Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit

I Introduction and Project Summary

Sharon Irving, Peter Polson, J.E. Irving

Institute of Cognitive Science
University of Colorado
Boulder, CO 80309

Technical Report 92-12

Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit

I Introduction and Project Summary

Sharon Irving
Peter G. Polson
J. E. Irving

University of Colorado
Institute of Cognitive Science
CB 344
Boulder, CO 80309

OVERVIEW

This paper describes a project funded by the Federal Aviation Administration to study training of the Control and Display Unit (CDU) in advanced commercial aircraft. Our initial impetus for choosing to focus on this device was the observation that pilots interacting with the CDU represented an instance of human-computer interaction: The CDU is the primary input device to the aircraft's Flight Management Computer. We draw on nearly a decade of research, starting with Card, Moran and Newell's description of the Model Human Processor and the GOMS model (Card, Moran & Newell, 1983), and leading to a rich body of work that followed (Kieras & Polson 1985, Olson & Olson, 1990; Bovair, Kieras, and Polson, 1990), giving clear outline to the emerging field of Human-computer interaction. This literature gave us the GOMS methodology and Cognitive Complexity Theory—the foundations of the training program developed under the project described herein.

Problems with the Current Flight Management Systems

We have seen in the past two or three years a spate of articles in the popular and research literature on pilot problems with the "glass cockpit". Such problems have been described as increased "mental workload", waypoint insertions errors, and other issues directly related to the operation of the Flight Management Computer (FMC).

One overriding theme manifests itself clearly in the work of these groups and individuals, and that is the concern over the use of automation to assist in or carry out tasks involved in flying an aircraft from Point A to Point B. Two aspects of the automation issue are cause for alarm among the FAA, other agencies, and the aviation research community: 1) the somewhat haphazard proliferation and implementation of automation in modern cockpits, and 2) the surprising lack of evidence that it has had the intended consequence of reducing pilot workload and error. Over the last decade, a renewed respect for the "person in the loop" as emerged and a new series of guidelines is taking shape to ensure that aircraft now on the drawing boards incorporate "human-centered automation" (Dept. of Transportation's National Plan, 1990; Billings, 1991).

It is difficult to evaluate the warnings and complaints about the "glass cockpit." Although any given incident, if studied carefully, can be shown to have multiple causes, as is characteristic of most serious aircraft accidents, there is a rapidly accumulating body of evidence in the form of pilot reports concerning errors in use of the FMC and other aspects of a highly automated aircraft (Hughes, 1989; Reingold, 1989; Wiener, 1988, 1989). Possible causes range from failures of vigilance to failures to remember the correct procedure to be carried out.

A recent summary of flight crew errors, incidents and accidents relating to the flight management computer (Eldredge, Mangold & Dodd, 1992) indicate that a large number are keyboarding errors (inputting data via the CDU), and errors of expectation or interpretation of FMC logic. Further, many of these FMC difficulties were attributed to inadequate training, according to the pilots reporting.

Irving and Irving (1990) have pointed out that increasing sophistication of the available underlying technology has resulted in a proliferation of separate automated warnings, displays and automation devices in a modern cockpit—all introduced incrementally. The Irvings have shown the sometimes untenable consequences of interacting systems which take little note of one another in the world of airline transport. Further, numerous other investigators note that there has been no detailed analysis of the cognitive demands placed on the skilled pilot by these new and sometimes incompatible systems. Significantly, there seems to have been no detailed analysis of either the tasks required or training procedures necessary to ensure acquisition of the critical skills by new pilots in these aircraft.

The FAA has instituted, as one response to the growing concern over pilot error and the glass cockpit as possible sources of airline transport mishaps, the Advanced Qualification Program (AQP). This somewhat ambitious program has as a broadly described goal the standardization and quantification of pilot training programs within the airlines. In order to carry out the various phases of AQP it will be necessary for the airlines to-among other things-conduct task analyses, institute instructor training curricula, and establish a database of pilots-in-training and end-level proficiency criteria. This primarily information-gathering effort is a herculean task, especially if it is carried out with little or no firm empirical and theoretical grounding.

Indeed the project described herein was inspired by a great many reports from both pilots and researchers that one particular device in advanced cockpits was cause for a considerable amount of grief and frustration. Pilots were finding that the very tasks it was designed to "automate" took more time to accomplish, and were subject to more errors and confusion. They frequently elected to "shut off" the automation and hand-fly the airplane, rather than deal with this inscrutable device.

These issues were the motivation to focus on the Control and Display Unit (CDU) in the project to be described.

I. The Tasks

Our work began with a detailed analysis of the cognitive skills involved in the use of the CDU. The analysis we used draws upon formal theories of human computer interaction (Card, Moran, & Newell, 1983; Kieras & Polson, 1985; Bovair, et al, 1990) which were originally developed in the context of office automation, e.g. word processors. The content and structure of our computer based training (CBT) for the CDU was derived from that task analysis. Both this task analysis and the training program are described in the second part in this series, GOMS Analysis (Irving, Polson & Irving, in preparation a). We proposed to explore the tentative analogy between the two contexts—office and "glass cockpit"—and hoped to show how the theoretical constructs and analytical methodology developed for the former are entirely appropriate to apply to the domain of piloting automated aircraft (Williges, Williges & Fainter, 1988).

The first goal of our research program, then, was to carry out a detailed, theoretically based analysis of the skills involved in using this interface. We proposed to use the GOMS methodology for this analysis, and to use, as a subject matter expert, a 737-300 Captain flying the line for a major airline.

Why GOMS?

A core set of theoretical constructs and analytical techniques has enabled researchers studying text editing and other office automation tasks to predict very large positive transfer effects. (Polson, 1988; Singley & Anderson, 1989). The GOMS model (Card et al., 1983; Kieras, 1989) is a formalism for representing in detail the knowledge necessary to perform routine cognitive skills. The acronym GOMS stands for Goals, Methods, Operators and Selection rules. A GOMS analysis takes as input Goals (what the user wants to accomplish), Selection Rules (the circumstances under which a given method should be used to accomplish a goal), along with a detailed description of the user interface, and produces as its output a detailed description of the cognitive operations and physical actions necessary to accomplish a goal.

Operations are elementary physical actions (e.g. pressing a function key), or cognitive operations not analyzed by the theory. Examples of the latter are perceptual operations, such as noting the representation of a recent route modification on the electronic map display; retrieving from memory that the ATC directive to intercept the so-and-so radial of such-and-such requires converting that radial to its reciprocal before entering the value into the CDU; or reading off a route element from the FPF and storing it in working memory.

A user's knowledge is organized into methods which are subroutines. Methods generate sequences of operations that accomplish specific goals or subgoals. A pilot might learn a method for carrying out the necessary steps on the CDU to install a hold, or intercept a leg to a waypoint.

Selection rules specify the appropriate conditions for executing a method to effectively accomplish a specific goal in a given context. Selection rules are compiled pieces of problem solving knowledge. The pilot knows that if she is given a directive to fly direct to the third downpath waypoint of the route of flight, the method of "Direct To on the route of flight" is the appropriate method to invoke.

An important point should be made about the GOMS model characterization of skills. First, routine cognitive skills are not simple skills. The GOMS model is intended to describe what Rasmussen (1983) has characterized as the skill-based and rule-based levels of knowledge. For example, the GOMS model is capable of representing situations in which pilots must choose among several alternative methods to accomplish a high level goal (e.g. carrying out an air traffic control directive), where the correct method depends on the details of the current situation. The methods themselves can have a complex internal structure, alternative courses of action dependent on context, and other considerations. As we stated earlier, selection rules are essentially compiled problem solving rules. The characterization of the complete set of Selection Rules for a skilled pilot would describe the details of how that pilot would go about using the FMC in the safe and efficient operation of the aircraft.

A thorough description of the GOMS task analysis will be found in Irving, Polson and Irving (in preparation a).

II. Training: The CBT and Cognitive Complexity Theory

A Cognitive Complexity Theory (Kieras & Polson, 1985; Bovair et al., 1990) model of a specific task using the CDU is a collection of very detailed rules. Each rule specifies a single step, a cognitive operation, or an overt action on the part of the pilot. CCT represents the knowledge described qualitatively in a GOMS model as a production system. Representing this knowledge as a production system makes possible accurate predictions about training times and about transfer of training (Polson, 1988).

Successful transfer is the critical issue in developing an effective training program. CCT models the transfer process using a common elements assumption (Thorndike & Woodward, 1901). The analysis of potential positive transfer begins with the development of production system models for both the training and transfer task. The analyst then identifies rules common to both models. The theory predicts that common rules learned during training will be transferred without any additional instruction and successfully executed in the transfer task. A major reason for development of the CBT under this proposal is the unique presentation to the learner of the rules as revealed by the GOMS analysis and the CCT model of the selected tasks. The other requirement for a specially developed CBT was the need to train these unique set of tasks completely out of context from all other cockpit tasks.

This analytical methodology enables us to identify good candidates for part task training. Thus, the required instrument readings not simulated in the part task trainer or interactions with other aspects of the flight automation obviously could not be successfully tutored. Such rules could not be acquired in the part task trainer, and so would not be expected to transfer to the full mission simulator. This is fairly obvious, but what is not so obvious is when such dependencies exist between successful use of the CDU and these other aspects of the flight automation. The very fine-grained analysis provided by Cognitive Complexity Theory enables us to identify such dependencies, and either simulate them successfully during training, or tutor these tasks when such interdependencies between various systems in the aircraft can be successfully simulated (such as in a fixed base or full mission simulator).

Our proposal for the initial study of 18 months limits itself to implementing a training program which specifically ignores all of the interactions found in the automation suite of the 737-300 airplane. To test the theory that a specific set of tasks associated with one device (the CDU) can be analyzed and trained out of context of the full mission environment, it was necessary to build a computer based training program dealing only with those tasks. Further, the tasks which are the focus of the analysis and training (lateral navigation—LNAV) are clearly an arbitrary subset of the full collection of activities carried out on the CDU. LNAV tasks are not the most troublesome for pilots—vertical navigation programming (VNAV) reveals evidence of greater conceptual difficulty. Yet, the focus on the specific set of tasks under this proposal was driven by the fact that proficiency in only the LNAV tasks must be demonstrated in the transition check ride.

The computer based training (developed with Authorware Professional™ on a Macintosh platform) is a single-path program with a severely constrained "simulation" of the functionality of the CDU. Experimental subjects (described below) are to be trained in the tasks, then demonstrate their proficiency in a transfer task to be carried out on a full mission simulator. To reduce the transfer distance from the Authorware CBT to the final transfer task, another, "second stage" simulation of the CDU was implemented in SuperCard™. This implementation provides more veridicality with the actual device than the Authorware CBT, allowing multi-path movement in executing tasks, and there is no tutorial text. Details of the SuperCard implementation can be found in McArthur and Irving, 1992 (in preparation, c).

Preliminary Results

Sixteen individuals have gone through the CBT program and SuperCard simulator. These individuals are all general aviation pilots, and thus have knowledge of terminology and procedures related to the air traffic control system in general, and flight planning in particular. None has had any exposure to an advanced automated aircraft—no "glass cockpit", and no flight management computer. There was very little variability in the performance of the subjects: all took about 3-1/2 hours to complete the Authorware CBT, and (usually a week later) an additional 45 minutes to an hour carrying out tasks—unaided—on the SuperCard simulator. Thus, they were trained in the basic operations of the device, then given airline dispatch paperwork (flight papers, weight manifest, etc.), told where they could find the parameters required for the pre-flight tasks, and asked to carry out the appropriate actions. Irving, Polson and Irving (in preparation c) provides complete details of the preliminary results of this project.

III: The Experiment

Under this proposal, for the actual experiment, airline pilots were used as subjects. A rather fortuitous situation exists for the selection of subjects which provides a measure of control over the most critical aspect of the training: the use of advanced automation in the cockpit. Boeing 737-300 airplanes differ from 737-200 aircraft almost exclusively in the implementation of "glass" and the flight management computer in the former. (In fact, the FAA issues one type rating for the two aircraft, and some airlines operate them as a common fleet. This research utilizes subjects from an airline which in fact operates the aircraft as separate fleets.) 737-200 pilots can be trained in the use of the CDU on the part-task devices used in this research project without ever having seen the inside of the more advanced aircraft, then transfer to the 737-300 full mission simulator for a transfer test.

The test is a real-time flight incorporating simulated air traffic directives which require subjects to carry out operations on the CDU, along with the other normal procedural duties of the pilot not flying (PNF) (checklists, etc.). Although the 737-300 includes the more advanced features of the "glass" and FMC, there are great similarities in the two cockpit environments. If another pilot handles pilot flying (PF) duties, experimental subjects will be performing very similar duties to those that are found in actual line conditions flying that pilot's own equipment. The scenario chosen for this transfer test involves some time compression, so tasks must be executed without undue delay.

What is a reasonable metric of success in this transfer test? Major airlines train transition pilots coming from non-"glass" aircraft to advanced equipment such as the 737-300 with its flight management computer. Most of these airline training programs use an integrated approach to training and sophisticated simulators in which to do this training. One airline uses a fixed base simulator with actual aircraft hardware and fully operational flight controls as part of its groundschool. Transition students learn CDU operation along with a myriad of procedural tasks and required systems knowledge. Pilots having completed such a transition course successfully (they passed the check ride), but who have had no line experience yet can be the metric against which the experimental subjects are compared. In this project, 737-300 transition pilots, also participating as PNF, get the same simulated flight with identical task requirements as the experimental subjects.

REFERENCES

Billings, C.E. (1991). *Human-centered aircraft automation: A concept and guidelines* (TM 103885). Moffett Field, NASA .

Bovair, S., Kieras, D., & Polson, P. (1990). The acquisition of and performance of text editing skill: a production system analysis. *Human-Computer Interaction, 5,* 1-48.

Card, S. K., Moran, T. P., & A., N. (1983). *The psychology of human-computer interaction* . Hillsdale: Erlbaum.

Dept. of Transportation, Federal Aviation Administration (1990). National Plan for Aviation Human Factors (Draft).

Eldredge, D., Mangold, S.J. & Dodd, R.S. (1992). A Review and Discussion of Flight Management System Incidents Reported to the Aviation Safety Reporting System. Final Report, October 1990-July 1991. Dept. of Transportation DOT/FAA/RD-92-2.

Hughes, D. (1989, August 7). Glass cockpit study reveals human factor problems. *Aviation Week & Space Technology*, pp. 32-33.

Irving, S., & Irving, J. (1990, April). *Workload and human-computer interaction issues in the automated cockpit/air transport environment.* Paper presented at The Workshop on Computer Human Interaction in Aerospace Systems, CHI '90, ACM, Seattle.

Irving, S., Polson, P.G., & Irving, J. (in preparation a). *Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit: Part II - GOMS Analysis* (in press).

Irving, S., Polson, P.G., & Irving, J. (in preparation c). *Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit: Part IV: Preliminary Results* (in press).

Irving, J.E., & Polson, P.G. (in preparation d). *Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit: PartV: A taxonomy of modes and interactions in advanced automated cockpits* (in press).

Kieras, D. E. (1989). Towards a practical GOMS model methodology for user interface design. In M. Helander (Ed.), *The Handbook of Human-Computer Interaction* (pp.135-156). Amsterdam: Elsevier.

Kieras, D. E., & Polson, P. G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies, 22,* 365-394.

McArthur, G. & Irving, S. (in preparation b). *Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit: Part III - Development of SuperCard CDU Simulator* (in press).

Olson, J. R., & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Human-Computer Interaction, 5,* 221-266.

Polson, P. G. (1988). The consequences of consistent and inconsistent interfaces. In R. Guindon (Ed.), *Cognitive Science and Its Application for Human-Computer Interaction.* NJ: Lawrence Erlbaum Associates.

Polson, P. G., & Kieras, D.E. (1985). A quantitative model of the learning and performance of text editing knowledge. *Proceedings of the CHI '85 Conference on Human Factors in Computer Systems.*

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, SMC-13,* 257-266.

Reingold, L. (1989, December). Behind those mistakes people make. Air Transport World.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge: Harvard University Press.

Thorndike, E. L., & Woodward, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8,* 247-261.

Wiener, E. (1988). Cockpit Automation. In E. L. Wiener, & D. C. Nagel (Ed.), *Human Factors in Aviation* (pp. 433-461). San Diego: Academic Press.

Wiener, E. L. (1989). Human factors of advanced technology ("glass cockpit") transport aircraft (No. 177528). NASA Ames Research Center.

Williges, R.C., Williges, B.H. & Fainter, R.G. (1988). In E. L. Wiener, & D. C. Nagel (Ed.), *Human Factors in Aviation* (pp. 433-461). San Diego: Academic Press.