Human Error in Critical Systems

Francesca Madeddu
803623@swansea.ac.uk

31st January 2015

Abstract

To err is human. We all have heard or maybe even pronounced this sentence at least once in our life. However, when developing a system, engineers and developers often seem to ignore this fact: the human error is not contemplate in any design choice so that when an accident happens, the first logical, easy but unfair consequence is that the human is blamed. This attitude is dangerous because it does not allow to investigate the real causes which led to the accident, making impossible to acquire the knowledge that is essential to prevent the accident to happen again. Therefore the situation is even more dangerous in case of critical systems, because of the threat in terms of consequences as injuries or human lives losses.

The aim of this paper is presenting the role of the human error in critical systems, explaining how it can be analyzed, prevented and finally reduced.

Contents

"Errare humanum est,
perseverare autem diabolicum"

Lucius Annaeus Seneca
4 B.C. to 45 A.D.

1 Introduction

Nurses overdose six patients causing multiple injuries or deaths. Three pilots die returning from an aerospace mission because of a malfunction of a hatch. A woman, incarcerate for shooting her three children, escape from a prison unhurt. Those three stories, excellently telled in the book "Set phases on stun" [7] have something in common: in fact all of them end with injuries, deaths or dangerous events and they are all caused at least in part by the human error, or to its incapacity to manage emergency situations. The first story is about the already famous case of the Therach-25, a machine to cure cancer through radiotherapy: between the 1985 and the 1987 a series of six accidents happens, and several patients were overdosed with 100 times the normal dose: three of them died. The reason why this happened was that the *feedback* of the machine used for the treatment was not clear enough, so that there was no way for the nurse to understand what was going wrong. The second story tells the famous tragedy of the Sojuz 11, an aerospace Soviet mission started in the 1971 and ended with the death of the entire crew: Vladislav Volkov (36 years), Georgy Dobrovolsky (43 years), and Viktor Patsayev (38 years) lost their life during the way back home. The hatch for the pressurization regulation was damaged when the space-ship left the space station, allowing the oxygen to pour out from the space-ship before it reached the atmosphere. One component of the crew, Viktor, realized at some point what was happening, and he tried with an emergency procedure to close the hatch. However, because the asphyxia process was already started, he had not enough strength to complete the procedure. Apparently, even if the hazardous situation was expected, and an emergency procedure was provided, no consideration had been taken about the human limits in the context of the emergency. The third and last story is about a woman, Diane Downs, imprisoned in a medium security prison in Oregon for shooting her three kids. A new alarm system was installed since a couple weeks earlier, and because it was even to much sensible, it was always in an emergency state: birds and even a strong wind were in fact able to trigger the sensor. Knowing that the guards were less alert because of all the false alarms, Diana decided to trying the escape by climbing over the fence. She incredibly succeeded thanks to the fact that, as she predicted, the guards did not care about the alarm soon enough, thinking of the umpteenth false alarm. Who can blame them? Even Aesop in the *"The Shepherd Boy and The Wolf"* fable already knew the power of that effect. In all the mentioned stories the human behavior seems to apparently have some kind of responsibility in the sequence of events which led to the final accident. However blaming the human appears sometimes more as a shortcut than as a real effort to understand the real reasons why the accident happened. The role of the human within a system becomes incredibly important when we are dealing with systems which failure can lead to series consequences. Fortunately during the last years more and more importance has being dedicated to the human aspect, however a complete awareness has still to be developed and too many times we can see accident's analysis concluded and tagged as "human fault".

The aim of this essay is to give a panoramic of what has been done in this sense so far. It is structured as following. In the *Section 2*, after briefly explaining why the human is always blamed and why this approach is not useful, it is introduced the background of the problem, defining what a critical system is, and giving some data related the cost of human error in critical systems with a special attention to healthcare and aviation. In the *Section 3* it is discussed the difficult relationship between the human and the machine, and how the arrival of automation process is affecting the human performance. *Section 4* and *Section 5* present models of accident and error which are indispensable tools to understand and evaluate the role of the human error in accident and therefore to prevent it. In the *Section 6* there are illustrated different approaches to manage the human error. Finally, in the last section, the conclusion is reported.

2 Human and critical systems

2.1 Blaming the human

As shown by several statistics, within the most famous is definitively the one from Shappel where the human is blamed for the 80% of the total cases [25], the human being is often blame for accident. However there are several reasons to explain why this happens, and Leverson provides us with some of them which I will be briefly summarize in the current section [15]. Firstly when looking at statistics we should always use our critical sense and as first point we should ask ourself who makes those statistics. In fact, usually, they are proposed by the same people who have all the interested in blaming the human. Also, what happens sometimes is that the human is blamed when no other reason can be found or when the reason is too embarrassing for the company. Another interesting point is that when errors result in accidents, are always investigated, studied and reported. However when human actions results in the recovery from emergencies situation, no kind of record is usually taken. This leads to think that the human always behaves in a wrong way and that he always takes the bad decision event if it obviously is not true. For example, an U.S. Airforce study reported that the crew was able to recover from 681 emergency situations due to the equipment failure making only 10 mistakes, which would be a proportion of 1.5% of human errors versus the 91% of computer errors. Accidents, therefore, are always evaluated *a posteriori* so that the human behavior is evaluated once what happened it is made clear. However the hindsight is very powerful and situations that seem very obvious to manage, are not so easy to understand while the emergency is on going. Also the prospective usually used to look at the human error is often wrong: in fact rather than say the humans cause accident, we should ask why they were not able to prevent it. In fact sometimes the operator is expected to do actions that are mentally of fiscally impossible to perform, or he is requested to intervene at the limits, when consequences will be serious and a lot of stress is involved. In these case, the designer or the management should be blamed.

At this point it appears clear that a change prospective is the key for preventing and mitigate errors. In fact the worst part in blaming the human is not that it is not right or fair, but that following this kind of approach we cannot analyze accident in a proper way and if we cannot understand the error than we cannot even prevent it.

2.2 Human factors in critical systems

A critical system can be defined as a system which failure may cause series consequences. When the consequences can be defined in terms of “death or injury to people, harm to the environment or economical loss” the system is *safety-critical* [5]; when the consequences are in terms of business, then the system is *business critical*; finally, when the consequences are in terms of failure of a goal-directed activity, the system is *mission-critical*.

2.2.1 A software engineering approach

The software engineering approach to the management of critical system includes several aspects. The system is investigated in terms of *software*, *hardware* and *liveware,* namely the human being. In fact the operator is treated and investigated exactly as all the other parts of the systems. In particular the focus is on the *environment* where the *interaction* between the operator and the system - so the *HCI* - takes place, on *protocols*, on *training* and on *cultural habits.* The aim of this approach is to reduce the probability of failure [5]. The idea of including the human aspect in the equation is quite recent, and as mentioned it is related to the growing awareness that a deep knowledge of the human is crucial in order to achieve safety.

2.2.2 Safety-critical system areas

Safety-critical systems are present in healthcare, aerospace, chemical industry, nuclear power station, traffic control (air, railways, roads) and military equipment. In the following paragraphs are presented a series of interesting statistics about the impact of the human error in some of those areas.

Aviation Since the 1950, despite the high number of accidents in the early history of aviation when airplane were still mechanically unsafe, accidents are drastically decreased as shown in figure 1 [28]. In fact today is it considered safer to flight on a commercial airplane than driving on a busy city street.

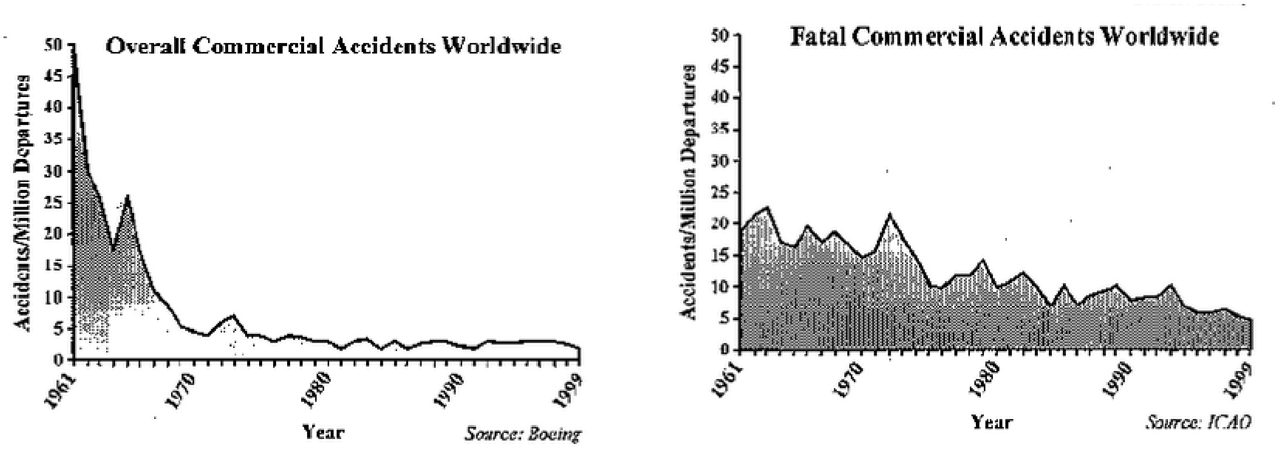


Figure 1: Overall (left) and fatal (right) commercial air carrier accidents worldwide 1961-99

However even if it is true that the number of accidents is decreased, during the last decades the decrease stopped. Therefore it seems that a further reduction is impossible: as Wiegmann and Shappel say "we have reached a point at which accidents may simply be the *cost of doing busines*s" [28]. The impossibility of reducing the number of accidents has consequences. For example in the military aviation, as shown in figure 2 [28], the cost of accident is sharply increasing. As remarked in the report of Shappel "while the aviation accident rate has declined tremendously since the first flights nearly a century ago, the cost of aviation accidents in both lives and dollars has steadily risen" [25].

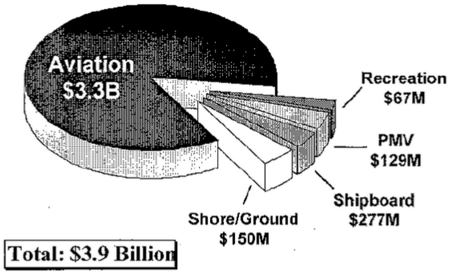


Figure 2: Monetary costs of accidents in the U.S. Navy/Marine Corps from fiscal year 1996 to 2000

The situation is critical also on the commercial side if we consider that, even if the rate is stable, the air traffic is expected to increase in the future. It is interesting to note that even if today the mechanical part is safer then in the past, the same effort has not been put in considering the human aspect.

Healthcare The healthcare is propably the sliest critical area, since it kills in a silent way through *complications* and *misadventures* and because it is not as spectacular as a plain crush, the attention over the problem has been underrated for a long time. Also many people involved in the healthcare system judge as acceptable loosing lives when “applying a imperfect science to sick people frequently in not ideal conditions [24]. However looking at the report made by Dhillon, shown in table 1, highlights the seriousness of the problem [8].

In a typical year around 100000 Americans die due to human errors. The financial impact of these errors on the US economy is estimated to be somewhere between $17 billion and $29 billion.
Operator errors account for more than 50% of all technical medical equipment problems.
A study of anesthetic incidents in operating rooms revealed that between 70% and 82% of the incidents were due to human errors.
Human error accounts for 60% of all medical device-related deaths or injuries reported through the Center for Devices and Radiological Health (CDRH) of the Food and Drug Administration (FDA).
In 1993, a total of 7391 people died due to medication errors in the United States alone.
The annual cost of medication errors is estimated to be over $7 billion in the United States.
In the emergency departments over 90% of the adverse events are considered preventable.
In the United States, the annual cost of hospital-based medication-related errors is estimated to be around $2 billion.
In 1984, a study examined the records of 2.7 million patients discharged from hospitals in New York and it was found that 25% of the 98609 patients who suffered from an adverse event was the result of negligence.

Table 1: Cost of the human error in healthcare

Perhaps even if a lot of literature about medical errors is available today, we still do not have a complete and objective picture of the situation. This happens because to interpreter the finding is hard due to the lack of a standardized nomenclature [14].

3 Humans and automated systems

Critical system today are complex systems, and a big part of them is automated. This fact has several consequences which will be analyzed in the following paragraphs.

3.1 History of automation

Starting from the industrial revolution in the 1760 the presence of machines tools in our word is constantly increased: it is true for industries, where automation has reached a pervasive presence, as for our everyday life, since today we still hardly do things manually. Automation made possible things that were unimaginable before. For example in the 1905 it was invented a glass bottle blowing machine which allowed two men working 12 hours shifts to produce 17,280 bottles in 24 hours against the 2,880 produced by six men in the same period of time [1]. Automation has deeply transformed the world where we are living in, allowing us to live in a culture so called "the good life": we have now better material goods, which can be delivered and shared more efficiently then ever in history so that we have greater choices and lower prices [2].

3.2 Human versus machine

At this point we would be tempt to ask: why is the human still needed? Why do not we leave in a completed automated world? The answer is that machines are perfectly fine in most of situations, they are more reliable, they execute orders without debate and they are very precise. However this straightforward behavior can represent a weakness in non standard situations: computers are not usually useful in case of emergency, at least not as much as humans because they can manage the situation in a better way thanks to their flexibility. Therefore, despite Groover defines automation as "the technology by which a process or procedure is accomplished without human assistance" [9] automation still needs the human presence.

3.2.1 Machine strength and weakness

In the book "Safeware" Leverson gives a list of seven myths related to the supremacy of software over mechanical systems [15]: she explains why we should not believe that the digital world is better than the analogical one. The comparison can be easily enlarged to the manual versus automatic systems. For example advantages usually related to automation include a higher production throughput, improvements in the goods quality, a reduction of the required labor and finally an enhancement in safety. In fact often the use automatic systems allows to remove workers from hazardous situations [9]. However as Leverson points out, even if workers can be move away from hazardous situations, it means that they loose familiarity and experience with those hazard situation, so that they will not be able to manage and understand the hazard properly exposing themselves to a even higher hazard. Even if it cannot be said that machines are always better than humans or the other way around in a general way, it is definitively true that each one fits better that the other given a specific context. Computers and automated systems are excellent for calculations and straightforward tasks. However because their behavior its pragmatically determined by some kind of algorithm produced *a priori*, it is not possible to cover all possible situations since not all conditions are foreseeable [15]. In fact, one of the most important problem when looking for safety in critical system, is that all possible cases cannot be completely covered or investigated because of the complexity which increase very sharply [10].

3.2.2 Human strength and weakness

Humans are not as good as computers in calculations, nor are straightforward. They often do not follow rules, by choice or by mistakes, even when that rules are simple and clear. However they have the incomparable advantages to be adaptable and flexible. Humans, unlike machines, are also capable to develop skills and performance patters through experience, and they can use problem solving and creativity to cope unusual situations [15]. It is the same coin, but different faces: the weakness of the human being, its incapacity to be perfect, is its strength when non standards situation and emergencies happen. In other words the fact he can make mistakes is the price we have to pay for its flexibility.

3.3 The role of human in automated systems

As mentioned before, during the last decades there has been a shift on the operator's role. At the beginning the operator used to make all decisions while today he usually only cooperates with the machine, but the tendency is to push versus automation even more. However, as seen, because of the lack in terms of flexibility of the computer, it is not desirable to exclude completely the operator's presence.

3.3.1 The impact of automation on the operator's job

The automation in theory should make the human job easier, unfortunately it is not true. In fact it has been proved that the introduction of automation made the operator's jobs more complex hence more prone to accident [15]. The main point is that automation only reduce the quantity of the job but does not affect on the improvement of the quality: because the job could be boring and repetitive the operator's attention decrease. Another point is that the operator is not involve anymore in the process, so he cannot develop the proper mental model, and he is less effective when actually needed because he does not have the complete vision of the system and a deeply knowledge about how it works.

3.3.2 Operator's role

According to Leverson the operator can cope different roles, as he can be monitor, backup or partner, basing on the level of cooperation which is required [15]. For each role there are latent issues which need to be fixed.

Monitor The operator acts as a monitor of the system and he does not execute all the main operations anymore. The main problem is it has been proved that automation usually makes this role unsuitably for humans for several reasons. Firstly the only information the operator knows about the system is the one the system shows: it means that if the design is not clear or if the information is not properly displayed or simply it is not enough, the operator could be enable to make the correct choice. This is particularly true during emergencies: in fact even if the information is usually displayed in a proper way during the correct operating of the system, things can quickly changes when something goes wrong. A typical example is the case when the operator is overwhelmed by the high number of alarms. Secondly, the distance between the operator and the process (the operator cannot hear, touch, experience the system as he was used to) does not allow him to create a complete mental model, nor to update it efficiently. As Brehmer says “the connection between the actions and outcomes is opaque, indirect, abstract” [6]. Thirdly, system failures could not be immediately visible to the operator, and it could stay silent for a while: this happens because automatic system are developed to cope and manage errors during early stages [4]. Finally, because the job required by the operator is reduced by the automation, it can result in a lower level of attention and vigilance which can lead to a too high confidence over the system. It is known in fact that it is very difficult for the operator to keep the attention alive for more than a certain period time if nothing happens [19]. Because of all the mentioned reasons the operator's performance can be seriously compromised in case of emergency.

Backup The operator can be relegate to a backup role, meaning he has to be available only to cope emergency situations. Several aspect can make the operator job harder. Firstly, if the system interface is not properly designed, the operator will not be able to understand what is going on and he will not manage the emergency. This could happen for example because the system is not transparent enough. To give an idea about how much this aspect is important, we can think that operators usually arrive in the control room from 15 to 20 minutes before the shift starts only to "get the feeling" about the current situation [15]. This is also known as *situational awareness*. Secondly, because as seen the operator is taken away from the heart of the process, he has a poor mental model of the system. Moreover, because he is called to intervene only in emergency situations, he usually has got a little experience and he is also prone to easily forget even the little experience and knowledge that he has. It is in fact proved that the time to access the long term memory is strictly related to the frequency it is accessed [3]. Finally, because designers often develop the system in order to manage emergency situation, they do not pay attention in the developing of system in the case the system will not actually capable to cover the emergency. As happens when the operator is used as a monitor, because of all the mentioned reasons, the operator can be not able to manage the emergency situation in a efficient way.

Partner Finally the operator and the system can be partner. In this case the most critical aspect is that there has to be made very clear which tasks are managed by the system and which one are on the responsibility of the operator. If it is not clear enough, the operator could ignore a task thinking that the system is taking care of it while it is not, leading to possible errors. Another problem which can raise when operator and computer are partner is that usually the computer takes the easiest part, leaving the operator with the hardest one [3]. Ironically, even if the operator job decrease, the risk increase.

3.4 Which conclusion?

We have seen both advantages and disadvantages of the human presence in automatic systems and it seems clear that, despite all the problematic situations the operator can face, his presence is still required: human capabilities are unique and at least until today they are not replaceable. However, because operators are subject to make mistakes, and because automation can make operators even more prone to error, system needs to be developed in order to manage, prevent errors or at lest to reduce consequences. In particular as Leverson says "the goal of the human machine interface design should be to preserve the human capability to intervene positively while making harmful intervention as difficult as possible (...) The design of the interface should be based not on functionalities but on operations, operators should be included in the designing process too" [15]. More important, the right approach when talking about human error should be that the human does not cause the accident but he is simply unable to prevent it.

4 Causation accident models in human factors

To understand the role of the human error in accidents, we need first to define what an accident is and what causation accident models are related to human factors. Leverson defines the accident as “an undesired and unplanned (but not necessarily unexpected) event that result in (at least) a specified level of loss” [15]. Models are important as they give us the tools to investigate, understand and talk about complex phenomena: models are essential to understand past accidents, and derive some knowledge that can be exploited to prevent the new ones.

4.1 Models classification

During the last century the investigation on the cause of accidents have moved from the hardware or equipment failures to a higher level, including all aspects within the organizational system. In fact as Rafferty points out “is it now widely accepted the accidents which occurs in complex technical systems are caused by a range of interacting human and systemic factors” [20]. According to the Rafferty classification there are three kinds of causation accident model: *sequential*, *epidemiological* and *systemic*. The sequential model, within the Henrich’s domino is probably the most famous, view accidents as the result of a series of events where the last one itself is the accident. The epidemiological model, within the most famous is probably the Reason’s one, treats accidents like the spreading of disease [21] and “describes the combination of latent conditions present in the system for some time and their role in unsafe acts made by operators at the so-called ‘sharp end”’ [20]. Finally, in the systemic model accidents are seen like an emergent property of the overall technical system [20].

4.2 The Domino model

Heinrich was the fist who suggested since the 1931 a real accident model focused on unsafe acts rather than unsafe conditions; it comes from the industrial safety and reflects the need to protect workers from hazardous situations [15]. Heinrich defined a list of five transitions, shown in figure 3, which starting from the analysis of the social environment lead to an injury.

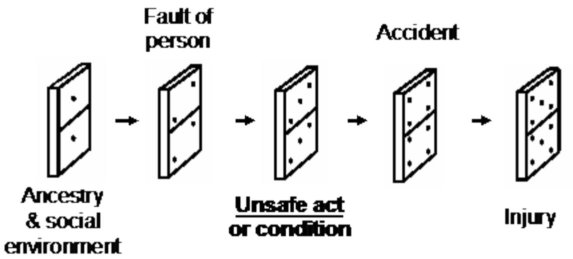


Figure 3: In the Domino model there are five transitions: (1) an ancestry or social environment, leading to (2) a fault of a person, which is the proximate reason for (3) an unsafe act or condition (mechanical or physical hazard) which results in (4) an accident, which leads to (5) an injury.

The idea of the domino if very effective since suggests that each domino’s fall is caused be its ancestry and falls to its son until the chain of events is finished, where the last event is the injury itself. According to Henrich if we want to improve the safety we have to act on the third level (unsafe act or condition).

4.3 The Swiss cheese model

The Reason's model, rather than focusing on the so called *sharp end* of the system, investigates on the interaction between the unsafe acts made by operators and the *latent conditions* of the systems which allow the operator to act unsafely. The unsafe actions are not just errors, they are errors performed in a hazardous situation. The latent conditions which allow the unsafe actions to perpetuate go along all levels of the organizational system, from the management to the operator and they can be from a unclear interface design to poor procedures, from an inadequate training to an equipment failure or unclear management directions [20].

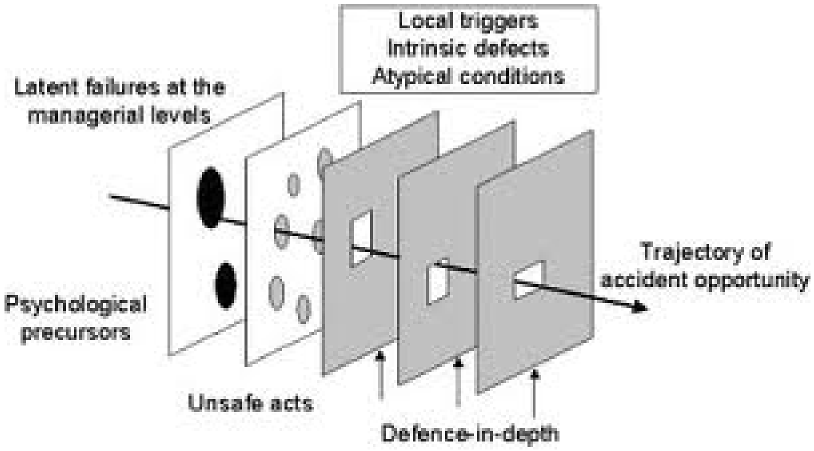


Figure 4: The dynamics of accident causation according to Reason's model

As for the Domino's model, also the Swiss Cheese image is very suggestive and appropriate to express the meaning of the model. Each level of the organizational system is a slice of cheese which can stop the series of action that lead to the accident: however each level has to be properly developed in order to accomplish this function. If it does not happen, the latent conditions (the holes in the cheese) allow the propagation of the error which finally arrive to its accomplish.

4.4 The Rasmussen's model

The Rasmussen's model is also very popular and it inspired the development of several accident analysis methods (such as the famous AcciMap). Rasmussen's model is based on the idea that "accidents are shaped by the activities of people who can either trigger accidental flows or divert normal work flows" [20]. As happens with the Reason's model, the Rasmussen's model instigates the causes of the accident along different organization levels, from operators to the management and the unsafe human acts are evaluated within this hierarchy. According to the model the safety of the system "is viewed as an emergent property arising from the interactions between actors at each of the levels" [20], where each actor has a specific responsibility in terms of causing and preventing hazardous situations. A critical aspect, according to this model, is the way those safety actions are propagate between all levels, to reach the bottom from the top or the other way around. This aspect is named as *vertical integration*. Again, similar to the Reason's model, where there were latent conditions, Rasmussen talk about accidents as they are *waiting for release*, and they are triggered by behaviors which deviate from standard.

5 Human error models in human factors

As for accidents, there are different error models, because no one by itself is capable to capture all salient aspects, so it is needed to use different models in relation to the context. The simplest way to model an error is treating it as an *external manifestation*, as explained in table 2. However this classification is too trivial and it does not give any interesting information to deeply understand the human error and therefore to prevent it.

Error	Description
Omission	not performing a function.
Commission	performing a function that should not have been performed.
Ignorance	not recognize an hazardous situation.
Wrong action	responding with the wrong action to a problem.
Inadequate action	responding with an inadequate action in a situation.
Poor timing	responding too early or to late.

Table 2: Classification of error as external manifestation

Therefore there are several more complex classifications, which will be presented in the following subsections: they can be based on tasks and environment, on cognitive aspects or on sociological aspects.

5.1 Task and environment models

Often in the human factor theory the human activity is described as a task. In this sense a task can be categorized as *simple*, *vigilance*, *emergency* or *complex* [15] so that the error is seen in relation to a specific kind of task.

Task Type	Description	Affecting factors
Simple task	simple task with little decision making	psychological stress, quality of engineering control and displays, quality of training and practice, presence and quality of written instructions, personnel redundancy
Vigilance task	regards the detection of signals	sensory modality, nature, strength and frequency of signal, unexpectedness, length of watch, motivation
Emergency response task	regards the behavior requested to respond to an emergency situation	history of the emergencies situation, quality of engineering control and displays
Complex task	defined sequences of operations which requires decision making	

Table 3: Classification of tasks according to the general behavior required

The most important limitation of these model is that it is mainly focused on the concept of task and on the environment, but it does not accurately take account of the human characteristics. Better model to describes those aspects are the cognitive mechanism models.

5.2 Cognitive mechanism models

Within the cognitive models two of the most important one are definitively Norman's model of slips and the Rasmussen's model.

5.2.1 Norman's model

Norman classifies errors in two main categories: "an error in the intention is called *mistake* while an error in carrying out the intention is called *slip*" [17] where the intention is defined as "the highest level specification of a desired action" [17]. Norman also classifies errors by causes defining them as *mode*, *description*, *capture* and *activation* errors.

Mode errors occurs when the operator performs an action that would be correct in a certain state but it is wrong in the actual one. This kind of errors happen because the state of the system is not clear, so they are strictly related to the design of the interface and to the feedback. An example of mode error which happened to me recently was at the food vending machine at the university: after having inserted the required amount of coins I pressed the button corresponding to the desired snack. After trying three times, each one resulted in a failure, I realized the machine did not have accepted some of the coins with the result that in the current state (more money were needed) the action of selecting a snack had no meaning.

Description errors are strictly related to mode errors and they are connected the presence of indicator in the system. Indicators are a way used by the designer to make the state of the system evident, however if they are not properly designed they can overwhelm of confuse the operator leading him to perform the wrong action. Description errors are in fact quit common when indicators or controls are too similar to each other. Description errors according to Norman occur when "there is insufficient specification of an action and the resultant ambiguity leads to an erroneous act, (...) which is usually closely related to the desired one" [17]. Description errors can be some time amusing. An example of description error happened to me lately while I was in the university cafeteria. I was going to drink my coffee, so I opened the sugar bag but, instead of pouring the sugar in the cup, I put the bag instead.

Capture errors occur when there is an overlapping between different actions which start in the same way and one is more frequent than others. A typical example given by Norman would be the use of the VI text editor where the user is used to write the sequence of command "`wq`" to write and close the document. Then it could happen that the user gives the command "`wq`" when in fact he only wants to write, meaning give the command "`w`" just because he is used to perform that sequence of action.

Activation errors are of two types and they happen when "inappropriate actions get performed and appropriate actions fail to get done" [17]. The first type usually occurs when the wrong sequence of actions is activate while the second type usually occurs when the operator cannot accomplish his goal because of the memory failure.

5.2.2 Rasmussen's model

Rasmussen et al. in the 1981 defined the *skill-, rule-, knowledge-based model* which describe errors in terms of *human task mismatch*. The model explains "*what* error occurred (external mode of malfunction), *how* it occurred (internal mode of malfunction) and *why* it occurred" [12]. Within the model the error is seen as a normal step in the learning process and its classification takes account of the mental process in which the error happens, in particular an error can be [27]:

skill-based happens when the operator is in a *autopilot* mode, meaning when the actions are so well known that he performs them with his unconscious. This happens sometimes to me when I am driving or walking home and I suddenly find myself home without even realizing it;

rule-based happens when the operator is requested to perform a series of actions according to certain rules and he makes the wrong action or does not perform any action at all;

knowledge-based happens when the operator knowledge or experience is not developed enough to face the situation. This could happen because of a poor training, or because of a physical or mental limitation of the operator.

5.2.3 Reason model

Reason proposed his own classification of error primarily based on the nature of the error itself, classifying it as *slip* (or lapse), *mistake* and *violation* [27].

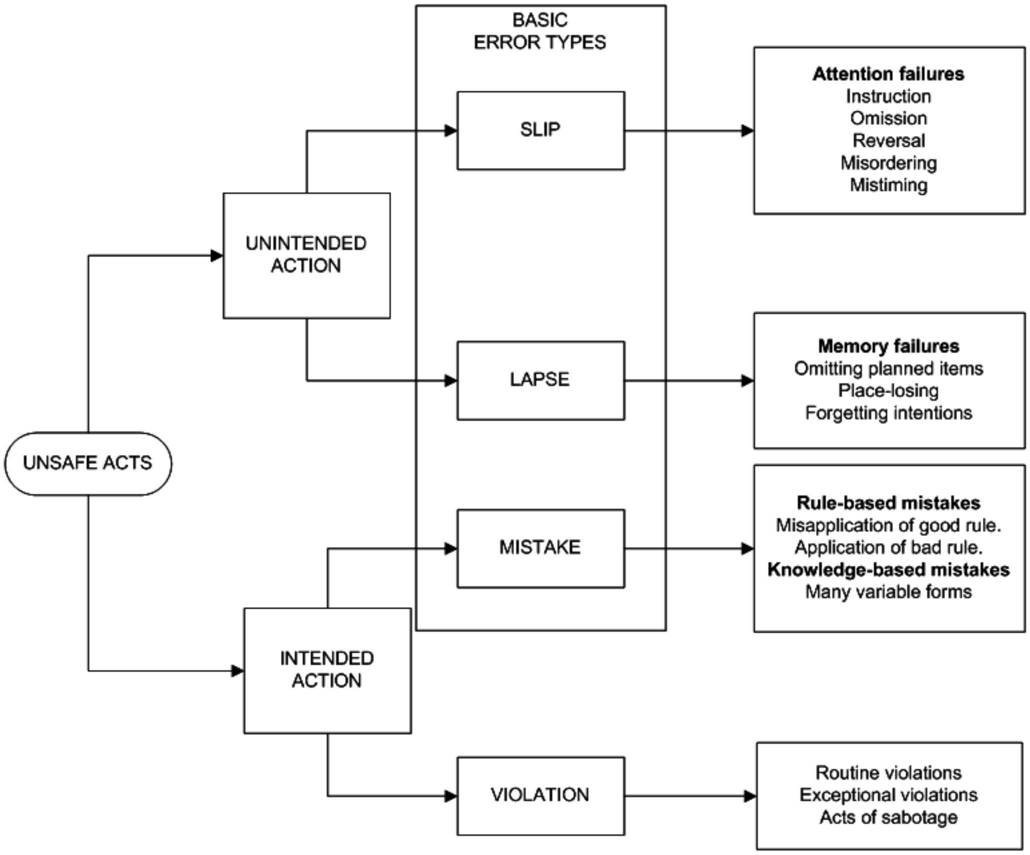
Slip or lapse occurs when the operator starts with the right intention but he performs the wrong action. Slips usually happen when our attention is focused on something else or we are distracted by something. The distraction could be internal, such as some other thought that is holding our attention, or external, like a sudden noise or a visual event. An example of slip which happens to me time to time is inverting the 'o' letter with the 0 number while typing.

Mistake occurs when the operator perform the wrong action but he is sure it is the right one.

Violation occurs when the operator voluntary decides to perform the wrong action, because he believes that the suggested action is wrong. Violations do not depends on a low level of attention, laziness or ignorance. A typical example of violation is the *reverse to the stereotype*: when the human being is overwhelmed by too much information, or he is exposed to too much stress, he is incline to ignore the given rules if they do not match with what he considers as standard. For example, in case of emergency it could be requested to an operator to decrease a value of a system variable by shifting up the handle of a control: however if he considers as standard shifting down a handle to decrease a value, he will probably prefer the standard action to the instructed one. Violations also occurs because the operator does not deeply understand the meaning behind a suggested rule so that he thinks he is justified in some way in ignoring the given rule. An example of this behavior happened in my university: the access to the computer lab is forbidden after 8pm, and clear signs are present on the door to warn students to respect the rules. However, because the reason, which would be a safety reason, is not really understood and shared by students, they keep staying in the lab even after 8pm. In fact, because they do not understand the reason, they cannot realize the importance of the consequences or their actions.

5.2.4 Relation between models

As seen each model emphasize on a particular aspect of the human behavior. Names can be sometimes deceiving since the same name is used by different models with different purposes and meanings. However a interesting summarizing picture comes from Reason.



,

Figure 5: Reason's summary of the psychological varieties of unsafe acts

Unsafe acts can lead to error through unintended or intended actions. In the first case we find what Reason defines as basic errors, which are the same Norman's errors. In fact in his model Norman does not consider violations or when mistakes are performed on purpose. Reason's lapses are what Norman defines as description errors.

5.3 Sociological model

Finally it would be naive to think that cognitive models of error are capable to describe all aspects of the human behavior: as Leverson says "error cannot be completely understood in isolation from the environment in which it occurs" [15]. Even if giving a sociological perspective of the error goes beyond the purpose of this essay, the reader who is interested in understanding the error deeply has to consider this aspect too.

5.4 Error probability

When talking about error, we may have the need to calculate the probability for a certain error to happen. However to define the error probability some limitations have to be considered [13]. Firstly, making prevision of errors is basically impossible: when looking at accidents after they happened they often seem easy to prevent, but this depend on the power of hindsight. However it seems reasonably to make prevision assuming that the error rate remains constant. Secondly, numbers are usually given by management or by people whose interest is to keep those numbers low, so there could be a conflict of interests. Finally, even if it is true that we can use the error rate to make assumptions, the probability is strongly connected to the stress level or other external conditions that could change in the course of time. The main reason why calculating the probability of the human error is needed is to evaluate risk. All risk assessment should in fact investigate not only the system causes, but it should account for the human presence too. This concept can be summarized in a simple formula:

$P_{tot} = P_{hw} + P_{sw} + P_{lw}$ where

P_{tot} is the total probability of the error to happen;

P_{hw} is the probability of hardware failure;

P_{sw} is the probability of software failure;

P_{lw} is the probability of human failure.

5.4.1 THERP

The THERP technique, which stays for technique for *Human Error Rate Prediction* is probably the best known and most widely techniques used to calculate the human error probability [23]. According to this technique the operator is treated exactly in the same way as he was a component of the machine so as a valve or a pump. The operator task is divided in atomic operations and each operation is evaluated singularly taking account of "the likelihood of detection, the probability of recovery, the consequence of error and a series of performance shaping factors such as temperature, hours worked, complexity, tools availability etc..." [13]. The THERP advantage is that it allow to consider each component separately, however it requires a lot of time and expertise effort.

5.4.2 TESEO

Another interesting techniques comes form Bello and Comumbori whose presented the *Technica Empirica Stima Errori Operati* where the probability of an error depends on the product of five factors illustrated in the following table 4 [23].

Factor	Description	Explanation	Probability
$k1$	type of activity	routine/not routine requiring close attention or not	$[0.001, 0.1]$
$k2$	a temporary stress factor for routine activities	assigned according to the time available	$[10, 0.5]$
	a temporary stress factor for non routine activities	assigned according to the time available	$[10, 0.1]$
$k3$	operator qualities	assigned according to selection, expertise and training	$[0.001, 0.1]$
$k4$	an activity anxiety factor	dependent on the situation: grave emergency, potential emergency or normal condition	$[3, 1]$
$k5$	an activity ergonomic factor	according to the quality of the micro-climate and plant interface	$[0.7, 10]$

Table 4: Factors of the TESEO technique

Kletz gives the following example : "Suppose a tank is filled once a day and the operator watches the level and closes a valve when it is full. The operation is a very simple one, with little to distract the operator while he is out on the plant giving the job his full attention" [13]. Then the following assignment could be considered:

$k1 = 0.001$, $k2 = 0.05$, $k3 = k4 = k5 = 1$ then

$k_{tot} = 0.005$

meaning the operator will fail 1 time every 2000.

6 How to prevent or minimize the human error

Managing the error in a critical system means mainly taking care of two aspects: *preventing the error* and *minimizing the consequences*. A typical example which implements this strategy would be: make dangerous action difficult to perform and in the case the action is performed at least give the operator a chance to undo that action. In the following paragraph will be presented some different approaches to the error managing which allow to implement these strategies.

6.1 Designing the human machine interface

The human machine interface, or HMI, works a crucial role in the human error management: in fact as seen most of errors the human does are due to his distance from the physical interface, resulting in the operator incapacity to get a proper mental model because not properly involved in the process, or are related to the poor implementation of the interface design. When developing a HMI for a critical system also the hazard analysis must be considered during all the developing process. In a standard process usually the usability engineers start with a task analysis and then they use prototypes and simulations to evaluate the system. However when dealing with critical systems a system hazard analysis should be used too: results retrieved from this analysis should be exploited to the design of the HMI itself. According to Leverson [15] there are four main strategies we can implement to obtain a better HMI: matching task to human characteristics, providing an appropriate feedback and giving a good training or clear instructions. In the next sections each strategy will be briefly explained.

6.1.1 Matching task to human characteristics

As Kletz suggests the design engineer should not try to change the human nature or persuade people to do not make errors, but he should accept men as they are and simply try to remove the possibility for them to make an error [13]. A possible approach to this philosophy consists in matching the task to the human characteristic instead of the other way around. This purpose can be achieved in several ways

Alertness As Leverson says "routing tasks tend to degenerate" [15]. If the operator has to perform the same activity and no mental challenge is involved, the risk is that his level of awareness and attention will drastically decrease. Therefore, as seen, automation seems to make the situation worse. To solve this problem the operator should not be assigned with boring tasks which require a low level of involvement and reasoning, such as the passive monitoring. A good way to avoid that it would be to engage the operator in challenging activities like production planning and management, quality control and system development. Another way is to develop the system in order to give the operator a certain level of decision making: this is possible by allowing to achieve the same goal in different way, so that the operator can solve problems and keep his mind active [11].

Designing for error tolerance As seen the error is a normal act of the human behaviors, and it can be considered as a stage of the learning process. The operator can in fact use experimentation to understand the model of the system and to keep it updated. In order to do that a proper feedback is essential. As Leverson highlight we say that an error occur only when human action are irreversible or non observable [15]. Besides providing a proper feedback, to be error tolerant, a system needs to offer a way to recover from the error: it can be achieved by designing the interface in such a way that the operator operator goal cab be splitted in several clear atomic steps where each step can be singularly reverted.

Task allocations When talking about tasks one of the first question which arise is how they should be allocated between the human and the computer. In fact when designing the HMI different styles of communication can be chosen, where each style reflects the role of the human within the system environment. Sheridan suggests the range of possibilities shown in the table 5 [26].

	Level of computer/human decision
1	Human does everything.
2	Computer tells human the option available.
3	Computer tells human the option available and suggest one.
4	Computer suggest an action and implements it if asked.
5	Computer suggest an action, inform human and implements if not stopped in time.
6	Computer selects and implements action if not stopped in time and then tells human if asked.
7	Computer selects and implements action and tells human if asked.
8	Computer selects and implements action and tells human if designer decides human should be told.
9	Computer selects and implements action without any human involvement.

Table 5: The Sheridan classification of a possible interaction between the human and the computer, where in the first level the human is completely in control while in the last one the computer is.

The simplest way would appear choosing for an automatic system, however as it has been explained along the last sections, automation always comes with a price in term of human performance and tendency to make errors, so when deciding for a computer-driven style of communication the consequences need to be carefully evaluated.

Considering the environment The way the operator acts is deeply influenced by the environment in which he is operating. For example, elements which usually lead to the error are

[15]: normal ranges not uniformly marked, different dials measuring the same quantity but with different scale (very common in medical devices), location of critical decimal points unclear o not following standards, labels and colors not meaningful on inconsistent. All those design errors must be avoided or fixed, and the operator should perform his actions in a clear and helpful environment.

6.1.2 Reducing safety-critical human errors

Safeguards, both procedural and actually implemented in the system, and and physical interlocks are a good way to prevent the human error [15]. However one of the most effective way to prevent error is to design the interface so that the error is physical impossible to be made. I have personally experienced the power of this principle in the launderette where I use to go: the washing machine payment interface is designed in such a way that it is almost impossible to insert the wrong coin. It is achieved by using constraints in fact the slots to insert coins have the same size of the required coin.

6.1.3 Presenting crucial information

What and how to present information to the user is one of the most important aspect in the design of the HMI. When displaying the information is very important to keep in mind all the cognitive aspect of the human being, in particular his limits. One of the most common mistake is to give too much information. Perhaps it usually happens in emergency situation when the operator should be instead provided with only the important one.

Alarms Alarms are used by automated system to detect an anomalous condition or event and to allow the operator to acquire awareness about that event [15]. The design of the alarm interface is usually ineffective if not dangerous for several reasons [28]: firstly alarms are usually designed not taking account that they will be used in an emergency situation, meaning the operator will be expose to stress and to other important conditions (act fast, serious consequences of each action, etc..); secondly, operators are not usually involved in the development process and this is a shame because they can really provide useful information; thirdly, as already mentioned, operators are usually overwhelmed with too much information. As cited by Mattiasson a study showed that in normal situation an operator does 3.1 actions for hour but the number increase to 53.3 per hour in case of emergency [16]. He also reported the example of a emergencies situation in a compressor trip to give the idea of what kind of information overwhelming the operator is expose to: the entire emergency last for 1.5, there were 392 alarms involved of which 254 only in the first hour; one alarm was triggered 118 times and the operator had to do 79 actions to recover from the emergency where the theoretical number of needed actions could have been 39. The HMI should be develop in order not to overwhelm the operator in such a way, especially considering that when the human being is overwhelmed it tends to not react at all, even if it is clear that the situation is very dangerous and than some action must be performed.

Feedback According to Norman [18] the feedback is needed by the operator to construct the proper mental model of the system and to keep it updated. Perhaps several operators highlighted that to monitor a system or to make a decision making job the same amount of information about the system is required. However often the automation hide those information. Leverson distinguished between thee kinds of feedback which are needed by the operator [15]. The *feedback about actions* is needed by the operator to understand if his action has been performed and what is the result. The *feedback to update the mental model* is the information the system provides about how the status is changing, so that the operator can update his mental model. Finally, the *feedback to detect faults* is needed to be aware of malfunctions in the system. Some designer think that using the operator in the same way as an alarm, meaning that he reacts only after the emergency has happened, is fundamentally wrong. They in fact think that the operator should be able to detect the problem in advance so that he can prevent and avoid the emergency, but this is possible only if the system gives to the operator the proper feedback.

6.1.4 Training and maintaining skills

As Keltz often highlight in his work, persuading the human or scolding him is not very useful because, with the exception of some cases such as for violations, the human does not make mistakes on purpose. So there are basically two way to prevent the error: *passive* and *direct*. The first one implies to act on the tools he uses, both software or hardware. The the second one is thought training and maintain skills: the theory of the control system and the design model must be explained deeply if the operator is expected to detected anomalies. According to Leverson there are three ways to do it [15]. By teaching the operators about safety features: the operators has to understand how to react to anomalies and which features are available in the system to prevent hazardous; by training for emergencies: operators need to have a deep understand of the design process and they have to be involved in the process so that they can use they creativity with proficiency when unexpected events happens; finally by simulators, even if someone thinks simulators are not very useful because they cannot simulate the stress which the operator is exposed to in a real emergency situation.

6.2 Solution based on the type of error

Kletz suggests that a proper prevention of errors is possible, in fact according to his idea for each typology of error it is possible to prevent or mitigate the error consequences through two main approaches: a better training or instructions or through changing the hardware or the software (CHAOS) [13]. The main idea of Kletz is that whenever the error occurs, it is not helpful at all spending time blaming the operator, since for most of errors it would not lead to any improvement. In case of errors due *slips* or momentary *lapse* of attention it is not of any help the exhortation or the punishment of the operator because the error is completely unintentional: a changing in the work situation is needed instead and the hardware or the software need to be changed. When *mismatches* occurs, as they are defined as the kind of error caused by the inability of the operator of perform a specific action since it requires something that is beyond his physical or mental possibility, the only solution in changing the hardware or the software. *Violations* are the only case in which persuasion can have a sense, since the operator choose deliberately not to follow the rules. However it would be more proficient to ask why the operator is not following the rules: it would lead, in most of cases, to the conclusion that retraining the operator is definitively a good choice. Helping him to understand the reason for the rules it will probably convince him to follow them. Finally, in case of *mist*akes, defined as errors due to poor training or instructions, it is needed to retrain the operator or to clarify the instructions. These errors are particularly sly because the operator thinks he knows but he doesn't. As mark twain said “what gets us into trouble is not what we don't know. It's what we know for sure that jut ain't so”. Kletz suggestions to preventing error are summarized in the table 6 shown below:

Error type	Description	Action required
Mistakes	Does not know what to do	Better training or instructions CHAOS
Violations	Decides not to do it	Persuasion / CHAOS
Mismatches	Unable to do it	CHAOS
Slips and lapses of attention		CHAOS

Table 6: Types of human error and the action required according to Kletz

7 Conclusion

Through the reading of the essay we have explored all the issues related to the presence of the human in critical systems: we have explained why the human is often blamed, why this attitude is wrong but nevertheless why the human presence is still fundamental. Accepting the human within the system, however, comes with a price. As Reasons said "systematic error and correct performance are two sides of the same coin" [22]. Because of this reason it is crucial to understand and prevent the human error. Fortunately, during the last years, the attitude towards the human error has changed and finally when accidents happen, the human is not always seen as someone to blame of everything and deeper investigations for back-tracking causes are performed. To do that, several models of accidents causation and of errors have been developed during the years. It is very important to know all of them, because each model takes over a specific aspect of the accident or of the error: only analyzing accidents and errors from all points of view we can have a clear and complete vision of what happened. Accidents are more than a software or an hardware failure and human errors are more than a misunderstood rule or an unclear management strategy. Human error cannot be just accepted and human cannot be just blamed. Moreover we have explored the difficult relationship between the human and the automated system, explaining why automation is not always a good thing, especially when we think about the implications and the consequences on the human behavior. Therefore, also the automation comes with a price.

The problem of the human error, as seen, is particular serious in case of critical systems. This happens because of the serious consequences in case of accident, in fact safety in these cases is crucial to prevent injuries and losses of lives. We have seen that, even if the technology, at least from a mechanical point of view, has reach a point where any improvement is almost impossible, there is still a lost of space for the improvement in the management of the human error. As we have seen, for example, in the case of aviation, even if the rate of accident has been constant during the last years, some precautionary measures have to be taken considering that the air traffic is going to grow. Awareness is another key concept, and it is definitively true in other critical system areas, such as the healthcare, where the problem and the impact of the human error in accident has remained silent for too long.

Maybe one day we will live in a world where computers are intelligent, so that we will reach a complete level of automation, and the human presence will not be required anymore. Until that day, until the day the human and the computer will not need to cooperate anymore, we have to accept the human error, where accepting does not imply a passive attitude towards the problem but it means we have to study and to apply methods to prevent the error and most important increase the awareness about its importance. Every system developed based on the assumption that the human behavior is error free is in fact destined to fail.

References

[1] Designates the Owens "AR" Bottle Machine as an International Historic Engineering Landmark. 1983.

[2] Majd Alwan and Mingjun Zhang. *Systems Engineering Approach to Medical Automation.* Artech House, 2008.

[3] Lisanne Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, 1983.

[4] William Bolton. *Control systems.* Newnes, 2002.

[5] Marco Bozzano and Adolfo Villafiorita. *Design and safety assessment of critical systems.* CRC Press, 2010.

[6] Berndt Brehmer. Development of mental models for decision in technological systems. *J. Rasmussen, K. Duncan & J. Leplat.(Eds). New technology and human error*, 1987.

[7] Steven Casey. *Set Phasers On Stun And Other True Tales Of Design, Technology, And Human Error: And Other True Tales Of Design, Technolo.* Aegean Pub Co, 1993.

[8] Balbir S Dhillon. *Human reliability and error in medical system*, volume 2. World Scientific, 2003.

[9] Mikell P Groover. *Automation, production systems, and computer-integrated manufacturing.* Prentice Hall Press, 2007.

[10] Dirk Helbing. *Managing Complexity: Insights, Concepts, Applications.* Springer Publishing Company, Incorporated, 1st edition, 2007.

[11] Gunnar Johannsen. *Analysis, design and evaluation of man-machine systems: proceedings of the IFAC/IFIP/IFORS/IEA conference, Baden-Baden, Federal Republic of Germany, 27-29 September 1982.* Pergamon, 1983.

[12] Barry Kirwan. *A guide to practical human reliability assessment.* CRC Press, 1994.

[13] Trevor A Kletz. *An engineer's view of human error.* IChemE, 2001.

[14] Linda T Kohn, Janet M Corrigan, Molla S Donaldson, et al. *To Err Is Human:: Building a Safer Health System*, volume 627. National Academies Press, 2000.

[15] Nancy G Leveson and Jorge Diaz-Herrera. *Safeware: system safety and computers*, volume 680. Addison-Wesley Reading, 1995.

[16] C Mattiasson. The alarm system from the operator's perspective. In *Human Interfaces in Control Rooms, Cockpits and Command Centres, 1999. International Conference on*, pages 217–221. IET, 1999.

[17] Donald A Norman. Design rules based on analyses of human error. *Communications of the ACM*, 26(4):254–258, 1983.

[18] Donald A Norman. The'problem'with automation: inappropriate feedback and interaction, not'over-automation'. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 327(1241):585–593, 1990.

[19] Dale Purves, Elizabeth M Brannon, Roberto Cabeza, Scott A Huettel, Kevin S LaBar, Michael L Platt, and Marty G Woldorff. *Principles of cognitive neuroscience*, volume 83. Sinauer Associates Sunderland, MA, 2008.

[20] Laura Rafferty, Daniel P Jenkins, Neville A Stanton, Guy H Walker, Michael G Lenné, et al. *Human Factors Methods and Accident Analysis: Practical Guidance and Case Study Applications.* Ashgate Publishing, Ltd., 2012.

[21] J Reason, E Hollnagel, and J Paries. Revisiting the «Swiss cheese» model of accidents. *Journal of Clinical Engineering*, 27:110–115, 2006.

[22] James Reason. Generic error-modelling system (GEMS): A cognitive framework for locating common human error forms. *New technology and human error*, 63:86, 1987.

[23] James Reason. *Human error.* Cambridge university press, 1990.

[24] Bill Runciman and Merrilyn Walton. *Safety and ethics in healthcare: a guide to getting it right.* Ashgate Publishing, Ltd., 2007.

[25] Scott A Shappel and Douglas A Wiegmann. The human factors analysis and classification system–HFACS. Technical report, US Federal Aviation Administration, Office of Aviation Medicine, 2000.

[26] TB Sheridan. Trustworthiness of command and control systems. In *3. IFAC/IFIP/IEA/IFORS Conference on Analysis*, pages 427–431, 1988.

[27] Geoff Simpson, Tim Horberry, and Jim Joy. *Understanding human error in mine safety.* Ashgate Publishing Limited, 2009.

[28] Douglas A Wiegmann and Scott A Shappell. *A human error approach to aviation accident analysis: The human factors analysis and classification system.* Ashgate Publishing, Ltd., 2012.