Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®
Int. J. Human-Computer Studies 60 (2004) 117–128

International Journal of
Human-Computer
Studies

www.elsevier.com/locate/ijhcs

# When mental models go wrong: co-occurrences in dynamic, critical systems

Denis Besnard[a,*], David Greathead[a], Gordon Baxter[b]

[a] Centre for Software Reliability, School of Computing Science, University of Newcastle upon Tyne, Claremont Road, Newcastle upon Tyne NE1 7RU, UK
[b] Department of Psychology, University of York, Heslington, York YO10 5DD, UK

Received 25 April 2003; received in revised form 16 September 2003; accepted 17 September 2003

**Abstract**

This paper highlights a psychological phenomenon affecting the accuracy of mental models. It occurs when two consecutive events happen as expected by an operator. Typically, such a situation reinforces the confidence in one's mental model. However, consecutive events can happen as a random co-occurrence, for reasons that actually differ from the ones believed by the operator. Nonetheless, because of the consistency between the environmental data and the operator's expectations, one event can be taken to be the cause of the other. When this false belief happens, the mental model is erroneously assumed to be valid. We discuss this phenomenon and its potential disastrous consequences using the example of a real commercial air crash. We finally address some implications for systems' design and support tools.


## 1. Introduction

During a presentation, one of the authors' colleagues using his laptop for displaying slides got interrupted when his computer's screen went blank. He hit the track pad of his laptop in case the latter had gone into sleep mode but the video signal did not come back. As he was using an adapter for the VGA display cable, he suspected the connection had come loose. After some seconds during which he

*Corresponding author. Tel.: 44-0-191-222-8058; fax: 44-0-191-222-8788.
*E-mail addresses:* denis.besnard@ncl.ac.uk (D. Besnard), david.greathead@ncl.ac.uk (D. Greathead), g.baxter@psych.york.ac.uk (G. Baxter).

1071-5819/$ - see front matter 
doi:10.1016/j.ijhcs.2003.09.001

tightened more firmly the adapter and the display cable together, the video came back on. The problem was solved and, hopefully, the adapter would now behave normally. Some minutes later, the same scenario happened again and our colleague performed the same actions as before. He hit the track pad but this triggered no signal. He then modified the position of the display adapter so that the weight of the cable would not pull on it. A couple of seconds later, the video was back on again. The third time the scenario happened, it became obvious, at least to everyone *except* the person using the laptop, that the manipulation of the display adapter and the video coming back on was pure coincidence. We discovered that the machine was going into sleep mode and needed about 6–8 s to exit it. It is likely that any action carried out just before the video signal came back would be considered as the solution.

Through this simple—yet real—example, we wish to highlight an interesting cognitive feature that is common across domains. Humans tend to consider that their vision of the world is correct whenever events happen in accordance with their expectations. However, two sequential events can happen as expected without their cause being captured. When this is the case, humans tend to treat the available evidence as exhaustively reflecting the world, erroneously believing that they have understood the problem at hand. These co-occurring events can seriously disrupt situation awareness when humans are using mental models that are highly discrepant to reality but nonetheless trusted.

In the next section, we consider how the mechanisms leading to the aforementioned error can be explained using the concept of mental models. We then assess the role of this error in dynamic, critical applications using the example of a commercial air crash (Section 3). We finally discuss some ways for improving the design of critical systems to avoid such problems (Section 4).

## 2. Mental models

Limitations in memory and processing capabilities mean that humans cannot handle the totality of the information displayed in their environment. Instead, they build representations that are meant to support behaviour (Rabardel, 1995). These mental models can take various forms (see Moray, 1996, for a taxonomy of the different types). Since we are describing cognitive activities in situ, we take mental models to be scarce, goal-driven images of the world that are built to understand the current and future states of a situation.

What characterizes best mental models is their incompleteness. Their content is only a partial representation of the environment and their scope is limited (Sanderson, 1990; Sanderson and Murtagh, 1990). They are essentially built from (a) the knowledge needed for pursuing a given goal and (b) some data extracted from the environment. The resulting image of the world is one in which the essential features of a problem are overemphasized whereas the peripheral data can be overlooked (Ochanine, 1978). In this respect, mental models are described as homomorphic representations of the world (Moray, 1987). They are simplified, cognitively acceptable versions of a too complex reality. In dynamic processes, it may be the case that mental models are correct at early stages of the interaction. Over time, with the evolution of this interaction and possible degradation of the situation (due to an emergency, for instance), the model gets simplified and becomes more based on correlation between system elements, such that only the best predictors of the system's states are taken into account (Baxter and Ritter, 1999).

Although building and maintaining mental models are core activities in the control of dynamic complex systems, human operators also have to perform several other critical tasks that affect system performance. They have to plan actions, control movements, exchange information with collaborators and so on. This complex combination of tasks has to be executed with a limited amount of cognitive resources. For this reason, operators tend to save resources whenever it is possible using mechanisms such as selective memory and heuristic, shortcut-based reasoning (Rasmussen, 1986). In the case of mental models, saving resources causes them to be built on the basis of partial pieces of evidence. However, this has to be seen as the consequence of cognitive limitations where problems are solved according to an intuitive cost-benefit trade-off. Since Simon's (1957) concept of bounded rationality, it is accepted that cheap adequate solutions are often preferred to costly perfect ones. In other words, people tend to satisfice rather than optimize, settling on a solution that is deemed good enough even though it may be sub-optimal.

The consequences of flawed mental models can be disastrous when operators are interacting with dynamic critical systems, e.g. commercial airplanes. Operators of these systems (i.e. pilots) are sometimes faced with unexpected problems, which they have to diagnose and resolve. This local troubleshooting activity, which is inserted in the more global objective of piloting the aircraft, involves the construction of an explanation in real-time. Factors such as limited cognitive resources, confirmation bias (Klayman and Ha, 1989) and time pressure can mean that pilots construct an erroneous explanation of such incidents. Flaws in mental models are detected when the interaction with the world reveals unexpected events. However, these inaccurate mental models do not always lead to accidents. Very often, they are recovered from. In this respect, error detection and compensation are significant features in human information processing. The weakness of mental models lies in their poor requirements in terms of validity: if the environmental stream of data is consistent with the operator's expectations, that is enough for the operator to continue regarding the mental model as valid. The understanding of the mechanisms generating the data is not a necessary condition.

We are not concerned here with how operators could build exhaustive mental models, as their incompleteness reflects a strong need for information selection. Rather, the issue of interest is to understand the conditions in which operators believe they have a good picture of the situation whereas the underlying causal mechanisms have not been captured. One of these conditions is the co-occurrence of events. The types of problem that can arise from this phenomenon are illustrated in the next section.

## 3. The Kegworth accident

The Kegworth air crash (Air Accidents Investigation Branch, 1989) was chosen to illustrate the problems that can be caused by co-occurring events, for two reasons. First, it is a well-known accident, allowing us to ground our discussions in a case that is now well-understood. Second, and most importantly, this crash offers a very good—although tragic—instance of the mechanism we wish to discuss in this paper. Although the Kegworth accident report was issued several years ago, the co-occurrence angle has not been studied to date. For these two reasons, we believe further progress can be made in understanding the psychological mechanisms that were causally involved in the crash. There is also some value in achieving this understanding since this mechanism is believed to be domain independent. Therefore, the Kegworth accident may provide data from which we can gain useful knowledge that can be generalized to human–machine interaction in the large.

On the 8th of January 1989, a British Midland Airways Boeing 737-400 aircraft crashed into the embankment of the M1 motorway near Kegworth (Leicestershire, UK), resulting in the loss of 47 lives. The crash resulted from the flight crew's management of a mechanical incident in the left (#1) engine. A fan blade detached from the engine, resulting in vibration (severe enough to be felt by the crew) and the production of smoke and fumes that were drawn into the aircraft through the air conditioning system. The flight crew mistakenly identified the faulty engine as the right (♯2) engine. The cockpit voice recorder showed that there was some hesitation in making the identification. When the captain asked which engine was faulty, the first officer replied *It's the le…it's the right one*, at which point the right engine was throttled back and eventually shut down. This action coincided with a drop in vibration and the cessation of smoke and fumes from the left (faulty) engine. On the basis on these symptoms, the flight crew deduced that the correct decision had been taken, and sought to make an emergency landing at East Midlands airport. The left engine continued to show an abnormal level of vibration for some minutes, although this seems to have passed unnoticed by the pilots. Soon afterwards, the crew reduced power to this engine to begin descent, whereupon the vibration in the engine dropped to a point a little above normal. Approximately 10 minutes later, power to the left engine was increased to maintain altitude during the final stages of descent. This resulted in greatly increased vibration, the loss of power in that engine and the generation of an associated fire warning. The crew attempted at this point to restart the right engine but this was not achieved in the time before impact, which occurred 0.5 nautical miles from the runway.

In addition to the crew's mistakes, several other factors contributed to the accident. When later interviewed, both pilots indicated that neither of them remembered seeing any indications of high vibration on the Engine Instrument System (EIS; see Fig. 1). The captain stated that he rarely scanned the vibration gauges because, in his experience, he had found them to be unreliable in other aircraft. It is also worth noting that the aircraft was using a new EIS which used digital displays rather than mechanical pointers. In a survey carried out in June 1989 (summarized in the accident report), 64% of British Midland Airways pilots

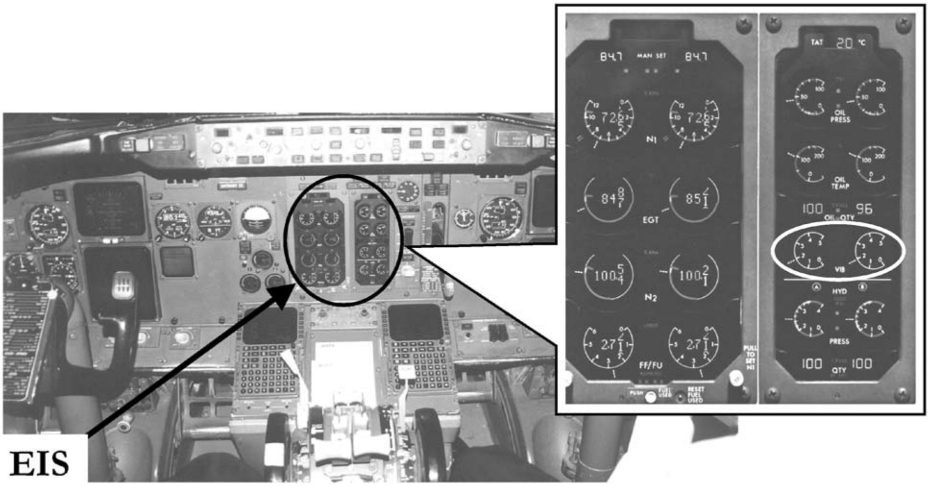


Fig. 1. A Boeing 737-400 cockpit. The EIS is located in the centre (© Pedro Becken. All rights reserved). The secondary EIS is magnified on the right-hand side of the picture. The vibration indicators are circled in white (photo from Air Accident Investigation Branch accident report).

indicated that the new EIS was not effective in drawing their attention to rapid changes in engine parameters and 74% preferred the old EIS. The secondary EIS, on which the vibration indicator was located followed standard design practice and hence did not include any audio or additional visual warning to indicate excessive readings.

As another contributing factor, the crew workload increased out of control. Some time after the ♯2 (working) engine had been erroneously shut down, the captain tried without success to stay in phase with the evolution of the incident. He was heard on the cockpit voice recorder saying: *Now what indications did we actually get (it) just rapid vibrations in the aeroplane—smoke*…. At this point, the crew were interrupted with a radio communication from air traffic control. Later, the flight service manager entered the flight deck and reported that the passengers were very panicky. This further distracted the flight crew and the captain had to broadcast a message of reassurance to the passengers. Both the captain and first officer were also required to make further radio communications and perform other duties in preparation for the landing. All of these actions affected the degree of control of the emergency.

Finally, it is worth noting that while both the captain and the first officer were experienced (over 13,000 h and over 3200 h flying time, respectively), they had only 76 h experience in the Boeing 737-400 series between them.

## 4. Discussion

In this section, we briefly address some general issues about mental models in complex dynamic systems (for a definition of these systems, see Brehmer, 1996 or

Cellier et al., 1997). Based on our analysis of the Kegworth accident data, we then make some general suggestions for improving the dependability of systems that are deployed in complex dynamic domains. We conclude the discussion by outlining some limits that need to be made explicit when studying human error.

There are some similarities between the particular flaws in mental models that we have identified using the Kegworth accident data, and some existing types of human error. Nevertheless, our approach can be distinguished from mode confusion (Crow et al., 2000; Leveson et al., 1997; Rushby, 2001; Rushby et al., 1999) in that the automation behaves as expected in our case. It can also be distinguished from fixation errors (De Keyser and Woods, 1990) in that we are concerned with the processes involved in the construction of mental models rather than their persistence.

### 4.1. Mental models in a dynamic world

Attempting to save cognitive resources causes mental models to be biased in such a way that partial confirmation is easily accepted. Instead of looking for contradictory evidence, people tend to wait for consistent data. This phenomenon, called confirmation bias (Klayman and Ha, 1989), has already been studied in human–machine interaction (e.g. Yoon and Hammer, 1988). The corollary of confirmation bias is that people overlook contradictory data. This is one explanation for the reinforcement of flawed mental models. In the case of the Kegworth accident, an erroneous decision coincided with a reduction in the level of the symptoms, which lasted for some 20 minutes. When it is compatible with the operator's expectations, this type of co-occurrence probably works against rejection of the existing mental model. It also makes it harder to integrate any contradictory evidence that may subsequently become available.

Operators are more likely to reject any information that is not consistent with their expectations, rather than update their mental model. The latter has a cost that operators cannot always afford in time-critical situations. In the end, data can be abusively reinterpreted to fit the model that operators have of a situation (Moray, 1987). This confirmation bias is probably the outcome of an economy-driven reasoning: following a line of least effort (Rasmussen, 1986), operators can treat random data as meaningful if it matches their vision of the world.

Operators can erroneously maintain as valid, representations that have already departed from a reasonable picture of the reality. In dynamic situations, one reason is that operators try to avoid the cost of revising their mental model as long as it allows them to stay more or less in control. In other words, they satisfice. Because mental models are constantly matched against the feedback from the process they control, they are fed with a constant stream of data. However, there exist situations where the feedback is discrepant from the operator's expectations. When this discrepancy provokes such a loss of control that required tasks cannot be run anymore, some costly revision of the mental model as well as diagnostic actions are needed (Rasmussen, 1993). This is a non-trivial task in dynamic situations such as piloting an aircraft in that some control is already lost and the crew is required to run, coordinate and share two processes at the same time. One is a rule-based control of the flight parameters: the plane must continue to fly. The other process is information gathering and integration. The potential work overload caused by this dual activity may explain why outdated (flawed) mental models are maintained even after the detection of some mismatches. Provided they can keep the system within safe boundaries, operators in critical situations sometimes opt to lose some situation awareness rather than spend time gathering data at the cost of a total loss of control (Amalberti, 1996).

Critical situations can be caused by the combination of an emergency followed by some loss of control. When this happens, there is little room for recovery. The Kegworth accident probably falls into this category. When the flight crew got around to thinking about checking the engine data, they were distracted by other more urgent tasks. The emergency nature of the situation and the emerging workload delayed the revision of the mental model, which ultimately was not resumed.

### 4.2. Implications for the design of dependable systems

The Kegworth crash highlights that the control of automation is a real dialogue between operators and machines. When this dialogue fails because information flow does not help situation awareness, events are likely to be processed in a sub-optimal manner. The following discussion will not focus explicitly on co-occurrences as they are a fine-grained mechanism as compared to the complexity and diversity of cognitive activities involved in the control of dynamic processes. Instead, we think a wider discussion is needed to assess more precisely the stakes of a more reliable interaction between operators and machines. We believe that human–machine interaction could be improved in two complementary ways.

#### 4.2.1. Training operators

Operators must be made more aware of human factors through training and education. Some psychological mechanisms can then become more obvious to the operators themselves and positively influence the perception that they have of their own performance. In aviation, for example, human factors have been an integrated part of pilot training for many years (e.g. Green et al., 1996). Nowadays, these training schemes include explicit consideration of cognitive psychology, decision-making and human error. Within the cockpit, these contents contribute to better communication, more efficient distributed decision-making and improved stress management. Looking at the broad picture, human factors must be regarded as one of the many elements that keep the commercial accident rate constant despite the increase in traffic. In the near future, it will be the case that more and more pilots, who will have been educated in human factors from the early stages, will contribute to an even higher degree to critical systems' dependability.

#### 4.2.2. Embedded agents

The automation must be designed to be aware of the operators by having some embedded knowledge of human reasoning as well as some screening functions (e.g. Boy, 1987; Rasmussen, 1991). This would allow machines to anticipate operator's decisions, provide more appropriate context-sensitive alarms and support for critical decisions. Expected benefits include the provision of some assistance in emergency situations before matters become too critical. Operators need more help in those situations for which they have not been trained, than on nominal settings. It implies that systems at large have to be designed in such a way that unexpected events can be recognized and appropriately handled by support tools.

Wageman (1998) argues that interfaces can typically flood operators with extra data at a time of the process (e.g. emergencies) where few resources are still available. From our point of view, we think it is precisely because operators' intentions are not captured by automated systems that over-information occurs. This issue has been developed by Hollnagel (1987) who proposed the concept of *intelligent decision support systems*, and further addressed by Filgueiras (1999). One way forward may be to design support tools that incorporate models of the system they are a part of. This would permit the automation to predict the future states it is going to enter given the inputs coming from the environment and the operator. Without this kind of assistance, operators will have to continue looking for data during critical phases of process control.

The use of embedded agents to support decision-making has already started in the aviation domain. The most notable example is probably Hazard Monitor (e.g. Bass et al., 1997). This system tracks user interactions and tries to match these against stored expectation networks of normative behaviour as a way of doing plan recognition. On the basis of what it detects, Hazard Monitor can then make suggestions about what the user should consider to do next. These suggestions are prioritized, and initially start as gentle reminders about something that needs to be done. The second level is a stronger reminder that it is getting more critical that a particular action be performed. Finally, as a last resort, Hazard Monitor can directly intervene to take the action. Even though Hazard Monitor is a sophisticated system, it still only deals with routine interaction. In other words, it is largely restricted to dealing with plans that would be used in normal circumstances. In the particular case of co-occurring events, it would not be able to offer any assistance to avoid the potential problems. Nonetheless, Hazard Monitor is a concrete example of a decision support agent that highlights a tangible research stream in critical systems' design (see Jennings and Wooldridge, 1998, for a more complete list of applications of intelligent agents).

We believe that research in intelligent agents can enhance the reliability of human-machine systems by improving the human–machine interaction in general. Agents need (a) to provide support to help operators make the right decisions and take the appropriate actions, and (b) to act as barriers when they try to perform erroneous actions. However, we are aware of some limitations. For instance, the notion of correctness itself is difficult to capture since it is highly context-dependent. Moreover, some totally unexpected failures, such as the one causing the crash-landing of a DC-10 in Sioux-City (NTSB, 1990), can make agents useless, if not undesirable. It could also be argued that agents based on inductive methods of reconstructing pilots' intentions can have flaws in data selection. Although the latter is a distinct possibility, these flaws may be caused by our current state of knowledge rather than by the principle of inferring mental models from interaction data. As suggested by Hollnagel and Woods (1999), more research is needed in cognitive ergonomics and embedded agents design. In the field of human–machine interaction, this would facilitate the integration of such mechanisms as construction, revision and decay of mental models. New research in neuroergonomics (e.g. Hancock and Szalma, 2003; Parasuraman, 2003) may help to provide more detailed information about how these processes occur.

### 4.3. Supporting operators towards aviation safety

Generally speaking, we see our recommendations as local measures belonging to a broader context where the operator still retains the overall responsibility as the final decision maker. But for operators' decisions to be accurate in emergency situations, for example, a fast and clear understanding of the situation is needed. We feel that dedicating some computing power to information gathering, analysis and prediction will feed operators' mental models with more useful data as well as supporting a better interaction between human agents and complex dynamic critical systems. This matters for critical operations such as aircraft piloting (see Dehais et al., 2003). The constant increase in commercial aviation traffic has not been followed by an increase in accident rate. Contributions to this rather positive state of facts are the steady technical improvement of modern aircraft as well as the genuine cognitive flexibility of flight crews when handling exceptions (Besnard and Greathead, 2002). Nevertheless, as reported by Amalberti (1996), a flat accident rate has persisted since the 1970s. This is why we think more efforts have to be invested in the reliability of the dialogue between operators and automation. We think, following Amalberti, that the limit to modern aviation safety now lies in the extent to which we can improve cooperation in the dialogue between automated systems and operators. This assertion undoubtedly extends beyond aircraft piloting and hits any critical system where humans have to take decisions.

### 4.4. Limits

We wish to emphasize that mental models can also fail for several reasons other than co-occurrence, such as complexity, lack of knowledge, and workload. We want to give co-occurrences the attention they deserve. They can lead to catastrophes but only account for a small portion of the failures of mental models.

Although we have focussed on the weaknesses of mental models, we also have to emphasize that human errors are not cognitive dysfunctions. Often, errors must be seen as marginal events caused by the same mechanisms that generate correct actions most of the time (Rasmussen, 1986). As a consequence, errors have to be considered in this paper as the side-effects of a cost/benefit driven reasoning process aimed at getting an optimal performance for the lowest mental cost (Amalberti, 1991; 1996).

## 5. Conclusion

Human performance is fallible. Paradoxically it remains one of the ironies of automation that operators are still required to intervene and fix the problem when the automation fails (Bainbridge, 1987). It is the inherent human traits of flexibility and adaptability that allow them to be able to do this. What we have started to argue here is that monitoring and control in human–machine systems needs to be considered more as a joint concept.

Our level of understanding of the ways and situations in which human performance can go awry continues to improve. In particular, in this paper we have shown how co-occurring events can lead to problems. The next step is to capture the details of this phenomenon in such a manner that they can be incorporated into the design of systems, using embedded intelligent agents, for example, as a way of increasing the dependability of the overall system in the face of increasing complexity.

## Acknowledgements

The funding for this study was provided as part of the Dependability Interdisciplinary Research Collaboration (DIRC)[1] which is funded by the EPSRC grant number GR/N13999. The authors wish to thank anonymous reviewers for useful comments.

## References

Air Accidents Investigation Branch, 1989. Report on the accident to Boeing 737-400-G-OBME near Kegworth, Leicestershire on 8 January 1989. Available online at http://www.aaib.dft.gov.uk/formal/gobme/gobmerep.htm

Amalberti, R., 1991. Modèles de raisonnement en ergonomie cognitive. In: Science et Défense 91, Sécurité des systèmes, Neurosciences et ergonomie cognitive.

Amalberti, R., 1996. La conduite de systèmes à risques. PUF, Paris.

Bainbridge, L., 1987. Ironies of automation. In: Rasmussen, J., Duncan, K., Leplat, J. (Eds.), New Technology and Human Error. Wiley, Chichester, UK, pp. 271–283.

Bass, E.J., Small, R.L., Ernst-Fortin, S.T., 1997. Knowledge requirements and architecture for an intelligent monitoring aid that facilitate incremental knowledge base development. In: Potter, D., Matthews, M., Ali, M. (Eds.), Proceedings of the 10th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems. Gordon & Breach Science Publishers, Amsterdam, The Netherlands, pp. 63–68.

Baxter, G.D., Ritter, F.E., 1999. Towards a classification of state misinterpretation. In: Harris, D. (Ed.), Proceedings of the Second International Conference on Engineering, Psychology and Cognitive Ergonomics. Ashgate, Oxford, pp. 35–42.

Besnard, D., Greathead, D., 2002. A cognitive approach to safe violations. Technical report CS-TR-776, University of Newcastle.

Boy, G., 1987. Operator assistant systems. International Journal of Man–Machine Studies 27, 541–554.

[1] Visit DIRC at www.dirc.org.uk.

Brehmer, B., 1996. Man as a stabilizer of systems. From static snapshots of judgement processes to dynamic decision making. Thinking and Reasoning 2, 225–238.

Cellier, J.-M., Eyrolle, H., Mariné, C., 1997. Expertise in dynamic systems. Ergonomics 40, 28–50.

Crow, J., Javaux, D., Rushby, J., 2000. Models of mechanised methods that integrate human factors into automation design. International Conference on Human–Computer Interaction in Aeronautics: HCI-Aero 2000, Toulouse, France, pp. 163–168.

Dehais, F., Tessier, C., Chaudron, L., 2003. GHOST: experimenting conflicts countermeasures in the pilot's activity. Proceedings of the 18th Joint Conference on Artificial Intelligence, Acapulco, Mexico, pp. 163–168.

De Keyser, V., Woods, D.D., 1990. Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In: Colombo, A. G., Saiz de Bustamante, A. (Eds.). Systems Reliability Assessment. ECSC, EEC, EAEC, Brussels, Luxembourg, pp. 231–251.

Filgueiras, L.V.L., 1999. Human performance reliability in the design-for-usability life cycle for safety human computer interfaces. In: Felici, M., Kannoun, K., Pasquini, A. (Eds.), SAFECOMP'99. Springer, Heidelberg, pp. 79–88.

Green, R.G., Muir, H., James, M., Gradwell, D., Green, R.L., 1996. Human Factors for Pilots 2nd Edition. Aldershot, UK, Ashgate.

Hancock, P.A., Szalma, J.L., 2003. The future of neuroergonomics. Theoretical Issues in Ergonomics Science 4, 238–249.

Hollnagel, E., 1987. Information and reasoning in intelligent decision support systems. International Journal of Man–Machine Studies 27, 665–678.

Hollnagel, E., Woods, D.D., 1999. Cognitive systems engineering: new wine in new bottles. International Journal of Human–Computer Studies 51, 339–356.

Jennings, N.R., Wooldridge, M., 1998. Applications of intelligent agents. In: Jennings, N.R., Wooldridge, M. (Eds.), Agent technology: Foundations, Applications and Research. Springer, Berlin, pp. 3–28.

Klayman, J., Ha, Y.-W., 1989. Hypothesis testing in rule discovery: Strategy, structure and content. Journal of Experimental Psychology: Learning, Memory and Cognition 5, 596–604.

Leveson, N., Pinnel, L.D., Sandys, S.D., Koga, S., Reese, J.D., 1997. Analysing software specifications for mode confusion potential. In: Johnson, C.W. (Ed.), Proceedings of a Workshop on Human Error and System Development. Glasgow, Scotland, pp. 132–146.

Moray, N., 1987. Intelligent aids, mental models, and the theory of machines. International Journal of Man–Machine Studies 27, 619–629.

Moray, N., 1996. A taxonomy and theory of mental models. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, Vol. 1, pp. 164–168.

NTSB, 1990. Aircraft accident report. United Airlines Flight 232. Mc Donnell Douglas DC-10-10. Sioux Gateway airport. Sioux City, Iowa, July 19, 1989. National Transportation Safety Board, Washington DC, USA.

Ochanine, D., 1978. Le rôle des images opératives dans la régulation des activités de travail. Psychologie et Education 2, 63–72.

Parasuraman, R., 2003. Neuroergonomics: research and practice. Theoretical Issues in Ergonomics Science 4, 5–20.

Rabardel, P., 1995. Les hommes et les technologies. Armand Colin, Paris.

Rasmussen, J., 1986. Information processing and human-machine interaction. North-Holland, Elsevier Science, Amsterdam.

Rasmussen, J., 1991. Technologie de l'information et analyse de l'activité cognitive. In: Amalberti, R., de Montmollin, M., Theureau, J. (Eds.), Modèles en analyse du travail. Mardaga, Liège, pp. 49–73.

Rasmussen, J., 1993. Diagnostic reasoning in action. IEEE Transactions on Systems, Man and Cybernetics 23, 981–992.

Rushby, J., 2001. Modelling the human in human factors. Invited paper, Safecomp 2001, Budapest, Hungary, pp. 86–91.

Rushby, J., Crow, J., Palmer, E., 1999. An automated method to detect potential mode confusions. Proceedings of the 18th AIAA/IEEE Digital Avionics Systems Conference, St. Louis, MO, USA.

Sanderson, P.M., Murtagh, J.M., 1990. Predicting fault diagnosis performance: why are some bugs hard to find? IEEE Transactions on Systems, Man and Cybernetics 20, 274–283.

Sanderson, P.M., 1990. Knowledge acquisition and fault diagnosis: experiments with PLAULT. IEEE Transactions on Systems, Man and Cybernetics 20, 225–242.

Simon, H.A., 1957. Models of man. Wiley, New York.

Wageman, L., 1998. Analyse des représentations initiales liées aux interactions homme-automatique (IHA) en situation de contrôle simulée. Le Travail Humain 61, 129–151.

Yoon, W.C., Hammer, J.M., 1988. Deep-reasoning fault diagnosis: an aid and a model. IEEE Transactions on Systems, Man and Cybernetics 18, 659–676.