Advanced Automated Glass Cockpit Certification: Being Wary of Human Factors

Rene Amalberti & Florence Wilbaux

Direction Générale de l'Aviation Civile

Summary

This paper presents some facets of the French experience with human factors in the process of certification of advanced automated cockpits. Three types of difficulties are described: first, the difficulties concerning the hotly debated concept of human error and its non-linear relationship to risk of accident; a typology of errors to be taken into account in the certification process is put forward to respond to this issue. Next, the difficulties connected to the basically gradual and evolving nature of pilot expertise on a given type of aircraft, which contrasts with the immediate and definitive style of certifying systems. The last difficulties to be considered are those related to the goals of certification itself on these new aircraft and the status of findings from human factor analyses (in particular, what should be done with disappointing results, how much can the changes induced by human factors investigation economically affect aircraft design, how many errors do we need to accumulate before we revise the system, what should be remedied when human factor problems are discovered at the certification stage: the machine? pilot training? the rules? or everything?).

The growth of advanced-automated glass cockpits has forced the international aeronautical community to pay more attention to human factors during the design phase, the certification phase and pilot training. The recent creation of a human factor desk at the DGAC-SFACT (Official French services) is a direct consequence of this.

The paper is divided into three parts. Part one debates human error and its relationship with system design and accident risk. Part two describes difficulties connected to the basically gradual and evolving nature of pilot expertise on a given type of aircraft, which contrasts with the immediate and definitive style of certifying systems. Part three focuses on concrete outcomes of human factors for certification purposes.

What Model for Human Error and What Links Between System Design and Accident-Risk?

The Goal of Aircraft Certification

The goal of aircraft certification is simple: guarantee that an aircraft fits the legal flight safety requirements when flown by qualified standard pilots who are as representative as possible of

Human Factors Certification of Advanced Aviation Technologies
Edited by J. A. Wise, V. D. Hopkin, and D. J. Garland
Copyright © 1994 Embry-Riddle Aeronautical University Press

end-users. The crews taking part in the certification campaign are used to bring systems into play and are at times as co-appraisers, but are never examined in their own right in the certification campaign. Thus, human errors observed during the certification campaign used to be classified in two categories. First, there are errors related to system-design: e.g., input errors, such as stick and throttle errors and inappropriate settings. These errors were taken into consideration in the certification process. Second, there are errors resulting from pilot attitudes, air-traffic control dialogues and clearances, or individual weaknesses related to general aviation know-how. These second type errors were not considered to be relevant for aircraft certification purposes.

Why Change the Current Procedures of Certification?

The present level of flight safety is very good. The risk of accident is about one per one million departures in industrial nations. However, this value has been virtually stable for fifteen years. It was attained before the growth of automated aircraft and before human factors became a target objective of FAA and the international aeronautical community.

In this context, why should the certification process bother with human factors?

The answer is twofold. First, advanced-automated aircraft have fewer dramatic failures, but the rate of accidents is not decreasing. Accident causes are increasingly due to the cognitive failure of crews. This introduces new problems of interference between system reliability and human reliability. In other words, the technical improvement of a system safety based on automation could result in a negative outcome for human reliability. Moreover, since these technical changes are barely related to the previous experience acquired on standard aircraft, certifiers, whether they are expert pilots or engineers, must ask human-factors specialists for scientifically-based assistance to evaluate systems better. Second, with a constant rate of accidents, any increase in traffic volume will result in an increase in the total number of accidents. Moreover, the negative impact of each (rare) accident on customers (passengers and also companies) is multiplied by modern information networks. Naturally this is undesirable, and the improvement of this situation is the stated objective of the aeronautical community. The only solution to remedy this situation is to continue to increase flight safety up to the present level. However, it is clear (given the current level) that gains will not be easy in previously examined domains and also that new domains will have to be taken into consideration. Human factors is one the most important of these new domains aside from the future of ATC communications.

Let us examine how these objectives impact the relationship between human error, system design and accident risk, and hence impact certification procedures. But of course the first problem to tackle is this of defining human error without ambiguity.

Wariness of the Definition of Human Error

The ergonomics literature has proposed numerous definitions and classifications of human error from process control. The dominant definition considers human error as a deviation from the norms, whether these norms are written or assumed from practice. This definition is both the easiest to use and the most debatable. Most recent cognitive ergonomics field-studies (Amalberti, 1992; deKeyser, 1986; Rasmussen, 1986) show that novices continuously interpret norms to make the job feasible with their limited resources and knowledge and that experts interpret norms still further with routines, shortcuts and violations. To sum up, norms are never respected, although they serve operators as references both to give and to limit degrees of freedom adapting to the on-going situation. Human operators are experts in piloting this derivative of norms to fit the situation requirements with minimum workload.

In contrast, another way to define error would be to consider error as a deviation from operator intention. This approach is probably less biased than considering error as a deviation to norms. However, operator intentions are hard to evidence, especially after a lengthy delay between action and analysis, and in the end human error still continues to be considered as a deviation from norms during the process of certification.

Experience acquired in participating as a neutral observer in the minimum crew campaign of advanced automated glass cockpits shows that ambiguities in human error definition lead the certification team, poorly trained in human factors, to make numerous mistakes in classifying and interpreting the crew errors. These mistakes are threefold. First, errors immediately corrected by the crew tend to be ignored. However, many of these errors point to incorrect system design, especially when they are repeated almost systematically. Second, deviations from norms are too often considered as errors although they are not. In many cases, they represent pilots' attempts to conserve resources or to manage the system and the task more conservatively. Third, some deviations where there are no specific procedural norms are ignored, although they potentially endanger the flight. This is typically the case of poor synergy and poor crew coordination which can result from the system design as well as procedure or input errors.

Modelling Relationships Between Human-Error and System Design

Ergonomics has always argued that systems should be designed primarily for end-users (human-centred design). This is a very basic and central value of ergonomics. Nevertheless, the concept of the human-centred system, and in a certain sense of a "good system," has significantly shifted over the last ten years with changes in technologies and ergonomics theories. Let us examine these changes.

A good ergonomics design has long been considered to be a design that prevents errors and facilitates good performance with as low a workload as possible. In the 1940's, the main interest was in unambiguous commands designed to minimized errors; e.g., confusion between gear and speed brakes (Fitts, 1947). This type of ergonomics, which is dominant in the USA, is termed *Classical Human Factors*. The basic philosophy draws on the central idea that human error is avoidable if the design respects human limitations and capacities, and this leads to the concept of "fault-preventing system design." It has been extensively and successfully applied to cockpit design and is currently used.

However, several factors have contributed to a recent decline of this approach: new technologies, new needs, and new ergonomics theories.

New Technologies and Classic Human Factors

Cockpit automation has enhanced flight performance in many domains such as precision approach, flight accuracy, engine performance, and pilots' situation awareness (with the introduction of map display). Automation has also mechanically reduced a great number of human errors resulting from improper power or stick settings and system handling simply because these tasks do not longer pilot-dependent.

However, the drawbacks of automation for human behaviour are as numerous as the advantages described above.

Cockpit automation and its consequences for cockpit layout have considerably reduced the benefits expected from a "simple human factors-based system-design."

The Flight Management System (FMS), with its undifferentiated and multiplexed keyboard, is a blatant example of this (Pelegrin, 1993; Sarter, 1992). This design is the source of many input errors in programming systems.

Criticisms have also been directed to information displays. The ability to display, aside from classical dials, much more information in various new forms (such as drawings and texts) has led designers to a series of poor ergonomics solutions regarding the capacities and limitations of human perception. Perception time tends to be increased with the use of textual information, perceptual feedback is reduced in peripheral vision (due to loss of motion of sticks and throttles and also due to cockpit architecture which requires the other crew member to move less), and auditory and kinesthetic sensations are also reduced (due to computer program smoothing system reaction to improve passenger comfort).

But of course the expected benefits of automation for ergonomics are elsewhere. Situation awareness has been improved with the use of map displays (MD), Primary Flight Display (PD) and ECAMs or EICAS.

To sum up, the new cockpit layout (glass cockpit) is assumed to enhance the pilot's situational awareness, but the solutions chosen to reach this goal are clearly to the detriment of classic sensory-motor human factors.

Evolution of Theories: Cognitive Ergonomics, Another Way to Consider Ergonomics

The change in goals for cockpit design have prompted new developements in ergonomics theories in the eighties. This is the domain of cognitive ergonomics, which draws heavily on the European ergonomics tradition. The value of such ergonomy is pilots' cognitive modelling focusing on their strategies and know-how. Numerous field studies have shown the advantages of this type of operator's cognitive model (Amalberti, 1992; Bainbridge, 1989; de Keyser, 1986; Hollnagel, 1993; Reason, 1990). The following section summarizes the main characteristics of one pilot's cognitive model.

Professional pilots generally have satisfactory procedural knowledge of their work domain and remarkable reasoning capacities, but they are resource-limited and cannot use all the knowledge and the reasoning capacities they would like to in time-related situations. The true task of pilots is to develop strategies to get the job done with respect to this resource-limitation bottle-neck.

Solutions Call for Planning, Anticipation and Risk Taking

Because their resources are limited, pilots need to strike a balance between several conflicting risks: an objective risk resulting from flight context (risk of accident) and a cognitive risk resulting from personal resource management (risk of overload and deterioration of mental performance).

To keep resource management feasible, the solution consists of decreasing outside risks, simplifying situations, only dealing with a few hypotheses, and schematizing reality. To keep outside risk within acceptable limits, the solution is to take as many preflight actions as possible in order to simplify the flight.

Any breakdown in this fragile and active equilibrium can result in unexpected situations, in which pilot performance may be decreased. Evidence shows that human errors result from internal characteristics of cognitive models and are not suppressible (Reason, 1990; Senders, 1991). Because resources-limitations force the pilot to make a series of compromises between what the situation should ideally require and what he is capable of doing, errors are the logical consequence. Moreover, expertise results from experiencing errors, (Anderson, 1985) and errors are generally profitable when the pilot receives immediate feedback from his errors.

New technologies confirm this general picture. As Wiener and Bainbridge point out (Bainbridge, 1989; Wiener, 1980), automation does not reduce the number of global errors, but merely changes error types. There are more routine errors and representation errors.

The consequences for ergonomics and certification purposes are twofold. First, the concept of "fault-tolerant system design" replaces the one of "fault-preventing system design." Fault-tolerant system design does not aim at limiting local errors but merely at improving pilot's awareness, giving as clear feedback as possible of error, and possibly correcting the immediate consequence of error when the flight is endangered; i.e., logical testing on FMS inputs or safety envelope of flight-laws (alpha-floor). Second, it is clear that in this theoretical framework it is no longer satisfactory to measure human performance as a simple error rate. More complex approaches are required to efficiently serve the certification process.

Should Error Analysis be Restricted to Human-Machine Interaction?

Standard certification procedures do not deal with crew errors which are not directly related to system design. However, new technologies could lead to a change in this position. The interdependency of any component of the aeronautical system, aircraft, crews, ATC, or maintenance, makes the analysis of causality between design and consequence of design much more complex than on a simple system. Any change in system philosophy influences the way operators carry out the task, even for actions not directly related to system interface. This is typically the case for crew coordination in glass cockpits.

Glass cockpit layout is generally assumed to make crew coordination more difficult. The reasons are threefold. First, as described above, communications require more and more central vision and active vocal dialogue to read the written information and to remedy the relative deprivation of sensory inputs. Second, new cockpit architectures, with independent access to information and commands, facilitate desynchronization. Pilots can display modes, change modes, or change parameters on their channel which the other crew member is totally unaware of. Third, problems of language emerge because more and more information is written. The standard language of aviation is English, although most pilots in the world have a different native language and do not speak perfect English (Pelegrin, 1993).

What emerges from these various difficulties is that many situations of poor coordination in a glass cockpit can be related to system design although they are not directly related to a specific action on the interface. This level of causality challenges the philosophy of the system and calls for complex corrections. It is easy to understand that designers are very reluctant to consider that these errors are related to system design and prefer to pass on the problem to trainers.

Relationship Between Human Error and Accident-Risk

Safety is a central concern of aircraft certification. System-failure classically serves to measure system reliability, and human error serves as an equivalent measure of human reliability. Measures could be quantitative or qualitative (type of error), but it is explicitly admitted that a good design and a safe system would provoke less errors than an unsatisfactory design, and would therefore result in fewer accidents.

This is only partially true. We have seen that human error is not totally avoidable. Moreover, it is important to remember that the accident-rate is 1 accident per million departures, and that there are over 5 human errors per flight which are not detected and corrected immediately by crews (Amalberti, unpublished report; this value comes from numerous flights made in 1992 on glass cockpits in the observer-position). Thus, even though human responsibility appears to have risen in glass cockpit accidents, the relationship between human error and accident is far from being trivial.

The key-point is that the relationship between system design, human-error and the risk of accident is not linear from great risks to no risk. Obvious bad system-design or/and unadapted regulations or training will cause numerous human-errors and will increase the risk of accident. However, even if the design, training and regulation are perfected, numerous human-errors and a non-decreasing risk of accident will remain. Without strong relations between the two arguments, remaining accidents are poorly linked to human-errors. They are better linked to a matter of circumstances, a dramatic combination of unexpected events in which human-error can occur but not seen as decisive factor. This picture specifically applies to rare accidents arising in the context of high reliability.

For certification concerns, this non-linearity between human-error and risk of accident can become a problem. The central concern is to have enough references on human-error theories to clearly separate what should be a "normal-rate and type of human error" from an "abnormal rate and type of human error" due to poor system-design, training or regulation. This is typically a domain in which human factors could improve the current process of certification.

The Evolving Nature of Pilot Expertise and the Immediate and Definitive Style of Certifying Systems

Although computer technology clearly enables software modifications during and after the end of certification, aircraft philosophy and most sub-system designs are considered to be stable and definitive at the beginning of the certification process. This is not the case for pilots' expertise. Most official pilots in charge of flying the aircraft during the certification process have less than 200 hours experience in the aircraft. This is far from having stabilized expertise in glass cockpit. Results from field experiments (Sarter & Woods, 1992; Pelegrin & Amalberti, 1993) all indicate that pilot expertise for flying an aircraft with a glass cockpit shifts significantly up to 800 flight hours and perhaps more.

These values are almost double the values observed for those required for expertise in a standard cockpit. The problem is that behaviours change with experience. Errors also change in nature.

The lengthy period required to stabilize expertise in glass cockpits creates difficulties for pilots becoming used to the system. Pilots cannot easily grasp the enormous possibilities of the system. With pilots of up to 400 flight hours, the main risks are overconfidence or excessive doubt concerning their own capacities which respectively result in engaging the system in unknown domains or hesitating to make the right choice of action.

Once pilots gain confidence with the system, routine errors and violations are multiplied. The risk of accident still exists but changes in nature. It is clear that some relationships between human error and system design will only emerge in experienced pilots.

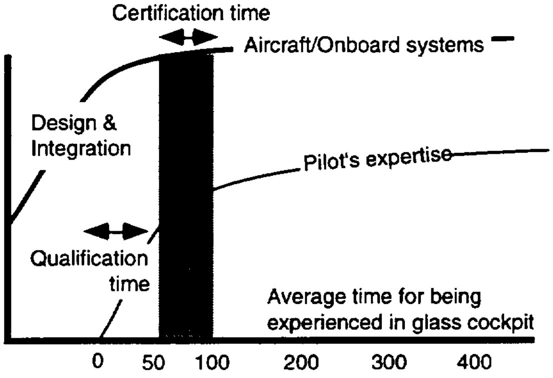

Figure 1

This picture raises a key-question on certification: do we test all aspects of system design and system safety with novices? If not, do we have to consider various levels of pilot expertise? A positive answer would result in envisaging a "double track certification," one initial test comparable to what is done currently and one operational complement administered after a few months of experience on the system.

Another solution would be to use official pilots who are already experts on similar machines (machine of the same family) to gain time and experience. In this case, novice official pilots in glass cockpits will also be required to represent this class of pilots and their specific problems.

Note that, in any case, the choice of official pilots who represent the future range of company pilots and the composition of crews is a key for efficient certification; e.g., take into account the pilots' level of experience, form a representative panel of pilots, avoid crews made of two captains, etc. This area could improve greatly in the future.

Practical Outcomes: How to Improve Certification With Human Factors Considerations

When to Bother With Human Factors in the Certification Process?

A good answer would be: "anytime the human factors perceptive gives a plus to the standard approach which is already being used." The objective of integrating human factors should not be to make a revolution in certification, but merely to support and improve the current way of doing certification.

The expected benefit is threefold. First, it must be in terms of the identification of undetected system weaknesses. Second, it is in terms of giving a rationale for pilots' problems and relationships to system design in the risk of accident. Last but not least, an aircraft can no longer be viewed as an isolated system. Certification typically concerns the integration of aircraft into operational conditions. Therefore, outcomes of certification must concern system-design as well as pilot training and regulations in order that sub-components of the macro-social system might be included in the evaluation. For this specific concern, psycho-sociology can effectively support certifiers to make the relevant decisions.

What to Certify?

The complexity and the novelty of systems lead one to consider that the integration of human factors in certification should overpass the simple evaluation of system-performance and the reliability of an end-product. It should extend to design-procedures, based on general principles which have proven to be efficient, and should begin with prototype evaluation at a period where changes are effectively possible.

What to Measure?

The starting point should be to analyse pilots' activities and detect human errors. However, the analysis of these human errors must vary according to the goal: assessing the risk of accident or testing pilot's ease with the system design. This is the reason why no unique classification of human error can figure out all questions raised by certification. Multiple classifications are required. Further analysis would concern the assessment and possible measurement of mental reasoning, mental workload, communications, crew coordination, to sum up all cognitive activities which serve pilots and set up a relevant representation of the on-going situation.

Who Makes the Evaluation?

We saw in previous sections of this paper that the selection of the panel of official pilots and engineers participating into the certification campaign is a key-factor in obtaining relevant results. One could suggest that this panel should be as large as possible to grasp a great variety of opinions, and also to avoid making certifiers co-designers due to a (too) long relation the certifier develops with designers. Whatever the panel, it seems useful to require a minimum human-factors background for people in charge of certifying.

What Limitations for Human Factors?

Many human factor aspects of cockpit automation are beyond the traditional certification process. We have seen that some of them could be easily better taken into consideration. Yet, numerous others which relate to psycho-sociology, work-organization, careers, trades, or companies will also be beyond human factors investigation during the certification campaign.

However, they should be crucial to system acceptability and risk acceptability, but this is another story.

Assuming that there is the integration of a human factors specialist in the certification process, another clear limitation should be the legal responsibility of this specialist in case of accident. Human factors cannot answer all the cockpit-problems, either because of the lack of knowledge or just because of the lack of time to apply relevant methodology. Therefore it shall probably be required to specify in writing what domains are relevant for human factors actions and what mix of responsibilities will draw on human factors specialists and on other certifiers.

What Should Change?

In most cases, the modifications in system-design required by the certification are small. The reason is obviously the financial cost. Therefore, when problems are observed, they tend to be solved by putting effort into pilot training and regulations.

Software technologies have changed this picture a little by introducing a greater level of flexibility, but it is clear that the underlying system philosophy remains unchanged.

Even though this is an acceptable outcome and even though experienced pilots rate the system as very good (this is the case of modern glass cockpit), human factors specialists worry about this increasing reliance on training solutions. What will occur for flight safety if we continue to produce opaque systems which require over 1,000 flight hours before pilots are experienced? Is this realistic?

Similarly, the presence of various generations of aircraft poses unsolved problems at this time: what will occur with pilots flying successively old and new aircraft? What will occur with multiqualified pilots flying almost identical aircraft with just a few differences, in particular as regards Cross Crew Qualifications (CCQ)?

In both cases, the questions overlap systems certification, introducing new types of questions to investigate and new constraints in forming the panel of pilots called for testing the system.

Systematic flight analysis of current modern aircraft is a fantastic tool to anticipate most troubles pilots will have with next technology. This is a central direction for improvements for all the aeronautical community, and it can be useful (for certification purposes) to ask the designer to take into account lessons from the previous design.

But again, any envisaged modification in a new machine will have to be investigated with a human factors perspective not really for itself, but for the possible negative consequences it will introduce when flying similar systems with bi-qualification.

Finally, one should remember that certifiers do not have to overpass the mandatory mission they are paid for (assessing that the system fits safety and minimum requirements). Once these minimum requirements are established and respected, designers will be free to create and certifiers will not have to officially judge a design in terms of being good or bad. In the context of a free market there is competition between manufacturers and the success or failure of a product remains a decision of customers.

Closing Notes

The international aeronautical community aims at introducing human factors more efficiently in the certification loop because of the desire to reduce the risk of accident and because of new technologies which have negatively impacted on human performance in a few domains. Analysis shows that this improvement cannot be made without a reconsideration of the concept of human error before (flight analysis), during, and after the certification phase (feedback and accident analysis).

One last and chronic source of misunderstanding in bridging knowledge between human factors specialists and engineers is that engineers superimpose human error and system failure upon one another. This makes no sense. Humans are intelligent and flexible. They can be perfectly adapted, whatever the complexity of situation and can ensure a very high level of safety. They learn from errors, cover billions of domains and can adapt to unknown domains. However, errors are always possible and always occur. These errors are poorly predictable and tend to occur at times, in areas and with people that nobody would have predicted. On the other hand, machines are rigid, unintelligent, and repetitive, and failures are predictable and curable. Because of their stability, machine reliability appears to be easily modeled and also more reliable than human reliability. The result is that engineers give a systematic priority to machines to the detriment of pilots because they feel this is the only way to improve and control safety.

All human factors findings show that they are wrong. Human and machine reliabilities are simply different and must work in synergy to reach a better level of reliability. Unfortunately, the solutions chosen at this time to increase system reliability interfere with human reliability and lower this human reliability.

Thus, it would not be realistic to discuss a very detailed point in the interface, although fundamental points are being ignored. A French maxim perfectly summarizes this point: "It is not good that the tree hides the forest."

The ideas expressed in this paper are aiming at launching debates. They are not firm directions decided upon SFACT, but only preliminary thoughts.

Future decisions of French official services will take into consideration, in addition to some ideas expressed in this paper and other technical ideas, all legal, international and sociotechnical aspects of the problems which have not been mentioned in the paper.

References

Amalberti, R. (in press). Safety in flight operations. In B. Wilpert, & Qvale (Eds.), *New technology, safety and systems reliability*. Hillsdale, NJ: L. Erlbaum.

Amalberti, R. (1992). Safety in risky process-control: an operator centred point of view. *Reliability Engineering & System Safety, 38*, 99-108.

Amalberti, R., & Deblon, F. (1992). Cognitive modelling of fighter aircraft's control process: a step towards intelligent onboard assistance system. *International Journal of Man-Machine Studies, 36*, 639-671

Anderson, J. (1985). Development of expertise. In Freeman (Ed.), *Cognitive psychology and its implications* (pp. 235-259). New York.

Bainbridge, L. (1989). Development of skill, reduction of workload. In Bainbridge & Quintinilla (Eds.), *Developing skills with new technology.* London: Taylor & Francis.

Bainbridge, L. (1987). Ironies of automation. In Rasmussen, Duncan, & Leplat (Eds.), *New Technology and Human Errors* (pp. 271-278). New York: Wiley.

de Keyser, V. (1986). Technical assistance to the operator in case of accident: some lines of thought. In Hollnagell, Mancini, & Woods (Eds.), *NATO Series* (pp. 229-254).

Fitts, P., & Jones, R. (1947). *Analysis of factors contributing to 460 "pilot error" experiences in operating aircraft controls* (Report TSE AA-694-12). Wright-Patterson Air Force Base, MA.

Hollnagel, E. (1993). *Reliability of cognition: Foundations of human reliability analysis.* Amsterdam: Elsevier.

Pelegrin, C., & Amalberti, R. (1993). *Pilot's strategies of crew coordination in advanced glass-cockpits; a matter of expertise and culture.* Second Flight Safety International Congress, Washington, DC.

Rasmussen, J. (1986). *Information processing and human-machine interaction.* Amsterdam: North Holland.

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

Sarter, N., & Woods, D. (1992). Pilot interaction with cockpit automation: operational experiences with the flight management system. *International Journal of Aviation Psychology*, *2*(4), 303-321.

Senders, J., & Moray, N. (1991). *Human error.* New York: L. Erlbaum.

Wiener, E., & Curry P. (1980). Flight desk automation: promises and problems. *Ergonomics*, *23*, 988-1011.

Wiener, E. (1985). Beyond the sterile cockpit. *Human factors*, *27*(1), 75-80.