

# **Flying Lessons: What can aviation investigations tell other disciplines about the human-computer interface?**

Blay Whitby  
School of Cognitive and computing Sciences  
blayw@cogs.susx.ac.uk  
**CSRP 533**

21/3/01

"Aviation in itself is not inherently dangerous. But to an even greater degree than the sea, it is terribly unforgiving of any carelessness, incapacity, or neglect."

Original author unknown, dates back to a World War II advisory.

## **Abstract**

Aviation has often been noted to be an ideal area for the examination of human cognitive performance. This paper contends that it also yields rich and reliable results for those interested in the study of human interaction with complex systems (often computer-based systems). A review is offered of how some of the major techniques of categorization and analysis developed in the aviation area, can inform HCI. In many ways, HCI can be seen as using investigative frameworks which have long been superseded in aviation.

## **Introduction**

Aviation contains such an embarrassment of riches for those interested in the human-computer interface (HCI), that it is important to begin by circumscribing the scope of the term. This paper will focus on a relatively limited part of aviation. It will, for example, not consider military aviation, though this is an area full of important results for HCI.

For the purposes of this paper, aviation will be taken to mean more or less the achievement of safe operation by public transport aircraft. The particular

focus will be on the general techniques used for analysing the interface between human beings and other complex systems. An attempt has been made to concentrate on aviation interface issues as most informative to HCI. Even with this limited definition, it is hard to review the vast amount of available material and so specific focus will be made on a series of very general 'lessons' which would be well drawn from aviation considered in this narrow sense. Many of these 'lessons' have been hard-won so ignoring them might well seem extremely parochial and ungracious.

It is important to make clear that this is not anything remotely like the claim that aviation has found all the answers to the problems of HCI. As is so often the case in science, answering one question tends to generate five further questions and the aviation field remains full of difficult HCI-related issues. In illustration of this, an example of a current aviation interface debate is given in the penultimate section of this paper.

The claim made is not the simplistic one that aviation studies have solved difficult HCI questions. It is rather that long-term changes in methodology which have been prompted by the need to improve aviation safety, have important lessons for other disciplines in general and HCI in particular.

## **Lesson 1 - Blaming the user is always easy and always worthless**

Most readers will have heard the expression 'pilot error'. It is a concept which is easy to grasp and easy to apply. Something with which readers may not be so familiar is the way in which this concept has been dropped as a worthless tool in the field of aviation.

The reasons for the abandonment of 'pilot error' are found in several long and complex historical processes. Firstly, there was the movement in aviation accident investigation away from attributing blame and towards preventing recurrence. This process took place in aviation during the 1950s and is thus long overdue in other areas. The abandonment of the 'blame model' is crucial to allowing the further steps in this process. This will be considered in more detail in Lesson 3.

The second reason for the abandonment of 'pilot error' is the aviation credo that *all* accidents (and incidents and failures) are due to a collection of human errors. If one considers the remainder of factors after removing pilot error, it should be clear that these must always involve human errors by managers, designers, maintenance personnel, controllers, governments, and others. One might, if one were still trapped in the 'blame model', introduce the concept of 'unforeseeable circumstances' at this point. However, once

one genuinely moves to the goal of preventing recurrence there can be no such thing as unforeseeable circumstances. The historical fact that they were not foreseen by the relevant people in this case is a human error which can now be rectified. To prevent recurrence one analyses the circumstances that obtained and makes the necessary interventions to prevent that particular set of circumstances (and maybe sets closely resembling it) from recurring. Of course, aviation had a strong advantage over many other areas of human activity in learning that blaming the user is unhelpful. Of course, pilots (like all human beings) are error prone. Pilots undoubtedly do make many errors. However, captured in grim aviation slogans such as 'They bury pilots with their mistakes' are at least two important truths. The first is that pilots are very strongly motivated not to make mistakes. The second is that they are highly unlikely to make a serious mistake more than once (in sharp contrast to spreadsheet users or doctors, for example). These two truths provided a motivation for looking beyond the dismissive 'pilot error'.

However, the advocates of 'pilot error' did not give up so easily. In commercial terms it is usually much cheaper to blame a pilot than to alter an aircraft design for example. Also there are many accidents where little else seems involved but an obvious error by the pilot. In these cases, relapse to the blame model is horribly tempting. Other people connected to the event are very strongly motivated to shout "it's not my fault" - this is a long documented problem with human beings. Even other pilots tend to say things like "Doing that was just plain stupid, I would never do that".

Overriding these human psychological defects is not easy. Commercial pressure comes in again because improving training, testing, information dissemination, and screening to prevent a recurrence may well be very expensive. Avoiding the inherent temptations of the 'blame model', requires fervent commitment by the advocates of the 'no recurrence model'. This second aspect of the first lesson - that the blame model cannot easily be dismissed - is perhaps more important than the rejection of the blame model itself.

On the other hand it should be clear, that unless one has a hidden agenda of accepting a certain level of accident or error, one must look to the factors behind a particular error. This lesson applies to the HCI area in spades. An interesting fact may help support this. A web search for the term "pilot error" turns it up in a computing, rather than aviation, context in about 30% of hits. The online dictionary of computing jargon defines "pilot error" thus: pilot error: n. [Sun: from aviation] A user's misconfiguration or misuse of a piece of software, producing apparently buglike results (compare UBD). "Joe Luser reported a bug in sendmail that causes it to generate bogus

headers." "That's not a bug, that's pilot error. His 'sendmail.cf' is hosed."

This is computing mimicking 1940s aviation attitudes. It's time to move on.

Of course, this sort of lesson is beginning to permeate other areas. It is not unknown for IT companies to claim that they have a 'no-blame culture'. Medicine in particular has deliberately sought to imitate aviation practice. What the aviation experience suggests is that this process will be extremely difficult and that much, much more than pious hopes will be required to move from a blame model. Determined conviction to eliminate the 'blame model' - perhaps taking decades to achieve - will be required. The aviation experience does, however, suggest that the benefits are well worth the effort.

## **Lesson 2 - Incidents, Accidents, and Failures don't have single causes.**

In aviation accident investigation it is an important canon that one should not speak of a single cause. Aviation investigators tend instead to use phrases such 'a pyramid of circumstances' or the 'links in the chain of events'. The thinking behind these phrases is that it is not the one or two events most recent events that produce an accident, but rather a steady accumulation of factors, many of which might seem innocent when considered alone.

In aviation, this view of the causation of accidents was developed hand in hand with the 'no-recurrence' approach of Lesson 1. That is to say that in order to prevent the recurrence of an accident it is far better to look at all (or as close as possible to all) the factors which contributed to that accident. It may well be that a particular 'link in the chain' is seen to be common to many superficially different accidents. In that case, the best way to prevent recurrence is to identify and that particular link. If one is satisfied with a glib single-cause explanation, this process can rarely get started.

Other disciplines may be able to take this lesson on board in a more direct fashion. However, in aviation it is very much part of a package with the other lessons described here. Consider, for example, the approach to pilot or user error described in the lesson 1. If an accident is described simply as pilot error, then the process of preventing recurrence stops at that point. Only by looking at the many contributory factors behind the error does a

'chain' become visible. This chain must be broken to prevent similar accidents occurring. Often the chain is long and has surprising links. Following such chains has led aviation psychologists to some very interesting conclusions (as well as some very obvious ones) about the ways in which humans are led to make erroneous judgements and mistaken decisions. It would seem, at first glance at least, that HCI could benefit from this sort of approach. Instead of looking for a single problem, analysis of chains could provide a more interesting pattern of results.

### **Lesson 3 - Categorization Matters**

The previous two lessons can be combined to yield a third concerning the importance of categorization. Following commitment to the no-recurrence model and the identification of long chains or circumstances leading up to a number of accidents or near-misses, it is natural to focus on what links if any these chains have in common. This becomes a major tool in identify where changes should be made in order to prevent recurrence.

One of the first reports commissioned into aviation accidents concluded the chief cause of accidents was the aircraft striking the ground. This is highly true but highly unproductive as a category. The quest to find a more useful categorization of accidents continued throughout the twentieth century and continues to this day. The benefits to safety, which flowed from the abandonment of the old 'pilot error' versus 'mechanical failure' categories, prompted a process of modification of all accident categories. Indeed, it is no exaggeration to say that categorization is the most important tool available in promoting aviation safety.

It is perhaps a little difficult to see how this could be. It is, in principle, an extension of the changes in approach briefly described in the first lesson. If someone, let us assume a pilot, has made an error and we wish to prevent a recurrence of that error, then we need to be able to classify all relevantly similar sets of circumstances and find ways of breaking the chain of events identified in this particular case.

A classic example was that of the 3-pointer altimeter. An example of this instrument is shown at Pic. 1a. Problems with this instrument were identified relatively early in the history of aviation. It is hard to read and errors of 1000 or 10,000 feet can easily occur. A graph of various time and error rates for various displays is shown at fig. 1. This is from Grether, (1949) and thus could not be considered a recent discovery. The problems with this type of altimeter display lead to the introduction of the servo altimeter which

included a digital readout. An example of this instrument is shown at Pic. 2. The differences in readability should be obvious. For the record the altimeter in Pic. 1 is showing a pressure altitude of 2790 feet and that in Pic. 2. 1320 feet.

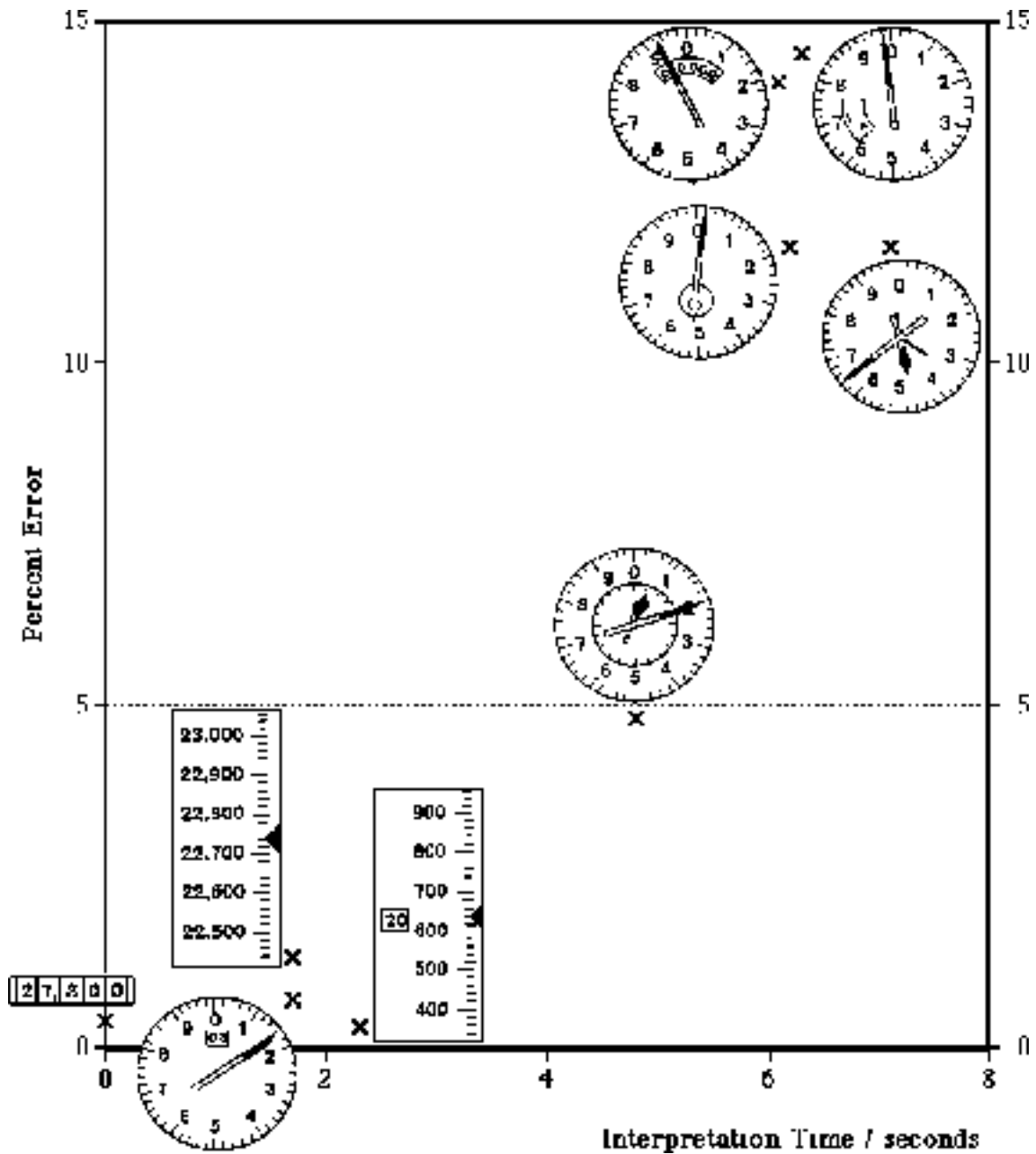
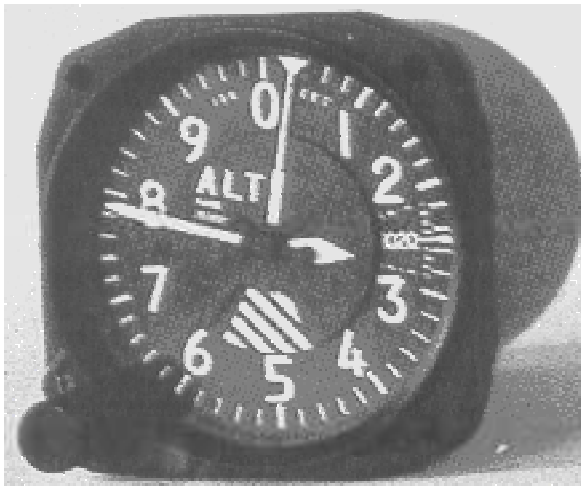


Fig 1. Interpretation times and error rates for various types of display.



Pic 1. The 3 -pointer Altimeter



Pic 2 The Servo Altimeter

Although the difference in readability was known, it would be hard to say what sort of effect this might have on flight safety were it not for the careful categorization of accidents in which misreading the altimeter might have been an important factor. If 'pilot error' had been used as a category for all these accidents then the problem with the altimeter readability would never have been recognized. A richer categorization such 'failure to monitor aircraft altitude' yields suggestions that are more interesting. (In practice even more detailed categorization would be attempted.) Without this, similar chains of circumstances would have continued to occur, whereas the

introduction of the servo altimeter would probably have broken the chain leading to the accident.

On most occasions, misreading of a three pointer would have caused no serious problem, of course. This shows the importance of gathering data from events that do not lead to accidents. Some of the most important ways in which this has been done in aviation are described in the next section.

A final ironic note to this lesson might be struck by returning to the highly unhelpful 'aircraft striking the ground' category mentioned at the start to this section. Probably the most recalcitrant category of aircraft accidents during the 1990s has been the CFIT (controlled flight into terrain). As the name implies this is the situation where an aircraft under full control and with no known problem flies straight into the ground. In spite of the servo altimeter and the second-generation of highly sophisticated GPWS (ground proximity warning system - which delivers a spoken warning or instruction to the crew) this remains a large category with little sign of decrease. The response of the accident investigators is therefore to subdivide. The largest sub-category of CFITs is that of ALAs (approach and landing accidents). The hope of the investigators is that this category may turn out to have more links in common than did CFITs. We have yet to see whether this will turn out to be the case, but it is a good example of the use of flexible categorization as a tool.

#### **Lesson 4 - It is better to collect wide-ranging data by constant monitoring, before analysing a particular problem.**

Adoption of the 'no-recurrence' model and the use of flexible categorization as an analytic tool naturally suggest a rich data approach. That is to say that more data - particularly on near-misses - will help to identify the similarities in the various chains of events which precede accidents. Collection of this rich data has proceeded and developed over the historical period with which this paper is concerned. The principle techniques employed have been anonymous reporting, monitoring and recording, and telemetry.

Some of the most effective tools developed in aviation involve the monitoring of 'near misses', 'reportable incidents', and the like. As with the other lessons described above there was powerful and continued resistance to many monitoring procedures. Thirty years ago, if a pilot or air traffic controller handled a potential disastrous situation without any injury or damage; they would feel resentful if the 'incident' were to be fully



investigated. They would probably feel proud that they had 'turned around' a difficult situation.

However, in the 'no recurrence model' it is just as important to examine near misses and the like as to examine accidents. No praise or blame is involved. This is a step in the process of categorization and of determining what sorts of combinations of factors can produce potential problems. Many other disciplines are in the habit of identifying potential problems by a process of guesswork. For example, software designers tend to think that they intuitively know what sorts of things users will find difficult. Maybe they are always correct in their intuitions. In aviation such intuitions have been replaced by real data.

Only when the 'blame model' has been thoroughly displaced is this sort of progress possible. Analysis of 'near misses' can then be seen by all concerned as a way towards the 'no-recurrence' goal rather than an attempt to blame anyone. It is worth repeating that this process does not happen easily. The attitudes of almost everyone involved with aviation had to change. This change took considerable time and effort.

One particularly useful innovation in aviation was the anonymous reporting of human factors in incidents and near misses. The UK system that provides for this is known as CHIRP (The Confidential Human factors Incident Reporting Programme). Professionally licensed pilots, cabin crew, air traffic controllers, licensed engineers, and approved maintenance organisations are encouraged to report any event or set of circumstances that involved human error and might have adverse consequences for flight safety. This has been in operation in the UK since 1982 and similar systems operate in many countries. Again, there was initial opposition. Some airlines, for example, thought this system would interfere with their disciplinary procedures. However, it has proved its worth many times over in practice. It is now administered by a charitable trust with extensive procedures that are designed to make the origins of reports untraceable. This system has produced a vast amount of useful data that combines well with the methods outlined above.

This sort of reporting is crucial in providing the data for identifying the points in the chain of circumstances that need to be modified as part of the 'no recurrence model'. Other methods of constant monitoring of operation are now being introduced in airline operation. Digital downlinking or telemetry of engine and airframe data during flight is becoming common practice. This trend is expected to continue into the human-factors area. One human-factors area where this is being investigated is in person-to-person communication.

The very nature of the HCI field makes it relatively easy to collect this sort of data. It is routine practice to monitor, for example, key strokes per hour in commercial organizations. As far as I know this is not generally used to provide data on potential HCI problems. A 'rich-data investigation' of interface performance might well yield different and potentially more interesting result than the existing techniques of small-scale observation. The aviation experience would certainly suggest that this would be a promising method of investigating, for example, usability problems.

### **Applying the Lessons: What, if any, are the differences between aircraft and software?**

Paradoxically, the best way to examine the differences between software and aircraft is to start from a point of similarity. One important point of similarity between modern public transport aircraft and large software systems is the degree to which highly automated aircraft are 'programmed' rather than flown. For a computer-literate audience it would be much more accurate to compare the operation of such aircraft to writing a macro for a spreadsheet or similar program. That is to say that a given route can be programmed and, after, obtaining the necessary clearances, the aircraft can be left to fly the entire route automatically, often including landing, touchdown and stopping on the runway.

The trend to complete automation has reached such a high degree in recent aircraft that some interesting human-factors conclusion may tentatively be drawn. The operation of such a highly automated aircraft perhaps more closely resembles the use of a computer program than it does the operation of conventional aircraft.

Such highly automated aircraft are bringing a new set of problems to aviation. It was initially thought the highly automated aircraft would be inherently both easier safer to operate than the previous generations of airliners. This has not turned out to be the case. The new generation of highly automated aircraft tend to require more crew training not less than the previous generation and are, as yet, not statistically safer. (This may be a temporary state of affairs resulting from the transition).

Applying the lessons outlined above, and seeking to use classification as an analytic tool, the flight safety community has identified a pattern of

incidents that have a certain amount of overlap with HCI. It is important to remember that flexible classification is a tool in aviation investigation and that these classifications are tentative.

The first is a class of accident that might loosely be called 'fascination with the technology'. In this class of accident, the crew seem to become overly preoccupied with a particular problem -often a relatively trivial technical problem - to such an extent that they temporarily 'forget' that they are flying an aircraft. In some cases, this preoccupation lasts long enough to cause problems or even accidents. Many explanations are on offer for this possible category of accidents including the general use of simulators for training and the way in which modern flight decks remove the actual hands-on task of flying leaving flight crews feeling less involved with the real-time aspects of their job.

The second is a phenomenon becoming known as 'mode confusion'. In this category of accident the flight crew of highly automated aircraft display confusion about which mode it operating in. In aviation jargon, they are 'behind the aircraft'. Typical cockpit voice recorder comments may be phrases like "What's it doing now?" This type of accident and incident would be probably still be attributed to inadequate training or similar problems were it not for an Airbus A330 crash at Toulouse in June 1994. This accident is worth examining in slightly more detailed terms because it is illustrative of the way in which the latest generation of airliners have generated truly novel problems and also because it particularly relevant to HCI. The aircraft concerned was brand new and under test at the Airbus base at Toulouse, before delivery to the customer. At the controls was Airbus' chief test pilot, Nick Warner. The test involved checking the aircraft's ability to handle an engine failure immediately after take off. This is a tricky manoeuvre for a human pilot but one which the FMS (Flight management System) of an Airbus A330 should be able to handle automatically. What should have happened is that the aircraft's 'pitch protection system' should have automatically kept the aircraft climbing safely on one engine.

What actually happened in this unfortunate accident was that the crew set the left engine to idle just after takeoff and waited as it lost speed. The pitch protection system should then have maintained a safe climb speed using the 'speed reference' function of the FMS. Unfortunately, a target altitude of 2000 feet had previously been entered into the FMS, which had switched from 'speed reference' to 'altitude acquire'. For special reasons 'pitch protection' does not function when the aircraft is operating 'altitude acquire mode'. It is not clear whether Captain Warner knew of this special feature of the software, but he realised too late that the aircraft's automatic systems were

not going to fly it out of this situation and was unable to recover in time by taking over manual control.

This accident has sent many ripples of concern through the aviation world. For many people, including several national aviation authorities, it marked the end of the belief that the latest generation of airliners are easier to fly than their predecessors. Some believe that the only safe option is to train crews to know and fully understand the details of the logic of the FMS computers. This is a considerable extra expense for the airlines and unpopular with the manufacturers.

Parallels with HCI should be obvious. One difficulty that may be affecting crews in 'mode confusion' accidents is the extremely arbitrary nature of software. Pilots may be able to use some sort of model of the physical characteristics of an aircraft to predict its behaviour. However, it is difficult, if not impossible, to acquire such predictive models of software. One either knows that pitch protection operates in this mode or one does not. Training crews in the underlying logic of automation is one method the aviation world has adopted to attempt to deal with this problem. It is, however, exactly the opposite direction to current trends in HCI and deserves detailed discussion.

## **A current interface debate**

One currently unresolved aviation debate concerns the degree to which a highly automated aircraft should resemble previous generations of aircraft at all. This is roughly equivalent to asking whether or not it is best to continue with the 'aircraft metaphor' when designing the interface.

Some illustrations may make this clearer. The first (pic.1) is a pilots view of the controls of a United Airlines Airbus 319-100. Some points to note in this photograph are the way in which the main features of the display are not many separate dials, but six computer-like screens. Even more relevant is the absence of any large control column, levers, or the like. These have been replaced by the small computer-style joystick visible at the sides of cockpit area. This aircraft is completely fly-by-wire which, in simple terms, means that it is flown by a bank of computers. Any input from the pilots is taken by the software as just another input to be handled.



Pic. 3

The computers can and do override or ignore the pilots if they consider their input to be inappropriate. The pilots cannot override the computers. An obvious sign of this change in the nature of flight control is the way in which the controls are now presented to the pilots. The small joystick presents the pilots' inputs to the aircraft.

In strong contrast is the flight deck of the Boeing 777-200, a photograph of which is shown below. (pic. 2). This aircraft is also completely fly-by-wire. As can be seen, the main displays are a bank of screens as in the Airbus A319. In contrast, however, to the small joysticks of the Airbus the Boeing 777 retains the appearance - what in computer circles would be called the 'look and feel' - of previous generations of aircraft. Substantial control columns are placed in front of each pilot. There was a time when such devices were necessary to physically control the aircraft. In the case of the Boeing 777, this is an illusion as they simply connect to small potentiometers at the base of the column. In practical terms, they are equivalent to the joysticks on the Airbus A319 flight deck. They simply convey the pilots' activities to the computer which actually flies the aircraft and may well ignore whatever the pilots are doing.

Other features to note in the Boeing cockpit are the throttles (the largest of the white levers on the centre console) which closely resemble real throttles



Pic 4

even though the engines are in fact controlled by the computers. On the Airbus the throttles are smaller and more like computer switches.

There is a great deal of debate in aviation circles as to the advantages and disadvantages of each of these approaches to the design of flight decks on automated aircraft. One of the issues which readers may find interesting include the way in which providing controls that resemble those in conventional aircraft may generate pilot behaviour which would be appropriate for conventional aircraft, but not for the present aircraft. One should remember that, for existing pilots at any rate, initial training will have been on conventional mechanically controlled aircraft. Thus one would expect them to acquire the sort of model of aircraft behaviour appropriate to conventional aircraft. It may be that such a model is inappropriate for fly-by-wire aircraft. A choice of what in HCI would be called an 'interface

metaphor' of conventional aircraft controls may be a factor in encouraging a misleading mental model of the aircraft.

Whether or not this actually happens, and whether or not it actually matters in practice remain open questions. It is perhaps overshadowed by the surprisingly high amount of training required to acclimatize experienced airline pilots to any of the highly automated aircraft of the present generation of airliners. If one accepts the assertions made in the previous section, then one might expect this to have some consequences. A fair summary of the debate would be to state that those who operate and fly each type of aircraft are great enthusiasts for the particular type of flight deck which they operate or fly.

Before coming to any hasty conclusion one should remember the lessons outlined above. There is no overwhelming evidence that either approach to flight deck design contributes to accidents or incidents. However collecting rich data on 'mode confusion' and 'fascination with the technology' incidents may well yield important clues on this. Again, classification is an important tool in dealing with this question. As data on 'mode confusion' is obtained from real incidents, from anonymous reports, and from laboratory experiments interesting conclusions may come to be drawn about which interface metaphor is better. These results would also seem to be significant for HCI in general.

## **Conclusions**

This paper began with a quotation stating that aviation is not inherently dangerous; this is perhaps the time to state that it is certainly not inherently safe either. It is a relatively new field and full of new challenges which require new techniques. Do the lessons outlined above work? Well they certainly worked in aviation. The fact that aviation has become, in three to four decades, such a safe way to travel is a tribute, I would argue, mainly to the way these lessons have been thoroughly learned and applied. Other fields, and I have mind medicine and law just as much as HCI, ignore such lessons at their peril.

## **Picture Credits**

Pic. 3 Justin Cederholm - Orlando/Tampa Aviation Photography

Pic. 4 Ulrich F.Hoppe

<http://ourworld-top.cs.com/ulrichhoppe/homepage.htm>

## **Bibliography**

Processes underlying human performance: using the interface, the bases of classic HF/E

Lisanne Bainbridge  
Department of Psychology, University College London

### **August 1995**

Faith N. (1996) Black Box: Why Air Safety is no Accident,

Grether, W.F. (1949). The design of long-scale indicators for speed and accuracy of quantitative reading. *Journal of applied Psychology*, 33, 363-372.

Hurst, R. & L. (ed) (1978) *Pilot Error the Human Factors*, Granada, London, 2<sup>nd</sup> edition.

Lyons, A. (1997) *The Commercial Pilot's Handbook*,

Weir A. (1999) *The Tombstone Imperative: The Truth about Air Safety*. Simon and Schuster UK, London.