Tenth International Symposium on Aviation Psychology
May 3rd–6th, 1999, Columbus, OH

"Why is it Doing That?" Two Perspectives of an Autoflight System

Michael Feary, Immanuel Barshi, Lance Sherry, and Martin Alkin

NASA Ames Research Center
Moffett Field, CA

ABSTRACT

Two preliminary studies explored differences in language and style between the design perspective and the operational perspective of autoflight systems. In the first study, responses of an expert pilot and an expert engineer to a set of incident reports involving an autoflight system were compared. In the second study, chapters describing an autoflight system in the system manufacturer's manual and a carrier's manual were compared. The qualitative differences found in the first study and the quantitative differences found in the second study suggest that some difficulties pilots experience when operating autoflight systems could be attributed to training materials which present the system in engineers' language and style rather than those of the pilots.

INTRODUCTION

Aircraft automation has been a hotly debated topic in the aviation community. Recent incident and accident reports cite problems with the use of automated systems as a causal factor in many events (see, e.g., FAA HFT Report, 1996). A number of studies have documented a range of issues concerned with the interactions between flight crews and automated flight systems (see, e.g., Parasuraman & Riley, 1997; Sarter & Woods, 1995). Due to a growing concern, recent studies have focused on the issues surrounding training for automation (see, Wiener, Chute, & Moses, 1999).

There are many possible reasons for the difficulties with pilot understanding of automated aircraft systems. It appears anecdotally, that one of these reasons may be a difference in perspective between the engineers who designed the system and the pilots who use it.

It is an engineering perspective which necessarily provides input to the user manual and the documentation associated with a system. Materials used for training on automated flight systems are prepared by the manufacturers of the system or the equipment. Such materials are written based on descriptions and specifications given by the design engineers. The engineering perspective also then influences the way training of the system is presented. While air carriers will translate the manufacturer's manual into more familiar terms/language for the use of their staff, it is possible that this translation is not extensive enough, and that the language of the manuals themselves add to users' (pilots) difficulties. Thus, one possible source of difficulties in learning such materials is differences in the language used by system designers and the language used by pilots. This study aims to determine whether or not a difference exists between the linguistic styles of pilots' and design engineers.

We have all experienced misunderstandings in communication between speakers of different languages. We expect such misunderstandings when the differences between the languages are obvious, as between the languages of different nations. Even within the same language, subtle terminology differences can lead to misunderstandings. In the aviation domain, which has a specialized vocabulary, these misunderstandings may have far reaching effects. This would imply that *if* there is a difference in the terminology, style or emphasis used by engineers and pilots, it could compound difficulties pilots have in understanding automation. To establish whether pilots and engineers do use a different aviation language we conducted two exploratory studies.

The first study examined the qualitative differences between the responses of an expert engineer and the responses of an expert pilot to a set of incident reports involving a particular aircraft type. The purpose of this study was to assess whether pilots and engineers do reach similar conclusions about an incident, but describe the same incident with different terminology, style and/or emphasis. The second study examined the quantitative differences between two descriptions of the autoflight system (the same system as the one examined in the first study) to assess the degree of translation that is currently present between manuals.

The first description was taken from the system manual produced by the manufacturer and formed the basis for materials used in the training of flight crews. The second description was taken from the manuals of an airline using the equipment in its operations.

STUDY 1

Study 1 examined the qualitative differences in approaches to aircraft automation between a pilot and an engineer through comparison of their descriptions of the same incidents.

METHOD

We selected 10 reports from NASA's Aviation Safety Reporting System (ASRS). These reports described problems with the autoflight system of a particular aircraft in such a way that the precise nature of the problem could not be ascertained. We gave these reports to two experts: a pilot, and an engineer. The pilot chosen has been a Flight Check Airman for many years, on the particular airplane discussed in the reports, and the engineer participated in the design of the autoflight system on that airplane. The two experts were asked to describe in their own words what happened in each of the incidents.

It was not our intention to see "who knows the system better". We compared the two sets of responses to identify systematic differences in language and style that could contribute to misunderstandings of the automation.

RESULTS

We found several qualitative differences between the responses of the expert pilot and those of the expert engineer. Even at a superficial glance it was easy to determine which response was written by which expert because the organizational style was very different. The expert pilot organized the response in a narrative fashion, constructing complete paragraphs, and coordinating the paragraphs into a coherent story. The expert engineer, on the other hand, used lists comprised of short statements (often preceeded by a bullet or a number), organized under headings and sub-headings (see Figure 1 for an example of two responses to the same incident report).

Another qualitative difference between the two sets of responses was the level of detail used. The example below demonstrates the engineer's use of numerical detail vis a vis the pilot's use of conceptual descriptors:

> "One possible explanation …… is that the aircraft was decelerating very rapidly during the descent (more than 0.05gs flight path acceleration)."

> "The 9000 foot altitude target was disregarded because the AFS was doing its version of a stall recovery."

The labeling of situations was another consistent difference. While the engineer focused on system states, the pilot referred to sequences of occurrences by labels such as "divide by zero" and "speed mode reversion", as in the two following examples:

> "A possible cause of this FMC malfunction is a known anomaly… sometimes the geometry can cause a *'divide by zero error'* and cause the FMCs to 'time out' or shut down unexpectedly."

> "If the aircraft was decelerating to target speed and … that caused the aircraft to be 5 knots under the FMS Vmin speed, the AFS will begin a programmed *speed mode reversion."*

An interesting difference between the responses of the pilot and those of the engineer was the pilot's ability to draw on broad experience with other aircraft types and the operational situation. The example above, in which the pilot refers to the aircraft performing "its version of a stall recovery" brings in the pilot's operational experience, as did other cases which included references to location-specific occurrences. The example below illustrates the pilot's perspective, gained through experience with other aircraft types and procedures at different companies:

> "There have been numerous reports of uncommanded climbs and other behaviors associated with sticking foreign objects such as checklists and flight plans in this area. One carrier has even warned pilots to not put things in this area for this very reason. Chances are no one moved the altitude selector knob. Many aircraft from the 'classic' fleet (such as the B-727) have a checklist storage here. The bracket for the GCP is not for that purpose."

STUDY 2

Study 2 examined the quantitative differences in approaches, displayed in training documentation for aircraft automation, between an airline and an avionics manufacturer.

METHOD

We selected the autoflight chapter from two manuals. One chapter was taken from the manual produced by the manufacturer of the autoflight system. The other chapter was taken from the manual produced by an air carrier using that system. The autoflight system is the same system that was the topic of the ASRS reports selected in Study 1.

Usually, the carrier's manual is a revised version of the manufacturer's manual. The revision path often takes the system manufacturer's manual to the aircraft manufacturer where the manual is revised once, and then on to the carrier where it is revised a second time. Manuals may go through several revisions during the lifetime of a given system or equipment. These revisions respond to changes in the company policy and procedures. The manual we selected had been through this process. That is, it is based on the system manufacturer's manual but was revised by the aircraft manufacturer and by the carrier.

Electronic versions of the two chapters were run through the grammar checker of a popular word processing program. This grammar checker produces the information given in Table 1 below.

RESULTS

Several measures displayed in Table 1 are of particular interest to our comparison. These include the total number of words, sentences, and paragraphs, the average number of sentences per paragraph, and the Flesch-Kincaid grade level readability index. Below, we discuss these measures.

The two chapters contain a similar number of words. These words are organized, however, in very different ways. As the total number of paragraphs in each chapter shows, the words used in the manufacturer's manual are organized into 401 paragraphs - less than half the total number of paragraphs in the carrier's manual (1004 paragraphs). Thus, on average, each paragraph in the manufacturer's manual is more than twice as long as those in the carrier's manual. Given that there is only an 11% difference between the total number of sentences in the two manuals (807 vs. 906), the manufacturer's manual must contain more sentences per paragraph than the carrier's manual. It does. The average number of sentences per paragraph in the manufacturer's manual is 2.9 compared with 1.6 in the carrier's manual.

Table 1 illustrates that, on average, the manufacturer's manual also has more words per sentence than the carrier's manual (X=19.7 vs. X=15.5 respectively). More words per sentence, more sentences per paragraph, and fewer paragraphs for the whole text mean a much more complex style which is harder to parse, comprehend, and retain. This conclusion is further illustrated by the readability index.

The Flesch-Kincaid Grade Level readability index indicates the estimated grade level for which a given text is suitable. However, this estimate is based on desired levels rather than actual students' ability. The average college student writes at a seventh grade level. In fact, most writing manuals, including those written for, and by, the U.S. military, recommend that documents for a general audience be written at a seventh grade level or lower.

As can be seen in Table 1, the readability index computed by the Flesch-Kincaid formula for the manufacturer's manual is 11.2, whereas it is 8.7 for the carrier's manual. The text in the manufacturer's manual is considerably more complex and demanding than the carrier's version of the same material. Regardless of the grade level of any given reader, and regardless of the inherited complexity of the material in question, a complex text is often more difficult to comprehend and retain than the same material expressed in a simpler text. The nature of automated systems is inherently complex and therefore the training material should be as easy to comprehend as possible. Again, it is important to remember that the grade level in question here refers to readers' actual reading ability rather than the number of years they spent in formal schooling.

DISCUSSION

The two exploratory studies presented here provide some insights into some of the differences in language and style used by pilots and engineers. If these findings are representative of the two communities, then some of the problems in understanding aircraft automation experienced by pilots could be explained by these language differences. Furthermore, these problems might be resolved if the language differences were to be reduced.

Although there are situations where it is not possible for pilots to know enough to predict precisely the behavior of the automated flight system (see, e.g., Degani & Heymann, 1999), there are many other situations where the system's behavior is predictable given the information available to the flight crew. Yet, pilots continue to make automation-related mistakes. For example, Feary and his collaborators (Feary, et al., 1998) showed that pilots’ prediction of autoflight system modes is error prone. While some of these errors were due to a lack of information available to the flight crews, some were not. One could argue that pilots should learn which information will be available to them during training. However, training methods aside, it is possible that the training materials present the information in ways that are not conducive to pilots' learning how the system operates.

From the limited data presented here, we can infer that the engineers who design the system and who provide the initial input to the manufacturer's manual, speak the language of engineering and bring the perspectives of engineers. This language is complex, expressed in sentences containing many words, and organized in paragraphs which contain many sentences. Moreover, concepts are organized in discrete items, in list-like arrangements, following a linear progression. The engineering perspective found in our data is detail-oriented, system-specific, and largely decontextualized from the broader flight situation.

The pilot perspective, on the other hand, brings in general flying experience beyond the specific system, and a conceptual framework which uses labeled situations as the basis for analogical reasoning. It is important to remember that pilots who receive training on autoflight systems and glass-cockpit aircraft often have extensive flying experience with non-glass aircraft; they naturally try to understand and operate the new equipment based on, or by analogy with, the old equipment with which they are more familiar (Baxter, 1998). Particularly under pressure, old habits and familiar patterns may be dominant even if erroneous. To prevent such situations it is therefore critical that training materials are presented in ways which are consistent with pilots' conceptual frameworks and perspectives. As discussed above, the engineering language in which a manual is written may not be compatible with the pilot's perspective and hence may be the source of some misunderstandings of automation operation.

To present the autoflight system in a manner that would be consistent with the pilot's perspective, a training manual could be written and organized in a more operational manner. Such organization would draw on the pilot's experience and would allow them to construct associative networks between systems they already know and the new system they will be operating. This kind of learning, based upon what has gone before, supports the development of expert performance (Ericsson & Charness, 1994).

Translation of descriptions of the system's operation occurs partially during revisions of the manual. It is interesting to note that in the process of adapting the system manufacturer's manual to their needs, a carrier simplifies the presentation of the text. It appears that Pilots tend to develop coherent narratives, rather than listing elements point-by-point, and draws extensively on examples. These two tendencies could be incorporated into training materials, along with efforts to further improve the general readability level of the text to make the content information more comprehensible to the pilot.

REFERENCES

Baxter, P. (1998). Old way/new way: Problems and strategies in behavioural training. Presented at the September meeting of the Human Factors Group of the Royal Aeronautical Society, London, UK.Degani, A., and Heymann, M. (1999). Pilot-Autopilot Interaction: A Formal Perspective. In R.S. Jensen (Ed.) *Proceedings of the Tenth International Aviation Psychology Symposium Conference.* Columbus, OH: The Ohio State University.

Ericsson, K. A., and Charness, N. (1994). Expert performance: Its Structure and Acquisition. *American Psychologist, 49*, 725-747

FAA (1996). Federal Aviation Administration Human Factors Team Report on: The Interfaces between Flight Crews and Modern Flight Deck Systems.

Feary, M., McCrobie, D., Alkin, M., Sherry, L., Polson, P., Palmer, E., and McQuinn, N. (1998). Aiding Vertical Guidance Understanding. Technical Memorandum No. NASA/TM-1998-112217. Moffett Field, CA: NASA Ames Research Center.

Parasurman, R., and Riley, V. (1997). Humans and automation: Use, Misuse, Disuse, and Abuse. *Human Factors, 39*, 230-253.

Sarter, N. B., and Woods, D. D. (1995). "Strong, Silent, and Out-of-the-loop": Properties of Advanced (Cockpit) Automation and their Impact on Human-Automation Interaction…CSEL Report 95-TR-01. Columbus, OH. The Ohio State University.

Sherry, L. (1995). A Formalism for the Specification of Operationally Embedded Reactive Systems. *Proceedings International Council of System Engineering.* St. Louis, MO.

Sherry, L., Kelley, J., McCrobie, D., Polson, P. G., Feary, M., Hynes, C.S. and Palmer, E. (1997). A Framework for the Design of Intentional Systems in Support of Cooperative Human-Machine Systems. In R.S. Jensen (Ed.) *Proceedings of the Ninth International Aviation Psychology Symposium*

Conference. Columbus, OH: The Ohio State University.

Wiener, E. L., Chute, R. D., and Moses, J. H. (1999). Transition to glass: Pilot Training for High-Technology Transport Aircraft. NASA contract report: 1999-208784. Moffett Field, CA: NASA Ames Research Center.

This incident seems to be nothing specific to the [aircraft type]. It deals with cockpit distractions, not making required callouts, and so forth. Perhaps there was some confusion with an altitude in the FMS versus and altitude sent in the GCP, but the distractions are the overriding factor here. To deal with the loss of the third set of eyes, automation has given us the capability to monitor the automatic flight. Essentially, we have two computer programmers monitoring the pilot, who is trained to follow behavior limited by very strict rules. The obvious method that the MD-11 was designed for in this situation was to program the crossing restriction in the FMS and allow PROF to make things happen. Since the copilot took PROF out of the loop, he changed the rules. Vertical speed will even override the GCP altitude if it is done during the level off. This mode requires constant attention of both pilots, and if there are cockpit distractions, it may not be an appropriate mode to use.	**Brief Description**: Aircraft descends to 10,000 ft when FMS Flightplan has 11,000 ft constraint **Summary:** - Operator failed to keep track of the level of automation engaged. Operator did not follow SOP **What happened:** - Descending with clearance to waypoint at 11,000 ft. Pilot entered 11, 000 ft constraint into the Flightplan - Aircraft descended to GCP alt = 10,000 ft. Pilots missed 1000 ft callout **Engineering explanation of behavior:** The behavior of the aircraft may be explained in a number of ways: 1. PROF was not engaged (pilot pulled alt knob after selecting the altitude). The avionics, correctly, controlled the aircraft to the GCP altitude at 10,000 ft. 2. PROF was engaged, but the aircraft was long (high) on the optimum pat and sequenced the waypoint with the 11,000 ft constraint above 11,000 ft. The aircraft continued the descent and leveled –off at the GCP Alt 3. Failure of avionics – no evidence to corroborate this supposition

Figure 1. Qualitative Results: Example comparison of Incident Analysis

Table 1. QuantitativeResults: Data produced by a grammar checker

	Airline Autoflight Chapter	Avionics Manufacturer's Autoflight Chapter
Counts		
Words	16647	17087
Characters	80194	85179
Paragraphs	1004	401
Sentences	906	807
Averages		
Sentences per paragraph	1.6	2.9
Words per sentence	15.5	19.7
Characters per word	4.6	4.8
Readability		
Passive voice	29%	34%
Flesch reading ease	59.7	49.3
Flesch-Kincaid grade level	8.7	11.2