Navigating through large display networks in dynamic control applications. In *Proceedings of the Human Factors Society*, 34th Annual Meeting, 1990.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE August 1993	3. REPORT TYPE AND DATES COVERED Contractor Report	
4. TITLE AND SUBTITLE Cognitive Engineering in Aerospace Application: Pilot Interaction with Cockpit Automation			5. FUNDING NUMBERS NCC2-592	
6. AUTHOR(S) Nadine B. Sarter and David D. Woods			8. PERFORMING ORGANIZATION REPORT NUMBER A-93108	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Cognitive Systems Engineering Laboratory Department of Industrial and Systems Engineering Ohio State University Columbus, OH 43210			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA CR-177617	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001			11. SUPPLEMENTARY NOTES Point of Contact: E. V. Palmer, Ames Research Center, MS 262A-4, Moffett Field, CA 94035-1000; (415) 604-6073	
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified — Unlimited Subject Category 03			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) <p>Because of recent incidents involving glass-cockpit aircraft, there is growing concern with cockpit automation and its potential effects on pilot performance. However, little is known about the nature and causes of problems that arise in pilot-automation interaction. In this paper, we report the results of two studies that provide converging, complementary data on pilots' difficulties with understanding and operating one of the core systems of cockpit automation, the Flight Management System (FMS). A survey asking pilots to describe specific incidents with the FMS and observations of pilots undergoing transition training to a glass cockpit aircraft served as vehicles to gather a corpus on the nature and variety of FMS-related problems. The results of both studies indicate that pilots become proficient in standard FMS operations through ground training and subsequent line experience. But even with considerable line experience, they still have difficulties tracking FMS status and behavior in certain flight contexts, and they show gaps in their understanding of the functional structure of the system. The results suggest that design-related factors such as opaque interfaces contribute to these difficulties which can affect pilots' situation awareness. The results of this research are relevant for both the design of cockpit automation and the development of training curricula specifically tailored to the needs of glass cockpits.</p>				
14. SUBJECT TERMS Human error, Automation, Flight safety			15. NUMBER OF PAGES 71	
			16. PRICE CODE A04	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

