The accident of flight 447 Rio-Paris: a case study for HCI research

Stéphane Conversy **Stéphane Chatty** **Hélène Gaspard-Boulinc** **Jean-Luc Vinot**

Université de Toulouse - ENAC
7 av. Edouard Belin, 31055 Toulouse, France
prenom.nom@enac.fr

ABSTRACT

On June 1st, 2009 flight AF447 from Rio to Paris crashed into the Atlantic Ocean. The safety and legal investigations concluded that human factors have played an important role in the accident. Observing that a number of elements from the report written by the French Office of Investigations for Civil Aviation Safety may be assimilated to known concepts from HCI, we propose to use the report as a case study for HCI research. After introducing the aeronautical vocabulary required to its understanding, we extract the HCI-related elements from the report, and assimilate, organize and translate them into conceptual frameworks from the Model of Action and Epistemology. We hope to foster further research aiming at a more formal modeling of the accident, or to foster the identification of possible improvements of the onboard systems.

Keywords: accident, aeronautics, HCI theories

ACM Classification Keywords: H.5.2 Information Interfaces and presentation: User Interfaces

INTRODUCTION

The F-GZCP Airbus A330 aircraft crashed into the Atlantic Ocean on June 1st, 2009 during flight Air France AF447 from Rio to Paris, with 228 casualties. The French Office of Investigations for Civil Aviation Safety (BEA, Bureau d'Enquêtes et d'Analyses) belonging to the Ministry of Transportation has published a report after a safety investigation on the circumstances of the accident [2]. As stated by the BEA, "its investigations are conducted with the sole objective of improving aviation safety and are not intended to apportion blame or liability". Nevertheless, the report aims at analyzing the causes and the chain of consequences that eventually led to the accident.

As is often the case, potential causes might be numerous. Aviation is a complex socio-technical system composed of multiple actors (national and internation agencies, manufacturers, airlines, training organism, pilots etc.). Therefore the investigation report includes sections on the course of the flight, on pilots and their behavior, on hardware and on weather conditions. The investigation notably relies on the recordings made on-board and retrieved from flight recorders.

The report seems to exclude any failure from embedded system, with the exception of the airspeed sensors called "Pitot probes". The failure of these probes is not a catastrophic event: the embedded systems have been designed to cope with such failure, and they actually behaved as expected by their designers. Thus, rather than a system failure, it is a combination of the behavior of these systems, the flight conditions and the reaction of the pilots that had led from the probe failure to a fatal issue. This makes this accident a relevant case study for research in HCI, especially because the report provide detailed elements.

This article aims at translating the investigation of the BEA into the concepts of the HCI community. The goal is to support the training of system designers, the assessment of how theoretical models faithfully account for the described phenomena and the research on pilot training.

AIRPLANE CONTROL

Before introducing the report, this section presents a summary on the control of the A330. Modern aircraft such as the A330 are controlled by human and automatic subsystems that interact together[1]. The mission of this hybrid system is to prevent the airplane from flying outside its flight envelope.

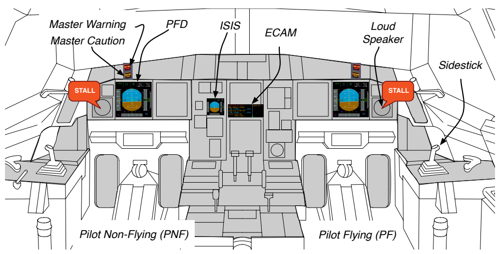

Figure 1. The cockpit and its components [1]

Steering

Two pilots are in commands (fig. 1), the flying pilot (PF, seated right in the cockpit) and the non-flying pilot (PNF). Each one has a number of input devices, notably a sidestick (isometric joystick on the side) that comes back to a centered position when let loose. The output devices consist of a number of screens, lights and loudspeakers. Among them are:

- the primary flight display (PFD) which displays the speed, the artificial horizon, level, heading (fig. 2); and the ISIS, a back-up system that display the same information;
- a monitoring display (ECAM), which displays notifications from subsystems (fig. 3) and lights Master Warning (red) and Master Caution (amber) that signal the respective level notifications of the ECAM;

- the stall warning, that signals visually (blinking Master Warning) and sonically (“cricket sound”, a loop sequence of four buzzes each lasting a few tenths of a second) and the vocal message “STALL”that the airplane is outside its flight envelope.

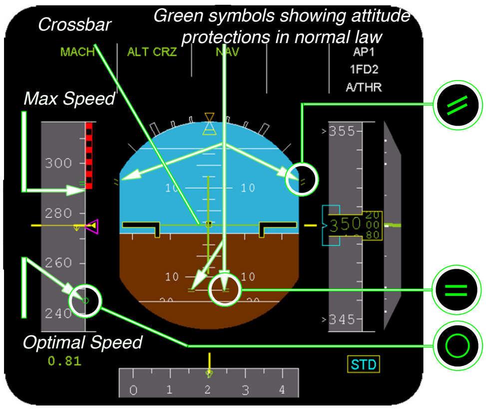

Figure 2. PFD in Normal Law

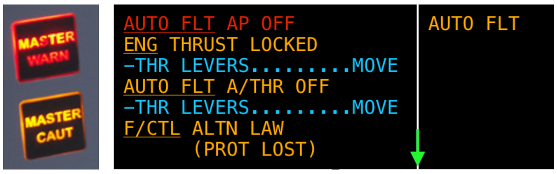

Figure 3. Master Warning & Caution (l.) and ECAM (r.)

Assistance systems

The assistance systems for steering involved in the accident are:

- flight control computers, which interpret pilots’ actions on the sidestick to move the surfaces of the plane. The aim of this subsystem if to make the flight more cost effective, safer, and more pleasant for passengers [1];
- the automatic pilot and thrust (resp. AP et A/THR), which aim at offloading from human pilots the tasks of reaching and maintaining the instructions input by the pilots (simple instructions such as heading or more complex ones such as an approach trajectory)
- the flight director (FD), which gives indications to pilots on actions to perform (nose up, nose down, to l., to r.) with a crossbar on the PFD.

The assistance systems have multiple modes of operation. Those of the flight control computers are called control laws. These laws define the use of automatic control, the transfer function between input devices and the physical systems of the plane, and the use of protections against instructions that would make the fly exit the flight envelope. The initiative to switch from a law to another one are from the pilots or from the automatic subsystems. Such automatic switches are triggered by outside events (e.g. speed lost). In the case of flight 447, the most interesting laws are the “normal”, “alternate” and “alternate 2” laws.

The outside parameters acquisition is performed by physical probes: Pitot probes measure the air pressure which is turned into the measure of speed; gyroscopes measure attitude pitch and roll; specialized probes measure the angle of attack.

OVERALL BEA CONCLUSIONS

In order to give to the reader an overview of the accident, we reproduce below the synopsis, the findings and the causes of the accident according to BEA. We selected the findings and causes that are linked to human-computer interaction and human factors.

Synopsis of the accident

At around 2 h 08, the crew made a course change of 12 degrees to the left, probably to avoid returns detected by the weather radar. At 2 h 10 min 05, likely following the obstruction of the Pitot probes by ice crystals, the speed indications were incorrect and some automatic systems disconnected. The aeroplane’s flight path was not controlled by the two copilots. They were rejoined 1 minute 30 later by the Captain, while the aeroplane was in a stall situation that lasted until the impact with the sea at 2 h 14 min 28. (p17)

Findings by BEA

The aim of the analysis was to determine the sub-group of the provisions that affected the expected behaviours and skills of the crews for the situation encountered. [...] Beyond the simple discovery of a psychologically probable, likely or plausible explanation of the behaviours recorded, this involved assessing the degree of specificity or generality of the behavioural responses recorded: are they specific to this particular crew, shared by all the airline’s crews, or can they be generalised to all crews? (p101)

Findings of the investigation:

- The aeroplane systems detected an inconsistency in the measured airspeeds. The flight control law was reconfigured to alternate 2B.
- No failure message on the ECAM clearly indicates the detection by the system of an inconsistency in measured airspeeds.
- The pilots detected an anomaly through the autopilot disconnection warning that surprised them.
- Although having identified and called out the loss of the airspeed indications, neither of the two copilots called the “Unreliable IAS” procedure.
- The Flight Directors did not disconnect. The speed displayed on the left PFD was incorrect for 29 seconds, that of the speed on the ISIS for 54 seconds and the speed displayed on the right PFD for 61 seconds at most. In less than one minute after autopilot disconnection, the aeroplane exited its flight envelope following inappropriate pilot inputs.
- The crossbars disappeared and then re-appeared on several occasions, changing mode several times.
- The approach to stall was characterised by the triggering of the warning then the appearance of buffet.
- In the absence of a display of the limit speeds on the speed tape on the PFD, the aural stall warning is not confirmed by any specific visual display.
- The stall warning sounded continuously for 54 seconds.
- Neither of the pilots made any reference to the stall warning or to buffet.

- The angle of attack is the parameter that allows the stall warning to be triggered; if the angle of attack values become invalid, the warning stops.
- By design, when the measured speed values are lower than 60 kt, the measured angle of attack values are invalidated.
- The aeroplane's angle of attack is not directly displayed to the pilots. (p197)

Causes (excerpts)

- The crew, progressively becoming de-structured, likely never understood that it was faced with a "simple" loss of three sources of airspeed information.
- In its current form, recognizing the stall warning, even associated with buffet, supposes that the crew accords a minimum level of "legitimacy" to it. [...] When crew action is expected, it is always supposed that they will be capable of initial control of the flight path and of a rapid diagnosis that will allow them to identify the correct entry in the dictionary of procedures. A crew can be faced with an unexpected situation leading to a momentary but profound loss of comprehension. [...] During this event, the initial inability to master the flight path also made it impossible to understand the situation and to access the planned solution. (p199)

Thus, the accident resulted from the following succession of events:

- Temporary inconsistency between the airspeed measurements, likely following the obstruction of the Pitot probes by ice crystals that, in particular, caused the autopilot disconnection and the reconfiguration to alternate law;
- Inappropriate control inputs that destabilized the flight path;
- The lack of any link by the crew between the loss of indicated speeds called out and the appropriate procedure;
- The late identification by the PNF of the deviation from the flight path and the insufficient correction applied by the PF;
- The crew not identifying the approach to stall, their lack of immediate response and the exit from the flight envelope;
- The crew's failure to diagnose the stall situation and consequently a lack of inputs that would have made it possible to recover from it.

These events can be explained by a combination of the following factors:

- The feedback mechanisms on the part of all those involved that made it impossible [...] to identify the repeated non-application of the loss of airspeed information procedure and to remedy this,
- Task-sharing that was weakened by incomprehension of the situation when the autopilot disconnection occurred,
- Incomprehension of the situation when the autopilot disconnection occurred,
- The lack of a clear display in the cockpit of the airspeed inconsistencies identified by the computers; (p200)

METHODOLOGY

The report provides a detailed description on the behavior of the cockpit human-machine interface (HMI), on the reasonings that the pilots may have performed using the information given by the embedded systems, and on the interactions between human and automatic subsystems. This analysis relies on verbal exchanges between pilots, the recording of the actions they performed and the responses of the subsystems. It also uses the results of post-accident simulations conducted to verify the behavior of the visual, auditory and haptic subsystems.

In order to translate the analysis into HCI concepts, we used the following methodology. We first extracted a number of excerpts from the BEA report that we evaluated as relevant for HCI. We then abstracted them into phenomena, from which we selected five (P1 to P5) that seem to both play a significant role in the analysis and constitute examples of application of available HCI models. We show how each of them can be linked to the analysis framework from HCI, HF and Ergonomics, in particular to the model of action from Norman [16], or linked to epistemology, in particular the concept of abduction [12].

In this work, we select the facts and analyses performed by BEA that are suitable to modelling with the corpus of theories from Human-Computer Interaction. This choice necessitates to take special care in reading this paper. First, the reader is invited to refer to the BEA report in case of doubt. Second, our selection may mislead the reader about the causes of the accident. The reader is invited to keep in mind that important facts are absent from the paper, because we could not translate them into HCI models. In particular, we limit ourselves to a specific phase of the flight, which begins at 2 h 10 min 05 with the freezing of the Pitot probes and the disconnection of the Automatic Pilot subsystem. Potential causes of the accident may have their roots long before this instant, and may be linked to the overall human-machine system, including pilot training or organization choices of the crew for this flight. However, the last minutes are those that pertain the most to HCI and its models.

P1: BAD DETECTION OF MODE CHANGE

In case of incident, PF and PNF are expected to take control of the plane stability, then to analyse the incident. Here, there is a doubt on the identification of law change.

Since the salience of the speed anomaly was very low compared to that of the autopilot disconnection, the crew detected a problem with this disconnection, and not with the airspeed indications. [...] For the same reasons relating to salience, it is likely that the crew had not yet perceived the reconfiguration to alternate law and the disconnection of the A/THR. (p172)

The crew nonetheless built an initial mental representation of the situation about ten seconds after the autopilot disconnection, based on their identification of a speed indication anomaly. However, they did not specify how many speed sources(21) were lost. The loss of airspeed indication was called out almost simultaneously by both pilots. (p175)

When one of the three speeds deviates too much from the other two, it is automatically rejected by the PRIM's and the voted value then becomes the average of the two remaining values. But if the difference between these two remaining values becomes too great the PRIM's reject them and the control law reconfigures to alternate 2. (p37)

There is however no explicit indication, apart from the red SPD LIM flag next to the speed tape (on the ECAM for example), of the level of alternate law that the aeroplane is in. The ECAM message associated with the reconfiguration to alternate law, of whatever type, indicates "PROT LOST". However, not all of the protections are lost, since the load factor protection remains available, and reduced protections can also exist. The precise identification of the consequences of a reconfiguration in alternate law is thus complicated. (p186)

These problems pertain to the management of the modes of an interface. A mode is a state of the interface in which the same user actions are interpreted differently than in other modes. Modes place two burdens on the users: the perception of the modes, and the memorization of the possible actions and their effects. The perception of the current mode is more difficult when the initiative of the mode change is with the automated subsystems and not with the human operators, and the risk of non-detection is higher. Previous research on glass cockpit aircraft has described mode errors as automation surprises [18]. In the case of flight 447, the pilots did not immediately perceive the change to the alternate 2 law (e.g. mode). They also did not infer the triggering of the change, which notably depends on the number of lost speed probes: one lost speed $\rightarrow$ alternate, two lost speeds $\rightarrow$ alternate 2.

HCI designers recommend to avoid modes as much as possible because they are sources of numerous errors [17, 20]. Nonetheless, a flight requires such a combinatorial complexity that it is difficult to avoid modes. It would be thus useful to understand more deeply the role of modes for complex systems: would it be possible to get rid of them, and if not how can we make the perception of changes more immediate and more reliable?

P2: ADAPTATION TO CONTROL LAW CHANGE

As soon as the autopilot was disconnected, the PF had to take over the steering of the airplane and adapt to the change of control law, without being aware of the change.

The PF was immediately absorbed by dealing with roll, whose oscillations can be explained by: A large initial input on the sidestick under the effect of surprise; The continuation of the oscillations, in the time it took to adapt his piloting at high altitude, while subject to an unusual flight law in roll (direct law). The excessive nature of the PF's inputs can be explained by the startle effect and the emotional shock at the autopilot disconnection, amplified by the lack of practical training for crews in flight at high altitude, together with unusual flight control laws. (p179)

In the case of the accident, the PF tried to control the roll, even if the amplitude of his inputs finally maintained these movements. The relatively strong nose-up inputs that he applied at the same time may have, among other hypotheses, have originated in a certain difficulty in integrating the various types of control laws and thus the difference in the type of handling inputs to adopt between the two axes. (p187)

The change of control law corresponds in HCI to the change of transfer functions. Transfer functions represent the relationships between user movements in control space (here the sidestick) and the result space (here the pitch of the airplane) [3]. To the best of our knowledge, there is no work on the performance of users at adapting "dynamically" to a change of transfer functions, be they conscious of the change or not. A better understanding of these performances and the induced cognitive workload could help prepare pilots to these changes, or to reconsider the design choices if need be.

P3: PERCEPTION DIFFICULTIES

BEA hypothesizes that during the first seconds after the loss of speed information, the pilots may have encountered difficulties to perceive and make sense of the pieces of information displayed at their intention.

The conditions of a night flight in IMC make it more difficult to monitor aeroplane attitudes (pitch attitude in particular). (p174)

A "signifier" is any perceivable indicator (visual mark, sound, touch, taste, force) that communicates an adequate behavior to a person [16]. Signifiers may be deliberate and intentional (e.g. a "PUSH" text on a door) or accidental. In the case of cockpits, the pilots use an artificial horizon in order to perceive the pitch attitude of the plane but can sometimes rely also on the outside view. For AF447, the night and the lack of lights over the ocean make it impossible to rely on this signifier.

The approach to stall on a classic aeroplane is always associated with a more or less pronounced nose-up input. This is not the case on the A330 in alternate law. The specific consequence is that in this control law the aeroplane, placed in a configuration where the thrust is not sufficient to maintain speed on the flight path, would end up by stalling without any inputs on the sidestick. It appears that this absence of positive static stability could have contributed to the PF not identifying the approach to stall. (p187)

Feedback is the signifier that communicates the result of an action: lighting of an indicator, sound of a click etc. The wired commands of the A330 minimize the pilots' physical efforts required to steer the plane. Thus, there is no haptic feedback on the sidestick linked to the physical phenomena ("positive static stability"), unless feedback is imitated with force-feedback devices. This may constitute one of the reasons identified by BEA why approach to stall was not identified by the pilots.

The PNF said “controls to the left”, took over priority without any callout and continued to handle the aeroplane. The PF almost immediately took back priority without any callout and continued piloting. (p23)

It is worth noting that the inputs applied to a sidestick by one pilot cannot be observed easily by the other one. (p174)

There is no feedback in the PNF’s sidestick of the PF’s actions on his sidestick. This pertains to Computer Supported Collaborative Work systems whose interactions are designed to foster teammates’ activity awareness, and in particular to *feedthrough* [7] (whose terminology is built on the model of the term *feedback* : “feed through the interface” toward a teammate as opposed to “feed back” to oneself).

Between the autopilot disconnection and the triggering of the STALL 2 warning, numerous messages were displayed on the ECAM. None of these messages helped the crew to identify the problem associated with the anomalous airspeed. Furthermore, the management of the priorities of the various messages resulted in a rapid changeover of the information displayed, which further complicated the crew’s analysis and understanding of the situation. (p188)

ECAM is a device that displays current “notifications” i.e. transient signifiers (switching from a non-perceptible state to a perceptible one) triggered by a sub-system (an automatic one in most cases) to be delivered to another subsystem (a human on in most cases). Norman seems to assimilate notification and feedback [16]. We diverge on this point and distinguish between the two: notifications are necessary in the case when actions triggered by the user have an effect in a future beyond the working memory, or in monitoring HMIs in which the physical world evolve independently from user’s actions. Without notifications, the user has to engage in a monitoring process that may be cognitively costly.

Notifications must catch users’ attention[5, 21]. Hence, they must employ a signifier that is compatible with human perceptive abilities, notably in a degraded context: apparition of a vibrating visual mark in peripheral vision with a large optical size, a sound louder than the environmental noises (motor, wind) and with a pitch variations, rumble of a sidestick, etc [5][9]. If multiple notifications occur simultaneously, they must be hierarchically organized by the device. In the case of the A330 cockpit, this functionality is provided by the choice of the auditory alarm to play (alternance between C-Chord and Stall), and by the display of messages in the ECAM screen with colored codes mapping the level of alarm.

The occurence of new notifications changes the display, notably the order of the messages along the Y dimension of the screen. Display changes occur instantaneously, which make it difficult for pilots to perceive the occurence of a change in peripheral vision [19], and which make it difficult to understand the reordering of messages even if they are staring at the ECAM [15, 19].

P4: STEERING OF THE PLANE

Beyond taking over during the first few seconds, the PF must take over the control of the plane in order to get back into a stable situation and a safe trajectory, before choosing an emergency procedure or exploring the problem at hand. Here the PF did not succeed in reaching a stable situation.

The first disturbances in speeds 1 and 2 occurred at about 2 h 10 min 04, causing the autopilot to disconnect, which was signalled by a visual and an aural (cavalry charge) warning. [...] Since the salience of the speed anomaly was very low compared to that of the autopilot disconnection, the crew detected a problem with this disconnection, and not with the airspeed indications. The crew reacted with the normal, learned reflex action, which was to take over manual control (indicated by the PF’s call-out “I have the controls”, acknowledged by the PNF). (p172)

Although the PF’s initial excessive nose-up reaction may thus be fairly easily understood, the same is not true for the persistence of this input, which generated a significant vertical flight path deviation. [...] There remain a number of possible explanations: The crew’s attention being focused on roll, speed or on the ECAM; The initiation, more or less consciously due to the effects of surprise and stress, of the action plan (climb) desired by the PF prior to the autopilot disconnection; [...] (p173)

The pitch attitude oscillations, in the seconds following the activation of the stall warning, reveal that the handling of the aeroplane was clearly very difficult and probably demanded the PF’s full attention. During this phase, the aeroplane symbol on the PFD was close to, but on average slightly above, the flight director horizontal bar. The PF likely attempted to track this crossbar as it changed without having integrated the change of longitudinal engaged mode. (p181)

Note: The "Vol avec IAS douteuse" procedure recommends disabling the FD, to prevent it from presenting cues that could potentially be irrelevant. (p181)

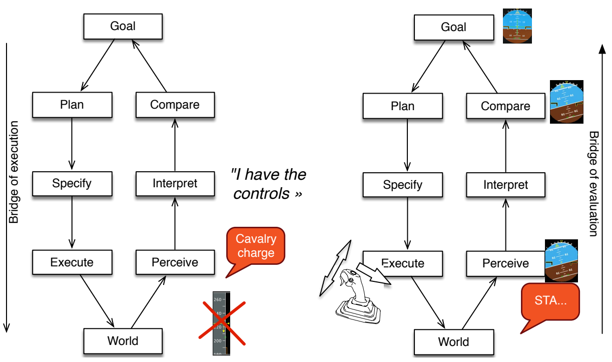

Figure 4. Action cycle for PF

In order to take back the control of the airplane, the pilot executes actions and evaluates the results of these actions by perceiving the surrounding environment. These processes can be modeled with the “action cycle” [16] (fig. 4). When people employ a device, they have to cross two "bridges"): the Bridge of Execution, when they try to figure out how to act, and the Bridge of Evaluation, when they try to figure out the consequences of their actions after performing them. Fig. 4 shows two instances of this model for the PF. The first cycle models the autopilot deconnection: the PF perceives it with the auditory C-Chord warning. The second cycle models the PF actions to control the flight path. The STALL 1 alarm plays truncated then stops, while the PF follows the indication of the Flight Director on the PFD: while he counteracts the roll, he makes a nose up.

P5: PROBLEM SOLVING (OR ABDUCTION)

The action cycle is a simplification of reality but nonetheless constitutes a useful framework to understand the how and the why of human actions in a particular situation. In particular, in case of unexpected behavior, users question the elements that pertain to concepts cited by Norman: signifiers, feedbacks, notifications. Notably, during these situations they aim at understanding the causes of the notifications that occur. Thus, pilots are expected to explore the nature of the problem and analyze it so as to identify the actions that could solve it, after applying potential emergency procedures.

> The intention is then that the crew will detect the anomaly, that they will possibly "make sense" of this detection [...] From the information available on the ECAM, the crew must analyse and confirm the type of failure before undertaking any failure processing action. [...] Applying rules assumes not only their knowledge, but also the recognition of their conditions of applicability, and therefore the correct identification plus a specific interpretation of the anomaly. The construction of a response by calling on experience assumes incorporation of the anomaly in the mental representation of the situation, which can go via its destruction/reconstruction, very wasteful in resources and time-consuming. In this way the correct perception of the situation by a crew, which enables the reliability and speed of diagnosis and decision to be improved, is linked not only to the way in which the situation is presented to this crew (interfaces, parameters) but also to their training and experience. (p101)

A process of abduction

The activity that the PNF must conduct is a process of abduction: "the process that allows to explain a phenomena or an observation from some facts, events or laws. This is similar to the medical domain for instance, where the final diagnostic explains the signs and symptoms of a patient by assuming a malfunction such as a disease of a fracture" [12]. BEA expresses the elements required to the completion of such a process:

> It is therefore necessary: That these signs [of the problem] be credible and relevant;
> That the available indications relating to the anomaly are very swiftly identifiable so that the possible immediate actions to perform from memory to stabilise the situation are triggered or that the identification of the applicable procedure is done correctly. In particular, it is important that the interfaces that usually carry anomaly information display, or at least allow, this initial diagnostic, given the minimum competence expected of a crew;
> [...] That there are no signals or information available that suggest different actions or that incite the crew to prior reconstruction of their understanding the situation. (p102)

Pilots thus seek to eliminate doubtful signifiers, or seek consistency between signifiers. They also seek to understand the consequences of their actions, and they even act to test hypotheses.

Signifiers credibility

The STALL warning triggered for the first time (STALL 1) as soon as the autopilot has disconnected for 2s.

> In these cases, the warnings are triggered by a local increase in the angle of attack; they are therefore transient and are generally expressed as truncated warnings (a synthesised voice sounds saying "STALL, STALL", sometimes incompletely). Previous events that have been studied (stall warning triggered in the context of a speed anomaly at cruise speed) show, however, that other crews have not reacted as expected to the proximity of the stall and had a tendency to consider the warning as spurious. For this reason, the behaviour of the AF 447's crew should be considered as liable to be reproduced as regards the lack of reaction to the STALL 1 warning. [...] such spurious triggering may be considered as inappropriate and likely to impair the overall credibility of a warning which is almost never encountered by crews during type rating, in flight or during training. (p189)

> The transient activations of the warning after the autopilot disconnection may have caused the crew to doubt its credibility. Furthermore, the fact that the flight director was advising a nose-up attitude may have confirmed the PF's belief that the stall warning was not relevant. During previous events studied, crews frequently mentioned their doubts regarding the relevance of the stall warning. (p180)

Signifiers Invisibility

> The crew is only informed of the consequences of the triggering of these monitoring mechanisms: disconnection of the AP and of the ATHR, transition to alternate law etc. No failure message is provided that identifies the origin of these other failures: in particular, the rejection of the ADR's and of the speed measurements. (p187)

Misunderstanding of the Consequences

> However, the PF may have assimilated the triggering of the warning as a consequence of the reduction in thrust, which he had applied four seconds earlier; he should then have applied full thrust to return to the earlier situation. A few seconds later, the PF said "I'm in TOGA, right?". Either he was unsure whether or not he had set the thrust controls to the TOGA detent, as he intended, or he did not understand why this action was ineffective in clearing the stall warning. This second case might therefore indicate that the PF had built an erroneous mental representation of the aeroplane's flight model, and that he had hoped that he could resolve the situation by applying TOGA thrust

at high altitude and a pitch attitude of twelve degrees, a strategy similar to that recommended at low altitudes. The fruitless result of his actions possibly heightened his mistrust of the warning. (p180)

Actions to test hypotheses

The disabling of the THRUST LOCK function by the PF indicates that he was searching for information. The PF may therefore have been overloaded by the combination of his immediate and natural attempts to understand the situation that was added to the already demanding task of handling the aeroplane. (p176)

Meanwhile, the PNF turned on the wing anti-icing system, after reading the ECAM, which suggests that at this point he may have considered there was a severe icing problem. The sound of ice crystals hitting the windshield, considered as rain by other crews, may have supported this perception of an associated risk. The symptoms perceived may therefore have been considered by the crew as anomalies to add to the anomaly of the airspeed indication, and thus indicative of a much more complex overall problem than simply the loss of airspeed information. (p176)

He also actuated the ATT/HDG rotary switch and called out this action ("I'm putting you in ATT..."). This change of inertial source, which with hindsight was not necessary, may indicate that his diagnosis of the failure was not completely defined. For him, the airspeeds indicated were inconsistent; he may not have excluded the possibility, however, that the inertial information was also inconsistent. After changing the ADR source, the PNF's "what is that" appears to indicate his total incomprehension faced with the result of this action, since the speed displayed on the right side was still erroneous. He appeared at this point to have been overwhelmed. (p177)

The pilots have performed multiple action cycles, during which they acted and evaluated the consequences of their actions. During the cycles, they also form and test hypotheses on the behavior of the system: they rebuild what is called a *conceptual model*.

Conceptual model of Stall Warning

A conceptual model is an explanation, usually highly simplified, of how something works [16]. For example, computer desktop interfaces rely on a conceptual model based on files and directories. Conceptual models bring an understanding of devices, predictions on how devices will behave and solutions when what is happening does not coincide with what has been planed. They are considered as the most important element for the usability of interactive systems[13, 10]. While they do not use this terminology, BEA mention something akin to conceptual models:

In the absence of reliable speed indication, an understanding of the physics of high-altitude flying, gained through training in the fundamental principles of energy conversion, equilibriums of forces, and lift and propulsion ceilings, could have considerably helped the pilots to anticipate the rapid deterioration in their situation and to take the appropriate corrective measure in time: initiate a descent. [...] Air France's Aeronautical Manual (MAC) describes in great detail, over 38 pages, the physics of high-altitude flight with real cases. This knowledge is also included in the theoretical teaching that is supposed to be provided at an advanced stage in the training of a future airline pilot (ATPL theory, type rating performance). The climbing flight path that was initially more or less deliberate on the part of the crew is likely a clue to the insufficient assimilation of these theoretical notions. (p183)

Conceptual models described in technical manuals may be complex. Nonetheless, they must be compatible with the cognitive abilities of the targeted users as well as the tasks they have to fulfill in a given context.

Inattentional deafness has been evoked as an explanation of the apparent disregard of the stall warning [6]. A number of sections of the BEA report rather suggest that one of the conceptual models used by the pilots was not sufficient to understand the situation.

Until the end of the flight, the angle of attack values changed successively from valid to invalid. Each time that at least one value became valid again, the stall warning re-triggered and each time the angle of attack values were invalid, the warning stopped. Several nose-down inputs caused a drop in the pitch attitude and the angle of attack, whose values then became valid, such that a clear nose-down input resulted in the triggering of the stall warning. It appears that the PF reacted, on at least two occasions, with a nose-up input, whose consequences were an increase in angle of attack, a drop in measured speed and consequently stopping the stall warning. Until the end of the flight, no valid angle of attack value was less than 35°. (p190)

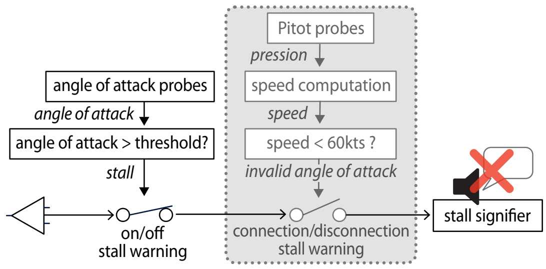

Figure 5. Conceptual Model of stall warning behavior (non-gray PF model, gray: model closer to reality)

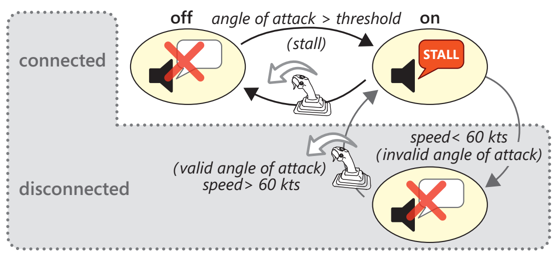

Figure 6. Conceptual model of states and transitions of the stall signifier (non-gray: PF model, gray: model closer to reality)

Fig. 5 presents two conceptual models of the behavior of stall warning: what might be the model used by the PF, and a model closer to reality that illustrates the lack of signifier in case of low speed. Fig. 6 shows the conceptual models of states and transitions of the stall signifier. The PF might have thought that the alarm had switched from state "off" to "on" when producing sound again, when it had actually switched from state "on-disconnected" to "on-connected". Both states "off-connected" and "on-disconnected" have the same signifier (silence), which makes them indistinguishable. The PF might have noted that the consequences of his actions are opposite to the predictions of his conceptual model: in general, nose-dive actions are supposed to exit from stall, while they have signaled here an entry in stall. Maybe inferring that his nose-down action triggers a stall, the PF stopped his action by making a nose-up, which switched the alarm off by switching its state to "on-disconnected" (and not to "off").

However, irrespective of the ergonomics of the warning, it is likely that the presentation of information that provides an overview of the aeroplane's situation (angle of attack, energy balance (kinetic and potential), flight envelope) would help pilots to "make sense" of the warning and to take the appropriate corrective action in time. To summarise, the following factors tend to diminish the performance expected from many crews: [...] The lack of any description of the functioning of the stall warning (a structure diagram or indications of threshold levels, for example) in the documentation; [...] (p189)

Consequently, the BEA recommends that EASA improve the feedback process by making mandatory the operational and human factors analysis of in-service events in order to improve procedures and the content of training programmes. (p212)

These findings suggest that, in degraded situations, the pilot profession requires skills that are similar to those of physicians, researchers and engineers: the ability to select facts, challenge one's mental model, build up a new plausible model, and infer a new plan of action and validation. Pilot training includes relatively advanced notions on flight physics and system behavior: why know these notions if not to use them when analyzing problems in non-nominal situations? However, how can we make sure that the available signifiers in degraded situations allow pilots to conduct a correct reasoning in a limited time? Should we define new pilot selection criteria before training, or should we define a specific training? Should we conduct research to provide them with representations of systems that more suitable to handle complexity, as is already done for engineers [8] or researchers [11]?

CONCLUSION

The match between the data and the available conceptual frameworks makes the BEA report a useful basis for HCI teaching. In particular, it may illustrate to future designers the importance of eliminating all factors that would lead to erroneous interpretations during the design of a human-machine system.

Our work may also open new research tracks. Despite ongoing efforts [4, 14], knowledge in HCI is still insufficient to formally isolate the whole set of failures of a human-machine system, notably those cited here. Hence, it is still difficult for an engineer to identify all problematic configurations. However, research in HCI may contribute to reinforce safety, by using the report as an inspiration source. In particular, this accident reminds us of the importance of eliminating factors that mislead users in their reasoning. Doing research on the abduction task that these pilots had to perform can provide other pilots with new skills to support their last-resort role: how can an HMI support a pilot during reasoning, as much as any tool that assists an engineer, a scientific or a physician?

REFERENCES

1. Airbus. *A330 Flight deck and systems briefing for pilots*, 1999.
2. Bureau d'Enquêtes et d'Analyse. Rapport final - Accident survenu le 1er juin 2009 à l'Airbus A330-203 immatriculé F-GZCP exploité par Air France vol AF 447 Rio de Janeiro - Paris. Tech. rep., République Française, Ministère de l'Ecologie, du Développement durable et de l'Énergie, 2012.
3. Casiez G. & Roussel N. No more bricolage!: Methods and tools to characterize, replicate and compare pointing transfer functions. In Proc. *UIST '11*, ACM (2011), 603–614.
4. Chandler F., Chang Y., Mosleh A., Marble J., Boring R. & Gertman D. Human reliability analysis methods: selection guidance for nasa. Tech. rep., NASA Office of Safety and Mission Assurance, Washington, DC, 2006.
5. Chatty S., Athènes S. & Bustico A. Votre attention s'il vous plait ! Éléments d'un espace de conception de signaux visuels. In *Actes de IHM*, Cépaduès Editions (1999).
6. Dehais F., Causse M., Vachon F., Régis N., Menant E. & Tremblay S. Inattentional deafness to auditory alarms in aviation: A safety issue. *Human Factors and Ergonomics Society*, in press.
7. Dix A. J. *CSCW - a framework*. 1994, 23–37.
8. Gaspard-Boulinc H. & Conversy S. Exigences d'utilisabilité pour les outils d'ingénierie des exigences. In *Proc. IHM*, ACM (2014).
9. Giraudet L., Imbert J.-P., Bérenger M., Tremblay S. & Causse M. The Neuroergonomic Evaluation of Human Machine Interface Design in Air Traffic Control using behavioral and EGG/ERP measures. *Behavioural Brain Research*, 293 (2015), In press.
10. Henderson A. & Johnson J. *Conceptual Models: Core to Good Design*, 1st ed. Morgan & Claypool Publishers, 2011.
11. Keim D., Andrienko G., Fekete J.-D., Görg C., Kohlhammer J. & Melançon G. Visual analytics: Definition, process, and challenges. In *Information visualization*. Springer, 2008, 154–175.
12. Lecourt D., Ed. *Dictionnaire d'histoire et philosophie des sciences*. PUF, 1999.
13. Liddle D. Bringing design to software. ACM, New York, NY, USA, 1996, ch. Design of the Conceptual Model, 17–36.
14. Martinie C., Navarre D. & Palanque P. A multi-formalism approach for model-based dynamic distribution of UI of critical interactive systems. *Inter. J. of Hum.-Comp. Studies 72*, 1 (2014), 77 – 99.
15. Mertz C., Chatty S. & Vinot J.-L. The influence of design techniques on user interfaces: the digistrips experiment for air traffic control. In *Proceedings of HCI Aero* (2000).
16. Norman D. A. *The Design of Everyday Things*, revised and expanded edition ed. Basic Books, USA, 2013.
17. Raskin J. *The Humane Interface: New Directions for Designing Interactive Systems*. ACM Press/Addison-Wesley, 2000.
18. Sarter N. B., Woods D. D. & Billings C. E. Automation surprises. In *Handbook of human factors and ergonomics 2*. 1997, 1926–1943.
19. Schlienger C., Conversy S., Chatty S., Anquetil M. & Mertz C. Improving users comprehension of changes with animation and sound: an empirical assessment. In *Proc. of Interact 2007, LNCS*, vol. 4662, IFIP, Springer Verlag (2007), 207–220.
20. Tesler L. A personal history of modeless text editing and cut/copy-paste. *interactions 19*, 4 (2012), 70–75.
21. Wickens C., Hollands J., Parasuraman R. & Banbury S. *Engineering Psychology & Human Performance*. Pearson Education, 2012.