Development of a Human Reliability Analysis (HRA) model for break scheduling management in human-intensive working activities.

Valentina Di Pasquale

Unione Europea

Ministero dell'Istruzione,
dell'Università e della Ricerca

UNIVERSITÀ DEGLI
STUDI DI SALERNO

Department of Industrial Engineering

Ph.D. Course in Industrial Engineering
(XXX Cycle)

Ph.D. Thesis in:
DEVELOPMENT OF A HUMAN RELIABILITY ANALYSIS (HRA) MODEL FOR BREAK SCHEDULING MANAGEMENT IN HUMAN-INTENSIVE WORKING ACTIVITIES

Supervisor
Prof. Salvatore Miranda

Ph.D. student
Valentina Di Pasquale

Co-Supervisor
Prof. W. Patrick Neumann

Ph.D. Course Coordinator
Prof. Ernesto Reverchon

Academic year 2017-2018

Alla mia grande famiglia,
vicina anche se lontana,
complicata e in continuo movimento
ma sempre AFFIDABILE…

Publications list

- Di Pasquale, V., Iannone, R., Miranda, S., & Riemma, S. (2013). An overview of human reliability analysis techniques in manufacturing operations. In *Operations Management*. InTech. 10.5772/55065.
- Di Pasquale, V., Miranda, S., Iannone, R., & Riemma, S. (2015). A simulator for human error probability analysis (SHERPA). *Reliability Engineering & System Safety*, *139*, 17-32. 10.1016/j.ress.2015.02.003.
- Di Pasquale, V., Miranda, S., Iannone, R., & Riemma, S. (2015). An HRA-based simulation model for the optimization of the rest breaks configurations in human-intensive working activities. *IFAC-PapersOnLine*, *48*(3), 332-337. 10.1016/j.ifacol.2015.06.103.
- Di Pasquale, V., Miranda, S., Iannone, R., & Riemma, S. (2015). Simulative analysis of performance shaping factors impact on human reliability in manufacturing activities. *27th European Modeling and Simulation Symposium, EMSS 2015,* 93-102.
- Di Pasquale, V., Miranda, S., Iannone, R., & Riemma, S. (2016). Integration of learning and forgetting processes with the SHERPA model. *IFAC-PapersOnLine*, *49*(12), 197-202. 10.1016/j.ifacol.2016.07.595
- Di Pasquale, V., Franciosi, C., Lambiase, A., & Miranda, S. (2016, December). Methodology for the analysis and quantification of human error probability in manufacturing systems. In *Research and Development (SCOReD), 2016 IEEE Student Conference on* (pp. 1-5). IEEE. 10.1109/SCORED.2016.7810093.
- Di Pasquale, V., Fruggiero, F., Iannone, R., & Miranda, S. (2017). A model for break scheduling assessment in manufacturing systems. *Computers & Industrial Engineering, 111,* 563-580. 10.1016/j.cie.2017.05.017.
- Di Pasquale, V., Miranda, S. & Neumann, W.P., (2017). The impact of aging on human error in manufacturing systems: a systematic review. *Proceedings of the Association of Canadian Ergonomists (ACE) 48th Annual Conference, Banff Centre, Alberta, July 31-August 3.*

- Di Pasquale, V., Franciosi, C., Iannone, R., Malfettone, I., and Miranda, S. (2017) Human error in industrial maintenance: a systematic literature review. *Proceedings of the XXII Summer School "Francesco Turco"*, Palermo (Italy).
- Di Pasquale, V., Miranda, S., Neumann, W.P., Setayesh, A. (2018). Human reliability in manual assembly systems: a Systematic Literature Review. IFAC Symposium on Information Control Problems in Manufacturing (INCOM 2018). *Under review*.

Summary

Figures Index

Tables Index

Abbreviations and acronyms list

ADS:	Accident Dynamics Simulator
AHP:	Analytic Hierarchy Process
APJ:	Absolute Probability Judgment
ASEP:	Accident Sequence Evaluation Program
ATHEANA:	A Technique for Human Error Analysis
BN:	Bayesian Network
CAHR:	Connectionism Assessment of Human Reliability
CES:	Cognitive Environmental Simulation
CF:	Contributing factor
CLO:	Thermal resistance index of clothing
CoCoM:	Contextual Control Model
COSIMO:	Cognitive Simulation Model
CPC:	Common Performance Conditions
CR:	Cloudiness Ratio
CREAM:	Cognitive Reliability and Error Analysis Method
EOC:	Commission error
EOO:	Omission Error
EPC:	Error Producing Conditions
ESAT:	Expert System for Task Taxonomy
ETA:	Event Tree Analysis

FMEA:	Failure Modes Effects Analysis
FTA:	Fault Tree Analysis
GT:	Generic Task
HAZOP:	Hazard and Operability Analysis
HCR:	Human Cognitive Reliability Model
HE:	Human Error
HEART:	Human Error Assessment and Reduction Technique
HEM:	Human Error in Maintenance
HEP:	Human Error Probability
HFE:	Human Failure Event
HR:	Human Reliability
HRA:	Human Reliability Analysis
IDA:	Influence Diagrams Analysis
IDAC:	Information, Decision, Action in Crew context
IDEF:	Integration Definition for Function Modeling
JHEDI:	Justification of Human Error Data Information
LFCM:	Learning and forgetting curves model
MERMOS:	Method d'Evaluation de la Realisation des Missions Operateur pour la Surete
MET:	Index of metabolic activity
MIDAS:	Man-machine Integration Design and Analysis
MMI:	Man - machine interface
MORT:	Management Oversight Risk Tree
NARA:	Nuclear Action Reliability Assessment
OATS:	Operator Action Tree System
PHECA:	Potential Human Error and Cause Analysis
PIF:	Performance Influencing Factors
PIPE:	Perception – Interpretation – Planning - Execution
PRA:	Probabilistic Risk Assessment
PROCOS:	Probabilistic cognitive simulator

PSA:	Probabilistic Safety Assessment
PSF:	Performance Shaping Factors
SHELL:	Software Hardware, Environment, Liveware, Liveware
SHERPA:	Systematic Human Error Reduction and Prediction Approach
SHERPA:	Simulator for Human Error Probability Analysis
SYBORG:	Simulation System for Behaviour of an Operating group
SLIM:	Success Likelihood Index Method
SLIM-MAUD:	Success likelihood index methodology, multi-attribute utility decomposition
SLR:	Systematic Literature Review
SPAR-H:	Standardized Plant Analysis Risk – Human
SRK:	Skill – Rule – Knowledge
THERP:	Technique for Human Error Rate Prediction
THI:	Thermohygrometric Index

Abstract

Human factors play an inevitable role in working contexts and the occurrence of human errors impacts on system reliability and safety, equipment performance and economic results. If human fallibility contributes to majority of incidents and accidents in high-risk systems, it mainly affects the quality and productivity in low-risk systems. Due to the prevalence of human error and the huge and often costly consequences, a considerable effort has been made in the field of Human Reliability Analysis (HRA), thus arriving to develop methods with the common purpose to predict the human error probability (HEP) and to enable safer and more productive designs. The purpose of each HRA method should be the HEP quantification to reduce and prevent possible conditions of error in a working context. However, existing HRA methods do not always pursue this aim in an efficient way, focusing on the qualitative error evaluation and on high-risk contexts. Moreover, several working aspects have been considered to prevent accidents and improve human performance in human-intensive working contexts, as for example the selection of adequate work-rest policies. It is well-known that introducing breaks is a key intervention to provide recovery after fatiguing physical work, prevent the growth of accident risks, and improve human reliability and productivity for individuals engaged in either mental or physical tasks. This is a very efficient approach even if it is not widely applied.

Starting from these research gaps, the thesis focuses on the development of a HRA model for the break scheduling management in human-intensive working activities. The Simulator for Human Error Probability Analysis (SHERPA) model provides for a theoretical framework that exploits the advantages of the simulation tools and the traditional HRA methods to model human behaviour, to predict the error probability for a given scenario in every kind of working system and to manage the work-rest policies. Human reliability is estimated as function of the performed task, the Performance Shaping Factors (PSF) and the time worked, with the purpose of considering how reliability depends on the time that a worker has already spent on his work. Knowing the HEP distribution allows to understand the nature of the factors that influence human performance and to intervene, from the

perspective of reducing the errors and their huge monetary costs, with re-design tasks or other interventions, such as the management of the worker's psycho-physical recovery through appropriate work-rest policies. SHERPA is not limited to the reliability assessment, as many existing HRA methods, but uses it in the operator recovery modelling and breaks scheduling management. The main focuses of SHERPA are the HEP quantification, the assessment of the impact of context via PSFs, the management of break scheduling through an economic model that identifies the best configuration among those possible to reduce errors and increase productivity and efficiencies. As shown in the case studies, SHERPA is able to predict the HEP, to assess the impacts of individual features and working environments on human reliability for every kind of working context. Furthermore, the model is useful in assessing the impact of different work- break policies, with different placement and duration of breaks, on human performance (HEP and recovery after the break) and on the overall system performance in terms of percentage of compliant performed tasks and economic results.

Introduction

Human error is here to stay (Kirwan, 1994). This perhaps obvious statement has a more profound implication if we consider how common human errors are in everyday life and in the working environment.

The vast majority of current catastrophes arises by a combination of many small events, system faults and human errors that would be irrelevant individually, but – when combined in a special time sequence of circumstances and actions – can lead to unrecoverable situations (Cacciabue, 1998). In recent years, there has been a decrease in accidents due to technical failures through high reliability of mechanical and electronic components, combined with a design suitable and technological developments of redundancy and protection, which have made systems more reliable. However, despite the possibilities of automation, human labour is still needed in many working environments and it is not possible to talk about system reliability without addressing the failure rate of all its components; among these components, "man" – because his rate of error changes the rate of failure of components with which he interacts.

The contribution of the human factor in the dynamics of accidents – both statistically and in terms of severity of consequences – is high. For this reason, wrong and inappropriate human actions are source of great concern and create efficiency and safety issues for every kind of working context. Although valid values are difficult to obtain, evidence in literature indicate that errors committed by man are responsible for 60–90% of the accidents; the remainder of accidents are attributable to technical deficiencies (Hollnagel, 1998; Griffith and Mahadevan, 2011). The percentage of incidents connected with human error in several industries is listed in Table 1. The accidents are, of course, the most obvious human errors in the industrial systems, but minor faults can seriously reduce the operation performance in terms of productivity and efficiency. In fact, human error has a direct impact on productivity because errors affect the rates of rejection of the product, thereby increasing the cost of production and possibly reduce subsequent sales. Human error affects the production cost because of internal costs (scraps, reworks, product and process revisions) and external costs (recall of products, loss of image and repair and replacement warranty).

Table 1: *Estimates of human error in various sectors as percentages of all failures (Griffith and Mahadevan, 2011).*

Sectors	**% Human Error**
Automobile	65
Heavy truck	80
Aviation	70-80
Jet transport	65-85
Air traffic control	90
Maritime vessels	80-85
Chemical industry	60-90
Nuclear power plants (US)	50-70

The evidence that human actions are a source of vulnerability for industrial systems gave birth to the Human Reliability Analysis (HRA), which aims at further examination of the human factor through the prediction of when an operator is more likely to make an incorrect action and which type of action is most probable (Hollnagel, 1996). The study of HRA is approximately 60 years old and has always been a hybrid discipline, involving reliability experts, engineers and human factors specialists or psychologists. Bell and Holroyd (2009) identified 72 human reliability related tools developed since the early 60s and classified them into three categories: first, second, and third generation. All these methods have the purpose of assessing the likelihood of human error – in working systems, for a given operation, in a certain interval of time and in a particular context – on the basis of models that describe, in a more or less simplistic way, the complex mechanism that underlies the single human action that is potentially subject to error. Despite the efforts of HRA experts to develop an advanced method, many of the limitations and problems of these approaches have not yet been resolved due to the complexity of human nature and the difficulty in predicting and simulating human behaviour.

At the same time, several aspects of the work were considered to prevent and/or reduce the number of accidents and incidents and to improve the human performance. Rest breaks are an aspect of considerable importance in this sense. Introducing breaks is a key intervention to provide recovery after fatiguing physical work to prevent growth of accident risks during working activities and improve human reliability (Dababneh, Swanson and Shell, 2001; Jansen, Kant and van den Brandt, 2002; Demerouti *et al.*, 2012). It is well-known that work–break configurations influence the performance of individuals and can result in different productivity levels for individuals engaged in either mental or physical tasks (Bechtold and Thompson, 1993). Finding an optimal distribution across time of work breaks has been a challenge in ergonomics and operational research for almost an entire century, and it has also engaged management scientists (Bechtold and Thompson, 1993; Aykin, 1996; Schafhauser, Musliu and Wild, 2009; Rekik, Cordeau and Soumis, 2010). To date the break scheduling problem has been addressed within the more general shift scheduling problem and numerous optimization algorithms and heuristic techniques have been proposed (Schafhauser, Musliu and Wild, 2009; Rekik, Cordeau and Soumis, 2010). None of existing methods considers human reliability in assessing worker performance due to the complexity of HRA approaches and given the difficulty of integrating this type of modelling in an exact algorithmic or heuristic technique. Furthermore, many of the studies in the literature have addressed the break scheduling problem only from the point of view of productivity. They do not address the problem of break management regarding the quality aspect, namely the impact of human errors on the system performance in terms of quality of the performed activities (e.g. non-compliant items and reworking). The impact of breaks, in fact, was investigated with respect on the loss of productivity, due to the decrease of work rate, without considering the effect on the human error probability.

Considering the current state of the art of HRA methods and the model for the break scheduling management, the two main research questions of the thesis are:

- The development of a HRA approach able to predict the human error probability (HEP), to assess the impacts of individual features and working environments, via PSF, on human reliability for every kind of human intensive working context.
- The study of impact of different work-break policies in order to assess the qualitative and quantitative effects of human reliability on the system performance and to identify the best work-break configuration.

A new HRA model is proposed in this thesis: the Simulator for Human Error Probability Analysis (SHERPA) model. SHERPA provides a theoretical framework that exploits the advantages of the simulation tools and the traditional HRA methods in order to model human behaviour and to predict the error probability for a given scenario in every kind of industrial system. Human reliability is estimated as function of the performed task, the Performance Shaping Factors and the worked time, with the purpose of considering how reliability depends not only on the task and working context, but also on the time that the operator has already spent on the work. The model is able to provide for the following functions:

1) Estimating human reliability, as function of time, work context conditions, physical and mental employee conditions and break scheduling.

2) Assessing the effects due to different human reliability levels, through evaluation of processes, activities or tasks performed more or less correctly.

3) Assessing the impact of environment on human reliability, via performance shaping factors.

4) Simulating a large numbers of break scheduling with several locations and duration of breaks, in order to assess their impact of different work-break policies on human performance (HEP and recovery after the break) and the overall system performance in terms of percentage of compliant performed tasks and economic results (e.g. profits, revenues, quality costs, rework costs and break costs).

Chapter I provides for a detailed and complete overview of the state of the art of HE taxonomies and HRA methods, beginning with the quantitative methods of the first generation and the qualitative methods of the second one and extending to the third generation HRA approaches and new dynamic HRA methods. Chapter II analysis the role of break in working field, giving an overview of the work-break literature, considering the impact of breaks on human performance (well-being, recovery, and risk) and the break scheduling problems. Chapter III presents the SHERPA theoretical framework based on the integration of traditional and simulative/dynamic HRA methods. Its logical foundations, the HRA principles, the evaluation and quantification of psycho-physical recovery and the break scheduling management system are described. Chapter IV presents a detailed description of the implementation of the theoretical model in the two simulation tools developed in Arena and Anylogic. Then the operating principles are discussed in some numerical experiments and case studies in Chapter V, where the simulation results are presented and analysed. Finally, the main findings and conclusion are discussed.

Chapter I: Human error and human reliability in human-intensive working activities

I.1 Introduction

Human errors in the workplace can have severe consequences such as accidents, malfunctions and quality defects in the performed task. The problem of human error is substantive, and many researchers have focused on trying to understand and evaluate the concept of human error (Reason, 1990; Hollnagel, 1993; Czaja, Nair and Salvendy, 2012) and considerable efforts have been made in order to predict human performance in working contexts. Human errors, in fact, can have several causes and produce different effects: if you know the probable causes, you can try to prevent it; if you know the error consequences you can try to limit them.

This Chapter provides for a detailed and complete overview of the state of the art of HE taxonomies and HRA methods, beginning with the quantitative methods of the first generation and the qualitative methods of the second one and extending to the third generation HRA approaches and new dynamic HRA methods. Furthermore, HE assessment and the applications of HRA methods in several human intensive working contexts, such as manufacturing systems, industrial maintenance and healthcare systems, are investigated.

I.2 Human errors and human reliability analysis

There are various definitions for human error. Hollnagel (1993) favoured the term erroneous action to human error, which he defined as “an action which fails to produce the expected result, and which therefore leads to an unwanted consequence”. Reason (1990) provides a broad definition proposing it as a generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some chance agency.

Two major approaches can be taken to characterizing human error: probabilistic and causal (Rouse and Rouse, 1983). The probabilistic approach is typically pursued by those who are interested in the human reliability aspects of risk analysis. In these analyses, human error is treated in a manner quite like that used for hardware failures. The use of this approach is often dictated by requirements that system reliability meet some specified level. On the contrary, the causal approach to characterizing human error assumes that errors are seldom random, and in fact, can be traced to causes and contributing factors which, once isolated, can perhaps be eliminated or at least ameliorated. Thus, the causal approach can be useful for evaluating and subsequently modifying system designs and training programs. This focus on why errors occur is rather different from the typical studies of human error which solely emphasize what occurs, a point of view that has received considerable criticism (Rouse and Rouse, 1983).

A core issue dealt with by many studies on human error is the classification scheme or taxonomy of error types. An effective taxonomy can be of value in organizing data on human errors and for giving advantageous insights into the ways in which errors are caused and how they might be prevented. A literature review revealed that no single taxonomy of human error is generally accepted for addressing all causal factors. Norman (1981) distinguishes between mistakes and slips, where a mistake reflects an inappropriate intention and a slip is an unintentional error that occurs when a person does an action that is not intended. Focusing on slips, Norman utilizes a human information processing perspective to develop a schema-oriented theory involving formation of intentions as well as activation and triggering of schema. The three-major category of slips are: (i) errors in the formation of intention (which considers the mode and description errors), (ii) faulty activation schemas (which considers capture errors, data driven, loss of intention and misordering of action components); and (iii) faulty triggering (which considers blends, intrusions of thoughts and premature triggering).

Rasmussen was the first to divide human behaviours into three levels: skill-based behaviour, rule-based behaviour, and knowledge-based behaviour (or the Skill-Rule-Knowledge SRK framework) (Rasmussen *et al.*, 1981; Rasmussen, 1982). The proposed Model of Internal Human Malfunction differentiates three basic levels of human performance:

- ✓ <u>Skill-based</u>, when automated actions follow an intention (sensory–motor behaviour), actions are routinely practiced and highly automatic. Conscious thought is used sporadically to verify progress.

- ✓ Rule-based, when there is a procedure or technique guiding the action, actions are a combination of conscious and unconscious processes to determine responses for situations that have been encountered before, either through experience or training. Unconscious pattern recognition matches the indicators of a problem to a potential solution.
- ✓ Knowledge-based, represented by actions developed to deal with an unfamiliar situation, actions require slow, demanding, and highly-error prone conscious thought to determine a response when other methods have proven unsuccessful.

Further, he distinguishes among causes, mechanisms, and modes of human error (Rasmussen, 1982). His overall goals include developing a comprehensive classification scheme for reporting events involving human error. The attention and conscious thought that an individual give to activities taking place decreases moving from the third to first level (Figure I.1).

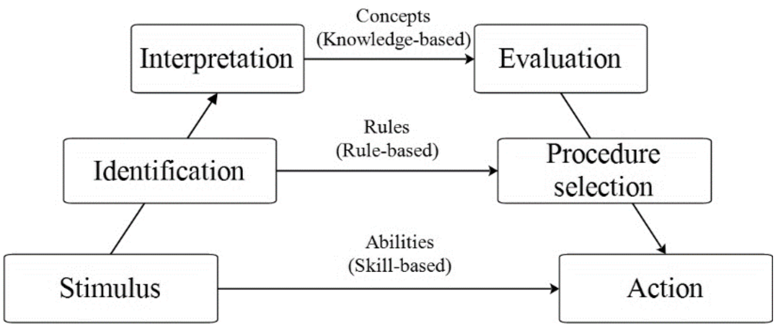

Figure I.1: *Rasmussen's SRK model.*

Human error has been extensively researched and classified in the reliability engineering and system safety field. Swain and Guttmann (1983) divide human errors into errors of omission (EOOs) and errors of commission (EOCs); EOOs are defined as the failure to perform an action, whereas EOCs are defined as unintended or unplanned actions.

Unlike Swan’s classification method, Model of Unsafe Acts (Reason, 1990) divides human errors in slips and lapses, when an execution failure or an omission occurs, and mistakes, which result from judgement processes used to select an objective, or the means to accomplish it. A slip is an incorrect execution with a correct intention; lapse is an unintended or unplanned action with a correct intention; and a mistake is incorrect execution with an incorrect intention. Reason also highlights an alternative behaviour from a social context, called violation, which emerges from an intentional deviation from operating procedures, codes of practice or standards. Errors differ from violations in that errors are unintended whereas violations are deliberate.

Despite there are many classifications, it is not easy arriving at a satisfying and unambiguous definition and classification of human error and literature provides little guidance on how to systematically classify an event into error categories. The error taxonomies previously described underline how human performance models can be used to predict human error. HEP is defined as the probability that a certain task within observation period was accomplished faulty, as a relative frequency (Bubb, 2005):

$$HEP = \frac{\textit{number of observed errors}}{\textit{number of the possibilities for an error}} \tag{1.1}$$

Human reliability (HR), instead, is the mathematical complement (HR=1-HEP), and it is the human ability to accomplish a given task under given conditions in a given time interval within the acceptance limits. The standard definition of human reliability is the probability that a person will perform according to the requirements of the task for a specified period and not perform any extraneous activity that can degrade the system (Hollnagel, 1998). Human reliability is, also, defined as the probability that each human component will perform successfully for an extended period (Czaja, Nair and Salvendy, 2012).

Due to the prevalence of human error and the huge and often costly consequences, its study has become an increasingly important research concern and an important focus within HRA, which has emerged as a well-defined discipline. The HRA goals defined by Swain and Guttmann (1983) in discussing the Technique for Human Error Rate Prediction (THERP) approach, one of the first HRA methods developed, are still valid: “the objective of HRA is to evaluate the operator’s contribution to system reliability. More precisely, the aim is to predict human error rates and to evaluate the degradation of human–machine systems likely to be caused by human errors in association with equipment functioning, operational procedures and practices which influence the system behaviour”.

Human Reliability Analysis is carried out, as part of the Probabilistic Risk Assessment (PRA), to identify and to quantify human actions and the associated impacts on structures, systems, and components of complex facilities, through the forecast of the events that can occur during the working activity. HRA has evolved into a discipline that encompasses theoretical and analytic tools needed for understanding how human actions and decisions are influenced by the system’s complexity and dynamics, the assessment of human errors that may arise during the work, and design interventions in the form of various barriers that can eliminate or mitigate these negative effects (Sharit, 2012). The purpose is to pursue quantitative estimates of human error probabilities during professional activity and their contribution to system risks. HRA typically encompasses three phases (Figure I.2), ranging from identifying error sources, to modelling these errors as part of a systemic analysis including hardware failures, to quantifying the HEPs (Boring, 2010).

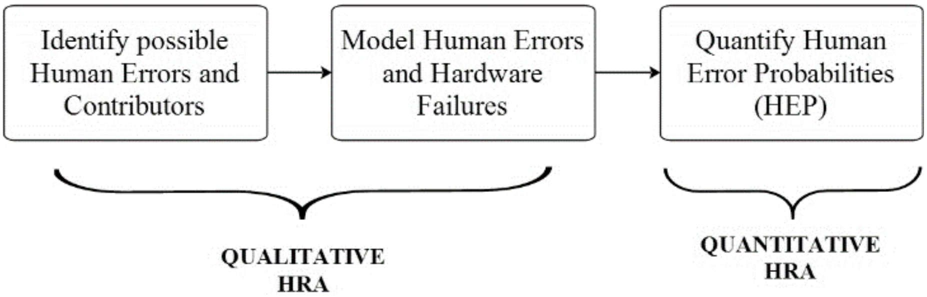

Figure I.2: *Three Phases of Human Reliability Analysis (Boring, 2010).*

The 10-step HRA process proposed by Kirwan (1994) deeply highlights the role of task and human error analyses in its earlier stages (Figure I.3).

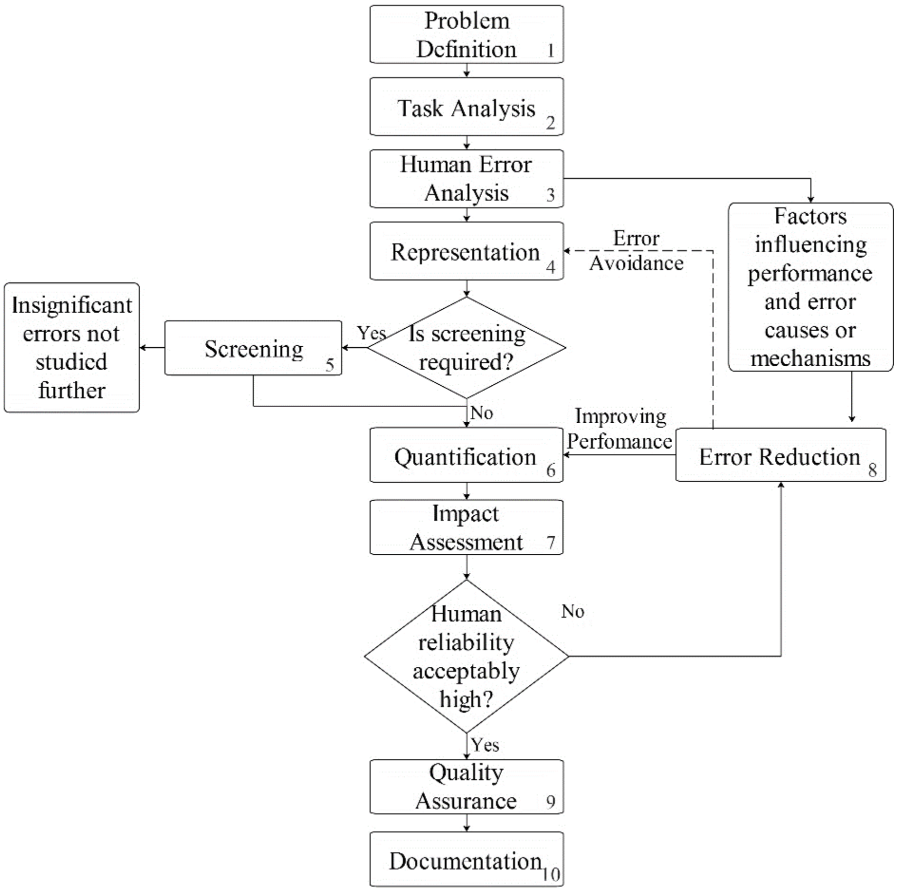

Figure I.3: *The HRA Process (Kirwan, 1994).*

In the traditional HRA process, task analysis is used to describe and understand the human interactions with the system. The results of this phase allow error identification through appropriate error taxonomy. The analysts first define human failure events (HFEs), which are analysed qualitatively and quantitatively, and then they assign relative nominal HEPs to events.

Nominal HEP is calculated on the basis of operator's activities and, to obtain a quantitative estimate of HEP, many HRA methods utilize performance shaping factors (PSF), which characterize significant facets of human error and provide a numerical basis for modifying nominal HEP levels (Boring, 2006). The qualitative analysis, in fact, determines the influencing factors that enhance or degrade human performance and that might lead to the failure of the activity. These influencing factors include the features of the plant and the PSFs; these last are determined by the individual characteristics of the human being, environment, organization, or activity. Their goal is to provide measures to account for human performance. There is no consensus to date on which PSFs should be used in HRA nor on the appropriate number of PSFs to include in an analysis. Some of the earliest HRA approaches adopted a single PSF, while a recent study commissioned by the US Nuclear Regulatory Commission (Good Practices for Implementing Human Reliability Analysis) identified the fourteen essential PSFs for HRA (Kolaczkowski *et al.*, 2005). This list of PSFs is not exhaustive but rather represents the minimum set of PSFs that should be considered in an HRA. There are numerous approaches to quantify HRA methods (Boring, 2015):

- Scenario matching methods: This approach, adopted by THERP (Swain and Guttmann, 1983), entails matching the HFE to the best fitting example scenario in a table and using the HEP associated with that template event as the basis for quantification.

- PSF adjustment methods: In methods, such as the standardized plant analysis risk-human reliability analysis method (SPAR-H) (Gertman *et al.*, 2005) or cognitive reliability and error analysis (CREAM) method (Hollnagel, 1998), PSFs modify the nominal error rates. The effects of the PSF on the HEP in SPAR-H are summarized in the following equation as follows (Boring, 2010).

$$HEP_c = HEP_n \cdot PSF = \begin{cases} 0 < PSF < 1 \rightarrow \; HEP_c < HEP_n \\ PSF = 1 \quad\; \rightarrow HEP_c = HEP_n \\ PSF > 1 \quad\; \rightarrow HEP_c > HEP_n \end{cases} \tag{1.2}$$

- Where HEP_c and HEP_n are the contextual and nominal HEPs, respectively. Each PSF can have both positive and negative effects on performance, respectively decreasing or increasing the overall HEP.

- <u>Expert estimation methods:</u> These tools provide a structured means for experts to consider how likely it is for an error to occur in a scenario.
- <u>Simulation based methods:</u> Although currently uncommon, these methods use cognitive modelling and simulation to produce a data framework that may be used in quantifying the likelihood of human error (Boring, 2007; Chang and Mosleh, 2007a, 2007b, 2007c, 2007d, 2007e).

I.3 State-of-the-art HRA methodologies

The first HRA methods date back to the early 60s in the field of missile applications and US defence, but most techniques for assessment of the human factor, in terms of propensity to fail, have been developed since the mid-'80s. The greatest development of these techniques has been closely linked to catastrophic events in high-risk industries, such as Seveso (1976), Bhopal (1983), Chernobyl (1986). These events came to understand how accidents are not only caused by technical causes or failures or human causes and failures interaction of several components: technological, human, organizational, in relation to each other and with the external environment in which the organization operates. The problem of human dependence has since raised its level of complexity. The methods implemented over time contain appropriate and enforceable procedures and user manuals for the identification of the most appropriate data and the application of the methodology, which invariably leads to the probability distributions and the uncertainty associated with human failure.

The development of human reliability methods occurred over time in three stages. HRA techniques or approaches can, in fact, be divided essentially into three categories: first, second and third generation. Figure I.4 shows the distribution of HRA methods over the years, starting from the early methods to the most recent developments.

I.3.1 <u>First generation HRA methods</u>

The first stage lasted twenty years (1970–1990) and was the first human reliability method generation that focused on human error probabilities and operational human error.

The first generation HRA methods have been strongly influenced by the viewpoint of probabilistic safety assessment (PSA) and have identified man as a mechanical component, thus losing all aspects of dynamic interaction with the working environment, both as a physical environment and as a social environment. First generation methods, in fact, include 35–40 methods for human reliability, many of which are variations on a single method. Many of these methods – such as THERP (Swain and Guttmann, 1983), Accident Sequence Evaluation Program (ASEP) (Swain, 1987), and Human Cognition Reliability (HCR) (Hannaman, Spurgin and Lukic, 1984) – have the basic assumption that the natural deficiencies of humans cause them logically to fail to perform tasks, just as is seen with mechanical or electrical components. Thus, HEP can be assigned based on the characteristics of the operator's task and then modified by performance shaping factors. In the first HRA generation, the characteristics of a task, represented by HEPs, are regarded as major factors; the context, which is represented by PSFs, is considered a minor factor in estimating the probability of human failure (Kim, Seong and Hollnagel, 2006).

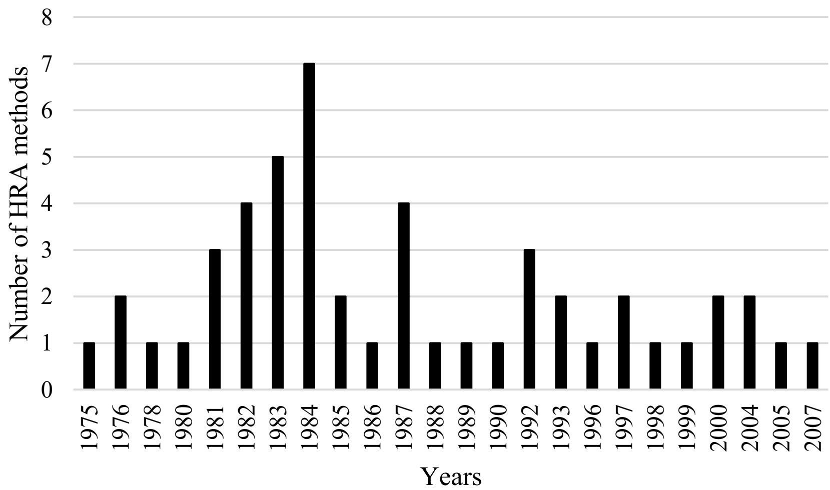

Figure I.4: *HRA methods timeline.*

Each approach of this generation focuses on quantification in terms of success/failure of actions, with less attention paid to in-depth causes and reasons of observable human behaviour, which for these techniques is borrowed from psychological studies in behavioural sciences (Cacciabue, 2004). These traditional approaches determine the human error probability by using established tables, human reliability models or expert judgment. The characterization of human failure modes is usually very simple, such as in terms of error of omission and errors of commission (Swain and Guttmann, 1983), which represent, respectively, the lack of realization of operations required to achieve the result and the execution of an operation, not related to that request, which prevents the obtainment of the result (Hollnagel, 1998).

The main characteristics of the methods can be summarized as follows:

- Binary representation of human actions (success/failure);

- Attention on the phenomenology of human action;
- Low concentration on human cognitive actions (lack of a cognitive model);
- Emphasis on quantifying the likelihood of incorrect performance of human actions;
- Dichotomy between errors of omission and commission;
- Indirect treatment of context.

Among the first-generation techniques, in addition to the methods already mentioned, there are: Absolute Probability Judgement (APJ) (Kirwan, 1994), Human Error Assessment and Reduction Technique (HEART) (Williams, 1985; B Kirwan, 1997), Justified Human Error Data Information (JHEDI) (Kirwan, 1994), Probabilistic Human Reliability Analysis (PHRA) (Bell and Holroyd, 2009), Operator Action Tree System (OATS) (Bell and Holroyd, 2009), and Success Likelihood Index Method (SLIM) (Embrey, 1986).

Among these, the most popular and effectively method used is THERP, characterized as other first generation approaches by an accurate mathematical treatment of the probability and error rates, as well as computer programs well-structured for interfacing with the trees for evaluation of human error of a fault event and trees (Boring and Blackman, 2007). The base of THERP is event tree modelling, where each limb represents a combination of human activities, influences upon these activities, and results of these activities (Griffith and Mahadevan, 2011).

First generation HRA methods are demonstrated with experience and use, not able to provide sufficient prevention and adequately perform their duties. The criticism of is that these approaches tend to be descriptive of events in which only the formal aspects of external behaviour are observed and studied in terms of errors, without considering reasons and mechanisms that made them level of cognition. These methods ignore the cognitive processes that underlie human performance and, in fact, possess a cognitive model without adequate human and psychological realism. They are often criticized for not having considered the impact of factors such as environment, organizational factors, and other relevant PSFs; and for not using proper methods of judging experts (Hollnagel, 1998; Mosleh and Chang, 2004). Swain remarked that “all of the above HRA inadequacies often lead to HRA analysts assessing deliberately higher estimates of HEPs and greater uncertainty bounds, to compensate, at least in part, for these problems” (Hollnagel, 1998). This is clearly not a desirable solution. Despite the criticisms and inefficiencies of some first-generation methods, such as THERP and HCR, they are regularly used in many industrial fields, thanks to their ease of use and highly quantitative aspects.

I.3.2 Second generation HRA methods

In the early 1990s, the need to improve HRA approaches interested many important research and development activities around the world. These efforts led to much progress in first generation methods and the birth of new techniques during the second phase (1990–2005), known as second human reliability method generation, focused on human performance factors and cognitive processes. Human performance factors are internal or external and in general are everything that influences human performance, like workload, stress, sociological issues, psychological issues, illness, etc. (Calixto, Lima and Firmino, 2013). These HRA methods have been immediately unclear and uncertain, substantially because the methods have been defined in terms of what should not be – that is, they should be as the first generation of HRA methods (Hollnagel, 1996).

While the first generation HRA methods are mostly behavioural approaches, the second generation HRA methods aspire to be of conceptual type (Chang and Mosleh, 2007e). The separation between generations is evident in the abandonment of the quantitative approach of PRA/PSA in favour of a greater attention to qualitative assessment of human error. The focus of the second generation shifted to cognitive aspects of humans, causes of errors rather than their frequency, study of factor interactions that increase the probability of error, and interdependencies of PSFs.

Second generation HRA methods are based on a cognitive model more appropriate to explain human behaviour. It is evident that any attempt at understanding human performance needs to include the role of human cognition, defined as "the act or process of knowing including both awareness and judgement" by an operator. From the HRA practitioner's perspective, the immediate solution to take into consideration human cognition in HRA methods was to introduce a new category of error: "cognitive error", defined both as failure of an activity that is predominantly of a cognitive nature and as the inferred cause of an activity that fails (Hollnagel, 1998). For example, in CREAM, developed by Erik Hollnagel in 1993, maintained division between logical causes and consequences of human error (Hollnagel, 1996). The causes of misbehaviour (genotypes) are the reasons that determine the occurrence of certain behaviours, and the effects (phenotypes) are represented by the incorrect forms of cognitive process and inappropriate actions (Mosleh and Chang, 2004).

Moreover, advanced cognitive models have been developed, which represent the process of logic operator and synthesize the dependence on personal factors (such as stress, incompetence, etc.) and by the current situation (normal conduction system, abnormal conditions, or even emergency conditions), and models of man–machine interface, which reflect the control system of the production process. One of the more widely used second generation techniques, CREAM (Hollnagel, 1998), has an operator model that is more significant and less simplistic than in the first generation methods; HEP is derived from four Contextual Control Modes (CoCoMs): Scrambled, Opportunistic, Tactical and Strategic. CoCoM assumes that human behaviour is guided by two basic principles: the cyclical nature of human cognition and the dependence of cognitive processes from context and working environment. The Standardized Plant Analysis Risk–Human Reliability Analysis method (SPAR-H) (Gertman *et al.*, 2005) is, instead, built on an explicit information processing model of human performance derived from the behavioural sciences literature. An information-processing model is a representation of perception and perceptual elements, memory, sensory storage, working memory, search strategy, long-term memory, and decision-making.

Additionally, second generation considers the context in which humans make errors and derive relative PSFs. A major difference between two generations can be simply stated as consideration of the PSF impact on operators. None of the first generation HRA approaches tries to explain how PSFs exert their effect on performance; moreover, PSFs – such as managerial methods and attitudes, organizational factors, cultural differences, and irrational behaviour – are not adequately treated in these methods. PSFs in the first generation were mainly derived by focusing on the environmental impacts on operators, whereas PSFs in the second generation were derived by focusing on the cognitive impacts on operators (Lee *et al.*, 2011). The context is carefully incorporated into the behavioural patterns, considering all the factors that may affect human performance. This is evident in SPAR-H, where its eight operational factors can be directly associated with the human performance model and show the human information processing model with which they are associated (Figure I.5). The PSFs of both generations were reviewed and collected in a single taxonomy of performance influencing factors for HRA (Kim and Jung, 2003).

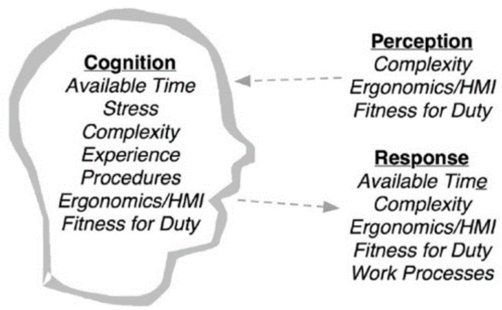

Figure I.5: *SPAR-H Performance Shaping Factors in the Information Processing Context (Whaley et al., 2011).*

Among the methods of the second generation can be mentioned: A Technique for Human Error Analysis (ATHEANA) (Cooper *et al.*, 1996), Cognitive Environmental Simulation (CES) (Woods, Roth and People, 1987), Connectionism Assessment of Human Reliability (CAHR) (Strater, 1996; Strater and Reer, 1999) and Méthode d'Evaluation De La Réalisation des Missions Opérateur Pour La Sûreté (MERMOS) (Bieder and Le Bot, 1998; Serdet and Le Bot, 2012).

Many proposed second-generation methods still lack sufficient theoretical or experimental bases for their key parts. Missing from all is a fully implemented model of the underlying causal mechanisms linking measurable PSFs or other characteristics of the context of operator response. The problem extends to the quantification side, where the majority of the proposed approaches still rely on implicit functions relating PSFs to probabilities (Mosleh and Chang, 2004). In short, some key shortcomings that motivated the development of new methods still remain unfulfilled. Furthermore, unlike first generation methods, which have been largely validated (B Kirwan, 1997), the second generation has yet to be empirically validated. There are four main sources of deficiencies in current HRA methods (Griffith and Mahadevan, 2011):

- Lack of empirical data for model development and validation;
- Lack of inclusion of human cognition (i.e. need for better human behaviour modelling);
- Large variability in implementation (the parameters for HRA strongly depend on the methodology used)
- Heavy reliance on expert judgement in selecting PSFs and use of these PSFs to obtain the HEP in human reliability analysis.

I.3.3 Third generation HRA methods

In recent years, the limitations and shortcomings of the second generation HRA methods have led to further developments related to the improvement of pre-existing methods. The third phase, started in 2005 and still in progress, is represented by methods that focus on human performance factor relations and dependencies. While some experts have focused on the development of third generation methods, others have carried out studies on the so-called dynamic HRA, as reported in the next section.

The Nuclear Action Reliability Assessment (NARA) (Kirwan *et al.*, 2004) method and the Bayesian networks (Droguett and Menêzes, 2007) are defined as the only current HRA tools of the third generation.

On one hand, NARA recaptures and improves the first-generation method HEART, trying to overcome some limitations of the same, while on the other hand, Bayesian networks use qualitative analysis, which emphasizes the importance of representing interactions between human actions and the dynamics between them.

I.4 Simulation and modelling for dynamic HRA methods

Cacciabue (1998) has outlined the importance of simulation and modelling of human performance for the field of human reliability. Specifically, simulation and modelling address the dynamic nature of human performance in a way that has not been possible in most HRA methods. Many efforts have been recently directed towards simulation, to assess human behaviour and calculate the reliability for the performed activity. Boring (2007) posits that the key to dynamic HRA is not in the development of specific methods but in the using of cognitive modelling and simulation to produce a data framework that may be used in quantifying likelihood of human error. A cognitive simulation consists of the reproduction of a cognition model using a numerical application or computation (Trucco and Leva, 2007; Leva *et al.*, 2009).

Simulator experiments can produce important basic information for HRA method development and data for informing the use of existing HRA methods. Simulators allow the study of variations in context and how this impacts human performance (Bye *et al.*, 2006). As depicted in Figure I.6, simulation and modelling may be used in three ways to capture and generate data that are meaningful to HRA (Boring, 2007):

- The simulation runs produce logs, which may be analysed by subject matter experts and used to form an estimate of the likelihood of human error. This approach builds heavily on expert estimation techniques that are commonly used in HRA.
- The simulation may be used to produce estimates of performance shaping factors, which can be quantified to produce human error probabilities based on specific HRA methods. For example, Boring (Boring, 2006) postulated a mapping of performance measures produced by the Man Machine Integration Design and Analysis system (MIDAS) to the eight influencing factors utilized by the SPAR-H method.
- A final approach is to set specific performance criteria by which the virtual performers in the simulation are able to succeed or fail at given tasks. A common performance criterion is time to complete a task, whereby failure to complete the task within a prescribed limit is

considered unsatisfactory performance. Through iterations of the task that systematically explore the range of human performance, it is possible to arrive at a frequency of failure (or success).

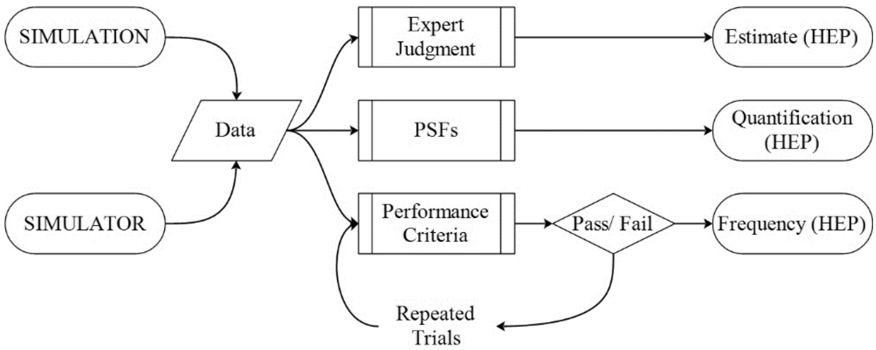

Figure I.6: *Simulation and modelling in Human Reliability Analysis (Boring, 2007).*

No modelling or simulation tool currently exists that completely or perfectly combines all elements of simulation-based HRA. Significant work is, however, already underway. A list of the main simulation projects and some of their main features are reported in Table I.1. Some of these, such as CES (Woods, Roth and People, 1987) and COSIMO (Cognitive Simulation Model) (Cacciabue *et al.*, 1992), have been developed in the nuclear field and are computer simulation methods that could potentially be useful, but no use or development is evident since the late 90s (Hollnagel, 1996). Unlike CES and COSIMO, the environment simulation MIDAS (Man Machine Integration Design and Analysis system) (Boring, 2006) was developed in 1986 in field of aerospace and aeronautic and has seen ongoing developments and applications over the years. Among the latest the integration efforts with HRA is the use of SPAR-H performance shaping factors (Boring, 2006).

Another system, the Information, Decision, Action in Crew context (IDAC) model (Chang and Mosleh, 2007a, 2007b, 2007c, 2007d, 2007e) combines a realistic plant simulator with a cognitive simulation system capable of modelling PSFs. The IDAC model is an operator behaviour model developed based on many relevant findings from cognitive psychology, behavioural science, neuroscience, human factors, field observations, and various first and second generation HRA approaches. Three generic types of operators are modelled: decision maker (e.g. shift supervisor), action taker (e.g. operators at the control panel), and consultant (e.g. resource experts in the control room). IDAC covers the operator's various dynamic response phases, including situation assessment, diagnosis, and recovery actions in dealing with an abnormal situation. Due to the variety, quantity and detail of the input information, as well as complexity of applying its internal rules, the IDAC model is best implemented through a computer simulation such as the Accident Dynamics Simulator (ADS) environment.

Trucco and Leva(2007) developed a new probabilistic cognitive simulator (PROCOS) for approaching human errors in complex operational frameworks (Trucco and Leva, 2007; Leva *et al.*, 2009). The PROCOS simulator attempts integration of the quantification of first generation HRA methods in safety assessment (e.g. THERP) with a cognitive evaluation of the operators involved in the context under examination. Its focus is mainly in conveying a quantitative result, comparable to those of a traditional HRA method and taking into account a cognitive analysis of the operator as well. The simulation model comprised two cognitive flow charts, reproducing the behaviour of a process industry operator. The model used for the configuration of the flow diagram that represents the operators is based on a combination of PIPE (Cacciabue, 1998) and SHELL (Software, Hardware, Environment, Liveware, Liveware; Edwards, 1998). The PIPE model is based on the cognitive functions: Perception, Interpretation, Planning and Execution. The two combined models allow for representation of the main cognitive processes that an operator can carry out to perform an action (PIPE) and describe the interaction among procedures, equipment, environment and plants present in the working environment, and the operator, as well as taking into account the possibility of interaction of the operator with other operators or supervisors (SHELL). As a further step, the simulator considers the evaluation of error management as part of the overall assessment from the same cognitive point of view. PROCOS does not imply the development of a detailed model for the operator–context interaction; the context is taken into account mainly through the input coming from the PSA framework to which it belongs, and through the use of performance shaping factors, as proposed in traditional HRA methods.

Table I.1: *Review of main simulators developed for simulating human behaviour in HRA field.*

SIMULATOR	TYPE	DESCRIPTION	FIELD
CES (Cognitive Environment Simulation) (Woods, Roth and People, 1987)	Quantitative and qualitative	It simulates the behaviour of a control room operator in a nuclear power plant in emergency scenarios. The purpose of CES is to imitate the way in which operators decide to respond, as a basis for quantification (Hollnagel, 1996). Developed using artificial intelligence programming.	Nuclear

SIMULATOR	TYPE	DESCRIPTION	FIELD
COSIMO (Cognitive Simulation Model) (Cacciabue et al., 1992)	Quantitative and qualitative	It simulates the behaviour of an operator reproduced through the Fallible Machine model by Reason (1990), combined with a model for the specific system to be considered. Study the operator actions in abnormal plant conditions (accident scenarios) in a nuclear power plant.	Nuclear
ADS-IDAC (Accident Dynamics Simulator-Information, Decision, and Action in Crew context model) (Chang and Mosleh, 2007a, b, c, d, e)	Quantitative	Developed for probabilistic prediction of the responses of the nuclear power plant control room-operating crew during an accident for use in probabilistic risk assessments (Chang and Mosleh, 2007e). The operator response spectrum includes cognitive, emotional and physical activities during the accident. Within the crew context, each individual operator's behaviours are simulated through a cognitive model under the influence of a number of explicitly modelled PSFs.	Nuclear
MIDAS (Man Machine Integration Design and Analysis system) (Boring, 2006)	Quantitative	An integrated suite of software developed to aid designers and analysts to apply human factor principles and human performance models to the design of complex human–machine systems in aviation. It can simulate the behaviour of a pilot for civil aviation or an air traffic controller. The model of the operator is based on Rasmussen's model (Rasmussen et al., 1981).	Aviation
PROCOS (Probabilistic Cognitive Simulator) (Trucco and Leva, 2007)	Quantitative	It supports human reliability analysis in complex operational contexts. It integrates cognitive human error analysis with standard hazard analysis methods (Hazop and event tree) by means of a semi static approach (Trucco and Leva, 2007; Leva et al., 2009). The simulation model comprised two cognitive flow charts reproducing the behaviour of a process industry operator. The simulator allows analysis of both error prevention and error recovery.	General

SIMULATOR	TYPE	DESCRIPTION	FIELD
SYBORG (Simulation System for Behaviour of an Operating group) (Takano, Sasou and Yoshimura, 1995)	Qualitative	It simulates a group of nuclear power plant operators. It needs input coming from a specific plant simulator. It highlights some possible combinations of operator errors and plant condition that can lead to accident sequences; it proposes different strategies to improve the collaboration within the group (Trucco and Leva, 2007; Leva et al., 2009).	Nuclear

Simulators implemented over time are, above all, cognitive simulators; their aim is to simulate operator or crew behaviour in terms of correct and incorrect actions. These simulations model the operator's thought processes and offer potentially powerful ways of determining how human operators will respond in emergency scenarios, typically in complex environments such as nuclear power plants. The cognitive simulators developed to date have been mainly used for qualitative analysis, and they have not found substantial applications in the quantitative risk assessment framework. The models are sometimes not easy to understand and therefore are not used by HRA specialists that have not been directly involved in their development.

I.5 Performance shaping factors

One of the undisputed assumptions in all HRA method is that the human performance depends on the conditions under which the tasks or activities are carried out (De Ambroggi and Trucco, 2011). In the HRA methods, conditions that influence human performance are often referred by term performance shaping factor (PSF), but also with alternative synonyms such as Contributing factors (CF), Individual related factor, Common Performance Condition (CPC), Error promoting condition (EPC), Error inducing factors or Performance influencing factors (PIF). They are used in qualitative approaches in order to identify contributors to human performance, while in quantitative ones, they are used to estimating a more realistic HEP. These contextual factors characterize significant facets of human error, and they are determined by the individual characteristics of the human being, the environment, the organization or the activity that enhances or decreases human performance and increases or decreases the likelihood of human error.

Their modelling and quantification is one of the most complex issues in the HRA field, to which many researchers recently are concentrating their efforts. While completing an HRA, an analyst may review a list of possible PSFs to identify possible sources of human error. The analyst may subsequently use predefined error rates associated with specific PSFs to determine a human error probability for a given task or situation (Boring, Griffith and Joe, 2007).

The first-generation HRA methods are less concerned with what people are likely to do than with whether they will succeed or fail (Lee *et al.*, 2011). None of these approaches consider explaining how the PSFs exert their effect on performance. PSFs such as managerial methods and attitudes, organizational factors, cultural differences, and irrational behaviour are not adequately treated in the first-generation. On the contrary, the second-generation considers the context in which humans make errors and derives PSFs based on these contexts. PSFs in the first-generation HRA methods were mainly derived by focusing on the environmental impacts on operators, whereas in the second one they were derived by focusing on the cognitive impacts on operators (Lee *et al.*, 2011)**.**

Within HRA, PSFs are often categorized as internal or external, corresponding to the individual vs. situational or environmental circumstances, respectively, that brings to bear on performance. The research literature divides the PSFs into two other categories: direct and indirect measures of human performance (Boring, Griffith and Joe, 2007). While some popular PSFs such as "time needed to complete a task" are directly measurable, other PSFs, such as "fitness for duty," can primarily be measured indirectly through other measures and PSFs, for example through fatigue measures.

Their definition and classification, although complex and variable, have been carefully detailed by researchers who have proposed over time numerous taxonomies, as reported in Table I.2. There has been a greater emphasis recently to catalogue ways in which PSFs might also enhance performance and to develop taxonomy of performance influencing factors for HRA of emergency tasks (Kim and Jung, 2003; Boring, 2010; Lee *et al.*, 2011). Kim and Jung (2003), for example, have collected and ordered the eighteen taxonomies in Table I.2 in a new series of PSF, consisting of about 220 detailed PSFs. These PSFs were classified the collated PIFs are classified into four main groups:

- **Human:** Personal characteristics and working capabilities of the human operator.
- **System:** Man-Machine Interface (MMI), plant hardware system, and physical characteristics of the plant process.
- **Task:** Procedures and task characteristics required of the operator.
- **Environment:** Team and organization factors, and physical working environment.

Table I.2: *Summary of PSFs taxonomies (Kim and Jung, 2003).*

<table>
<tr><th colspan="2">PSF TAXONOMIES FOR HUMAN ERROR ANALYSIS</th></tr>
<tr><td>CSNI taxonomy (Rasmussen *et al.*, 1981)</td><td>PSF taxonomy (Bellamy, 1991)</td></tr>
<tr><td>THERP (Swain and Guttmann, 1983)</td><td>Influencing factors (Gerdes, 1997)</td></tr>
<tr><td>HEART (Williams, 1985)</td><td rowspan="2">K-HPES (Kim, 1997)</td></tr>
<tr><td>PHECA (Whalley, 1987)</td></tr>
<tr><th colspan="2">PSF TAXONOMIES FOR HRA</th></tr>
<tr><th>HEP quantification</th><th>Analysis of commission errors</th></tr>
<tr><td>SLIM (Embrey, 1986)
PLGSLIM (Chu, Musicki and Others, 1994)</td><td>Macwan’s PIF taxonomy for errors of commission (Macwan and Mosleh, 1994)</td></tr>
<tr><td>INTENT (Gertman *et al.*, 1992)</td><td>Julius’ PIF taxonomy for errors of commission (Julius *et al.*, 1995)</td></tr>
<tr><td>STAHR (Philips *et al.*, 1990)</td><td rowspan="2">ATHEANA (Cooper *et al.*, 1996)</td></tr>
<tr><td>HRMS (Barry Kirwan, 1997)</td></tr>
<tr><th>Evaluation of the global context and analysis of errors</th><th>HRA database</th></tr>
<tr><td>CREAM (Hollnagel, 1998)</td><td>Taylor-Adams’ PSF taxonomy for CORE-DATA (Taylor-Adams, 1995)(Taylor-Adams, 1995)</td></tr>
<tr><td>INCORECT (Kontogiannis, 1997)</td><td>Rogers’ PSF taxonomy for CORE-DATA (Kirwan, Basra and Taylor-Adams, 1997)</td></tr>
</table>

On the other hand, the interrelationships between PSFs gain much attention from the HRA community. Despite continuing advances in research and applications, one of the main weaknesses of current HRA methods is their limited ability to model the mutual influence among PSFs, intended both as a dependency among the states of the PSFs dependency among PSFs influences on human performance, as shown in Figure I.7 (De Ambroggi and Trucco, 2011). Very different conceptual and analytical models are proposed for describing how these factors exert their influence on the human error probability; indeed, if a PSF influences human performance it is crucial to account for how this influence comes about. Several studies argued that the dependency between PSFs should be included in the quantification of HRA and suggested that the Bayesian network (BN) would be a promising technique because it can describe the casual relationship between PSFs (Groth and Swiler, 2013).

Some HRA methods – such as CREAM, SPAR-H, and IDAC – try to provide guidance on how to treat dependencies at the level of the factor assessments but do not consider that a PSF category might depend on itself and that the presence of a specific PSF might modulate the impact of another PSF on HEP; therefore, they do not adequately consider the relationships and dependencies between PSFs (De Ambroggi and Trucco, 2011). The study of De Ambroggi and Trucco (2011), instead, deals with the development of a framework for modelling the mutual influences existing among PSFs and a related method to assess the importance of each PSF in influencing performance of an operator, in a specific context, considering these interactions.

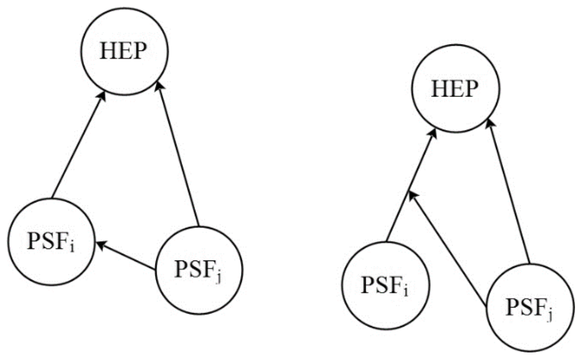

Figure I.7: *Possible types of dependency between PSFs: (A) dependency between the states/presence of the PSFs and (B) dependency between the state of PSFj and the impact of PSFi over the HEP.*

Another limitation of current HRA methods is the strong dependence on expert opinion to assign values to the PSFs; in fact, during this assignment process, subjectivity plays a significant role, causing difficulties in assuring consistency. To overcome this problem and obtain a more precise estimation, Park and Lee (2008) suggest a new and simple method: AHP– SLIM. This method combines the decision-making tool AHP – a multicriteria decision method for complex problems in which both qualitative and quantitative aspects are considered to provide objective and realistic results – with success likelihood index method (SLIM), a simple, flexible method of the expert judgement for estimating HEPs (Park and Lee, 2008). Therefore, through a type of HEP estimation using an analytic hierarchy process (AHP), it is possible to quantify the subjective judgement and confirm the consistency of collected data.

I.6 Shortcomings and limitations in HRA methods

Despite the efforts of HRA experts to develop an advanced method, many of the limitations and problems of these approaches have not yet been resolved due to the complexity of human nature and the difficulty in predicting and simulating human behaviour. Currently, no methodology has a consensus, and most of them have not been very attractive to the practitioners and managers due to the complexity of the techniques developed and the lack of information that allows implementation in a comprehensive manner. Over 70 human reliability tools were developed since 1960 for the same aim: human error quantification. Every method has the same purpose but uses different methodological frameworks, priority, operator models and performance shaping factors. HRA methods and simulation tools, proposed over the years, have not always been particularly useful to the purpose for which they were developed.

The review processes (Griffith and Mahadevan, 2011) demonstrated that HRA criticism may be classified into key issues:

1) lack of empirical data for model development and validation;
2) model’s theoretical basis (including taxonomy and concept’s specificity),
3) definition and use of PSFs with heavy reliance on expert judgment in selecting PSFs, and use of these PSFs to obtain the HEP,
4) large variability in implementation (i.e. HRA parameters are different depending on the method used);
5) HRA quantification.

In particular, the quantification method is weak, and the quantitative results are unsubstantiated since many methods pay attention only to the responses of humans in accident scenarios. Other tools, such as THERP, include levels of detail that may be excessive for many assessments. The existing HRA tools allow very thorough evaluations of human behaviour in high-risk environments but can be resource intensive and time-consuming.

Furthermore, HRA approaches have been mainly developed for high-risk contexts (e.g., aviation or nuclear power plants) wherein only the typical accident scenarios are considered. Methods, as THERP or CREAM, were born as approaches for nuclear power plant, considering only the typical accident scenarios in this context. In the same way the major HRA simulation tools, seen in previous section, are adapted to specific field, such as aviation and control rooms of nuclear power plants. For this reason, the use of methods and simulators in other working areas is strongly restricted. Traditional HRA approaches need several efforts to be applied in different fields such as manual assembly or manufacturing systems or medical context (Schemeleva *et al.*, 2012; Yang *et al.*, 2012). The specificity of the models can also be considered a weak point since it means that they are difficult to be applied to task analysis different from the one they have been developed for. Mosleh and Chang (2004) have analysed the limitations of the existing HRA methods and outlined the guidelines for future methods, emphasizing the importance of having methods that:

- identify human response (errors are the focus),
- estimate response probabilities,
- identify causes of errors to support development of preventive or mitigating measures.
- have explicit role for ‘context’ both in error identification and probability estimation;
- be applicable by different users for different problems;
- be traceable, consistent and repeatable.

I.7 Human error in manual assembly systems

Human error in manual assembly systems affects system reliability, safety and is one of the most important causes of quality defects. The assembly process is often the final stage of the production process, which implies that the products have a lot of accumulated value hence making errors is expensive at this point that the products have a lot of accumulated value hence making errors is expensive at this point of the product life-cycle (Claeys *et al.*, 2015). In particular, the major part of the active manufacturing workforce is currently involved in assembly line systems (Claeys *et al.*, 2015). Assembly errors are associated with worker’s capabilities such as knowledge and skills; psychophysical fatigue; task parameters such as workload and repetitiveness; and the work environment (Elmaraghy, Nada and Elmaraghy, 2008). The occurrence of human errors in manual assembly line can be affected by several factors, such as (Mura, Dini and Failli, 2016):

- <u>Assembly system factors</u>: workplaces with high repetitiveness of tasks, high noise and poor ergonomics can cause both mental and physical stress and reduce the attention of the operator.
- <u>Product factor:</u> over time products with many or similar components can cause an increase in the number of errors; the increasing variety of products was also identified as the main cause of the complexity perceived by an operator in carrying out his tasks.

- Operator factors: worker's memory, mental and physical abilities, skills, training level and experience are some factors that determine the probability of mistakes during the assembly phase.

General error mode includes tasks performed non-sequentially, use of the wrong part/object, application of the wrong force and other types (Michalos, Makris and Chryssolouris, 2013), which can cause accidents, quality defects or delays. Accidents, as the most obvious kind of errors, are easily traceable, however, minor faults can dramatically reduce the operation performance and increase remarkably production time, cost, rework and scrap rate. Moreover, it may cause significant loss in the quality image and global profitability of the company.

The assembly errors and the application of HRA techniques in this field has focused the attention of researchers only in recent years. Most of the papers that presented empirical evidence on the relationship between human reliability and assembly system, focusing on the assembler performance or the application of new and existing HRA techniques to assembly tasks were published in the last 5 years.

Several scholars applied existing HRA technique, such as CREAM (Schemeleva *et al.*, 2012; Yang *et al.*, 2012; Wang, Zhang and Xue, 2014), THERP (Bubb, 2005) and ESAT - Expert System for Task Taxonomy (Kern and Refflinghaus, 2011, 2013, 2017; Neumayr *et al.*, 2015). For example, Schemeleva *et al.* (2012) use CREAM to create a simulation model capable to reproduce a real automobile assembly line with a high degree of details and to assess the qualitative and quantitative lack of operator's assurance. Kern et al. (2011, 2013, 2017), instead, create an assembly specific HRA-model based on ESAT method to evaluate potential human error rates quantitatively in advance associated with cost and time early analysis (Kern and Refflinghaus, 2011, 2013, 2017; Neumayr *et al.*, 2015). This new method allows quantifying potential human error rates in assembly operations before the start of production and it allows comparing planning alternatives under time and cost aspects early.

Original approaches to assess human reliability and quantify human error probability, not considering the typical HRA principle but focusing on some specific features of the manual assembly tasks are also proposed in literature. Baez *et al.* (2014) address operator's failing behaviour and develop a human reliability model by using Cox's Proportional Risk Model. The model describes the behaviour of the rate of human error, considering the effect of the operational environment and the time of the shift, namely the time passed before the errors are reported. The authors showed that cognitive and psychosocial risk factors (stress, motivation, memory and personality) have a significant influence in error occurrences of 120 assembly line operators of an electronic company in Mexico (Baez *et al.*, 2014). Saptari, Leau and

Mohamad (2015) analyse specifically the effect of different parameters, like time pressure, working position, component bin position and gender, obtaining that time pressure is the most significant parameters followed by working position and gender (Saptari, Leau and Mohamad, 2015). Finally, Givi, Jaber and Neumann (2015), instead, developed a new human reliability model that estimates the human error rate while performing an assembly job under the influence of learning–forgetting and fatigue–recovery and can anticipate how and when an error occurs, dynamically measuring the human error rate and reliability with time.

Current state-of-the-art underline that a prospective analysis of human reliability in the manual assembly systems until now has been neglected in literature, even if the variability of human behaviour and worker performance remain a pressing and relevant issue in this field. Nevertheless, the research results showed that HRA methods, both the first generation and the latest dynamic-based ones, which were developed for the high-risk industries, can be applied with success to manufacturing industries and its assembly systems.

I.7.1 Modelling of error consequences on assembly systems

The most serious problem for HRA approaches is the scarcity of empirical data on human performance (including data on basic human error probability and the effect of individual and contextual factors that impact on human performance) for model development and validation (Liu and Li, 2014). A wide range of domains have provided source data for studies of human error: aircraft; nuclear power plants and process control; ships and everyday routines including highly skilled tasks such as typing (Rouse and Rouse, 1983). The manufacturing systems, and in particular assembly systems, have become objects of study and application of HRA techniques only in recent years and for them data collection methods are not always applicable. Data collection methods, in fact, fall into two categories (Bubb, 2005):

- ✓ Directly observable in the human actions: this is possible by observing worker's activities except in artificial experimental situations; however direct observation is usually not available in practice.
- ✓ Indirectly observable in the result: in this case the deviation of result from the demanded quality is assigned as error. In practice an accident is unambiguously the case of exceeding of limit of acceptance. Therefore, accident research is an essential source to get basic data of human error.

Data collection in manufacturing systems to feed human reliability seems to have more severe constraints. The effort to directly collect human data is time-resource consuming and accuracy of the collection method is very difficult to assess. In this field, human error translates and manifests itself not only in the form of accidents, which in fact represent a small portion, but also in different consequences as listed in the proposed taxonomy. Therefore, data sources should belong to different type of human error consequences:

1) **Non-compliants:** is the failure to satisfy a requirement, a need or expectation that can be expressed, implied or obligatory. Non-compliants consist of rejected items and reworks offline.
2) **Incidents and injuries:** derive from the aggregation of human actions with physical system behaviours.
3) **Machine's failures:** are systemic and not random due to design errors, omissions, wrong applications, but also to improper operation, incorrect use of the equipment, and more others.
4) **Machine's slowdowns or delay:** deceleration caused by re-working in-line or increase of processing times.
5) **Latent errors and near-miss:** the first may lie dormant within the system for a long time, only becoming evident when they combine with other factors to breach the system's defences (Reason, 1990); whereas near miss event is a potential hazardous condition where the accident sequence was interrupted (Andriulo and Gnoni, 2014).

The first two classes are more easily measurable and attributable to man. There are, in fact, several methods in the literature developed to distinguish causes of accidents between human error or technical and organizational factors as well as the management of non-compliance allows to know the causes of scraps or reworks offline. The other classes, instead, are more complex to identify, quantify and above all attribute to operator.

I.8 Human error in healthcare systems

The activities carried out in the healthcare sector are characterized by a strong human component: human operator can make a mistake that, in this case, has both a social implication, from the point of view of patient safety, and economic, from the point of view of the costs generated (Cuschieri, 2000).

The consequence of an error is of crucial importance and the spectrum varies from no consequence to serious and fatal. For example, in endoscopic surgery, a surgeon may exert too much tenting force during use of an electrosurgical knife with inevitable follow- through of the hook knife once the tissue is cut. It is a matter of luck, where the hook knife stops or impinges - mid-air (no consequence), into bowel or large vessel (serious consequence). Thus, avoidance of all errors underlies safe execution and, in this respect, inconsequential errors in surgery are rare. Therefore, it is essential the use of human error estimation techniques as a tool for risk analysis and decision support. The strongest influence of HRA approaches has been on the analysis of serious clinical incidents in healthcare, which have drawn on the critical incident technique, root cause analysis and other methods (Lyons *et al.*, 2004). In the last few years, however, there has been growing interest in a wider range of safety and reliability techniques used in other industries.

Lyons *et al.* (2004) performed a full literature review of HRA techniques in healthcare. This produced a brief list of fourteen primary HRA techniques (Table I.3), which have either had practical application in healthcare or which were well-established elsewhere and had potential application. Most of these techniques are based on an initial task analysis and a task simulation to identify a list of the potential errors that could occur associated with this task. Quantification is usually based on either fault trees or event trees, which provide the basis for quantification.

Table I.3: *HRA methods applied in healthcare sector (Lyons et al., 2004).*

TECHNIQUE	DEFINITION	APPLICATION HEALTHCARE
Absolute Probability Judgement (APJ)	Experts provide for their judgement on the likelihood of specific human error and this information is gathered mathematically for inter-judge consistency.	None reported in healthcare.
Barrier Analysis	Barrier analysis is used to examine the defences and controls that have been put in place to protect something or someone from harm, their effectiveness and suggestions for improvements.	-
Change Analysis	Tool used to analyse the effect of process changes, considering the differences between normal practice and incidents.	Applied to the process of care that leads to patient incidents.
Cognitive Reliability and Error Analysis Method (CREAM)	CREAM puts emphasis on defining and analysing the causes of human errors. The theoretical background of CREAM is the classifications of error modes and elements of humans, technology, and organization.	Not yet applied in healthcare.
Event Tree Analysis (ETA)	An event tree is a tree-like diagram that splits according to escalation and recovery events as well as an operator's choices between responses at each stage. Usually	Ambulance treatment of patients with suspected

TECHNIQUE	DEFINITION	APPLICATION HEALTHCARE
	the probability of given branches is calculated providing the expected probability of each outcome.	
Failure Modes Effects Analysis (FMEA)s	A FMEA is a systematic method of identifying and preventing product and process problems before they occur. This involves using a team of multidisciplinary experts to evaluate the process, what failures could occur and the severity and probability of the effects and what actions can reduce these effects.	Reducing risk in blood transfusion; Intravenous drug infusions; improving a drug distribution system; drug prescription in wards.
Fault Tree Analysis (FTA)	A fault tree is a tree diagram using AND/OR logic which is used to examine how an incident occurred or could occur due to contributing factors and events.	Potential exposure risk for radiotherapy staff; medication error; medical device failure.
Hazard and Operability Analysis (HAZOP)	HAZOP involves a team of multi-disciplinary experts evaluating processes using the application of guidewords – such as “task not done”, “task done too late”, “task done too much”.	Medical imaging; cervical screening.
Human Error Assessment and Reduction Technique (HEART)	HEART is used to quantify error probability by applying weighting factors associated with error producing conditions to the relevant generic error probability associated with the types of task being examined.	Widely used in industry but not yet applied in healthcare.
Influence Diagrams Analysis (IDA)	Influence Diagrams are a means of modelling and quantifying the effects of a number of contributory factors and human actions on outcome.	Medical decisions, surgical problem solving.
Management Oversight Risk Tree (MORT)	MORT involves the applications of a toolbox approach to analyse incidents in terms of the adequacy of the safety management measures already in place. This involves the use of a fault-tree like structure to look at what happened, why it may have happened then examines these concepts in terms of systems and	Not yet applied in healthcare.

TECHNIQUE	DEFINITION	APPLICATION HEALTHCARE
	organizational failures and precursor events.	
Paired Comparisons	This is similar to the absolute probability judgement except the experts are provided with task descriptions with known error probabilities to use as a baseline	None used in healthcare.
Systematic human error reduction and prediction approach (SHERPA)	SHERPA is a comprehensive technique involving task analysis. SHERPA identifies error modes. (not done, partially done, too little) and "psychological error mechanisms" – the thought processes that may fail or lead to errors, potential for recovery from error, the consequences of error and error reduction strategies.	Errors in endoscopic surgery.
Technique for human error rate prediction (THERP)	THERP is a total methodology for human reliability analysis – from task analysis, development of event trees to error quantification. Like HEART, for quantification, this involves the use of nominal human error probabilities adapted by the relative effects of Performance Shaping Factors to determine success and failure probabilities as well as looking at the effect of recovery effects.	Widely used in industry but not yet applied in healthcare.

Most analyses have gone little further than the relatively simplistic incident decision trees. HEART and THERP, for instance, are both well-validated error analysis and quantification techniques and whilst they have been primarily applied in the nuclear industry, the detailed level of behaviour that they have considered makes them at least conceptually useful to apply in healthcare. HEART uses an estimation of error based on the familiarity and complexity of the task modified by estimates of the influence of "error- producing conditions" such as time shortage, stress or ambiguity in the required performance standards. Although many of HRA techniques (e.g. THERP & HEART) rely on expert judgement to assign probabilities of error to the task being performed, it has been found that the reliability and accuracy of these judgments made by trained human factors personnel is incredibly accurate. These methodologies take performance shaping factors into account. There are situational, contextual or environmental factors that may impact on an individual or system and make errors more or less likely to occur. Onofrio, Trucco and Torchio (2015) developed an ad hoc taxonomy of Influencing Factors for surgery (Table I.4).

Table I.4: *Influencing Factors in Surgery Applications.*

Influencing factors	Valence
Noise & back- ground talk not related to the task	+/-
Safety Culture and Safety Climate	+/-
Standardization	+/-
Equipment, HMI and space design	+/-
Communication and team- work	+/-
Experience and Team Training	+
Fatigue	-
Leadership	+/-
Staffing and team member familiarity	+/-
Workload	-

HRA techniques might be used, for instance, in the design of surgical instruments; in decisions about the labelling of dangerous drugs; in designing a system of double checks for drug administration; in the design of work processes such as booking appointments or patient flow in Accident and Emergency; in identifying the factors that lead to high stress and liability to error in clinicians; and in the analysis of the range of factors involved in a serious incident and in the subsequent implementation of safety solutions across a clinical department or healthcare system. Specific applications are proposed in literature. Cox, Dolan and Macewen (2008) describe the application of HRA as a tool to quantify errors that occur during small incision cataract surgery. Malik, White and Macewen (2003) describe the nature of active skill-based errors occurring in endoscopic dacrocystorhinostomy surgery. A human reliability analysis methodology was used to assess surgical error from observational capture data. The breadth of application of HRA techniques suggests that the potential application of these techniques is very wide, encompassing design of equipment and procedures, organization of work processes, the manner in which tasks are carried out and the wider, less obvious, factors that contribute to error and patient harm.

I.9 Human error in industrial maintenance

The maintenance process is essential for a safe and reliable system and efficient performance of devices in different work environment, such as nuclear power plants, aviation, chemical plants, offshore facilities, manufacturing systems or other type of industries. Dhillon and Liu (2006) reported the impact of human errors in maintenance as found in the literature as a pressing problem. In fact, although the equipment reliability has significantly improved, and the processes are becoming more and more automated, yet human factor continues to be fundamental and maintenance tasks, that are expected to be perfect, are vulnerable to human error. Human error in maintenance tasks may lead to incorrect decisions, actions, or checks and it is influenced by a variety of individual and contextual factors with a wide variability in the success of interventions. There are several reasons for the occurrence of human error in maintenance, like (Dhillon and Liu, 2006):

- ✓ Complex maintenance task;
- ✓ Inadequate or improper work tools;
- ✓ Poor equipment design;
- ✓ Poorly written maintenance procedures;
- ✓ Poor work layout;
- ✓ Outdated maintenance manuals;
- ✓ Fatigued maintenance personnel;
- ✓ Poor job environment (e.g., lighting, humidity, and temperature);
- ✓ Inadequate training and experience.

Bao and Ding (2014) show that, from maintainers perspective, HE number accounts to 91% and the most significant types of error are inspection and installation of system components. For these reasons, the assessment of the likelihood of human error is essential in maintenance field. The type of human error, its consequences, the main individual and contextual factors and their impact has been investigated trough a Systematic Literature Review (SLR) concerning human error in maintenance (HEM), following the guidelines presented by Neumann et al. (2016) and Pires *et al.* (2015). Systematic literature reviews aim at structuring a certain research area and synthesizing research findings, following a clearly defined, rigorous and reliable approach that allow presenting objective and reproducible results (Hochrein and Glock, 2012). This search aims to identify peer-reviewed papers that presented evidence on the relationship between human performance and maintenance activities. Four research questions were addressed in this study: (1) What are the industrial sectors mainly investigated in the field of interest? (2) What are the main causes and contributing factors that lead to human error in maintenance? (3) What are the main HEM consequences? (4) How HE is evaluated and integrated within the maintenance management? The SLR was carried out through the listed steps below:

- **Identification of research databases and keywords definition:** two scientific databases (Scopus and Web of Science) were used and a set of keywords, structured in two distinct groups, was prepared for these databases: Group A, which includes “human error”, “human reliability analysis”, “human reliability assessment”, “human error probability”; and Group B which includes “maintenance”. The final keywords list used to search consisted of all possible combinations of keywords from Groups A and B using the Boolean operators.

- **Literature search and paper selection through specific exclusion criteria:** only articles in English and published in peer-reviewed journals or conferences between 1997 and 2017 were screened. After running the search on the two databases, all papers were uploaded into a database manager (i.e., Mendeley) and all duplicates were removed. The selection process was divided into two phases. The first selection phase was the reading of the title, abstract, keywords. In this screening stage, articles were classified as included, excluded and undefined according to the specific exclusion criteria described below:

✓ No full text is available;

✓ Articles presenting only one of the main key concepts (maintenance and human error);

✓ Papers do not establish a link between maintenance and human error.

✓ HEM is a secondary aspect than the main purpose of the paper.

- The second stage included the reading of the full text of the papers selected in the previous stage and therefore a definitive assessment based on the 2nd, 3rd and 4th exclusion criteria.

- **Analysis process and information extraction strategy:** the analysis was performed using a pre-determined systematic methodology based on specific criteria to extract and structure the information:

✓ industrial sectors;

✓ methodologies for HE analysis;

✓ types and typical HEs in maintenance;

✓ error contributing factors;

✓ maintenance policies;

✓ maintenance error consequences.

The total number of studies resulted from the database search was 576. After the first screening stage, 120 articles were identified as relevant. Among them, 63 papers were selected after the second screening stage. Table I.5 reports the full list of the 63 selected papers.

Table I.5: *Papers selected trough the Systematic Literature Review concerning human error in maintenance.*

ID	REFERENCE	ID	REFERENCE	ID	REFERENCE
1	(Aalipour, Ayele and Barabadi, 2016)	22	(Hameed, Khan and Ahmed, 2016)	43	(McDonnell *et al.*, 2015)
2	(Abbassi *et al.*, 2015)	23	(Hayama *et al.*, 2011)	44	(Mc Leod and Rivera, 2009)
3	(Achebo and Oghoore, 2010)	24	(Heo and Park, 2010)	45	(Mc Leod and Rivera, 2011)
4	(Asadzadeh and Azadeh, 2014)	25	(Hobbs and Williamson, 2002)	46	(Mc Leod and Rivera, 2013)
5	(Bao and Ding, 2014)	26	(Hobbs and Williamson, 2003)	47	(Nicholas, 2009)
6	(Bao *et al.*, 2015)	27	(Hobbs, Williamson and Van Dongen, 2010)	48	(Noroozi *et al.*, 2013)
7	(Bozkurt and Kavsaoglu, 2010)	28	(Islam *et al.*, 2016)	49	(Noroozi, Khan, *et al.*, 2014)
8	(Carr and Christer, 2003)	29	(Islam *et al.*, 2017)	50	(Noroozi, Abbassi, *et al.*, 2014)
9	(Castiglia and Giardina, 2013)	30	(Khalaquzzaman et al. 2010a)	51	(Okoh, 2015)
10	(Chen and Huang, 2013)	31	(Khalaquzzaman et al. 2010b)	52	(Papic and Kovacevic, 2016)
11	(Chen and Huang, 2014)	32	(Khalaquzzaman *et al.*, 2011)	53	(Rankin *et al.*, 2000)
12	(Chiodo, Gagliardi and Pagano, 2004)	33	(Kim and Park, 2008)	54	(Rashid, Place and Braithwaite, 2013)
13	(Chiu and Hsieh, 2016)	34	(Kim and Park, 2009)	55	(Rashid, Place and Braithwaite, 2014)
14	(Dhillon and Kirmizi, 2003)	35	(Kim and Park, 2012)	56	(Razak, Kamaruddin and Azid, 2008)
15	(Dhillon and Liu, 2006)	36	(Kovacevic *et al.*, 2016)	57	(Sheikhalishahi, Azadeh, et al. 2016)
16	(Dhillon and Shah, 2007)	37	(Kumar and Gandhi, 2011)	58	(Sheikhalishahi, Pintelon, et al. 2016)
17	(Dhillon, 2009)	38	(Kumar, Gandhi and Gandhi, 2015)	59	(Sheikhalishahi et al. 2016)
18	(Dhillon, 2014)	39	(Latorella and Prabhu, 2000)	60	(Singh and Kumar, 2015)
19	(Emami-Mehrgani *et al.*, 2016)	40	(Lawrence and Gill, 2007)	61	(Su, Hwang and Liu, 2000)
20	(Geibel, Von Thaden and Suzuki, 2008)	41	(Liang *et al.*, 2010)	62	(Wang and Hwang, 2004)
21	(Gibson, 2000)	42	(Lind, 2008)	63	(Zhou *et al.*, 2015)

Different industrial sectors were identified showing that most of the papers are related to aviation (38%), nuclear industry (24%) and oil and gas offshore facilities (11%).

Various methods and approaches to measure human reliability or human error were found. Most of them are based on the HRA theoretical principles, which aim to identify the causes and sources of human errors and to pursue quantitative HEP estimates during professional activity. For example, Islam *et al.* (2017) developed a monograph for assessing the likelihood of human error in marine operations that can be applied for instant decision-making. Kim and Park (2012) introduced human error analysis procedures for a predictive HE analysis when maintainers perform test or maintenance actions based on a work procedure or work plan. Each procedure consists of three steps: analysis of basic error potential, evaluation of possible impacts on the system, and identification of deficient work context or PSFs. Noroozi, Khan, *et al.* (2014) presented a revised version of HEART methodology to assess the effects of cold on the likelihood of human error in offshore oil and gas facilities. Instead, other papers applied the existing HRA techniques to real case studies for estimating human error probabilities, validating their consistency through the comparison of the obtained results (Castiglia and Giardina, 2013; Aalipour, Ayele and Barabadi, 2016) or integrating the HEP estimate within maintenance management methodologies (Asadzadeh and Azadeh, 2014; Bao *et al.*, 2015; Hameed, Khan and Ahmed, 2016; M Sheikhalishahi *et al.*, 2016). For example, Abbassi *et al.*(2015) integrated the Success Likelihood Index Method (SLIM) with the Technique of Human Error Rate Prediction (THERP) for the HEP assessment in an offshore condensate pump maintenance task; whereas (Aalipour, Ayele and Barabadi, 2016) compared three common HRA methods (HEART, SPAR-H and Bayesian Network) during the maintenance tasks in a cable manufacturing company in Iran. Among these 22 papers, the most common HRA methods are: SLIM (36%), THERP (23%), HEART (23%) and the Bayesian Network (3%).

Other methodologies, not based on HRA principles, were developed over the years in order to quantify and integrate HE in maintenance management. Carr and Christer (2003) analysed the delay-time modelling of inspection maintenance, incorporating HE existence in the form of fault injection and evaluating HE impacts on system reliability or maintenance decisions. Chiu and Hsieh (2016) established a new analytic process for investigating latent human error and provided a strategy for analysing human error using fuzzy TOPSIS. Kumar and Gandhi (2011) applied graph theory for quantifying HE in maintenance activities modelling HE influencing factors and their interactions/ interrelationships based on a fuzzy cognitive map methodology.

From analysis of methodologies is evident that work environment, organization and individual features are considered as the major contributors to human error and are the key factors in analysing performance of maintainer. HRA methods use the PSFs for enhancing or degrading the HEP (Aalipour, Ayele and Barabadi, 2016; Hameed, Khan and Ahmed, 2016; Islam *et al.*, 2016, 2017), whereas the other methods consider these factors as HE influencing or contributing factors (Gibson, 2000; Dhillon, 2009). Moreover, SLR results underline that accidents are the most evident HEM consequence in terms of safety, while other consequences were not deeply analysed in literature respect to the modelling of HE and contributing factors. The human performance, in fact, can affect also the system reliability, the frequency of maintenance interventions, and the length of intervention time. For example, only Dhillon et al. (2003, 2007) considered the impact of human errors on the system availability and on the probabilities of system being in unsafe working states; while (Achebo and Oghoore, 2010; Azadeh, Asadzadeh and Seif, 2014; Bao *et al.*, 2015) evaluated how HE can impact on system reliability.

The SLR results provide for a wide overview in the field of interest shedding light on relevance of considering HEM and its non-negligible effects on the systems.

I.10 Impact of ageing on human error in manufacturing systems

Population ageing is acknowledged as a global trend, and this trend affects the working population. Increasing longevity and declining fertility rates are shifting the age distribution of populations in industrialized countries toward older age groups (Anderson and Hussey, 2000; United States General Accunting Office, 2003). The International Labour Organization (ILO) has estimated that by the year 2025, the proportion of individuals over the age of 55 years will be 32% in Europe, 30% in North America, 21% in Asia, and 17% in Latin America (Ilmarinen, 2001). This demographic change has a significant impact on various dimensions of society, including the available workforce age structure. In Europe, the working population age trends indicate that the oldest age group (55–64 years) will expand by about 16.2% (9.9 million) between 2010 and 2030, whereas all the other age groups show a declining trend (Fritzsche *et al.*, 2014; Boenzi, Digiesi, *et al.*, 2015; Kenny *et al.*, 2016). For this reason, currently great attention is being paid to the age from scientific community, policy-makers and business leaders (Harper and Marcus, 2006; Ilmarinen, 2006; Thun, Größler and Miczka, 2007; Silverstein, 2008; Backes-Gellner, Schneider and Veen, 2011; Boenzi, Mossa, *et al.*, 2015; Börsch-Supan and Weiss, 2016).

Older workers have more serious, but less frequent, workplace injuries and illnesses than younger ones. Ageing is, also, associated with a progressive decrement in various aspects of human capabilities (motor, cognitive and sensory aspects), that may lead to increase of human errors during the working activity. However, older workers can often compensate for age-related losses with relatively age stable strategies and skills related to their experience, expertise, or learning ability.

The impact of age on HF capabilities has been quantified, as reported in many studies in the psychological, gerontological and medical disciplines that examine how various abilities change over an individual's lifetime (Chan, Tan, & Koh, 2000; Crawford, et al. 2010; De Zwart, Frings-Dresen, & Van Dijk, 1995; Salthouse, 2010; Salthouse, 2012; Shephard, 1999; Silverstein, 2008). From a general psycho-physiological perspective, ageing means a progressive and universal deterioration of the various physiological systems. Changes in physical work capacity have often concentrated on the cardiovascular and musculoskeletal systems, body structure, and some important sensory systems (Ilmarinen, 2001). Studies about ageing workers demonstrated that the functional capacities, mainly physical, show a declining trend after the age of 30 years, and the trend can become critical after the next 15–20 years, so that from 45 to 64 years old there is a significant decrease of their capacities, both physical and cognitive ones.

To date, the ageing theory has been widely discussed in the literature from a physiological point of view. But little is known about the impact of age on HE, despite the inevitable role of ageing workers and human errors in manufacturing systems. In fact, chronological age impacts on human reliability and the occurrence of human errors strongly influences system reliability and safety, equipment performance and economic results. A systematic literature review, following the guidelines presented by Neumann et al. (2016) and Pires *et al.* (2015), has been conducted using three scientific databases (Scopus, Web of Science and Engineering Village) to identify peer-reviewed papers that presented evidence on the relationship between ageing and human performance in manufacturing systems. To investigate this relationship, it is necessary to consider research from several disciplines. A set of keywords was prepared for the databases (Table I.6). Groups A, B, C and D list keywords related to age, human error, industry type and human field, respectively. The final keywords list used to search consists of all possible combinations of keywords from Groups A, B, C and D using the Boolean operators to make the relationship (AND) and the sum of words (OR) (e.g., Age AND Error AND Manufacturing AND Human).

Articles that had the searched keywords in its title or abstract and were published between 1996 and 2017 were screened. As restrictions, only articles in English, published in peer-reviewed journals or conferences and with available full text were considered. After running the search on the three databases, all articles were uploaded into a database manager (i.e., Mendeley) and all duplicates were removed.

Table I.6: *Set of keywords used in the systematic search of Engineering Village, Scopus and Web of Science.*

KEYWORDS			
A	**B**	**C**	**D**
Age (*)	Error	Manufacturing (*)	Human
Older (*)	Reliability	Industry (*)	Worker
Senior (*)	Failure	Production (*)	Workforce
Elder (*)	Performance	Assembly (*)	Employee
	Slip		Operator
	Lapse		
	Mistake		
	Mismatch		

Papers identified by the systematic review went through two selection processes. The first selection (as a result of reading of the title, abstract, keywords) excluded: a) articles presented only one of the main key concepts (age and human error); b) they did not establish a link between age and human error; c) they were not related to manufacturing environments. The second stage included the reading of the full text and a definitive assessment as function of the exclusion criteria. The references from selected papers were examined as a further source of papers in a "snowball" approach. The selected papers were analysed through a pre-determined framework, to achieve the research objectives, based on these specific criteria to extract and structure the information:

- ✓ Publication year;
- ✓ Type of contribution (Development of method/methodology/model, State of the art, Proposition of framework, Other type of contribution);
- ✓ Research method (Experimental Research; Simulation, Case study, Literature review, Other type of research method);
- ✓ Demographic features;
- ✓ Type of human error analysed;
- ✓ HF capabilities link to age and human error;
- ✓ Age effect on system performance.

The database search for the systematic literature review resulted in 6521 possible articles. First and second screening and the additional snowball searching led to a final set of only 21 studies. Two articles are conference proceedings, whereas the others are published on scientific journals in the engineering, medical and social sciences areas. The final set of papers includes empirical studies and in-depth literature review that aim to establish a link between age and human error. The limited number of selected studies and empirical data in literature is due to the main challenge related to the study of the link between ageing and human error in working context. In fact, human errors are hard to measure directly in manufacturing contexts, because they may cause a quality defect, a productivity loss but also a latent error, which is complex to identify. Furthermore, HEs are influenced by many individual and contextual factors, that may additionally modify the assessment of ageing impact. True experimental work on ageing is not possible because age levels cannot be manipulated. Sophisticated theoretical frameworks and modelling techniques are required to reach valid inferences about age effects and age changes.

The analysis process provided for evidence of age and human error relationship. The main results show a significant correlation between the human error rate and the operator's age; such correlation is a function of the psycho-physical workload (Börsch-Supan & Weiss, 2016; Fritzsche, et al., 2014; Haji Hosseini, et al., 2012; Pennathur, et al., 2003). The analysis of papers shows that age is highly associated with HE in a way that no simple linear decreasing effect exists, and a variety of mediating factors come into play. This outlines the relevance of considering the non-negligible effects of ageing workforce on system performance. Several age-related HF capabilities (vision and hearing loss; decrement of working memory, attention, reaction and response time; physical decline) that affect worker performance and its reliability, have arisen from the systematic review. However, the decrease of HF capabilities is sometimes compensated by the experience, that allows to better manage the performed tasks and reduce the number of human errors (Mehta and Agnew, 2010).

Furthermore, human errors, due to ageing impact, affect industrial operations in terms of safety and system performance. In particular, eleven of the total paper describe the impact of human error on system performance (productivity, quality, efficiency), while 7 papers address the safety issue (occupational accidents, slips, trips and falls) with reference to age.

Chapter II: Break scheduling management

II.1 Introduction

Introducing breaks is a key intervention to provide recovery after fatiguing physical work to prevent growth of accident risks during working activities and improve human reliability (Dababneh, Swanson and Shell, 2001; Jansen, Kant and van den Brandt, 2002; Demerouti *et al.*, 2012). The selection of adequate work-rest policies through the introduction of appropriate breaks is a very efficient approach even if not very applied, because it is well-known that work–break configurations influence the performance of individuals and can result in different productivity levels for individuals engaged in either mental or physical tasks (Bechtold, Janaro and Sumners, 1984).

This chapter analysis the role of break in working field, giving an overview of the work-break literature, considering the impact of breaks on human performance (well-being, recovery, and risk) and the break scheduling problems.

II.2 Psycho-physical effects of continuous work

Irregular or continuous working hours can have negative consequences for human health and well-being due to stress that interferes with psychophysiological functions and social life. Continuous work has several negative consequences, such as physical and mental fatigue, health problems, stress, decreased concentration resulting in reduced productivity and an increased risk of accidents and injuries in the workplace. The shift scheduling and the conditions of the working environment influence many aspects of human family and social life, determine the daily habits and rhythms, modify the biological clock and can ultimately generate problems, such as sleep disturbances, which hinder the natural recovery process. In the short term, individuals may experience symptoms similar to jet lag, such as fatigue, insomnia and difficulty falling asleep, as well as gastrointestinal malfunction, reduced mental abilities and performance efficiency. In the longer term, rhythmic disorders may eventually translate, often in combination with other factors, in the manifestation of a wide range of disorders and diseases (Knutsson, 2003; Rouch *et al.*, 2005). Working for prolonged periods of time subjects the human body to excessive stress and exposes the operator to an elevated risk of disturbances to the circulatory system or heart disease.

The work performance decrements have generally been attributed to the concept of fatigue or stress at work. A tired or stressed employee is also an unreliable employee. Fatigue can be considered as a global concept, which may take several forms including sleepiness as well as mental, physical and/or muscular fatigue depending on the nature of its cause and can be defined as "a biological drive for recuperative rest" (Williamson *et al.*, 2011). The muscular fatigue has been linked to the decline of performance, the increased reaction times, the slowing of the sensory abilities and the reductions in motor control and force fluctuations (Perez *et al.*, 2014), while the mental fatigue results in a high psychological discomfort (nervousness, tiredness, dizziness and headache).

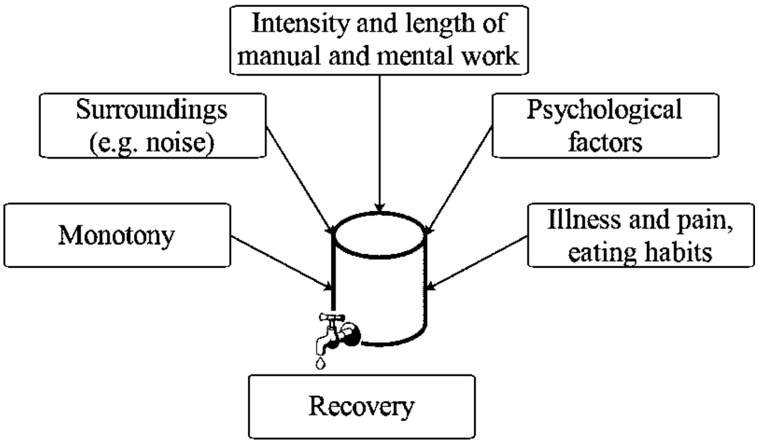

Figure II.1: *Fatigue model of Grandjean (1968).*

Fatigue induced by work derives from prolonged activities, but also from psychological, socio-economic and environmental factors that influence the mind and body. Grandjean (1968) compares fatigue to the level of a liquid present in a box that is continuously filled by the monotony of tasks, the environment, the intensity and duration of manual and mental work, and by psychological and physical factors, and which can be emptied only from recovery or rest, as shown in Figure II.1.

Pimenta *et al.* (2014) details a non-invasive approach on the monitoring of fatigue of a human being, based on the analysis of the performance of his interaction with the computer. The collected data cover a 30days-period of computer use and for each of them four periods were distinguished to evaluate the effect of circadian rhythm and fatigue in one working day: (1) the start of the day, when the user is mentally fresh; (2) immediately before lunch break; (3) after lunch break; and (4) the end of the day, when the individual is most fatigued. The study was carried out based on the measures of different variables: Time between keys (time span between two consecutive keys); Error for key (error in pressing a key); Mouse acceleration; Mouse velocity; Distance between two consecutive clicks; Click duration; and Average excess of distance.

The analysis of the interaction of each individual with the computer shows that during the day fatigue involves a decrease in efficiency in the use of the mouse and keyboard, as well as a gradual and constant decrease in the speed of the mouse and a consequent increase the temporal distance between two consecutive clicks, as shown in Figure II.2.

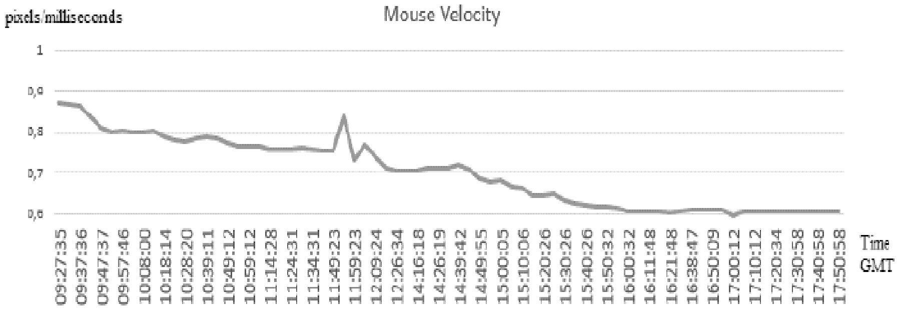

Figure II.2: *Fatigue causes a gradual and consistent decrease in the mouse velocity over the day (Pimenta et al., 2014).*

The potential impact of long work hours on health and safety is a major concern that has resulted in various work hour regulations. Continuous work can, in fact, be associated with specific pathological disorders, such as headache, stomach ache, cardiovascular disease and others that can have a secondary impact on productivity and absenteeism. Furthermore, attention must be paid to work-related stress. According to European Agency for Safety and Health at Work (2009) about 22% of European workers experience work stress, which, is the cause of about 60% of lost working days. A European worker, on average, is absent for about four and a half days a year from work due to health problems (Parent-Thirion *et al.*, 2007).

Likewise, prolonged work-hours are risky for the safety of operators (Tucker, Folkard and Macdonald, 2003; Folkard and Lombardi, 2006). Tucker, Folkard and Macdonald (2003), for example, assess the increase in the risk of injuries because of work sessions of growing length, in which no breaks were made. After first working hours, the relative risk of accidents or injuries in the workplace increases by 33% compared to the first half hour with no risk. This probability increases up to 108% in the case of two hours of continuous work.

The performance decreases, human errors, quality losses and health related problems associated with employee unreliability may translate into huge monetary costs for companies. There are different methods which may be used to improve the human performance and to reduce errors. Many studies have focused on ergonomic interventions for improving musculoskeletal health and postural comfort (Westgaard and Winkel, 1997; Battini *et al.*, 2017) while others have focused on the impact of industrial shift systems with particular attention to long work hours and night shifts (Smith et al., 1998; Åkerstedt, 2003; Folkard and Lombardi, 2006; Caruso, 2014). Rest breaks are a further aspect of considerable importance, as describe in the next section.

II.3 Rest breaks

Jett and George (2003) defined the rest break as “planned or spontaneous suspension from work on a task that interrupts the flow of activity and continuity”. Breaks can be formally planned by organizational practices (e.g., coffee and lunch breaks) or can be informally instituted by workers themselves. It may be noted that the work preferences, related to timing and length of breaks, are not equal for everyone. For instance, some people may schedule breaks at regular intervals throughout the day, whereas others may take breaks at random times throughout the day and follow a configuration of seemingly unproductive days punctuated by a highly productive day (Jett and George, 2003).

Rest periods involve multiple and important positive functions for the person being interrupted, including stimulation for the individual who is performing a job that is routine or boring, opportunities to engage in activities that are essential to emotional well-being, job satisfaction, and sustained productivity and time for the subconscious to process complex problems that require creativity. In addition, the regular breaks seem to be an effective way to control the accumulation of risk during the industrial shift. They are recommended to prevent the accumulation of risk of accidents during the activities supported and results of laboratory tests and field give strong support to these recommendations (Tucker, Folkard and Macdonald, 2003). Nonetheless, they can potentially be disruptive to the flow of work and the completion of a task, because they can result in loss of available time to complete a task, a temporary disengagement from the task being performed, the procrastination (i.e. excessive delays in starting or continuing work on a task) and the reduction in productivity (Jett and George, 2003).

Traditionally, work breaks have been the subject of an exclusively sectoral discipline, marked both by collective bargaining and by some laws aimed at protecting certain categories of workers, such as, for example, video terminal workers, minors and drivers.

The Italian legislator has established with the article n.8 of the legislative decree n.66/2003 a general regulation, that is, a minimum protection threshold valid for all workers, leaving to the collective bargaining the primary and punctual regulation of the temporal, modal and salary profiles of the work breaks. This article stats: "If the daily working time exceeds the limit of six hours, the worker must benefit from a rest break, whose procedures and duration are established by collective labour agreements, for recovery of psycho-physical energies and possible consumption of the meal also in order to reduce the monotonous and repetitive work ". In the absence of collective bargaining, a break cannot be less than ten minutes and it can be undertaken at any time of the work shift. The discipline referred to in article n.8 must necessarily be coordinated with the rules and provisions established to protect certain categories of workers:

- workers, who use video terminals for at least twenty hours per week, are entitled to a 15-minute break every two hours of continuous application to the video (Article175, legislative decree. No. 81/2008);
- children and adolescents cannot work more than four and a half hours without interruption; if this working period is exceeded, an intermediate break of one hour will be mandatory. Collective agreements, subject to authorization by the territorial labour departments, may reduce this break period to half an hour, if these are not unhealthy and dangerous works (Article 20, Law No. 977/1967);
- domestic worker is entitled to a convenient rest period during the day and to no less than eight consecutive hours of night rest (Article 8, Law No. 339/1958).;
- the working time of the personnel involved in the transport of goods or people must be interrupted by 30-minute intermediate rest periods, if the total hours worked are between six and nine hours, 45 minutes if greater than nine hours (Article 5, Legislative Decree No. 234/2007).

II.3.1 Breaks impact on the human performance: well-being, recovery, risk.

Most research on breaks has focused on the long-term consequences of extensive breaks such as sabbaticals (Davidson *et al.*, 2010), vacations (Fritz and Sonnentag, 2006), weekends (Fritz and Sonnentag, 2005; Ragsdale *et al.*, 2011), and evenings (Demerouti *et al.*, 2009). Whereas most studies on daily recovery focus exclusively on the engagement in off-job activities that may reduce fatigue and restore physiological and psychological readiness.

Little is known about recovery from short breaks that occur during the working day. The relatively few studies that directly address breaks indicate that people need occasional changes in the time of work or an oscillation between work and recreation, particularly when they are fatigued or are working continuously for an extended period (Dababneh, Swanson and Shell, 2001; Jett and George, 2003). The lines of research typically examined focus on different aspects as the frequency, the timing, and the length of the breaks**,** or the activity undertaken during the rest period (doing physical exercises, socializing, napping etc.).

Surprisingly, a small number of studies have examined the function of recovery both at work and home. The recovery experience refers to the degree to which individuals perceive that the breaks they take help them to restore energy resources. Demerouti *et al.* (2012) examined the recovery experience after breaks at work and psychological detachment from work when being at home by investigating the role of recovery at work in the process of energy replenishment. The authors distinguish between two types of recovery: recovery during work, which takes place when the stressor factors are present, and recovery after work, which occurs when the stressor factors are absent (Demerouti et al., 2012). All examined relationships are summarized in Figure II.3.

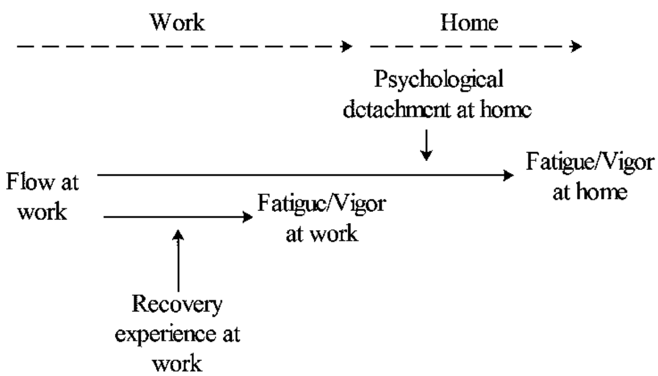

Figure II.3: *Hypothesized relationships (Demerouti et al., 2012).*

Results of the multilevel analysis indicated that recovery at work and detachment from work moderated the relationship between flow (specifically, the enjoyment component) and after-work energy. An association between need for recovery from work, fatigue, and psychological distress in the working population was also observed in (Jansen, Kant and van den Brandt (2002). Need for recovery was higher in men than in women and in the higher age groups, as others have found (Mohren, Jansen and Kant, 2010).

Evidence has recently highlighted that the beneficial effects of rest breaks on strain and mood are influenced by the nature of the activity undertaken during the breaks (Tucker and Folkard, 2012). Experimental field studies found that rest breaks were more likely to enhance subsequent mood if they involved respite activities (e.g., napping, relaxing, socializing) rather than chores (e.g., working with customers, running errands, and work preparation) (Tucker and Folkard, 2012; Mathiassen *et al.*, 2014).

The primary domain for exploring the benefits of within-day work breaks is ergonomics because of its role in preventing musculoskeletal problems, although systematic reviews suggest that there is only limited evidence of their effectiveness in this regard (Brewer *et al.*, 2006; Kennedy *et al.*, 2010).

Many studies have focused on computer-based tasks. Researchers in this area have focused on standard and micro breaks as a means to alleviate musculoskeletal discomfort and strain associated with prolonged or repeated office-related tasks. Galinsky *et al.* (2000), Mclean *et al.* (2001), Balci and Aghazadeh (2004) found positive effects depending on the time between rest breaks and musculoskeletal outcomes. In Mclean's. (2001) study, the authors examined the benefit of micro breaks by investigating myoelectric signal behaviour, perceived discomfort, and worker productivity while individuals performed their usual keying work. Participants were randomly assigned to one of the three experimental groups: micro breaks at their own discretion, micro breaks at 20-min intervals, and micro breaks at 40-min intervals. It was determined, with p-value equal to 0.05, that micro breaks had a positive effect on reducing discomfort in all areas studied during computer terminal work, particularly when breaks were taken at 20-min intervals. Similarly, Balci and Aghazadeh (2004) investigated three different work-rest schedules (60-min work/10-min rest, 30-min work/5-min rest, 15/micro breaks four from each hour in addition to a 14-min break after 2 h) considering two types of task (cognitive task and data entry). The results indicated that the effect of the work-rest schedule was significant on various perceived discomfort categories and the performance of the participants, and the author suggested that the 15/micro break schedule is preferable to the longer and infrequent rest break schedules considering upper extremity discomfort, eye strain, speed, accuracy, and performance of the participants.

Worker productivity takes advantage of short rest breaks. Balci and Aghazadeh (2004) reported that the performance in data entry tasks with the 15/micro break schedule was 18% higher than the 30-min work/5-min rest schedule and 24% higher than the 60-min work/10-min rest schedule. Henning *et al.* (1997) and van de Heuvel, de Looze and Hildebrandt (2003) combined physical exercises with breaks in order to study their effect on human performance. Productivity growth and discomfort reduction were achieved with two 5-min rest breaks with exercises in addition to the normal rest breaks both in the mid-morning and mid-afternoon during an 8-h work day. Exercise breaks also improved workers' well-being and eye, leg, and foot comfort (Henning *et al.*, 1997).

The impact of frequent short rest breaks on productivity and well-being has also been investigated in the manufacturing field. Dababneh, Swanson and Shell (2001) tested two rest break policies, both of which provided 36 minutes of extra break time over the regular break schedule (30-min lunch and two 15-min breaks), in a meat-processing plant. In the first rest break configuration, the further break has been scheduled with 12 breaks of 3-min evenly distributed over the workday (3-min break for every 27 min of work). In the second schedule, workers were given four 9-min breaks evenly distributed over the workday (9-min break every 51 min of work). Results showed that neither of the two experimental rest break schedules had a negative effect on production rate, and the 9-min break schedule improved discomfort ratings for the lower extremities.

Several studies have examined the impact of rest breaks during a shift on injury or accident risk (Folkard and Tucker, 2003; Tucker, Folkard and Macdonald, 2003; Folkard and Lombardi, 2006; Tucker *et al.*, 2006; Tucker and Folkard, 2012). They agree that risk is reduced in the first half-hour following a rest break and that this effect is similar across all three shifts. The number of injuries within each of the four 30-min periods between breaks was calculated, and the risk in each 30-min period was expressed relative to that in the first 30-min period immediately following the break. Results are shown in Table II.1, and it is clear that injury risk rose substantially and approximately linearly between successive breaks such that risk had doubled by the last 30-min period before the next break.

The trends over subsequent half-hours varied, possibly reflecting the extent to which the work was either self-paced or machine paced. It would therefore appear that the beneficial effects of rest breaks may be relatively short lived in at least some work environments. Tucker *et al*. (2006) analyse the trend in work-related injuries in relation to the timing of rest breaks in two separated studies. Risk increased from the first to the second half-hour of continuous work and then remained relatively constant in the third half-hour. In some data, there was also a decrease in risk in the period leading up to the end of a work period. There was a sharp decline in reported injuries toward the very end of a shift, but otherwise, the observed trends did not differ between successive periods of continuous work or among morning, afternoon, and night shifts. However, no direct epidemiological evidence exists for the effect of rest breaks on the trend in risk as a function of time-on-task.

Table II.1: *Frequency (% of total per period) of accidents per half-hour for each work period and relative risks for all periods combined (Tucker et al., 2003).*

Time on task (min)					
	0-29	**30-59**	**60-89**	**90-119**	**Total**
Period					
1	23 (13%)	41 (23%)	50 (29%)	61 (35%)	175
2	28 (16%)	30 (18%)	47 (28%)	65 (38%)	170
3	35 (19%)	43 (24%)	50 (28%)	53 (29%)	181
All periods	86 (16%)	114 (22%)	147 (28%)	179 (34%)	526
Relative risk	1 (reference)	1.33	1.71	2.08	

II.4 Break scheduling management

Proper design of work–rest schedule that involves frequency, duration, and timing of rest breaks may be effective in improving workers' comfort, health, and productivity. Break scheduling problems emerge in many working contexts where rest period is indispensable due to features of the tasks to be performed. These features include the requirement of high concentration during extended periods of time, continuous work in front of computer monitors, or other monotonic and exhaustive activities. Typically, break scheduling problems arise in call centres, security checking, or assembly lines. A major problem in this field, from both a research and practical perspective, has been with respect to the appropriate technique for the development of effective work-rest policies that can be described by the number, timing, and duration of rest periods.

In literature, the break scheduling problems have hardly been addressed on their own, but they are part of the most famous shift scheduling problem, which has received a lot of attention in the operations research literature. Shift scheduling problems, in fact, deal with the assignment of employee starting and finishing times, and possibly the placement of relief and meal breaks within each shift in order to maximize work output per unit time or minimize the costs of assigning an employee to alternative shifts (Aykin, 1996; Rekik, Cordeau and Soumis, 2010). The validity of approaches to scheduling breaks development over the years has been limited by the assumption of optimality of complete recovery, exclusion of rest break penalties, or restriction to a single break (Bechtold, Janaro and Sumners, 1984).

The first work-rest model was developed by Eilon in 1964 to determine the optimal length and placement of one break over a finite time horizon for a single employee for a general work rate *r (t)*, which was a decreasing function of time. Gentzler, Khalil and Sivazlian (1977) developed a multirest break model for an infinite time horizon based on the assumption that full recovery was optimal. Starting from this incorrect assumption and assuming linear performance decay during work and linear recovery of work-rate performance potential during rest, the selection of the optimal number, duration, and placement of rest breaks over a single finite time horizon became a mixed-integer quadratic programming problem in Bechtold *et al.* (1984). This model was applied in experimental settings observing productivity improvements of around 8% for a mental task and around 3% for a physical task. Results suggested that it is likely that breaks of a given length may be more effective if taken earlier in the time horizon than when they are evenly spaced. Bechtold and Thompson (1993) extended this earlier research by considering the choices of placement for and during a single rest period that must be taken simultaneously by all employees in a work group through an appropriate model formulated as a mixed-binary, cubic programming problem. Aykin (1996) considered a more general shift scheduling problem with multiple breaks and disjoint break windows and developed an integer programming model for optimal shift scheduling with multiple rest and lunch breaks and break windows, which reduces the number of variables compared to the set-covering formulation, typically used in the scheduling problems. Rekik, Cordeau and Soumis (2010) extended this formulation incorporating two other forms of flexibility: fractionable breaks and work stretch duration restrictions. This provides the possibility of fixing only the total duration of breaks that must be given within a shift without specifying which break length comes in which position. Experimental results prove that using fractionable breaks may yield, for some instances, a considerable saving of workforce.

In addition to the exact methods, the meta-heuristics such as min-conflicts-based local search algorithm, or memetic algorithm have been presented in literature for breaks scheduling. Schafhauser, Musliu and Wild (2009) proposed a memetic algorithm to obtain solutions of improved quality for the break scheduling problem for supervision personnel. This algorithm consists of the selection, crossover, and mutation of three standard operators and is hybridized with a min-conflicts search. Initial solutions are constructed with break patterns already fulfilling constraints representing labour rules and ergonomic criteria. For every iteration, the genetic operators generate a pool of different solutions from the previous generation, and the best solutions are further optimized by the local search procedure. Wild and Musliu (2010) improved the previous method by proposing a new memetic representation, a new crossover and selection operator, and a penalty system that helps to select memes that have a better chance to be improved by a local search. Di Gaspero *et al.* (2010) devised a hybrid strategy that combines a local search method for determining the shifts with a constraint programming model for assigning breaks. This model has shown to be very practical for the local search to find legal break assignments that optimize over/under staffing.

Quantitative models for optimal rest period scheduling were developed with work rate function as basic component. The work rate function defines the performance level from the end of one rest period to beginning of the next rest period, representing the individual fatigued state. The processes of works output decay during work periods and recovery of work rate potential during rest breaks are modelled as linear functions of time (Bechtold *et al.,* 1984).

Therefore, researchers in this field have traditionally concentred on the experimental approach to determine optimal work-rest schedules for specific tasks and under specified environmental conditions, considering human performance in terms of a generic work rate function. They have considered constrains as minimum break time, location of breaks, maximal working time without breaks in order to optimize the number of workers assigned to every shift and their work-rest policy (Aykin, 1996) or to maximize labour productivity, as measured by output per unit time (Bechtold, Janaro and Sumners, 1984). The work rate performance is often modelled as a linear function without a detailed analysis of human reliability trend during the work shift and its qualitative effects on system performance (e.g. non- compliant items and reworking).

None of existing methods, in fact, considers human reliability in assessing worker performance due to the complexity of HRA approaches, as underlined in Chapter I, and given the difficulty of integrating this type of modelling in an exact algorithmic or heuristic technique. Moreover, many of the studies in the literature have addressed the break scheduling problem only from the point of view of productivity. They do not address the problem of break management with regard to the quality aspect, namely the impact of human errors on the system performance in terms of quality of the performed activities (e.g. non-compliant items and reworking). The impact of breaks, in fact, was investigated with respect on the loss of productivity, due to the decrease of work rate, without considering the effect on the human error probability.

Despite continuing advances in research and applications, work breaks are not taken into proper consideration, and there are ongoing efforts to create systems that better manage the business in various areas. The literature review, in fact, has pointed out the almost total lack of systems for the management of work breaks in an automatic manner. The only exception is the software that stimulates workers at video terminals to take frequent breaks and recommend performing exercises during breaks. The validity and effectiveness of this type of software has been demonstrated by several studies (Mclean *et al.*, 2001; van de Heuvel, de Looze and Hildebrandt, 2003). Van Den Heuvel's (2003) study evaluated the effects of work related disorders of the neck and upper limbs and the productivity of computer workers stimulated to take regular breaks and perform physical exercises with the use of an adapted version of WorkPace, Niche Software Ltd., Mclean *et al.* (2001), instead, examined the benefits of micro-breaks to prevent onset or progression of cumulative trauma disorders for the computerized environment, mediated using the program Ergobreak 2.2.

Chapter III: Simulator for Human Error Probability Analysis: theoretical framework

III.1 Introduction

HRA literature highlights the severity of the human unreliability at work and the need to evaluate and to quantify it for reducing the errors and improving the productivity. Despite the impact of human errors in industrial systems and the development of numerous HRA and break scheduling approaches in literature, they have still many limitations, as previously seen. The purpose of each HRA method must be to assess human behaviour and to quantify HEP, in order to reduce and prevent possible conditions of human error in a working context. Existing methods, as previously seen, do not always pursue this aim in an efficient way, but every method or simulator has its own strength. As well as, work breaks are not taken into proper consideration, and there are ongoing efforts to create systems that better manage the break scheduling in various areas, especially in manufacturing.

This chapter presents a new HRA model that exploits the advantages of the simulation tools and the traditional HRA methods to predict the likelihood of operator error, for a given scenario, in every kind of industrial system or other type of working environment. The aspiration for the Simulator for Human Error Probability Analysis (SHERPA) model is not that it be a new HRA method in the extensive list of existing ones, but that it provides a theoretical framework addressing the problem of human reliability in a different way from most HRA methods. SHERPA focuses on the quantitative aspect to obtain a significant numerical result in terms of HEP and combines the HR assessment with the management of work-rest policies. The most important objective of the work has been to realize a model for the evaluation of human reliability that can obtain useful information about human reliability for every kind of work task.

SHERPA can be used in the preventive phase, as an analysis of the possible situations that may occur and for the evaluation of the percentage of non-compliant performed tasks due to human error and in post-production to understand the nature of the factors that influence human performance in order to reduce errors.

Human reliability is estimated as a function of the performed task, influencing factors (PSFs), and time worked, with the purpose of considering how reliability depends on the task and working context as well as on the time that the workers have already spent at their work. Knowing the HR distribution allows intervening from the perspective of reducing errors with re-design tasks or other interventions such as the management of the worker's psychophysical recovery through appropriate break configurations.

The proposed HRA-based model is, in fact, addressed to the break scheduling problems through the hypothesis that breaks allow the mental and physical recovery and lead to improvements of human reliability. The positive break impact on human reliability is a function of break time, location of break during the shift, recovery speed and type of performed activities. Rest breaks have also a negative aspect due to increased idle time that corresponds, for example, to a decrease of productivity in a manufacturing context. For this reason, SHERPA is based on an economic model, that allows to assess both positive and negative break effects and to compare their impact on the system performance, considering for example the cost of lost production due to break and the quality costs related to operator errors. The model can be adapted to alternative set of constraints (minimum number of breaks and minimum time guaranteed by legislation or internal union agreements, maximum hours of continuous work and other possible constraints), assigned in the initialization phase of the system as inputs. SHERPA can then evaluate the effect of every work-rest policy, defined as acceptable for the system under consideration, with the aim of identifying the best configuration among those possible. The model is able to provide for the following functions:

1) Estimating human reliability, as function of time, work context conditions, physical and mental employee conditions and break scheduling.
2) Assessing the effects due to different human reliability levels, through evaluation of processes, activities or tasks performed more or less correctly.
3) Assessing the impact of environment on human reliability, via performance shaping factors.
4) Simulating a large numbers of break scheduling with several locations and duration of breaks, in order to assess their impact of different work-break policies on human performance (HEP and recovery after the break) and the overall system performance in terms of percentage of compliant performed

tasks and economic results (e.g. profits, revenues, quality costs, rework costs and break costs).

5) Determining optimal breaks scheduling, identifying for each case: the number, the location in the shift and the optimal break time.

The proposed model was not created for a particular industry or application and therefore can be easily applied to contexts that vary widely. For example, the module can equally represent manual maintenance activity, manual assembly tasks, medical task in a surgery room etc., by varying the input variables such as performed task, level of contextual factors, or physical and mental employee condition and by modelling the specific system considering all the working context features. Simulators and tools similar to the one proposed do not exist today, either from the theoretical point of view or from the point of view of the analysis carried out.

This chapter presents the SHERPA theoretical framework based on the integration of traditional and dynamic HRA methods. Its logical foundations, the HRA principles and rules, the evaluation and quantification of psycho-physical recovery and the break scheduling management system are described.

III.2 Notation

The following notations will be used in this chapter:

$HEP_{nominal}$:	nominal human error probability.
$HEP_{contextual}$:	contextual human error probability.
HR:	human reliability ($1 - HEP_{contextual}$).
$R(t)$:	nominal reliability function.
k, β, and α:	shape and scale of Weibull distribution.
τ:	length of the transitional phase of human adaption.
$PSF_{composite}$:	performance shaping factors composite.
PSf_x:	assigned multiplier for each PSF.
F_x:	multiplier value of sub-factors.
W_x:	weight of each sub-factor.
THI:	thermohygrometric index.
CLO:	thermal resistance index of clothing.
MET:	index of metabolic activity.
T_a:	environmental temperature.

$TNOW_{deg}$:	current simulation time.
UR:	relative humidity.
$T_{perceived}$:	perceived temperature.
$T_{operating}$:	operating temperature.
E:	illuminance on the horizontal plane.
A, B and C:	constants for different sky conditions.
α_s:	solar height.
E_e:	horizontal surface placed outside, shielded by direct solar irradiation and exposed to the light coming from heavenly vault.
η:	daylight factor.
F_u:	utilization factor.
M:	maintenance factor.
N:	number of light sources.
L:	luminous flux.
A:	area to be illuminated.
LEX, d:	individual exposure level to daily noise.
LEX, w:	weekly individual noise exposure level.
r:	recovery factor.
ω:	recovery rate.
R:	profit.
P:	price/value added of the processing.
CF_{STD}:	standard fixed costs.
CV_{STD}:	standard variable costs.
T_c:	processing time.
C_r:	reworking costs.
c_b:	rest break costs.
T_b:	rest break time.
T_r:	reworking time.
T_{total}:	total time of processing that considered the time increment linked to possible rework.
P_r:	reworking probability.
i:	working cycle index.
q:	number of units worked in every cycle.
T_1:	theoretical time to produce the first unit.
$\hat{T}_1$:	equivalent time for the first unit of the forgetting curve.
LR:	learning rate.
b:	learning slope.
f:	forgetting slope.

x:	amount of output that would have been accumulated if interruption did not occur.
B:	minimum time for total forgetting.
$A(8)$:	equivalent weighted acceleration in frequency referred to eight hours of work.
a_v:	value of the vector sum acceleration of the components detected on the three axes.
T_e:	total daily duration of vibration exposure expressed in hours.
a_{vi}^2:	vector sum of the frequency weighted acceleration relative to the i-th operation.
T_i:	the exposure time relative to the i-th operation expressed in hours.

III.3 Structure and logic of the SHERPA model

The SHERPA theoretical model was developed according to the current state of the art of HRA methods and break scheduling problem. The operator's recovery and the work - break policies management were based instead on the state of the art previously presented.

Three HRA elements converge into the model:

- the task classification in one of the generic tasks proposed by HEART method (Williams, 1985);
- the PSFs analysis of the SPAR-H method (Gertman *et al.*, 2005);
- the dynamic implementation using computer simulation (Boring, 2007).

The SHERPA theoretical framework has been described with the technique IDEF0 (Integration Definition for Function Modelling), which is a widely-used technique for the structured analysis and design of systems developed through the Air Force's integrated computer aided manufacturing program (Presley and Liles, 1995). The four elements (inputs, outputs, controls, resources) to the IDEF0 functional model are shown in Figure III.1, where the activity box is the SHERPA model. Inputs are represented by the arrows flowing into the left-hand side of the activity box, and they are the entities, which equally represent the pieces to be processed or the physical/mental activities to be performed by the worker.

The model reproduces the employee's work during a whole shift, quantifying the reliability and error probability that moved on the outputs of the system, represented by arrows flowing out the right-hand side of the activity box. HEP is estimated here as function of performed task, performance shaping factors and worked time, with the purpose of considering how reliability depends on the task and on the working context, as well as on the time that the operators have already spent at their work. SHERPA determines as outputs the number of compliant, non-compliant, and rework entities. These outputs are calculated considering in a first approximation the contextual HEP, in this way the human error represents non-compliant entities. The concept of quality defects and scraps is not limited to manufacturing processes, but extends to a wider range of working environments, ranging from services to medical field. Further outputs are the HEP distribution and the economic results. The available outputs allow a clear and direct assessment of how the system reacts to change in the given break scheduling, as well as to change in environmental and psychophysical conditions. The arrows flowing into the top portion of the box represent constraints or controls on the activities: the HRA and the recovery principles; the influencing factors (PSF), namely the contextual factors and the physical and mental employee conditions; and the assigned work-rest policy. Finally, the resources, represented by arrows flowing into the bottom of the activity box, are the mechanisms that carry out the activity.

The main activity box has been decomposed into more detailed levels of analysis, through the four sub-models shown in Figure III.2 and analysed in detail in the next sections. The four sub-activities (entities entry, HR quantification, process simulation, entities exit) are supported by the operating logic shown in Figure III.3.

The starting analytical basis for the assessment of human errors in SHERPA is the determination of HEP, followed by quantification of PSF influences on the initial value of HEP. This module receives, as input, the type of activity and generic task, each of which is connected to an appropriate probability distribution that describes HEP as a function of time. Different scheduling of breaks can be assigned and can be simulated in the shift, considering that a break determines the worker's recovery and the consequent increase in reliability.

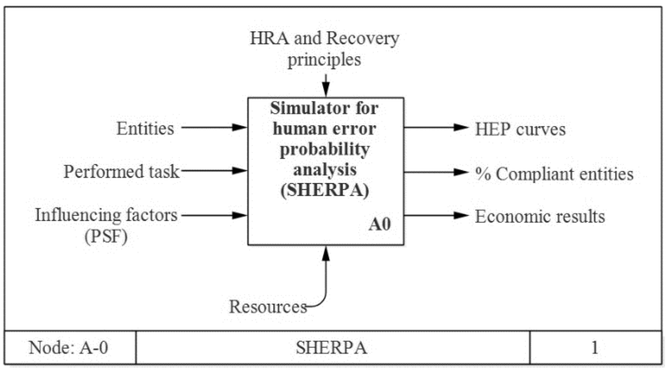

Figure III.1: *IDEF0 representation of SHERPA simulator.*

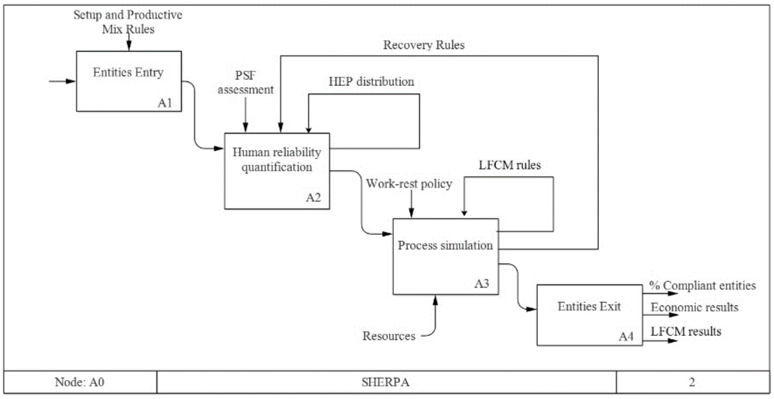

Figure III.2: *SHERPA decomposition overview.*

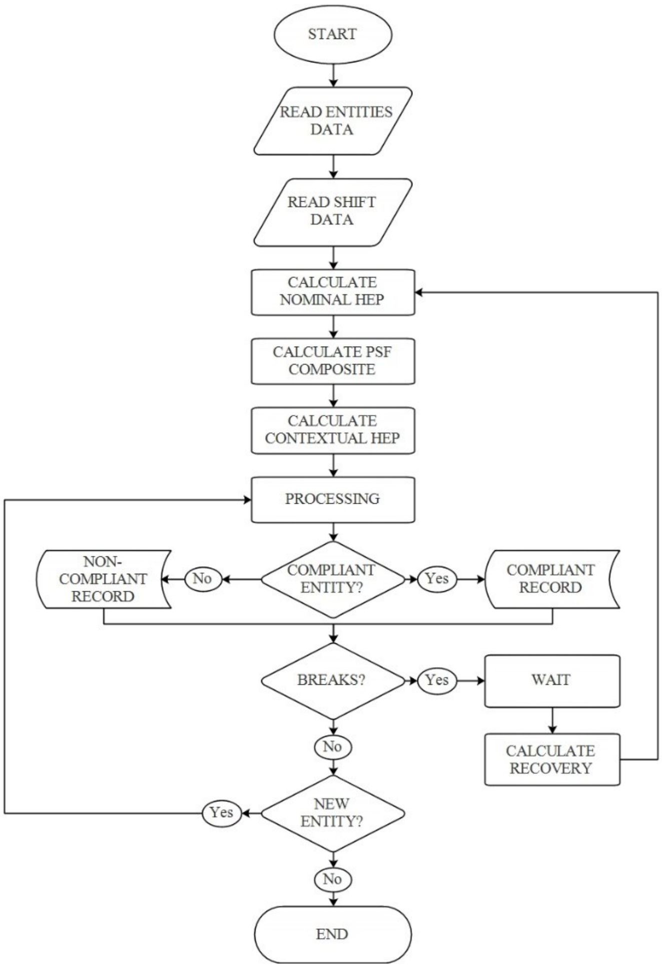

Figure III.3: *Logical architecture of SHERPA model.*

As explained in detail later, the proposed model can manage a pool of break configurations, which are included into three main groups: no break in the shift (continuous working), fixed work-rest policy or automatic management. Furthermore, the Learning and Forgetting Curve Model (LFCM) module allows evaluating the impacts of the learning and forgetting phenomenon on the processing time.

III.3.1 Entities entry

The entities in entrance represent many working contexts because they can equally simulate a work piece, a document to be drafted, or in general, a task to be performed. The model manages in the same manner all the typologies, recognizing in the case of product mix the needs of setup. In this phase, the model follows the flowchart in Figure III.4. A set of technical data (type of performed task, processing time, setup time, time for rework) and economic data (product price, fixed and variable costs) is allocated to each entity in the first step.

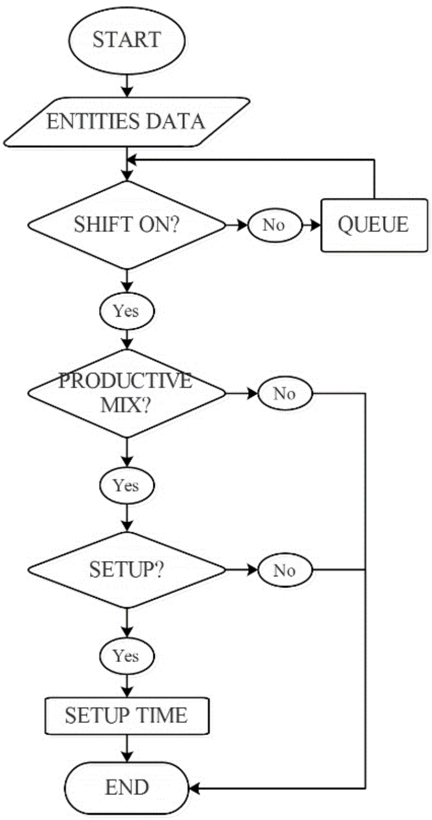

Figure III.4: *Input logic of the entities with and without productive mix.*

III.3.2 Human reliability quantification

The second phase is addressed to the nominal and contextual HEP quantification, which is the first step in any HRA approach, as reported in Figure III.5. The preliminary analysis of the model requires advance knowledge of the probability with which an operator can make mistakes, and therefore assumes probability distributions of HEP as functions of time and type of operation to perform, which describe the variations in human performance. The flowchart illustrates the process of HR quantification and its main phases. Nominal and contextual HEP and the PSF composite are quantified for each entity and are representative of each performed task, as described hereafter.

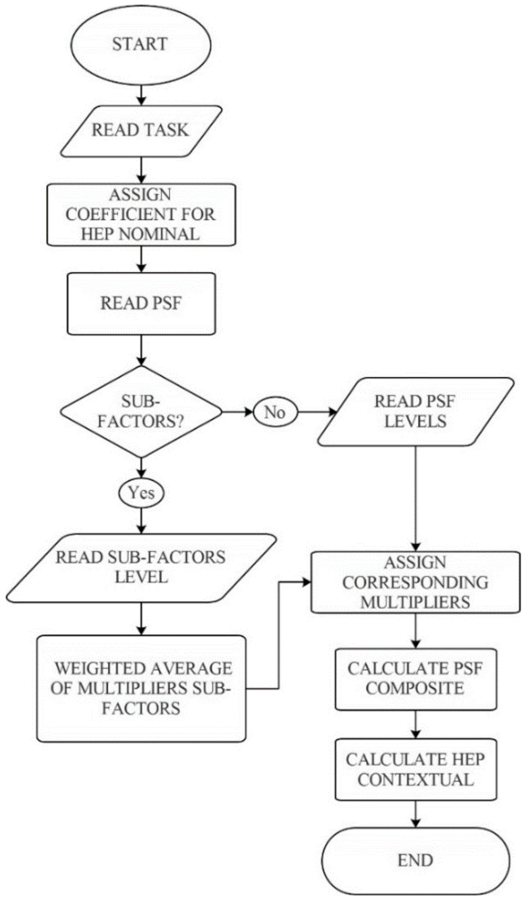

Figure III.5: *The HRA process in SHERPA model.*

The nominal HEP, independent of the presence of influencing factors, is a function of the performed activity and worked time. The Weibull probability distribution is presented by Giuntini (2000) as the best distribution to describe the error probability and to characterize the human reliability process. It is adapted in the proposed model to take into account the natural process of adaptation for a typical human for a given operation that results in a lower reliability in the initial part of the shift as follows:

$$\begin{cases} HEP_{nominal}(t) = 1 - k \cdot e^{-\alpha\cdot(1-t)^{\beta}} & \forall\ t \in [0; \tau] \\ HEP_{nominal}(t) = 1 - k \cdot e^{-\alpha\cdot(t-1)^{\beta}} & \forall\ t \in\ ]\tau;\ \infty[\end{cases} \tag{3.1}$$

where t is the time worked by an employee; k, β, and α change the scale and shape of the curve for the six generic tasks used in the model; and τ is the length of the transitional phase of human adaption. The function has also been assumed to have a minimum value of error probability in τ (imposed as the first hour of processing) and a maximum value at the eighth hour of work during an eight-hour shift (Figure III.6).

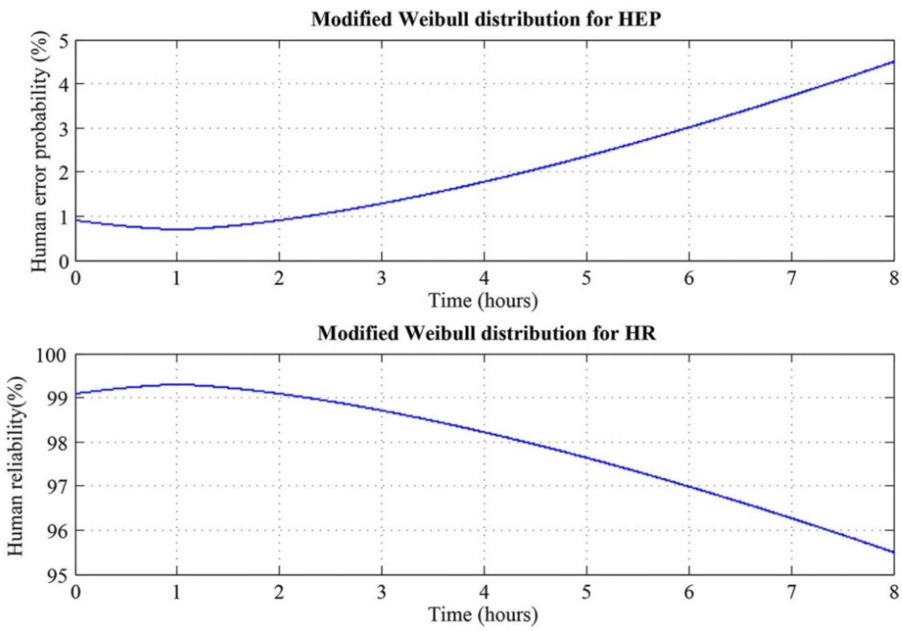

Figure III.6: *Trend in modified human error probability and human reliability.*

As noted above, the first generation HRA methods, such as THERP and HEART, focused on the quantification of nominal HEP. The second generation does not give significant importance to the formal quantification of HEP, but often uses standard values, as in the case of SPAR-H, to allow greater focus on the influence of PSFs. For SHERPA, the best choice was the

HEART method, designed to be a quick and simple method applicable to any situation or sector in which human reliability is important. HEART uses eight generic categories (GT) to classify operator tasks, but only six have been chosen for the proposed model (Kirwan, 1996). The categories shown in Table III.1 can represent a wide range of work activities from simple to more complex ones, from ones with a very high error rates to those more reliable, thanks to the presence of automatic systems of supervision. This range of activities allows the module to apply the model to very different working environments without any kind of restrictions.

SHERPA uses six generic categories to classify the type of performed task, derived by the HEART, and each of them is connected to an appropriate probability distribution that describes nominal HEP as a function of time. For each category, Figure III.7 shows the performance of probability of human error.

Table III.1: *Coefficient values for the six generic tasks.*

Generic task		Limitations of unreliability for operation	k	α	β
1	Totally unfamiliar	35% ÷ 97%	0.65000	0.1660762	1.5
2	Complex task requiring high level of comprehension and skill	12% ÷ 28%	0.88000	0.0108352	1.5
3	Fairly simple task performed rapidly or given scant attention	6% ÷ 13%	0.94000	0.0041785	1.5
4	Routine, highly-practiced	0.7 ÷ 4.5%	0.99300	0.0021068	1.5
5	Completely familiar, well-designed, highly practiced, routine task	0.008% ÷ 0.9%	0.99920	0.0004838	1.5
6	Respond correctly to system command even when there is an augmented or automated supervisory system	0.0001% ÷ 0.09%	0.99991	4.813*10^-5	1.5

Furthermore, according to the SPAR-H method, the tasks are divided into action (implementations of actions / processes simple or complex) and diagnosis (interpretation of system status and decision-making in case of need). The working context and employee state, instead, are taken into account through the PSFs of the SPAR-H method. The performance shaping factor is determined by the individual characteristics of the human being, the environment, the organization or the activity that enhances or decreases human performance and increases or decreases the likelihood of human error. PSFs allow all environmental and behavioural factors that affect human performance to be taken into account. While many HRA methods have often proposed numerous PSFs, even as many as fifty, SPAR-H attempts to provide a reasonable coverage of the influence spectra of human performance in a reasonable minimum number of PSFs.

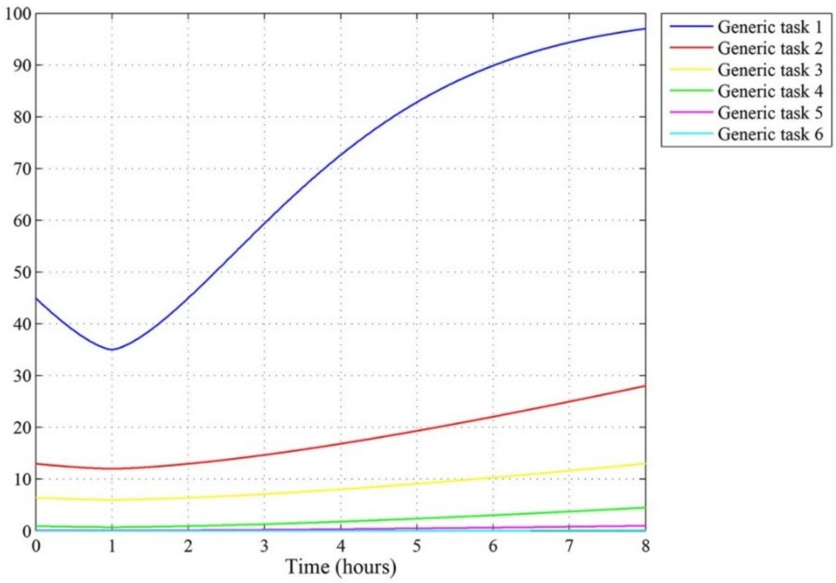

Figure III.7: *HEP trend for every generic task.*

The eight PSFs are the following: available time; stress; complexity; experience and training; procedures; cognitive ergonomics; fitness for duty; and work process. These eight PSFs are among the most used in second generation HRA methods. Several studies have attempted to evaluate the discrepancies between the influencing factors used by different approaches. Boring (2010) focused its attention on the SPAR-H method, noting that despite the variability of the factors used by other first and second-generation methods, the eight PSFs can largely cover environmental and individual factors, in the wake of the most used methods like CREAM and HEART. A study commissioned by the US Nuclear Regulatory Commission in 2005 and titled *"Good Practices for Implementing Human Reliability Analysis (NUREG-1792)"* identified 15 essential PSFs for HRA (Kolaczkowski *et al.*, 2005). Table III.2 shows a comparison between the performance shaping factors found in Good Practices, in the SPAR-H, CREAM and HEART methods.

Table III.2: *Performance Shaping Factors comparison (Kolaczkowski et al., 2005).*

<table>
<tr><th>SPAR-H</th><th>CREAM</th><th>Good Practices</th><th>HEART</th></tr>
<tr><td>Available time</td><td>Available Time</td><td>Time Available</td><td>2</td></tr>
<tr><td rowspan="3">Stress and stressors</td><td>Number of Simultaneous Goals</td><td>Workload/ Time Pressure/ Stress</td><td rowspan="3">29,33</td></tr>
<tr><td>Working Conditions</td><td rowspan="2">Environment</td></tr>
<tr><td>Time of day</td></tr>
<tr><td>Complexity</td><td>Number of Simultaneous Goals</td><td>Complexity</td><td>10</td></tr>
<tr><td>Experience and training</td><td>Adequacy of Training and Preparation</td><td>Training and experience</td><td>1, 6, 9, 15, 20, 24</td></tr>
<tr><td rowspan="5">Ergonomics</td><td rowspan="2">Adequacy of HMI and Operational Support</td><td>Instrumentation</td><td rowspan="5">3, 4, 5, 7, 13,14, 15, 23,26, 32</td></tr>
<tr><td>Human-System Interface</td></tr>
<tr><td rowspan="3">Working Conditions</td><td>Accessibility/Operability of Equipment</td></tr>
<tr><td>Need for Special Tools</td></tr>
<tr><td>Special [Equipment] Fitness Needs</td></tr>
<tr><td>Procedures</td><td>Availability of Procedures/Plans</td><td>Procedures and administrative controls</td><td>11,16, 17,28, 32</td></tr>
<tr><td>Fitness for duty</td><td>-</td><td>-</td><td>30,35</td></tr>
<tr><td rowspan="3">Work processes</td><td>Adequacy of Organization</td><td>Available Staffing</td><td rowspan="3">21, 25, 31,36, 37</td></tr>
<tr><td rowspan="2">Crew collaboration quality</td><td>Communications</td></tr>
<tr><td>Team/Crew dynamics</td></tr>
</table>

Unlike most of the HRA methods, SPAR-H recognizes that a number of PSFs can have both positive and negative effects on performance. As shown in Figure III.8, the probability of error increases with the growth of the negative influence of the PSFs, while, on the contrary, the error probability decreases as the positive influence of the PSFs grows. When the influencing factor represents a positive effect, it corresponds to a value less than one; therefore, the multiplication of a nominal HEP with this value is used to decrease the overall HEP. When the PSF, instead, represents a negative effect, it corresponds to a value greater than one and the multiplication of a nominal HEP with this positive integer serves to increase the HEP.

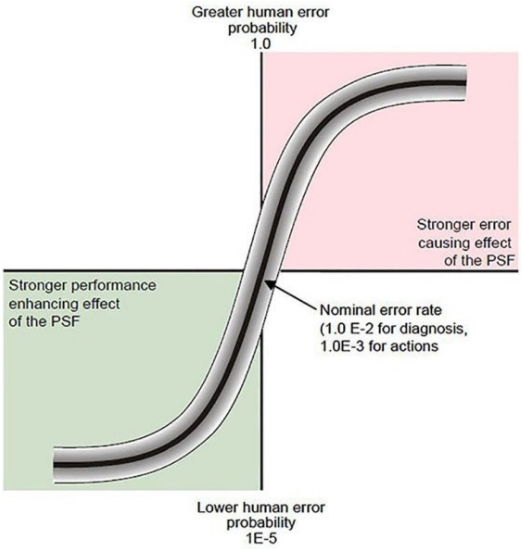

Figure III.8: *Ideal mean HEP as a function of the influence of performance shaping factors.*

Nominal HEP is thus modified by these eight PSFs using the following adjustment factors:

$$HEP_{contextual}(\mathrm{t}) = \frac{HEP_{nominal}(\mathrm{t}) \cdot PSF_{composite}}{HEP_{nominal}(\mathrm{t}) \cdot (PSF_{composite} - 1) + 1} \tag{3.2}$$

where $PSF_{composite}$ is calculated as

$$PSF_{composite} = PSF_1 \times \ldots \times PSF_x \times \ldots \times PSF_8 \tag{3.3}$$

where PSf_x is the assigned multiplier for each PSF. The strength of SPAR-H is in providing a guide for assigning numerical weights for the PSFs; the multiplier values for every PSF are reported in Table III.3.

Table III.3: *PSF multipliers for action and diagnosis.*

SPAR-H PSFs	PSF Levels	Multipliers Action	Multipliers Diagnosis
Available Time	Inadequate Time	P(failure)=1	P(failure)=1
	Time available = time required	10	10
	Nominal time	1	1
	Time available > 5 x time required	0.1	0.1
	Time available > 50 x time required	0.01	0.01
	Insufficient information	Nominal	Nominal
Stress/ Stressors	Extreme	5	5
	High	2	2
	Nominal	1	1
	Insufficient information	Nominal	Nominal
Complexity	Highly complex	5	5
	Moderately complex	2	2
	Nominal	1	1
	Obvious diagnosis	-	0.1
	Insufficient information	Nominal	Nominal
Experience/ Training	Low	3	10
	Nominal	1	1
	High	0.5	0.5
	Insufficient information	Nominal	Nominal
Procedures	Not available	50	50
	Incomplete	20	20
	Available, but poor	5	5
	Nominal	1	1
	Diagnostic/symptom oriented	-	0.5
	Insufficient information	Nominal	Nominal
Cognitive Ergonomics	Missing/Misleading	50	50
	Poor	10	10
	Nominal	1	1
	Good	0.5	0.5
	Insufficient information	Nominal	Nominal
Fitness for Duty	Unfit	P=1	P=1
	Degraded Fitness	5	5
	Nominal	1	1
	Insufficient information	Nominal	Nominal
	Poor	5	2
Work Processes	Nominal	1	1
	Good	0.5	0.8
	Insufficient information	Nominal	Nominal

The multiplier values were attributed by analysts of the method, on the basis of several studies carried out on nuclear power plants. In order to align the evaluation of PSFs in our model, we standardized the multipliers shown in Table III.4. This standardization changes the values of the multipliers for each generic task compared to the average value of the class, maintaining them equal to a nominal level. The multiplier value is attributed to some PSFs as a direct input of the level (e.g. experience is directly established as a low, nominal or high level). For other PSFs, the final value of the multiplier is obtained from the weighted average of the multipliers assigned to the single sub-factors, where the weight is assigned by the rating analysts. Thereafter, identified levels and influencing factors are considered for each PSF.

Table III.4: *Modified multipliers due to standardization for each Generic Task (GT).*

SPAR-H Multipliers	GT 1	GT 2	GT 3	GT 4	GT 5	GT 6
50	21.00	26.00	28.00	34.00	56.00	82.00
20	8.40	10.40	11.20	13.60	22.40	32.80
10	4.20	5.20	5.60	6.80	11.20	16.40
5	2.10	2.60	2.80	3.40	5.60	8.20
3	1.26	1.56	1.68	2.04	3.36	4.92
2	1.01	1.04	1.12	1.36	2.24	3.28
1	1	1	1	1	1	1
0,8	0.34	0.42	0.45	0.54	0.90	0.99
0,5	0.21	0.26	0.28	0.34	0.56	0.82
0,1	0.04	0.05	0.06	0.07	0.11	0.16
0,01	0.004	0.005	0.006	0.007	0.011	0.016

III.3.2.1 Available time

Available time, as a PSF term, can be misleading. In the assessment of the Available time, SPAR-H does not look solely at the amount of time that is available for a task (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). Rather, it looks at the amount of time available relative to the time required to complete the task. Available time refers to the amount of time that an operator or a crew has to diagnose and act upon an abnormal event. The time available can take on six levels, both positive and negative:

- <u>Inadequate time</u>: If the operator cannot diagnose the problem in the amount of time available, no matter what s/he does, then failure is certain.

- Barely adequate time= 2/3 the average time required to diagnose the problem is available.
- Nominal time= on average, there is sufficient time to diagnose the problem.
- Extra time= time available is between one and two times greater than the nominal time required and is also greater than 30 minutes.
- Expansive time= time available is greater than two times the nominal time required and is also greater than a minimum time of 30 minutes; there is an inordinate amount of time (a day or more) to diagnose the problem.
- Insufficient information.

Insufficient information is always present as alternative for all eight performance shaping factors and represents the situation where information is insufficient for assigning a level to a PSF. *Insufficient Information* is quantified with the same value as *Nominal*.

III.3.2.2 Stress and stressors

Stress specifically refers to the level of undesirable conditions and circumstances that impede the operator in completing a task (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). Note that the effect of stress on performance is curvilinear—that is, some small amount of stress can enhance performance, and in the context of SPAR-H should be considered nominal, while high and extreme levels of stress will negatively affect human performance. The degradation of performance is the key point when assigning high or extreme stress levels. For stress, as well as for the complexity and work processes, the value of the multiplier is determined by the presence of more sub-factors. In these cases, the overall PSF value is given by the weighted average of the sub-factor multipliers with respect to the weights that can be reset or assigned from time to time during the insertion of the input. Several environmental and behavioural sub-factors contribute to identify the multiplier:

1) Mental stress;
2) Pressure time;
3) Workplace;
4) Circadian rhythm;
5) Microclimate;
6) Lighting;
7) Noise;

8) Vibration;
9) Ionizing and non-ionizing radiation.

Each of these contributes to the calculation of the total PSF stress through the formula $PSF_{stress} = F_1 \cdot W_1 + \cdots + F_x \cdot W_x + \cdots + F_9 \cdot W_9$, where F_1 is the level assigned to one of the nine sub-factors listed above and $W1$ is the weight of each sub-factor between 0 and 1. The weights must respect the condition $\sum_{\mathrm{i}=1}^{9} \mathrm{W_i} = 1$.

The assignment and definition of levels of stress or stressors is identical across action and diagnosis and action tasks:

- Extreme: a level of stress in which the performance of most workers deteriorates drastically. This is likely to occur when the onset of the stressor is unexpected, and the stressing situation persists for long periods. This level is also associated with the feeling of threat to one's physical well-being or to one's professional status and is considered to be qualitatively different from lesser degrees of high stress.
- High: a level of stress higher than the nominal level (e.g., multiple instruments alarm unexpectedly and at the same time; continuous noise impacts ability to focus attention on the task; the consequences of the task represent a threat to plant safety).
- Nominal: the level of stress that is strategic to good performance.
- Insufficient information.

The values of the multipliers for mental stress, pressure time and workplace are assigned directly, as reported in Table III.5.

Table III.5: *Stress levels and multipliers.*

Sub-Factors	Stress levels	Multipliers SPAR-H	Weight
Mental stress	Extreme	5	Variable
	High	2	
	Nominal	1	
	Insufficient information	1	
Pressure time	Extreme	5	Variable
	High	2	
	Nominal	1	
State workplace	Extreme	5	Variable
	High	2	
	Nominal	1	
	Insufficient information	1	

The other factors must be treated individually because the level assignment is more complex and depends on a combination of physical, environmental and plant-related factors.

III.3.2.2.1 Circadian Rhythm

Many studies have been conducted regarding circadian rhythms in human performance and the importance of biological and physiological rhythms has been demonstrated by the recent Nobel Prize for Medicine 2017 won by Jeffrey C. Hall, Michael Rosbash and Michael W. Young. These rhythms serve to induce individuals of the need to find somewhere safe to sleep at the end of their day, and to become active and productive as they awake from sleep in the morning (Monk *et al.*, 1997).

Several physiological variables (temperature, cortisol and melatonin), humoral variables (vigour and well-being) and performance variable (hand dexterity, search speed, reasoning speed and accuracy, vigilance speed and hits) were found strongly related to circadian rhythm (Monk *et al.*, 1997). Two different types of task can be considered (Folkard and Monk, 1980; Folkard and Rosen, 1990; Monk *et al.*, 1997; Carrier and Monk, 2000; Folkard, Lombardi and Spencer, 2006):

1) <u>Task 1 (action):</u> require immediate information processing and constant attention; they involve little memory (long-term memory) and are characterized by a considerable stress due, in many cases, to the speed required by the execution of the task. They therefore include manual and repetitive operations;
2) <u>Task 2 (diagnosis):</u> involves both information processing and memory; these types of tasks are often referred to as working memory tasks, that is, memory work tasks. They therefore include operations of an intellectual type, typical of office work.

The tasks have a different trend of performance during the day and the trends have been used to model the value of the multiplier as shown in Figure III.9 and Figure III.10.

III.3.2.2.2 Microclimate

Microclimate consists of a set of physical and environmental parameters that characterize the local environment, not necessarily confined, and that with individual parameters, such as metabolic activity and clothing, determine the thermal exchange between the environment itself and individuals that operate there. Microclimate control in the workplace is one of the fundamental aspects that allow people to work in conditions of well-being and comfort, avoiding thermal stress.

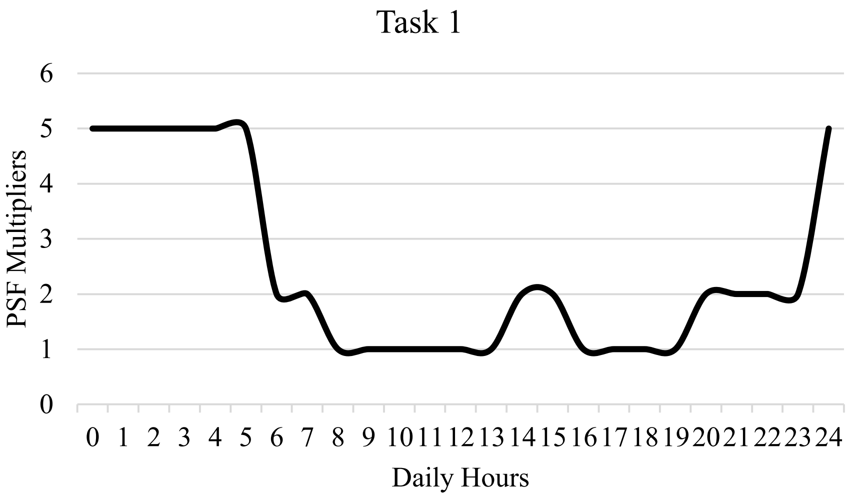

Figure III.9: *Automatic trend of circadian rhythm multiplier for Task 1 (action).*

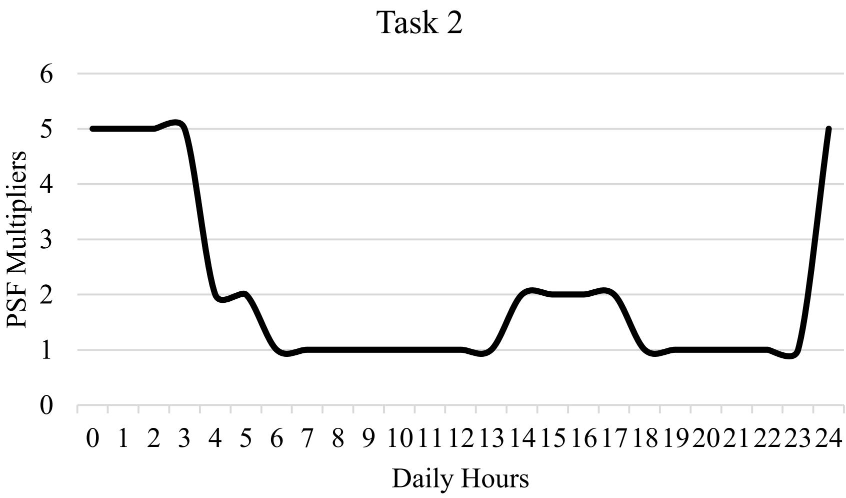

Figure III.10: *Automatic trend of circadian rhythm multiplier for Task 1 (diagnosis).*

One of the basic conditions for comfort is the preservation of thermal neutrality, through a physiological response of the thermoregulation system, namely the situation in which a person does not feel too cold or too hot. The thermoregulation mechanisms, that allow the humans to maintain constant the internal temperature, are:

- natural or involuntary:
  - ✓ vasomotor physiological activity;
  - ✓ behavioural activity: sweating or shiver;
- artificial or voluntary:
  - ✓ clothing;
  - ✓ modification of environmental conditions.

The thermohygrometric conditions of the operator in the working environment are modelled in SHERPA considering four variables:

- air temperature (°C);
- relative humidity (%);
- physical activity carried out (class of metabolic activity - MET);
- thermal insulation clothing (CLO).

These variables allow to assess the level of thermohygrometric comfort according to the process shown in Figure III.11. The first step evaluates the environmental temperature.

In the working environments in which thermoregulatory systems are present, the temperature remains constant and equal to the value on which the thermoregulatory system is regulated, whereas in the absence of the latter the seasonal temperature daily trends are taken into account. Table III.6 reports the average, maximum and minimum temperatures for the four seasons and the times in which the maximum and minimum temperatures are recorded.

Table III.6: *Values of seasonal temperatures.*

Season	Average temperature (°C)	Maximum temperature (°C)	Time	Minimum temperature (°C)	Time
Winter	9,5	13,5	14:30	5,5	2:30
Spring	16,9	22	16:00	11,8	4:00
Summer	24,6	29,7	16:30	19,5	4:30
Autumn	16,4	20,9	15:00	11,9	3:00

Starting from these values the sinusoidal curves, which represent the temperature daily trend, day have been obtained (Figure III.12). The winter curve was obtained considering that the maximum temperature is recorded at 14:30 and the minimum temperature at 2:30 and that the average temperature is equal to 9.5 ° C:

$$T_a = 9{,}5 + 4\sin(TNOW_{deg} - 127{,}5) \qquad (3.4)$$

The spring curve was obtained considering that the maximum temperature is recorded at 16:00 and the minimum temperature at 4:00 and that the average temperature is equal to 16,9 ° C:

$$T_a = 16{,}9 + 5{,}1\sin(TNOW_{deg} - 150) \qquad (3.5)$$

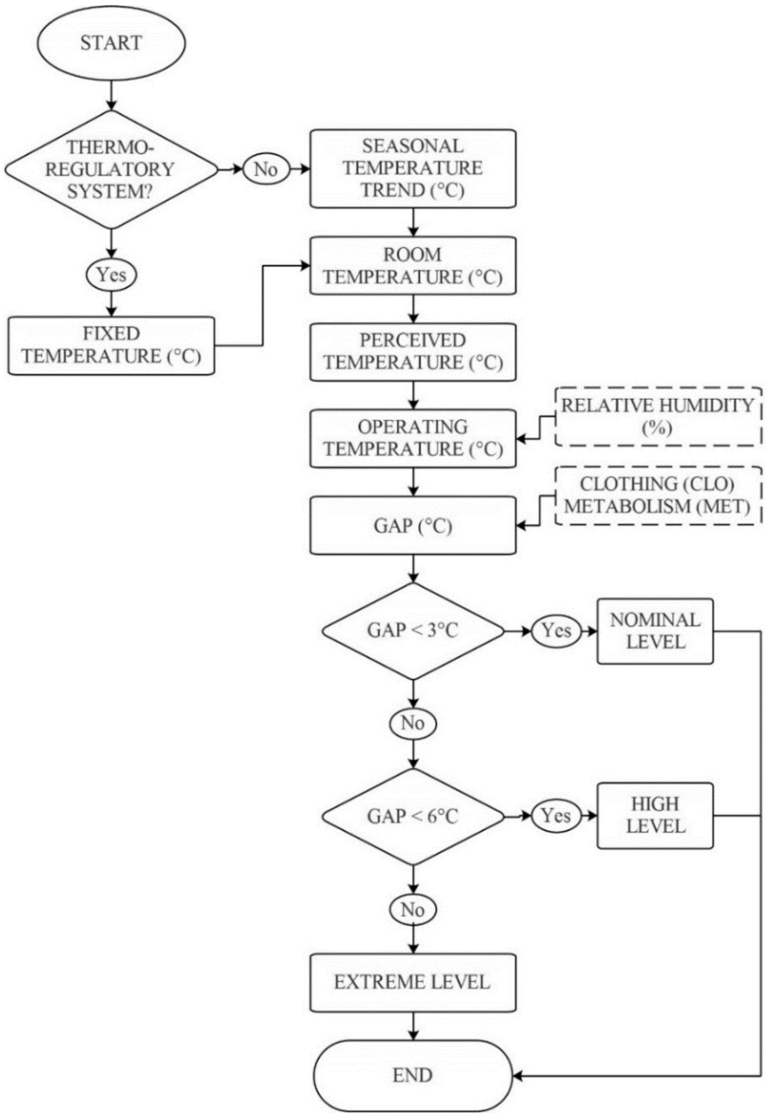

Figure III.11*: Assessment process for the microclimate factor.*

The summer curve was obtained considering that the maximum temperature is recorded at 16:30 and the minimum at 4:30 and that the average temperature is 24,6 ° C:

$$T_a = 24{,}6 + 5{,}1\sin(TNOW_{deg} - 157{,}5) \tag{3.6}$$

The autumn curve was obtained considering that the maximum temperature is recorded at 15:00 and the minimum at 3:00 and that the average temperature is equal to 16,4 ° C:

$$T_a = 16{,}4 + 4{,}5\sin\left(TNOW_{deg} - 135\right) \tag{3.7}$$

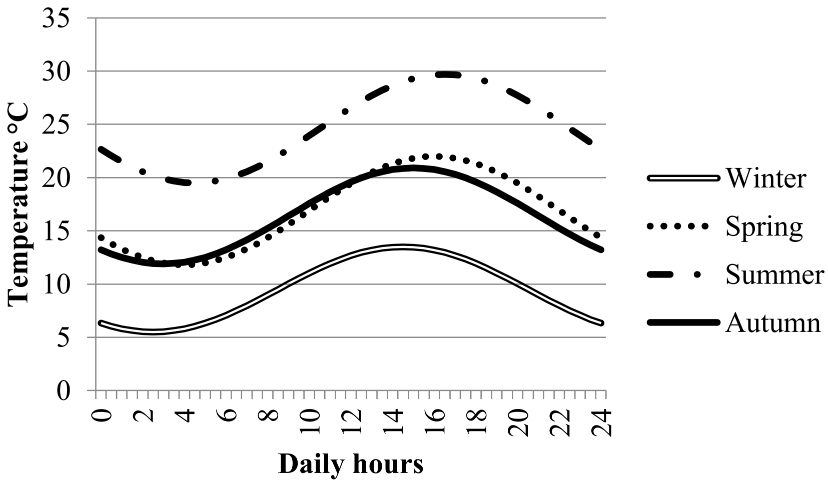

Figure III.12: *Daily temperature trend.*

Perceived temperature is then determined using the Thermohygrometric Index (THI) on the basis of environmental temperature and relative humidity, as follows:

$$THI(°C) = T_{perceived} = T_a - (0{,}55 - 0{,}0055 \cdot UR) \cdot (T_a - 14{,}5) \tag{3.8}$$

where T_a is the environmental temperature expressed in °C and *UR* is the relative humidity expressed as a percentage. Thanks to the THI index, the temperature perceived by the operator is obtained. This is then modified to take into account metabolic activity and clothing through the formula:

$$T_{operating} = T_a - 3 \cdot (1 + CLO) \cdot (MET - 1{,}2) \tag{3.9}$$

where CLO represents the thermal resistance index of clothing and MET the index of metabolic activity. Metabolic energy, also called metabolic rate or metabolism, depends on muscle activity. Normally, muscle activity is transformed into thermal energy; during heavy physical work this transformation can be limited to 75%; in fact, if a person climbs a mountain, part of the energy used is stored in the body in the form of potential energy.

Traditionally, the metabolism is measured in met (1 met = 58.55 W/m^2 of body surface area). An adult has a body surface of 1.7 m^2; therefore, under conditions of thermal comfort, with a level of activity equal to 1 met, it will have a metabolism and therefore a dispersion of energy equal to about 100W. Our metabolism reaches the minimum during sleep (0.8 met) and its maximum during sports activities, where the value of 10 met is often reached. Table III.7 shows some metabolic energy values related to typical activities, taken from the UNI EN ISO 7730: 2006 standard.

Table III.7: *Classes of metabolic activity and generated power.*

Type of activity	Metabolic power	
	W/m^2	met
Rest	58	1
Light activity seated (office, diving, laboratory …)	70	1,2
Light activity up (laboratory, light industry…)	100	1,7
Moderate activity (work on machine…)	117	2
High activity (heavy engineering…)	175	3
Very high activity (intense activity next to limits…)	290	5

Clothing reduces the dispersion of energy from the human body and it is classified according to the level of thermal insulation provided. The unit of measurement usually used for the thermal resistance of clothing is CLO (1CLO = 0.155m^2 ° C / W). Table III.8 shows typical values of the thermal resistance of some typical clothing, reported by the UNI EN ISO 7730: 2006 standard. These values are generally measured by using appropriate heated manikins.

Perceived and operating temperature are compared through the Gap quantification. The Gap is identified as the difference in absolute value between the perceived and the operating temperature. The assignment and definition of levels of the microclimate sub-factor for stress performance shaping factor is based on these values:

1) <u>Extreme level</u>= Gap> 6 ° C;

2) High level= 3 ° C <Gap <6 ° C;

3) Nominal level= Gap <3 ° C.

Table III.8: *Types of thermal clothing and relative thermal resistance.*

Type of clothing	**Description**	**Thermal resistance**	
		CLO	**M^2K/W**
Typical tropical clothing	T-shirt, shorts and sandals	0,1	0,02
Lightweight summer clothing	Shirt with short sleeves or t-shirt, light trousers and shoes	0,5	0,08
Lightweight clothing or work	Shirt or shirt in cotton, pants and shoes Lightweight suit and shoes	0,7	0,11
Indoor winter clothing	Short-sleeved underwear, shirt, jacket, heavy pants and shoes Short-sleeved underwear, sweatshirt, heavy pants and shoes	1	0,16
Outdoor winter clothing	Short-sleeved underwear, shirt, jacket, coat, heavy pants and shoes Short-sleeved underwear, overalls, overalls with heavy padding, heavy-duty hat and shoes	1,5	0,23
Special clothing	Special suits	2	0,31

III.3.2.2.3 Lighting

The lighting of a work environment must be such as to satisfy basic human needs. A good lighting needs, in addition to the required lighting, other qualitative and quantitative requirements:

- good visibility: to correctly perform a certain activity, the object of the vision must be perceived and unequivocally recognized with ease, speed and accuracy;
- visual comfort: the whole visual environment must satisfy physiological and psychological needs; it guarantees workers a feeling of well-being, which, indirectly, also contributes to creating a high level of productivity;

- safety: the lighting conditions must always allow safety and ease of movement and a prompt and safe discernment of the dangers inherent in the work environment.

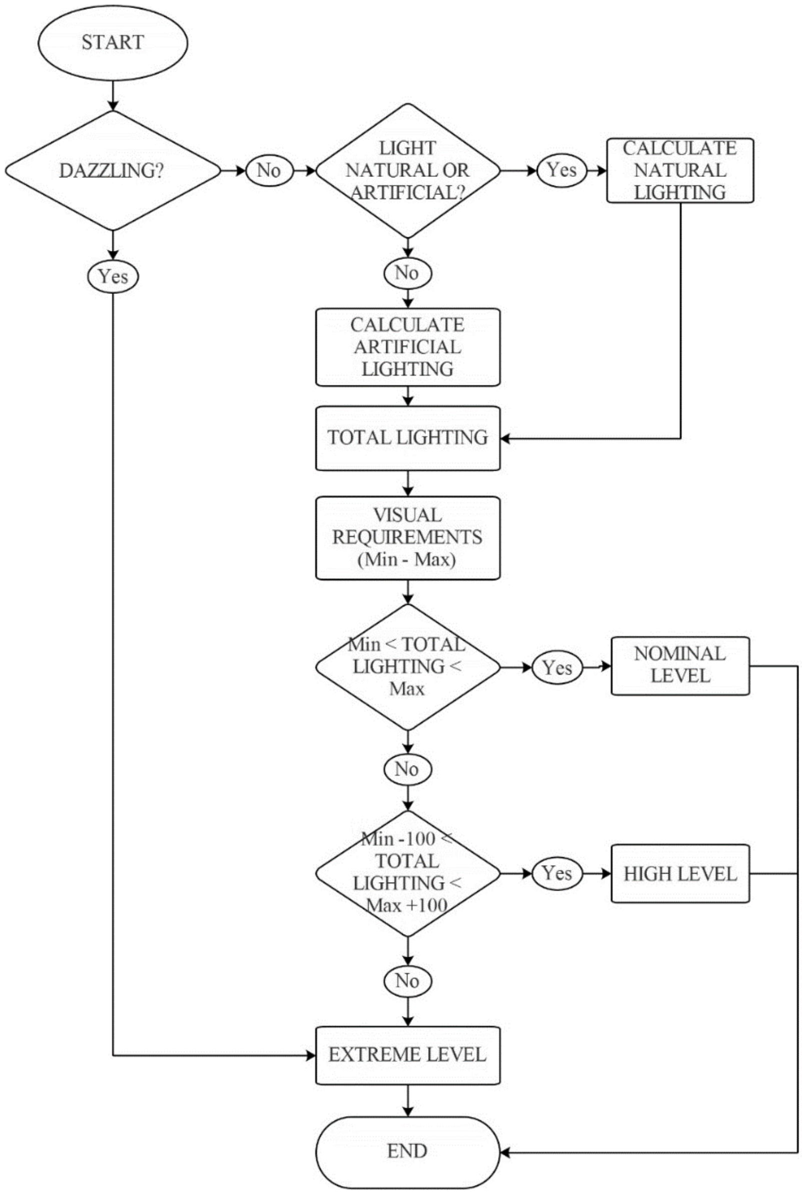

Figure III.13: *Evaluation process for the lighting factor.*

The notion of comfort is subjective and is therefore difficult to define, whereas lack of comfort is certainly easier to circumscribe; it is linked to a sensation of visual discomfort caused by the presence of a strong contrast of luminance in the visual field or by dazzling phenomena. The main requirement, required by Legislative Decree. 81/08 for lighting workplaces (Annex IV, Article 1.10), is that workplaces must have sufficient natural light. In any case, all rooms and workplaces must be equipped with devices that allow adequate artificial lighting to safeguard workers' safety, health and well-being. Based on this requirement SHERPA model considers two possible scenarios:

1) Natural lighting combined with artificial lighting.
2) Artificial lighting.

In both cases the evaluation of the lighting factor during the operator's activities is based on the comparison between the illumination required by the visual task under examination and the illumination provided by the lighting system of the workplace. The evaluation process is shown in the following flowchart (Figure III.13).

IESNA model is used to evaluate the amount of natural illumination in different periods of the year (Cucumo *et al.*, 1996). This model considers three types of sky: clear, partly cloudy and cloudy sky, that are identified with the cloudiness ratio (CR), defined as the ratio between diffused radiation and global hourly radiation in the horizontal plane. The illuminance is calculated as follows:

$$E = A + B \cdot (\sin \alpha_s)^C \tag{3.10}$$

where E is the illuminance on the horizontal plane, α_s the solar height and A, B and C are constants whose values are shown in Table III.9.

Table III.9: *Values of constants A, B and C for different sky conditions.*

Sky	**A**	**B**	**C**
Clear	0,8	15,5	0,5
Partly Cloudy	0,3	45,0	1
Cloudy	0,3	21,0	1

The solar height trend is modelled through linear regression, referring to the values assumed during each season in the industrial area of Salerno, as follows:

1) Winter: $\alpha_s = 0{,}0039t^4 - 0.205t^3 + 3{,}3097t^2 - 16{,}339t + 18{,}985$
2) Autumn: $\alpha_s = 0{,}004t^4 - 0.208t^3 + 3{,}2847t^2 - 15{,}551t + 17{,}235$

3) Spring: $\alpha_s = 0{,}0068t^4 - 0.3561t^3 + 5{,}6325t^2 - 25{,}971t + 28{,}257$
4) Summer: $\alpha_s = 0{,}0069t^4 - 0.3592t^3 + 5{,}6843t^2 - 26{,}333t + 28{,}735$

A representation more realistic and close to the real lighting conditions in the workplace is obtained considering the percentage reported in Table III.10 and derived from the meteorological historical series.

Table III.10: *Allocation between the different types of seasonal level sky.*

	Clear sky	Partly Cloudy Sky	Cloudy Sky
Autumn	22%	33%	45%
Winter	28%	39%	33%
Spring	33%	43%	24%
Summer	48%	32%	20%
Total	33%	37%	30%

The internal illumination is then derived applying a mean daylight factor. The daylight factor is a parameter suitable to characterize, from the lighting point of view, in the case of a source of natural light, the environment under study. The ratio or daylight factor η, is defined as:

$$\eta = \frac{E}{E_e} \tag{3.11}$$

where E is the illuminance in a point of the environment and E_e is the illuminance which, at the same time, would assume a horizontal surface placed outside, shielded by direct solar irradiation and exposed to the light coming from whole (unobstructed) heavenly vault. The daylight factor η is often expressed in percentage units. In the case of the working environment, a value of 0.02 is assumed.

The second step consists in calculating the lighting provided by the artificial lighting systems in the workplace, considering:

1) type of artificial light sources (Table III.11);
2) number of artificial light sources;
3) area to be illuminated (A);
4) maintenance factor (M), equal to 0.6;
5) utilization factor (F_u), equal to 0.5.

Knowing the type and number of light sources installed in the work environment, it is possible to calculate the illuminance, using the inverse formula applied by the total flow design method:

$$E = \frac{F_u \cdot M \cdot N \cdot L}{A} \tag{3.12}$$

where E is the illuminance of the plant expressed in lux, F_u and M are respectively the utilization and maintenance factors, N is the number of light sources, L the luminous flux emitted by each light source and expressed in lumens and finally A is the area to be illuminated in square meters. In the case of combined natural and artificial lighting the total value of illuminance is given by the sum of the natural and the artificial one.

Table III.11: *Types of light sources available for a lighting system.*

Artificial light source	Watt	Luminous flux (lm)	Colour temperature	Rendering index
Incandescent lamp	100	1380	2700	100
Tubular fluorescent lamp T5	54	4450	4000	85
Compact fluorescent lamp	18	1200	4000	82
Sodium vapor lamp high pressure	150	15000	2100	25
Metal halide lamp	400	40000	4140	90
Mixed light lamp	500	14000	4100	49
LEDs tubes	30	3300	6500	85

This is the first value necessary for the evaluation of the illuminance factor. It must be compared with the limits imposed by the law for the different visual tasks, where the visual task refers to the set of visual elements (dimensions of the structure, contrast and duration) that relate to the work carried out. The values specified in Table III.12 are the average illuminances maintained necessary to ensure visual comfort.

The quantification of the level of the lighting factor is carried out through the following criteria:

- If the total illuminance level is between the maximum and the minimum value required by the visual task, the <u>nominal level</u> is assigned;

- If the total illuminance level is between the maximum value increased by 200 lux and the minimum value decreased by 200 lux the high level is assigned;
- If the total illuminance level is higher than the maximum value increased by 200 lux or lower than the minimum level decreased by 200 lux, the extreme level is assigned.

Another aspect taken into consideration in the assessment of lighting is the dazzling phenomenon. Dazzling is the visual sensation produced by surfaces that determine high luminance gradients within the visual field and can be perceived as harassing or debilitating glare. In the case of known and present dazzling phenomena the extreme level of stress is always assigned.

Table III.12: *Visual tasks and levels of illumination required for each one.*

Tasks and visual requirements	Illumination value required for the task
Tasks with simple visual requirements	<300
Tasks with medium visual requirements	300-600
Tasks with visual requirements of precision	600-900
Tasks with difficult visual requirements	900-1200
Performing visual tasks very precise	1200-1500
Tasks with special visual requirements	>1500

III.3.2.2.4 Noise

Noise in the workplace has become one of the most important problems among those included in occupational hygiene. A sound that causes an unpleasant, annoying or intolerable sensation can be defined noise. The sound is measured in decibels, as regards sound pressure, and in hertz, with regard to frequency. There are two quantities to measure the continuous noise exposure value and which can be compared to the legal limits:

- the individual exposure level to daily noise (LEX, d): weighted average value of the noise exposure levels for a nominal working day of 8 hours.
- the individual exposure level to weekly noise (LEX, w): weighted average value of the daily noise exposure levels for a nominal week of 5 days of 8-hours.

Legislative Decree 195/2006 sets the following noise limits:

1) Exposure limit value (8 h working day): 87dB (A).
2) Higher value of action (8 h working day): 85 dB (A).
3) Lower value of action (8 h working day): 80 dB (A).

SHERPA considers the noise impact on HEP, as follows:

1) Extreme level: noise level above 85 dB;
2) High level: noise level between 60 and 85 dB;
3) Nominal level: noise level below 60 dB.

III.3.2.2.5 Vibrations

The vibrations, according to the physical definition, are mechanical oscillations generated by waves of pressure that are transmitted through solid bodies. Depending on the physio pathological effects on humans, the vibrations are divided into three main frequency bands:

- 0-2 Hz: low frequency oscillations, generated by means of transport (land, air, sea).
- 2-20 Hz: medium frequency oscillations, generated by machines and industrial plants.
- > 20-30 Hz: high frequency oscillations; generated by a wide range of vibrating tools that are widespread in the industrial field, involving many work activities, from the simplest to the most sophisticated.

The vibrations are characterized by three other parameters closely related to each other:

- amplitude of the displacement (expressed in cm),
- speed (expressed in cm / sec),
- acceleration (expressed in m/sec^2 or multiples of g, gravity acceleration: 1g = 9.8 m/sec^2).

Acceleration is the most important parameter for the assessment of the body's response to vibrations. Legislative Decree 81/2008 distinguishes between vibrations transmitted to the arm-hand system and those transmitted to the whole body. Exposure to vibrations to the hand-arm system is generally caused by hand contact with the handgrip of hand tools or hand-driven machinery. Mechanical vibrations transmitted to the whole body pose risks to the health and safety of workers, in particular back pain and trauma to the spine. It is known that various work activities carried out on board transport or handling means expose the body to vibrations or impacts, which may be harmful to the exposed subjects. The current legislation requires that the exposure to vibrations in both cases is evaluated by calculating the equivalent weighted acceleration in frequency referred to eight hours of work, which is calculated using the following formula:

$$A(8) = a_V \sqrt{\frac{T_e}{8}} \tag{3.13}$$

where T_e is the total daily duration of vibration exposure expressed in hours and a_v the value of the acceleration considering the components quantified on the three axes:

$$a_v \left(m/s^2\right) \;=\; \left(\left(a_{wx}^2 + a_{wy}^2 + a_{wz}^2\right)^{1/2}\right) \tag{3.14}$$

If the worker is exposed to different vibration values, such as when using several mechanical means during the working day, the daily vibration exposure A (8), in m/s^2, is obtained by expression:

$$A(8) = \left[\frac{1}{8}\sum_{i=1}^{N} a_{vi}^2 T_i\right]^{\frac{1}{2}} \tag{3.15}$$

where a_{vi}^2 is the vector sum of the acceleration relative to the i-th operation and T_i is the exposure time relative to the i-th operation expressed in hours. Through the Vibration Database, available online, the exposure values to vibrations produced by the machines commonly used in the industrial field can be identified. Based on these limits in SHERPA the vibration factor levels are assigned according to the following criteria:

- Hand-arm system:

1) <u>Extreme level</u>: equivalent vibration level (calculated as per norm) above 5 m/s^2;
2) <u>High level</u>: equivalent vibration level (calculated as per norm) between 2.5 and 5 m/s^2;
3) <u>Nominal level</u>: equivalent vibration level (calculated as per norm) lower than 2.5 m/s^2;

- Whole body:

1) <u>Extreme level</u>: equivalent vibration level (calculated as per norm) above 1 m/s^2;
2) <u>High level</u>: equivalent vibration level (calculated as per standard) between 0.5 and 1 m/s^2;

3) Nominal level: equivalent vibration level (calculated as per norm) lower than 0.5 m/s^2;

Furthermore, the vibrations are not naturally present in all working contexts and for this reason insufficient information level is assigned in lack of them.

III.3.2.2.6 Ionizing and non-ionizing radiation

The ionizing radiations have enough energy to be able to ionize the atoms (or molecules) with which they come into contact. They are divided into two main categories: those that produce ions in a direct way (the particles α, β- and β +;) and those that produce ions in an indirect way (neutrons, γ rays and X rays). Ionizing radiation is generated by nuclear, artificial or natural reactions, from very high temperatures such as plasma discharge, through the production of high-energy particles in particle accelerators, or due to acceleration of charged particles by the fields electromagnetic products produced by natural processes, from lightning to supernova explosions. If ionizing radiation affects biological tissues, it can cause various types of health-related damage:

1) deterministic somatic damage: degeneration of organic tissues;
2) stochastic somatic damage: include leukaemia and solid tumours;
3) stochastic genetic damage: genetic mutations and chromosomal aberrations.

Legislative Decree 230/95 imposes different limits for exposure to ionizing radiation:

- Operator not exposed: effective dose absorbed (msV/year) <1;
- Operator exposed in category B: 1<absorbed effective dose (msV/year) <6;
- Operator exposed in category A: effective dose absorbed (msV /year)> 6.

On the basis of these limits, the levels of the ionizing radiation sub-factor are assigned as follows:

- Nominal level= unexposed operator;
- High-level= category B;
- Extreme level= category A.

Non-ionizing radiation refers, instead, to those forms of radiation not able to cause the breaking of electronic bonds of matter and which lead to the formation of pairs of particles having opposite charge. They refer to any type of electromagnetic radiation which, instead of producing charged ions by passing through matter, excites only the movement of an electron to a higher energetic state. Several biological effects are observed for different types of non-ionizing radiation. The electric, magnetic and electromagnetic fields present in the environment have the property of penetrating deeply into biological materials. The penetration thickness decreases with the frequency of the fields: it is maximum at the low frequencies, of the order of centimetres in the radio frequency range, millimetres in the microwave region.

These fields interact with the charged particles present in the exposed system and exert forces on them which can alter the original charge distribution. The main effect of the interaction of radio frequencies and microwaves with a living system is represented by a transfer of energy, in the form of heat, with an increase in the local temperature or in the whole system. Furthermore, the possibility of dispersing heat is of importance; for the human organism the best heat exchanger is represented by the blood, for this reason the less vascularized organs or apparatuses are more susceptible to damages from electromagnetic radiations as they are not able to redistribute the heat received from an external source.

The regulation n.36 / 2001 defines the exposure limits for electromagnetic fields:

1) High frequencies:
    - Limit value of the electric field: 20 V/m;
    - Attention value of the electric field: 6 V/m.
2) Low frequency:
    - Limit value of the electric field: 5 kV/m;
    - Attention value of the electric field: 0.5 kV/m.

Based on these limits, SHERPA assigns the levels of the non-ionizing radiation sub-factor: <u>nominal level</u> to values lower than those of attention of the electric field, <u>high level</u> for values between the limit and the attention level and, finally, <u>extreme level</u> for values above the limit value.

In the proposed module the evaluation is done only in case there is sufficient information. In the presence of both types of radiation, the total assessment of the level is given by the average of the levels assigned to ionizing and non-ionizing radiations.

III.3.2.3 Complexity

Complexity refers to how difficult the task is to perform in the given context; it considers both the task and the environment (Gertman *et al.*, 2005;

Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). A more difficult task has a greater chance for human error. Similarly, a more ambiguous task has a greater chance for human error. The complexity is one of those PSFs that Boring (2007) defines as indirect, as it cannot be measured directly. For this reason, the value of the complexity cannot be assigned directly but relies on input from several elements (Table III.13):

1) General complexity;
2) Mental efforts required;
3) Physical effort required from type of activity;
4) Precision level of the activity;
5) Parallel tasks.

Each of these contributes to the calculation of the overall complexity PSF through the formula $PSF_{complexity} = F_1 \cdot W_1 + \ldots + F_x \cdot W_x + \ldots + F_5 \cdot W_5$, *w*here *F1* is the level assigned to one of the five factors listed above and *W1* is the weight of each factor between 0 and 1. The weights must respect the condition $\sum_{i=1}^{5} W_i = 1$.

III.3.2.4 Experience and training

This PSF refers to the experience and training of the operator involved in the task (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011).

Years of experience of the individual or crew, and whether the operator/crew has been trained on the type of accident, the amount of time passed since training, the frequency of training, and the systems involved in the task and scenario are included in this PSF. In SHERPA, the data on the training and experience of the operator are inserted directly according to following levels:

- <u>Low</u>= less than 6 months of relevant experience and/or training. This level of experience / training does not provide the level of proficiency and profound understanding necessary to properly perform the required tasks; it does not provide adequate practice in these tasks or does not expose individuals to various abnormal conditions.
- <u>Nominal</u>= more than 6 months of relevant experience and/or training.
- <u>High</u>= extensive experience; a demonstrated master. This level of experience / training provides operators with extensive knowledge and practice in a wide range of possible scenarios.
- <u>Insufficient information.</u>

Table III.13: *Complexity levels and multipliers.*

Factors	Complexity levels	Multipliers SPAR-H	Weight
General complexity	Highly complex: very difficult to perform	5	Variable
	Moderately complex: somewhat difficult to perform	2	
	Nominal: not difficult to perform	1	
	Obvious diagnosis: diagnosis becomes greatly simplified	0,1	
	Insufficient information	1	
Precision of activity	High level of precision	5	Variable
	Moderate level of precision	2	
	Nominal level of precision	1	
	Insufficient information	1	
Mental efforts required	High degree of memory, high understanding of technical drawings and specifications	5	Variable
	Moderate degree of memory, moderate understanding of technical drawings and specifications	2	
	Minimal mental effort	1	
	Insufficient information	1	
Parallel tasks	Multiple run two or more machines	5	Variable
	Many different tasks on the same machine (machining, setup...)	2	
	Processing on the same machine	1	
	Insufficient information	1	
Physical efforts required	- High activity (heavy engineering...) - Very high activity (intense activity next to the limits	5	Variable
	Moderate activity (work on machines...)	2	
	- Light activity seated (office, driving, laboratory - Light activity up (laboratory, light industry...)	1	

III.3.2.5 Procedures

This PSF refers to the existence and use of formal operating procedures for the tasks under consideration (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). Common problems seen in event investigations for procedures include situations where procedures give wrong or inadequate information regarding a particular control sequence. Another common problem is the ambiguity of steps. Levels used for this PSF in SPAR-H:

- <u>Not available</u>= the procedure needed for a particular task or tasks in the event is not available.
- <u>Incomplete</u>= information is needed that is not contained in the procedure or procedure sections; sections or task instructions (or other needed information) are absent.
- <u>Available, but poor</u>= a procedure is available, but it is difficult to use because of factors such as formatting problems, ambiguity, or such a lack in consistency that it impedes performance.
- <u>Nominal</u>= procedures are available and enhance performance.
- <u>Diagnostic/symptom oriented</u>= diagnostic procedures assist the operator/crew in correctly diagnosing the event. These procedures allow operators to maintain the plant in a safe condition, without the need to diagnose exactly what the event is, and what needs to be done to mitigate the event.
- <u>Insufficient information.</u>

III.3.2.6 Cognitive ergonomics

Ergonomics refers to the equipment, displays and controls, layout, quality, and quantity of information available from instrumentation, and the interaction of the operator/crew with the equipment to carry out tasks (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). Aspects of the human-machine interface are included in this category, as well as the adequacy or inadequacy of computer software. SPAR-H was born in the nuclear field, so ergonomics is mainly oriented to the interaction of a human or group to the instrumentation typical of a control room, such as the display and control buttons. In other kinds of industries, this PSF focuses instead on the ergonomics of the workplace and the equipment used. The multipliers are as follows.

- <u>Missing, misleading</u>= lack of ergonomic design for the workstation.

- Poor= low level of ergonomics limited to single workstation.
- Nominal= average level of ergonomics limited to single workstation.
- Good= ergonomic workplace design for both the posture that the tools used for.
- Insufficient information.

III.3.2.7 Fitness for duty

Fitness for duty refers to whether or not the individual is physically and mentally suited to the task at hand (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). This PSF includes fatigue, sickness, drug use (legal or illegal), overconfidence, personal problems and distractions and also includes factors associated with individuals, but not related to training, experience or stress (which are covered by other PSFs). Levels used in SHERPA are:

- Unfit= the individual is unable to carry out the required tasks, due to illness or other physical or mental incapacitation (e.g. having an incapacitating stroke).
- Degraded Fitness= the individual is able to carry out the tasks, although performance is negatively affected. Mental and physical performance can be affected if an individual is ill, such as having a fever. Individuals can also exhibit degraded performance if they are inappropriately overconfident in their abilities to perform.
- Nominal= the individual is able to carry out tasks; no known performance degradation is observed. Nominal should also be used when the analyst judges the PSF as not a performance driver.
- Insufficient information.

III.3.2.8 Work processes

Work processes refer to aspects of doing work, including inter-organizational factors, safety culture, work planning, communication and management support and policies (Gertman *et al.*, 2005; Boring and Blackman, 2007; Blackman, Gertman and Boring, 2008; Whaley *et al.*, 2011). How work is planned, communicated and executed can affect individual and crew performance. If planning and communication are poor, then individuals might not fully understand the work requirements. Work processes also include any management, organizational or supervisory factors that may affect performance. In this case, the value of the PSF also may not be assigned directly but may be based on multiple elements that are input:

1) Communication and integration in team work;
2) Work processes.

The multipliers are as follows:

- Poor= insufficient integration into team, bad or conflictual relationship, poor communication between different shifts / insufficient management of work processes.
- Nominal= sufficient integration into the team, professional relationship, good communication between different shifts / good management of work processes.
- Good= excellent integration into the team, none type of conflict, excellent management of work processes.
- Insufficient information.

Each of these contributes to the calculation of the total PSF work processes through the formula $PSF_{work\ processes} = F_1 \cdot W_1 + F_2 \cdot W_2$, *w*here *F1* is the level assigned to one of two factors listed above and *W1* is the weight of each factor between 0 and 1. The weights must respect the condition $\sum_{i=1}^{2} W_i = 1$.

III.3.3 Process simulation

The third SHERPA sub model (Figure III.2) represents the process modelling and it allows to simulate the execution of human working activities, taking into account the features of the process, the HEP, and the assigned breaks scheduling. The selection of the input variables such as performed task, level of contextual factors, or physical and mental employee condition, allows to model the specific system considering all the working context and worker features. SHERPA can manage three different rest breaks policies:

- no break in the shift (continuous working);
- fixed breaks (several timing and length of rest period);
- automatic break scheduling management.

The operating principles are displayed in the flowchart (Figure III.14), where the logic of process simulation and breaks policies management are represented. The absence of breaks (yellow box) corresponds to the mere reproduction of the process. The activities are simulated based on the model inputs, technical and economic data (processing and setup times, number of entities, workplace conditions and many others) and on the corresponding HR distribution, taking into account the hours of continuous work already carried out by the worker. The contextual HEP value, output from this block, consents to quantify in the next phase, the non-compliant percentage and to evaluate the overall worker performance. In this case, SHERPA represents the activities performed by a single operator during the work shift, estimating in real time the reliability curves and the impact of the operator's performance on the system under examination.

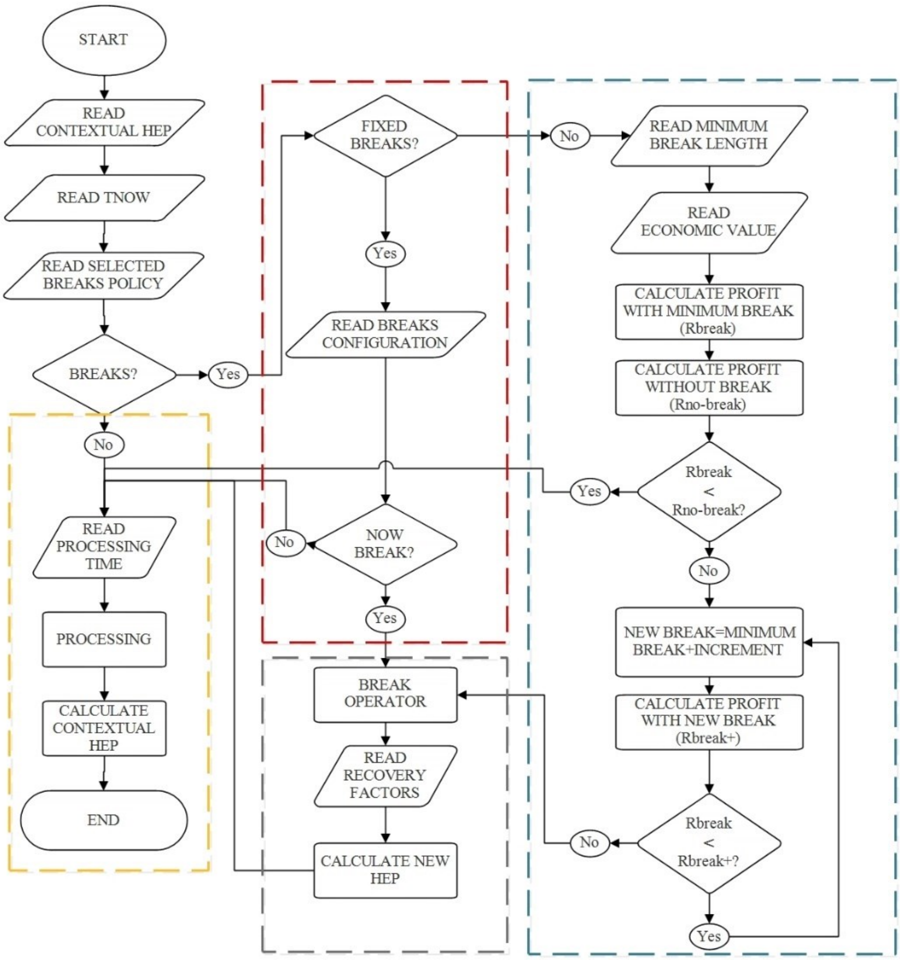

Figure III.14: *Logic framework of the break configurations management.*

The fixed break (red box) or automatic (blue box) sections involve the break scheduling management through the hypothesis that breaks allow the mental and physical recovery and lead to improvements of human reliability. The worker's recovery (grey box) is modelled as a function of the break length and of the type of activity carried out. Recovery modelling and break scheduling management are described in the following sections.

III.3.3.1 Recovery modelling

Despite few studies on the issue, rest breaks literature has shown a great influence of breaks, both micros and macros, on the human performance. Rest break allows, in fact, the worker's physical and mental recover (Dababneh, Swanson and Shell, 2001; Jett and George, 2003; Demerouti *et al.*, 2012).

The literature review on breaks impact on human performance allows to model the worker's recovery and the effect in terms of HEP. After a break, in fact, the human reliability curve is reported to a previous moment with lower level of HEP; this new distribution of nominal HEP is function of length of same break (Figure III.15). This means that after a break, more or less extended, the HEP nominal value changes, due to the effect of the recovery. This value is lower than the HEP when there is no break. Naturally, recovery depends on the activity performed, the characteristics of the operator and the duration of the break.

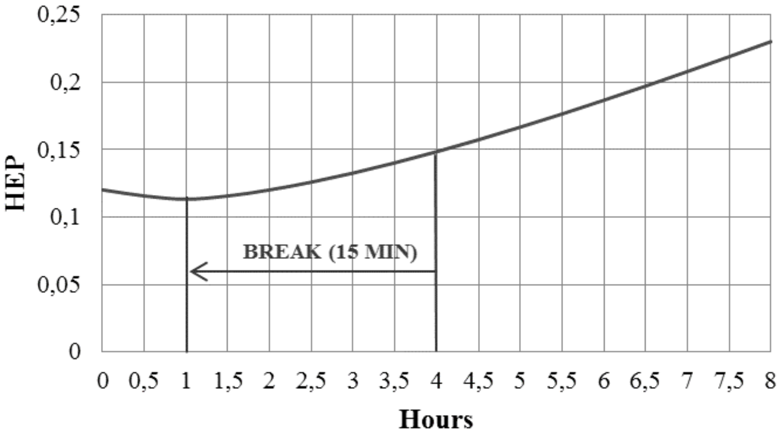

Figure III.15: *Break impact of human error probability (HEP).*

The recovery factor (r_p), that takes the Wright learning curve as a reference, is expressed by the exponential function:

$$r_p = e^{-\omega T_b} \qquad (3.16)$$

in which T_b is the break length (in hours) and ω is the recovery rate. This coefficient represents has been hypothesized for four general categories, which describe different working activities, as listed in Table III.14 and Table III.15. The recovery index for the four cases has been quantified by imposing, as a boundary condition, that the maximum/full recovery is achieved with a break of optimal duration and for r_p equal to 0.1 by setting this rate at three levels (slow, medium and fast). In this way the maximum break time is limited: in fact, the r_p curve tends asymptotically to a null value with increasing break time; full recovery would be obtained in correspondence to a break of infinite time. Below the value of r_p of 0.1, it has been assumed that the pause not have substantial effects on the recovery of the operator.

Table III.14: *Types of activities and corresponding parameters for the optimal allocation of the breaks (recovery rate).*

Activity	Recovery rate (ω)		
	Slow	**Medium**	**Fast**
Sedentary activities (office, laboratory)	2.76	4.61	13.82
Activity light, standing (laboratory, light industry)	2.51	3.95	9.21
Medium activity, standing (work machines)	2.30	3.45	6.91
Activities heavy (heavy work machines)	1.97	2.76	4.61

Table III.15: *Types of activities and corresponding parameters for the optimal allocation of the breaks (optimal time).*

Activity	Optimal time (min)		
	Slow	**Medium**	**Fast**
Sedentary activities (office, laboratory)	50	30	10
Activity light, standing (laboratory, light industry)	55	35	15
Medium activity, standing (work machines)	60	40	20
Activities heavy (heavy work machines)	70	50	30

Recovery rate needs an in-depth analysis as it is not possible to attribute a break, whose duration is valid for any type of activity carried out (whether physical or mental), for any age and/or gender of the worker. Recovery speed and type of physical activity (Table III.14) are not the only parameters to be considered for the recovery rate. Physiological job demands, age and gender are, in fact, factors that strongly determine the individual need of recovery. Table III.16 summarizes the main parameters that have been chosen to represent all the factors that influence the need for recovery of the operator in the workplace.

The analysis of the scientific literature on need for recovery has allowed defining and model the optimal break times according to the previously mentioned parameters. Firstly, the percentages of increase and decrease of the need for recovery between the different age groups were determined and applied with respect to the 36-45 age group which was considered as a reference (Mohren, Jansen and Kant, 2010).

Table III.16:*Parameters of recovery rate.*

Parameters	Levels	Values
Recovery speed	3	Slow (S), Medium (M), Fast (F)
Physical activity	4	Sedentary activities Activity light, standing Medium activity, standing Activities heavy
Psychological Job Demands	3	Low, Medium, High
Age groups	5	18-25, 26-35, 36-45, 46-55, 56-65
Gender	2	Male, Female

Likewise, variations were identified with respect to the three different tasks that describe the load of cognitive work that the operator perceives, being engaged in an activity that requires expenditures of mental energy and psychic efforts (Mohren, Jansen and Kant, 2010; Mathiassen *et al.*, 2014). The quantification results for male are reported in Appendix A. The results thus obtained were then adapted to the female gender in relation to the already known recovery needs (Mohren, Jansen and Kant, 2010). The optimal break duration for women is obtained starting from that of men, increasing or reducing it according to the percentage reported in Table III.17 and derived by Mohren, Jansen and Kant (2010).

Table III.17: *Female need for recovery.*

Age Groups	18-25	26-35	36-45	46-55	56-65
% increased or reduced need for recovery	+ 1.73%	-6.4%	-8.9%	-5.96%	+4.75%

The estimated of recovery times, obtained from the literature and reported in Appendix A, aims to be more realistic and relevant to the working reality in which the operators exercise a profession that requires physical and mental efforts. Recovery times obtained decrease with the increase in the psychological burden of work borne by the operator: beyond the expectations of an association rather linear between the increase in recovery and the increasing difficulty of cognitive work, it is possible to observe a greater recovery when the operator detaches from an activity that has committed him from the physical point of view to dedicate himself to one that requires different efforts, of a psychic nature and of shorter duration, allowing the body to recover the muscular energy expenses and accelerate the recovery process.

The HR improvement, due to the rest period, is modelled considering that the human reliability curves after a break is reported to a previous moment with a lower level of HEP and this new distribution of nominal HEP is a function of the length of same break, as previous defined. The recovery factor impact on the nominal HEP distribution is as follows:

$$HEP_{nominal}(t) = 1 - k \cdot e^{-\alpha \cdot (T \cdot r_p - 1)^{\beta}} \quad (3.17)$$

where T is the time in which the operator resumes its activity after the break and r_p represents the level of recovery, the two parameters α and β change the scale and shape of the curve for each generic task.

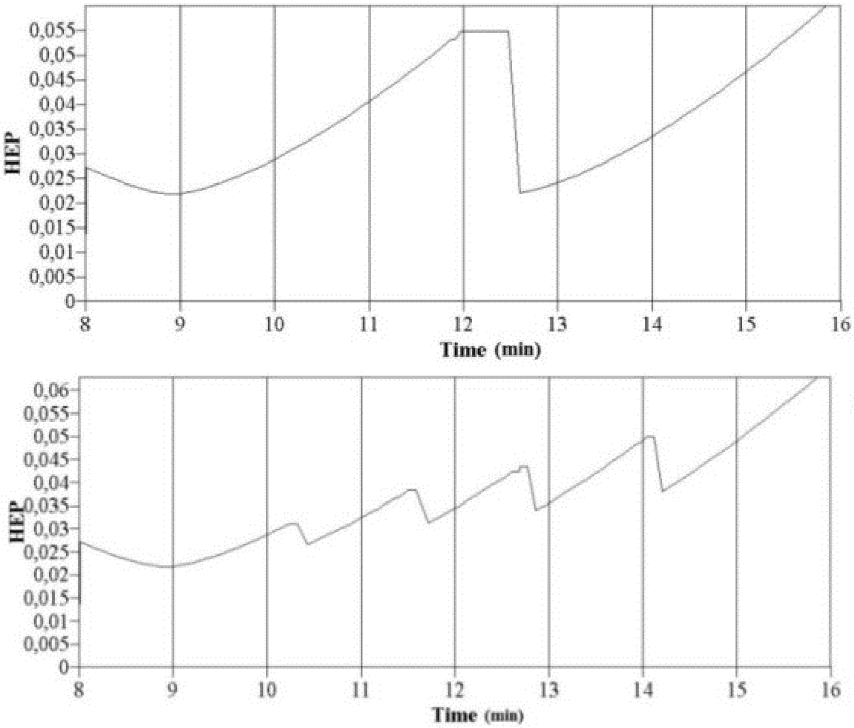

Figure III.16: *The human error probability distribution with two break configurations: a) one break of 20 min after 240 worked min; b) four breaks of 5 min the first after 132 worked min and the others every 72 min.*

Figure III.16 shows the error probability curves of an eight-h shift in two given scenarios, which display action mechanism of the breaks on reliability curves. In the first case, a break of 20 min is assigned to half shift and it allows the almost full recovery, decreasing the average HEP. In the same way, four breaks of 5 min, distributed in the shift, modify the average human reliability. The impact of the break length on the worker's recovery is evident comparing the two HEP curves: a longer break (20 min) leads to higher recovery than shorter breaks (5 min).

III.3.3.2 Break scheduling management

SHERPA addresses the break scheduling problems through the hypothesis that breaks allow the mental and physical recovery and lead to improvements of human reliability, as described in the previous section.

As shown in Figure III.14, for the modelling of the fixed breaks scheduling, the model takes into account the exact chosen work-rest configuration and then it recalculates the nominal HEP before performing the processing phase. In this way, HEP is modified due to the recovery of the operator, which naturally leads to a more or less evident performance improvement based on the break time, number of break and position of the breaks in the work shift.

Furthermore, knowing HR and HEP curves, SHERPA applies also an economic algorithm to automatically manage break scheduling, that allows to assess both positive and negative break effects and to compare their impact on the system performance. The algorithm is able to decide: whether to do or not the break for the operator before starting the current task and the optimum break time. The positive break impact on human reliability is a function of break time, recovery speed and type of performed activities. However, the break benefit is countered by increased idle time of employees during the rest period, that is modelled as a cost of lost production due to break and it corresponds, for example, to a decrease of productivity in a manufacturing context. The model directly links the contextual HEP to the number of non-compliant entities, as described in Section III.3.5. The model can be adapted to alternative set of constraints (minimum number of breaks and minimum time guaranteed by legislation or internal union agreements, maximum hours of continuous work and other possible constraints), assigned in the initialization phase of the system as inputs.

SHERPA can then evaluate the effect of every work-rest policy, defined as acceptable for the system under consideration, with the aim of identifying the best configuration among those possible. The algorithm (Figure III.14) takes into account the cost of lost production due to the break and the quality costs related to operator errors. In particular, before starting a new processing, the algorithm compares:

- The profit obtained at the end of next processing if the operator continues his work activities without break (Rno-break).
- The profit obtained at the end of next processing if the operator does a break of minimum length before (Rbreak).

The profit per unit is given by:

$$R = (HR + HEP \cdot P_r) \cdot P - CF_{STD} - CV_{STD} - HEP \cdot C_r \cdot P_r - c_b \cdot T_b \qquad (3.18)$$

where HR is the operator reliability; P is the item price; HEP is the probability of failure (1-HR); CF_{STD} is the standard fixed cost; CV_{STD} is the standard variable cost; P_R is the probability of recovery; C_R is the cost of recovery; cb is the cost of breaks per minute; and Tb is the break time in minutes. The break cost per minute is related to the loss due to the failure current work piece. It is therefore expressed as

$$c_b = \frac{((HR + HEP \cdot P_r) \cdot P - CV_{STD} - HEP \cdot C_r \cdot P_r)}{T_{total}} \qquad (3.19)$$

where T_{total} is the total time of processing in minutes that considers the time increment linked to possible rework:

$$T_{total} = (HR + HEP) \cdot T_c + HEP \cdot P_r \cdot T_r \qquad (3.20)$$

where T_c is the processing time; T_r is the time required for the reworking, defined as percentage increase of the processing time; and P_R is the probability of recovery.

Starting from the minimum break time given as input, the algorithm compares the two profits (Rbreak-Rnobreak) and it decides if the break is convenient. If the minimum break is convenient SHERPA increases the break one minute by one until it remains economically convenient (Rbreak+>Rbreak). In this way SHERPA provides the possibility of determining the optimal breaks scheduling.

III.3.4 Learning and forgetting Curves Model (LFCM)

The Learning and Forgetting Curves Model (LFCM), according to what proposed by Jaber and Bonney in 1996, has been implemented as a further module in the SHERPA simulator.

The learning curves are based on the clear improvements that occur when the workers learn how to do a job through the production of more and more units, decreasing the production time per unit (Azizi, Zolfaghari and Liang, 2010). This phenomenon is observed by the decrease in production time per unit as operators gain experience by producing additional units (Nembhard and Osothsilp, 2001). The learning impact on the system performance changes when the operator does a break of sufficient length and the forgetting process starts to take place. In this case the production time of the first unit after the break tends to be longer than the production time of the last unit before the break (Nembhard and Osothsilp, 2001).

The impact of these processes on the performance of repetitive tasks has been widely studied and applied in various sectors, like manufacturing, healthcare, energy, information technologies, education, design and banking (Anzanello and Fogliatto, 2011; Jaber and Glock, 2013; Grosse, Glock and Müller, 2015). Knowing how humans learn in production systems and how learning and forgetting affects the performance of the production processes is important for several reasons (Anzanello and Fogliatto, 2011). Considering learning in production planning may contribute to a significant reduction in total costs or to an improvement in the productivity (Jaber and Glock, 2013). Furthermore, it could have a positive effect on the human error rate. The experience gained in performing a repetitive task involves, in fact, a decrease in the production time which could improve the human reliability, considering for example the higher available time for the execution of the task. A learning curve is a mathematical description of workers' performance in repetitive tasks; in fact in every repetition the workers tend to demand less time to perform tasks due to familiarity with the operation and tools (Anzanello and Fogliatto, 2011). The Wright's learning curve (1936) is the earliest and the most popular model observed in an industrial setting that expresses an exponential relationship between direct man-hour input and cumulative production in the form of:

$$T_q = T_1 \cdot q^{-b} \tag{3.21}$$

where T_q is the time to produce the qth unit, q is the production count, T_1 is the theoretical time required to produce the first unit, and b= -ln(LR)/ln(2) is the learning slope, where LR is the learning rate (Jaber and Bonney, 1996). Typical *LRs* according to Givi (Z S Givi, Jaber and Neumann, 2015) are shown in Table III.18. Equation (3.21) shows that with growing production the unit time decreases by a constant percentage each time the quantity doubles.

At first this phenomenon has been studied individually but it is strongly correlated with the forgetting process. In intermittent production runs, there is a break of sufficient length that some learning accumulated in producing q units in the previous lots is not retained when a new run starts up (Jaber and Bonney, 1996). Hence, the production rate at the recommencement would not be as high as when the production ceased. The increase in time to produce the first unit in the next production run depends on the length of the interruption and the time to produce the qth unit which is when the interruption occurred.

Table III.18: *Learning rates for different tasks (Givi, Jaber and Neumann, 2015).*

TYPE OF WORK	LR%	INDUSTRY	LR%
Assembly	84-85	Aerospace	85
Prototype assembly	65	Complex machines	75-85
Clerical ops	75-85	Construction	70-90
Inspection	86	Electonix mfg	90-95
Machining	90	Machine shop	90-95
Welding	85-90	Shipbuilding	80-85

* Table III.18 was adopted from Crawford (1944).

To take forgetting effects into consideration, a handful of theoretical, experimental and empirical mathematical forgetting models have been developed, with no unanimous agreement among researchers and practitioners on the form of the forgetting curve (Jaber and Sikström, 2004). Carlson and Rowel describe the forgetting by a negative decay function comparable to the decay observed in electrical losses in condensers. Their curve is one of the most widespread models and it is expressed in the form of (Jaber and Bonney, 1996):

$$\hat{T}_x = \hat{T}_1 \cdot x^f \qquad (3.22)$$

where $\hat{T}_x$ is the time for the xth unit of lost experience of the forgetting curve, x is the amount of output that would have been accumulated if interruption did not occur, $\hat{T}_1$ is the equivalent time for the first unit of the forgetting curve, and f is the forgetting slope (Jaber and Bonney, 1996).

Starting from equations (3.21) and (3.22), (Jaber and Bonney, 1996) developed the learning and forgetting curves model (LFCM). The LFCM assumes that there is one learning curve and one forgetting curve, with the forgetting curve having its slope and intercept adjusted after each production break. The exponent of the power forgetting function depends on the total forgetting time, the learning slope and the amount of equivalent units of cumulative production accumulated by the point of interruption (Jaber and Sikström, 2004):

$$f_i = \frac{b(1-b)\log(q_i)}{\log(C+1)} \qquad i = 1, 2, 3 \ldots \qquad (3.23)$$

where $0 \leq f \leq 1$, which varies in every cycle, is the forgetting slope after interruption in cycle *i*, *b* is the learning rate, *q* is the number of units produced in cycle i up to the point of interruption and C=$B/T(q)$ is the ratio of B, the minimum time for total forgetting, to *T (q)*, the amount of time required to produce q units. As shown in Figure III.17, at the point of interruption in cycle *i,* the curves have the same value. After every break of length t_b the numbers of units remembered at the beginning of cycle *i+1* is determined by the equation:

$$\alpha_i = q_i^{(b+f_i)/b}(q_i + s_i)^{-f_i/b} \qquad i = 1, 2, 3\ldots \qquad (3.24)$$

where (q_i+s_i) is the sum of q units produced in the cycle i and s is the total number of products that could have been produced in cycle i if production was not interrupted:

$$(q_i + s_i) = [\frac{1-b}{T_1}T_b + q_i^{1-b}]^{\frac{1}{1-b}} \qquad i = 1, 2, 3\ldots \qquad (3.25)$$

where T_b is the length of the interruption that has occurred after cycle *i.* Therefore, the time to produce the first unit in the next production run is:

$$T_{q+1} = T_1(\alpha + 1)^{-b} \qquad i = 1, 2, 3\ldots \qquad (3.26)$$

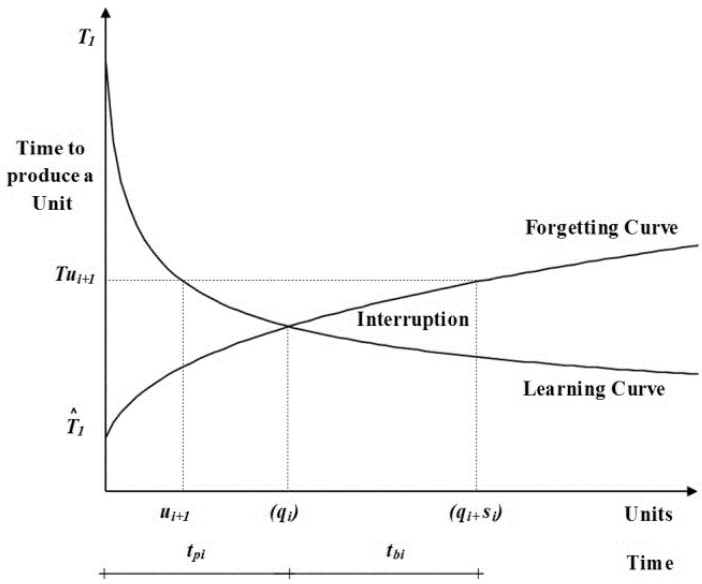

Figure III.17: *The curves at the end of cycle i: tpi is the time in production to produce qi units in cycle i; tbi is the length of the interruption period cycle i; si is the potential additional quantity that would be produced if no interruption occurred (Jaber and Bonney, 1996).*

The LFCM has been proven to conform with most requirements of learning and forgetting models and it benefits from easy calculations (Jaber and Bonney, 1997; Jaber and Sikström, 2004). The LFCM could be easily used in simulation models of work activities that involve repetitive tasks in order to assess the impact of these processes on the system performance and on the processing time, as reported in literature.

LFCM module in SHERPA allows evaluating how the human reliability and the learning and forgetting processes impact on the system performance. The human reliability assessment identifies the human error probability (HEP) and the rate of non-compliant performed task, while the LFCM algorithm allows the quantification of the decrease in the processing times, taking into account both the single shift and more consecutive working days. During a single shift the forgetting effect is caused by the rest breaks and it is less evident due to their short time, while between two consecutive working days the interruption is longer, and it has a greater impact.

Figure III.18 shows the logic of the LFCM algorithm. The input data are: the type of the performed task; the production time of the first unit T_1; the time of total forgetting B; the length of the break T_b and the chosen time horizon T_{end}. The first algorithm step provides the selection of the learning slope b as function of:

- learning rate LR of the working sector ($b_{moderate}$);
- performed task with more or less learning complexity ($b_{slow}=b_{moderate}+5\%$; $b_{fast}=b_{moderate}-5\%$).

The following phases allow to model the curves for all the possible cycles i separated by breaks of variable lengths. The production times, under the learning phenomenon for each cycle i, are quantified according to the equation (1):

$$T_{q+1} = T_1(q+1)^{-b} \tag{3.27}$$

where the learning slope b is determined as previously described; T_1 is an input data and q is the units counter. The q counter is initially set equal to zero and it represents the number of experience units, i.e. the units remembered before each new activity. In addition to the units counter, the counter of work cycles i and the counter of *Tcycle*, which identifies the time worked without breaks in the cycle i, are initialized. If breaks do not occur in the working activity, the algorithm calculates the improved production times until it reaches the constrain *Tend*.

The forgetting phenomenon takes place when a break of length T_b occurs, involving the decrease of the gained experience. It determines the decrease of the remembered units, when the operator starts to work in the cycle i+1, and the increase in the production time compared to the time achieved in the last processing of the previous cycle i. In this case the processing time is quantified through the equation (3.26) where α is determined as previously shown by the equation (3.24). The α-value is attributed to the q counter changing the q value assigned in the previous cycle.

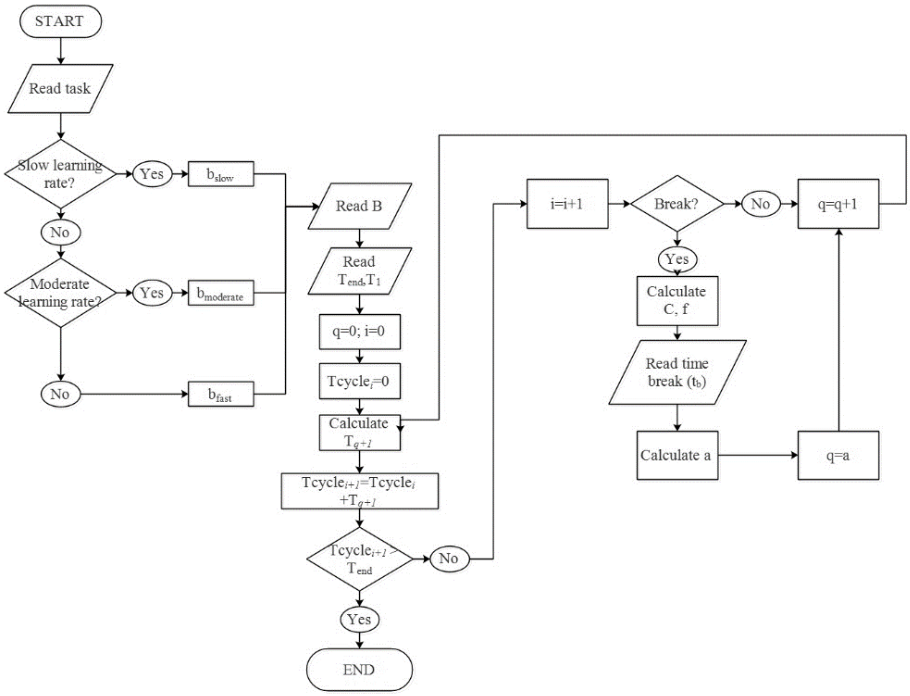

Figure III.18: *The LFCM algorithm.*

The LFCM algorithm is modelled into the theoretical framework of SHERPA. The LFCM module assumes, on the basis of the findings previously discussed, that the performance improves with the increase in cumulative production (or alternatively, number of repetitions of a specific task). The performance deteriorates, instead, when learning sessions are separated by breaks that result in knowledge depreciation or forgetting (Jaber & Sikstrom, 2004).

LFCM is adopted taking into account both the single shift and more consecutive working days. During a single shift the forgetting effect is caused by the rest breaks and it is less evident due to their short lengths, while between two consecutive working days the interruption is longer and has a greater impact (Figure III.19). The learning rate is fixed according to the working area for which the simulation is performed and, thanks to this value, the levels of learning slope (slow, moderate and fast) are quantified as reported in Table III.19.

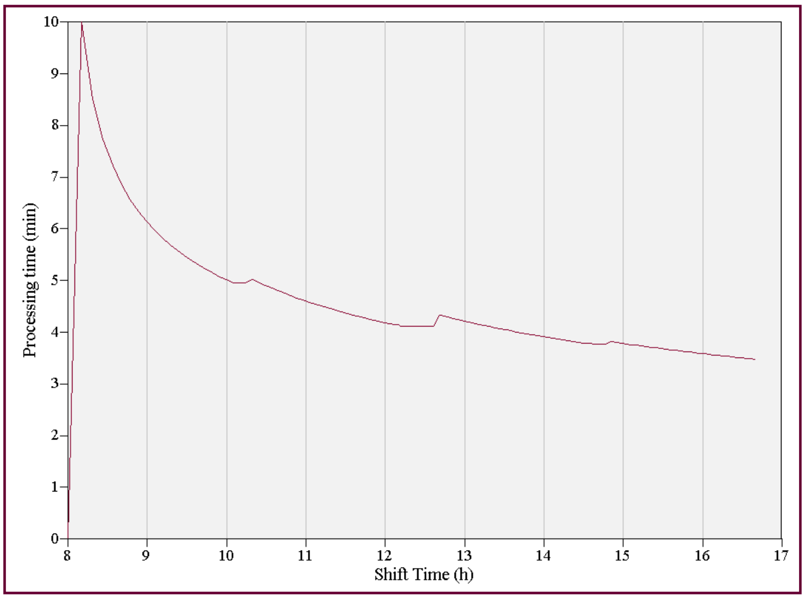

Figure III.19: *Learning and forgetting effects in a work shift.*

Table III.19: *Generic task and learning slope b.*

Generic task used in SHERPA		Learning slope b
1	Totally unfamiliar	b_{slow}
2	Complex task requiring high level of skill	b_{slow}
3	Fairly simple task performed rapidly	$b_{moderate}$
4	Routine, highly-practiced	$b_{moderate}$
5	Completely familiar, well-designed, highly practiced, routine task	b_{fast}
6	Respond correctly to system command even when there is an automated supervisory system	b_{fast}

III.3.5 Entities exit

The main SHERPA outputs are compliant, non-compliant, and rework entities. These categories are derived from the forecast of HEP on the basis of the performed activity of the period when the process is carried and of the contextual and individual conditions. This concept of quality defects and non-compliant entities is not limited to manufacturing processes, but extends to a wider range of working environments, ranging from services to medical field.

As shown in the flowchart (Figure III.20), each entity in the output from the system receives the compliant percentage and function of the error probability to overturn the human performances on the system ones. The reworking entails an increase of the processing time, as in Figure III.20. Finally, economic results in terms on profit (eq. 3.18) are quantified after each simulation.

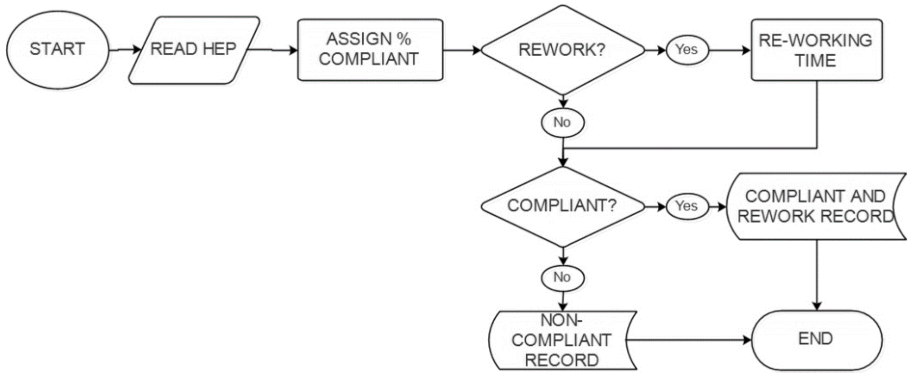

Figure III.20: *Entities exit.*

Chapter IV: Simulation tools

IV.1 Introduction

The theoretical framework has been implemented as simulation template in Arena 14.0© and Anylogic. It allows a lot of different scenarios to be simulated easily without consuming a lot of time by changing the type of activity, the influencing factors, and especially the break configurations. The simulator can be used to assess the impact of different rest breaks scheduling in every working context and conditions, moreover it can dynamically analyse a whole shift identifying the moments of the highest operator unreliability and automatically managing the breaks in order to reduce errors and to increase productivity and efficiencies.

IV.2 Simulator features and structure

The logical model of the simulator has been designed independently of the current implementation in the Arena and Anylogic environments. The simulator is designed to represent the activities performed by a single operator during the work shift, estimating in real time the reliability curves, the effect of recovery due to the work breaks and the impact of the operator's performance on the system under examination.

The main feature of the simulation template is the possibility to be specialized and configurable for different working contexts by modifying the inputs through suitably implemented dialog boxes. Another important feature is to make possible its integration in any simulation model from the simplest (one operator and repetitive processing) to the most complex (more operators interacting on complex activities).

The SHERPA structure is presented again in Figure IV.1. The simulator can be considered divided into two parts. A first "physical" part, in which the working process (*Process Simulation*) is performed taking the incoming entities and returning in output compliant, non-compliant and rework entities.

The distinction between success or failure of the performed activity is, instead, based on the logical part of the model (*Human reliability quantification*), which allows the HEP quantification, as described in Chapter III, and is also a function of the assigned break scheduling. The four modules in Figure IV.1 correspond to the four main sections implemented in the simulator.

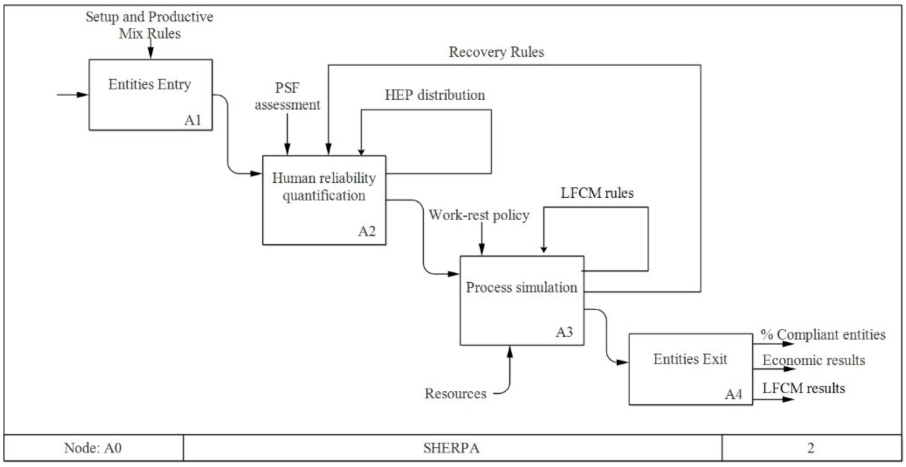

Figure IV.1: *SHERPA decomposition overview.*

Entities entry, that includes the implementation of the user interfaces to the definition of the inputs (worker data, performed activity data, PSF, break scheduling). The user interfaces are built to allow the rapid and guided insertion of all the variables. Particular importance is assumed by the levels of the PSF that follow those defined theoretically. Choosing one level over another allows you to easily modify the simulated scenario. Human reliability quantification allows, instead, to assign the nominal distribution parameters for HEP and the value of PSF multipliers and to quantify HEP contextual. Process simulation, instead, involves the implementation of the three alternatives break scheduling, of the LFCM module and finally of the entities exit to quantify the HEP impact on system performance.

The designed simulation template has been implemented both in Arena and in Anylogic. The simulation tools were implemented through several steps:

1) Design and implementation of the user interfaces;
2) Implementation of the logic model: entities entry, human reliability quantification, process simulation with break scheduling module, LFCM module and entities exit.
3) Check errors and debug form.

The two simulation environments, Arena and Anylogic, allowed to implement the theoretical model of SHERPA respectively in a template and in an agent resource. These simulation tools offer numerous advantages to simulate a large number of scenarios without being resource intensive or time consuming. In particular, Anylogic makes it possible to check the trend of all real-time model variables, and it is indicated for more complex models to evaluate their settings and to allow their validation. While the Arena template provides a tool easy to use and effective as a decision support system.

IV.3 Arena SHERPA Template

The model for the evaluation of human reliability and the management of breaks in the workplace, described in the previous chapter, was firstly implemented through the Arena 14.0® software, a widely used application for the simulation of manufacturing-oriented systems. Arena Simulation, a product of the US company Rockwell Automation, is an advanced simulation software that provides an interactive environment for the construction, animation, verification and analysis of simulation models. The basics of the Arena language are (Kelton, Sadowsky and Sadowsy, 2006):

- Entities: objects that flow through the system, such as customers, parts, vehicles, information, or logical elements, etc. Through the system, entities can change state, be influenced by other entities in the system and in turn affect system performance. They are dynamic objects within the simulation (they are usually created, they move within the system and then are released).
- Code: waiting areas where the movement of entities is temporarily suspended.
- Resources: system components that must be allocated to entities, such as machines, operators, switchboards, etc. An entity commits resources when it is available and releases it when it has finished processing.
- Attributes: are a common characteristic of all entities, but with a specific value that can change from one entity to another. Attributes make it possible to individualize entities, such as the type of processing, arrival time, etc.
- Variables: represent values that describe the status of the system or process, such as the number of available machines, the number of setups, etc. A system can have several variables, but each is unique, all entities can access a variable but cannot change.

The Arena template appears as a block flowchart that can be used in various types of simulation models and allows the assessment of human error probability without excessive time consumption Figure IV.3. The template is shown in Figure IV.2. It can be inserted in the processing flow and it simulates the task performed by operator during a whole shift, considering the time required to complete the task as well as calculating the operator's human reliability (HR) and human error probability (HEP).

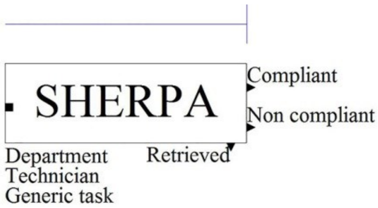

Figure IV.2: *SHERPA template user interface.*

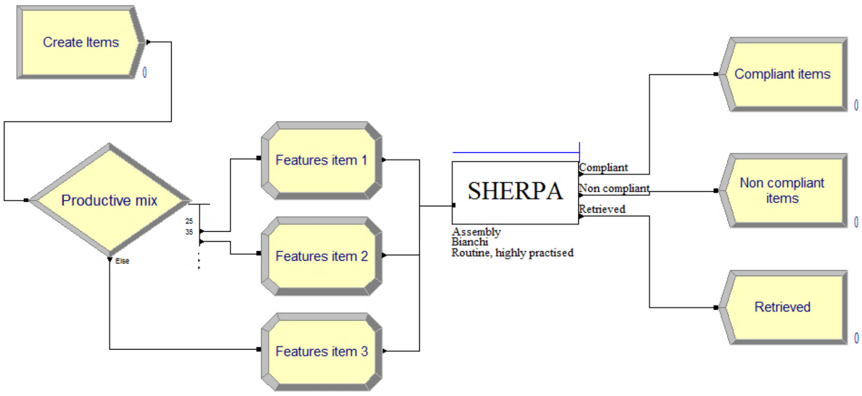

Figure IV.3: *SHERPA template used in a simulation model.*

IV.3.1 Design and implementation of the dialog boxes

In this first phase all the dialog boxes were designed and implemented through the Dialog Design Window. These windows and the related operands are described below. The main SHERPA dialog box, shown in Figure IV.4, allows access allows to initialize the template by entering all the information related to the scenario to be simulated; thus, every input can be defined. The SHERPA dialog box provides access to several other windows for data entry. Figure IV.5 shows the sub-dialog boxes that are linked to the main dialog of the HRA module. The initial dialog box allows access to all the forms for entering the data necessary for the simulation. Scrap recovery operation, as well as LFCM, are not mandatory and can be selected through the specific check box.

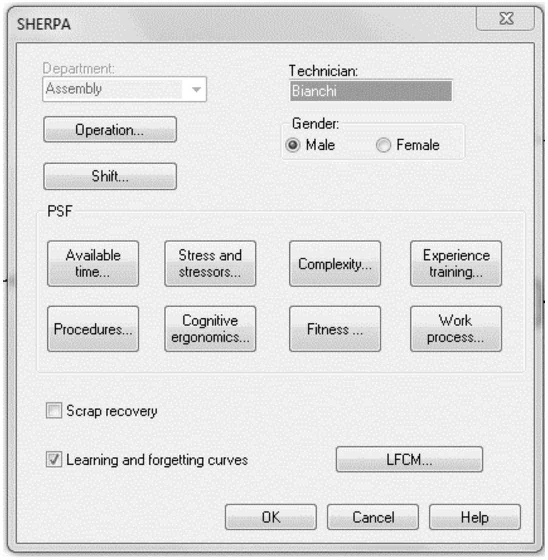

Figure IV.4:*Main dialog SHERPA template.*

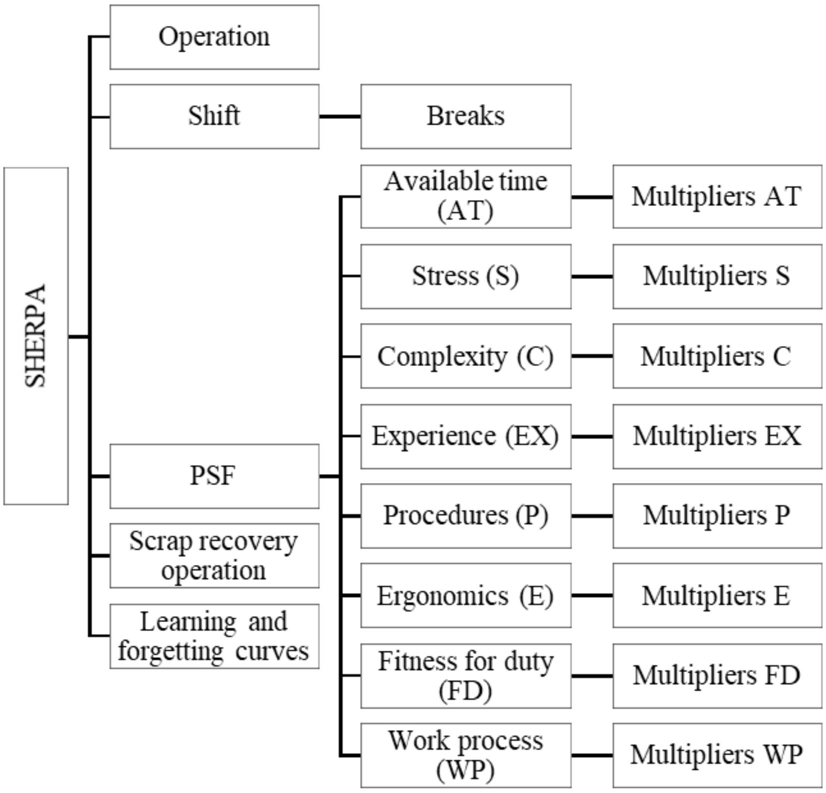

Figure IV.5: *Connections between sub-dialog boxes and dialog man.*

Each dialog box is defined by several operands. The operands are particular objects that allow the definition of a variable, of its default value, of the type of variable and of the representation in the data entry form (Kelton, Sadowsky and Sadowsy, 2006). In each window you can enter different types of operand, the most important are shown in Table IV.1.

Table IV.1: *Operand typologies.*

Toolbox Controls Type	Description
Text	Used to show a line of text on the dialog box
TextBox	Used to insert a user input or to display text
ComboBox	Used to view and edit data
RadioButtonGroup	Used to display a set of two or more mutually exclusive choices. The choices are presented in an array of buttons
CheckBox	Used to indicate if a particular condition is enabled or not. The possible choice is Yes / No or True / False
DialogButton	Used to insert a button that opens another window
HiddenOperand	Used to define a hidden object

These objects can be enabled or hidden using switches. A switch is a construct that allows you to define variations of (Kelton, Sadowsky and Sadowsy, 2006):

1) Fields displayed in a dialog box;
2) Logic of the model and the elements that are created during the simulation;
3) Animation objects displayed in the user view of a module.

The operands of the main SHERPA dialog box, shown in Figure IV.4, are defined in Table IV.2 with the following information:

- ✓ Name of the operand.
- ✓ Type of operand.
- ✓ Switch presence: indicates the presence of a switch assigned to the operand. The switch allows you to hide or disable the operand when required.
- ✓ Brief description of the operand function.

- ✓ Notes: insert further information on the considered operand.

Table IV.2: *Operand properties of the HRA dialog box.*

Name	Type	Switch	Description
HRA	Main dialog	No	Definition initial dialog box
Department	Combo box	No	Input of the production department
Technician	Textbox	No	Input of technician's name who performs work
Gender	Radiobuttongroup	No	Choice gender of the operator
Operation	Dialog button	No	Entry to secondary dialog box operation
Shift	Dialog button	No	Entry to secondary dialog box shift
Scrap recovery	Check box	No	Choice of the possible scraps recovery operation
Scrap recovery operation	Dialog button	swscrap recovery01	Entry to secondary dialog box scrap recovery opertion
Learning and forgetting curves	Check box	No	Choice of the possible learning and forgetting curves
LFCM	Dialog button	swlfcm01	Entry to secondary dialog box LFCM
Available time	Dialog button	No	Entry to secondary dialog box available time
Stress and stressors	Dialog button	No	Entry to secondary dialog box stress and stressors
Experience training	Dialog button	No	Entry to secondary dialog box experience and training
Fitness	Dialog button	No	Entry to secondary dialog box fitness for duty
Complexity	Dialog button	No	Entry to secondary dialog box complexity
Procedures	Dialog button	No	Entry to secondary dialog box procedures
Cognitive Ergonomics	Dialog button	No	Entry to secondary dialog box ergonomics and HMI
Work processes	Dialog button	No	Entry to secondary dialog box work processes
Entry item	Hidden operand	No	Defines the input connection to the template
Good items	Hidden operand	No	Module output (good items)

Name	Type	Switch	Description
Recovered items	Hidden operand	No	Module output (recovered items)
Scrap	Hidden operand	No	Module output (scraps)

The information about the task to be performed can be inserted into the dialog box Operation (Figure IV.6). The window allows the selection of the type of activity and the generic task performed by the operator. Such information is necessary to identify the nominal human reliability curve in order to quantify the total HR in the simulation. The operand contents are summarized in Table IV.3.

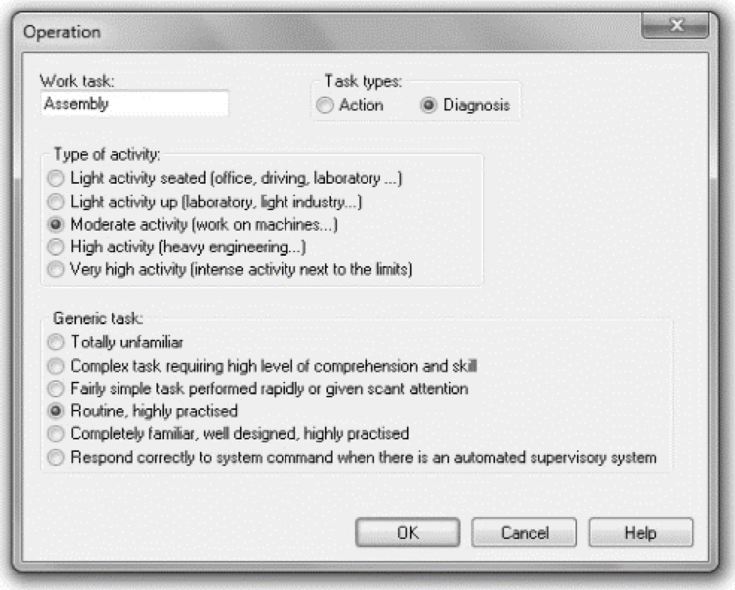

Figure IV.6: *Dialog box for operation data entry.*

Table IV.3: *Operand properties of the dialog box operation.*

Name	Type	Switch	Description
Work task	Text box	No	Definition work task to be performed
Task types	Radiobuttongroup	No	Choice between action and diagnosis
Generic task	Radiobuttongroup	No	Choice of generic task from the list
Workload	Radiobuttongroup	No	Choice of generic task from the list

In the proposed simulation module, the length shift in hours and the start time of the same can be inserted, in order to be able to carry out the simulation in all possible working conditions. Figure IV.7 shows the window in the three possible configurations that it can assume, based on the choice of the type of break scheduling. The configurations identify the three possible alternatives break scheduling (Figure IV.7):

1) No break, continuous shift: in this case neither "scheduled breaks" nor "automatic management of breaks" are chosen;

2) Scheduled breaks: by choosing "scheduled breaks" the main dialog *Breaks* is shown, in which the information on the breaks is inserted (described below);

3) Automatic management of the breaks: the module is thus automatically selected by specifying only the length (expressed in minutes) of the minimum break.

The operand contents are summarized in Table IV.4.

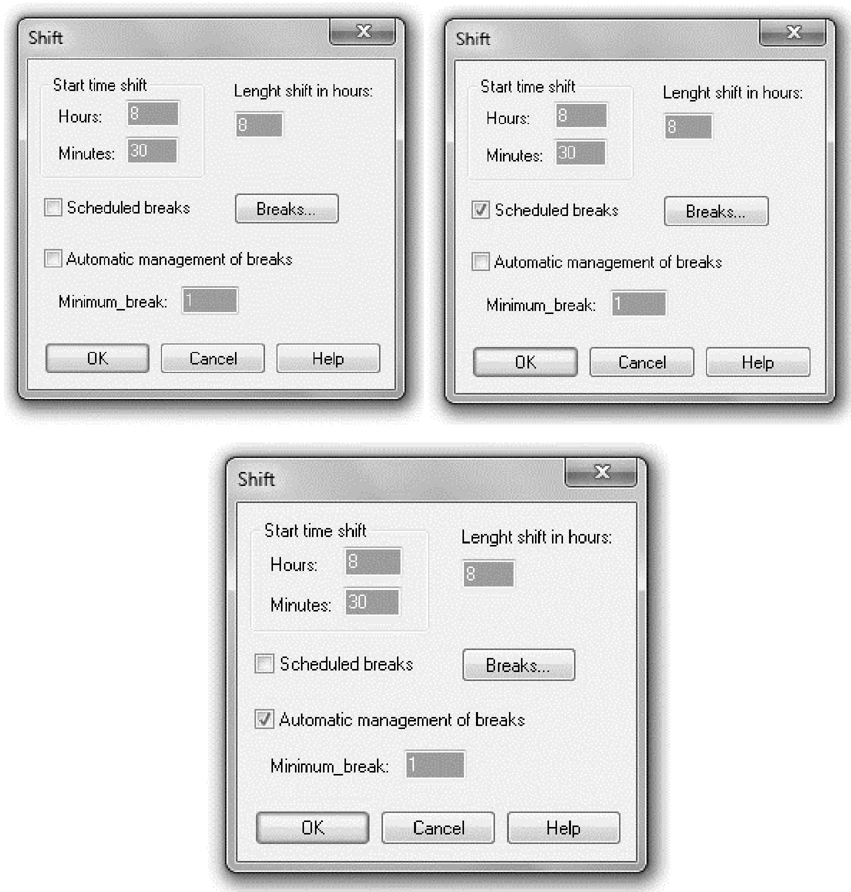

Figure IV.7: *Dialog box shift data entry.*

Table IV.4: *Operand properties of the dialog box shift.*

Name	Type	Description	Notes
Hours	Text box	Input work schedule (Hours)	Min 0 Max 23
Minutes	Text box	Input work schedule (Minutes)	Min 0 Max 59
Length shift in hours	Text box	Input length shift (Hours)	Min 4 Max 12
Scheduled breaks	Check box	Choice of scheduled breaks	Required
Breaks	Dialog button	Entry to secondary dialog box breaks	Required
Automatic management of breaks	Check box	Choice of automatic management of breaks	Required
Minimum break	Text box	Input minimum duration for break	Min 0 Max 59

The secondary dialog box break (Figure IV.8) allows entering the scheduled work-rest configuration in two different ways:

- Fixed length: breaks have the same duration, and they are assigned at regular intervals;
- Variable length: up to six breaks can be assigned with different lengths and at variable intervals.

Table IV.5 shows all the operand of the box and its properties.

The other operands of fundamental importance for the evaluation of human reliability are those related to the eight PSFs of the SPAR-H method, analysed in parallel with the logic developed in the following section.

It should be noted that the attributes necessary for the HRA module are not only those entered via the dialog boxes, but some of them must be defined outside the module and assigned as attributes to the entities entering the system. These, required in various sections of the module, such as setup and break scheduling management, are: setup time; job; processing time; price/value added of the processing; standard fixed costs and standard variable costs.

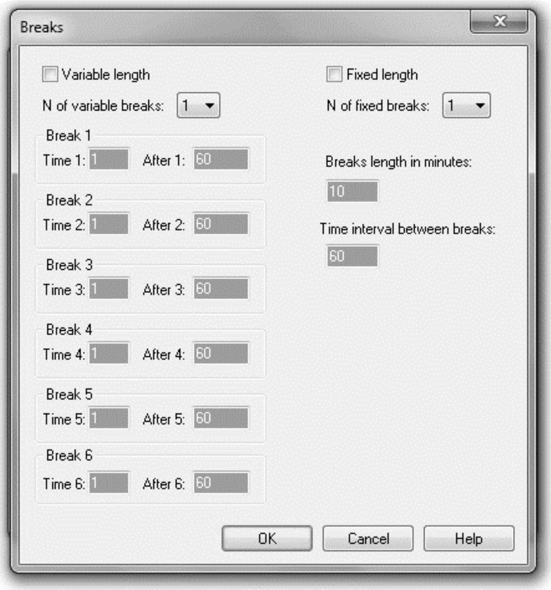

Figure IV.8: *Dialog box breaks data entry.*

Table IV.5: *Operand properties of the dialog box breaks.*

Name	Type	Description	Notes
Fixed length	Check box	Choice type of break (with fixed length)	-
N of fixed breaks	Combo box	Choice number of breaks	Min 1 Max 7
Breaks length in minutes	Text box	Input duration time	Min 0 Max 59
Time interval between breaks	Text box	Input time between two successive breaks	Min 0 Max 480
Cost break per minute	Text box	Input cost break per minute	Min 0
Variable length	Check box	Choice of type of break (with variable length)	
N of variable breaks	Combo box	Choice number of breaks	Min 1 Max 7
Time 1	Text box	Input duration time	Min 0 Max 59
After 1	Text box	Input time between two successive breaks	Min 0 Max 480

IV.3.1.1 Performance shaping factors dialog boxes

A specific dialog box for each PSF has been designed that makes it possible to choose and select the PSF level or to enter the information necessary to calculate the multiplier.

The available time, for example, has six levels, both positive and negative, that differ between action or diagnosis, as shown in Figure IV.9. In the available time dialog box, there is a dialog button that allows to modify the values of the multipliers (Figure IV.10). This type of window is present for every PSF, it allows to use the default values of the multipliers, but also to modify them to suit different operational situations. Table IV.6 shows all the operands of the two windows.

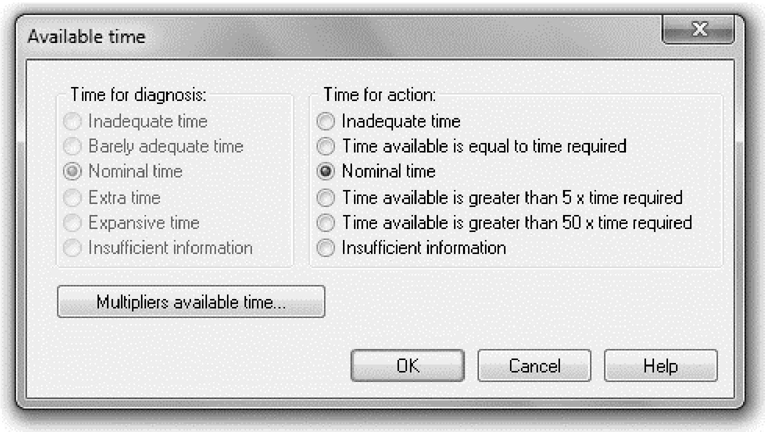

Figure IV.9: Dialog box available time data entry.

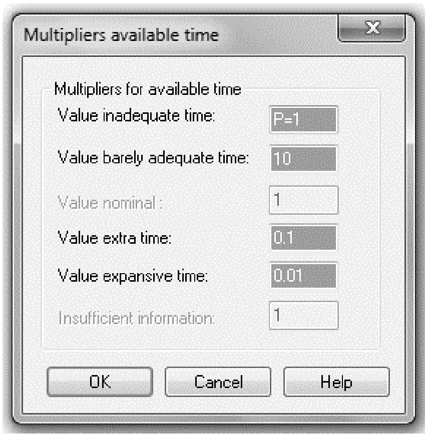

Figure IV.10: *Dialog box multipliers available time.*

Table IV.6: *Operand properties of the dialog box available time.*

Name	Type	Switch	Description
Time for diagnosis	Radiobutton group	swDiagnosis	Choice of level from the list
Time for action	Radiobutton group	swAction	Choice of level from the list
Multipliers available time	Dialog button	No	Entry to secondary dialog box multipliers
Value inadequate time	Text box	No	Input multipliers
Value barely adeguate time	Text box	No	Input multipliers
Value nominal	Text box	No	Input multipliers
Value extra time	Text box	No	Input multipliers
Value expansive time	Text box	No	Input multipliers
Insufficient information	Text box	No	Input multipliers

Unlike the available time, as explained in Chapter III, stress factor includes several sub-factors: mental stress; pressure time; workplace; circadian rhythm; microclimate; lighting; noise; vibration and ionizing and non-ionizing radiation. The presence of different factors makes the stress window richer and more complex as shown in Figure IV.11.

The levels of circadian rhythm, mental stress and pressure time can be selected directly from the dialog box (Figure IV.11). The remaining factors have been designed separately because the quantification of the level depends on multiple values and different situations. Appendix B reports the dialog boxes of all the sub-factors of stress. Table IV.7 shows all the operands in the dialog box stress and stressors.

Table IV.7: *Operand properties of the dialog box stress and stressors.*

Name	Type	Switch	Description	Notes
Circadian rhythm	Radiobutton group	No	Choice of level from the list	
Weight circadian rhythm	Combo box	No	Choice of weight	Min 0 Max 1
Mental stress	Check box	No	Selection of mental stress factor	

Name	Type	Switch	Description	Notes
Level mental stress	Radiobutton group	Swmental stress01	Choice of level from the list	
Weight mental stress	Combo box	Swmental stress01	Choice of weight	Min 0 Max 1
Pressure time	Check box	No	Selection of pressure time factor	
Level pressure time	Radiobutton group	Swpressure time01	Choice of level from the list	
Weight pressure time	Combo box	Swpressure time01	Choice of weight	Min 0 Max 1
Other stressors	Group box	No	Entry to secondary dialog box multipliers	

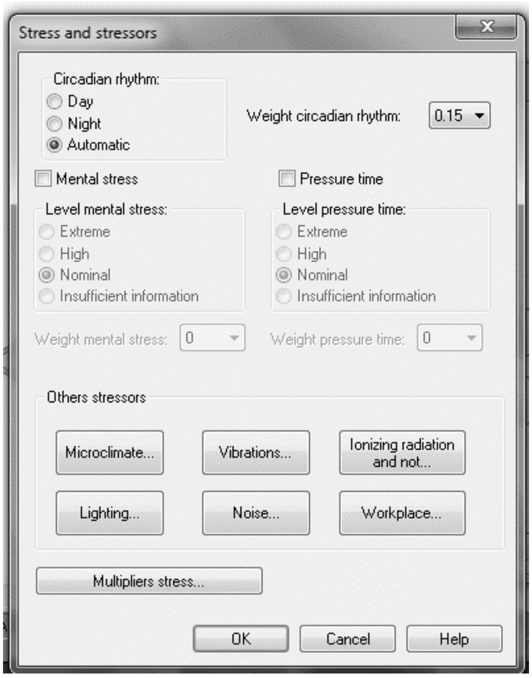

Figure IV.11: *Dialog box stress data entry.*

Figure IV.12 shows the user interface for the PSF complexity. The level of physical effort required is obtained directly from the operand Type of activity present in the Operation window. For each sub-factor the user can choose between different levels, explained in the clearest possible form, to ensure a good modelling of contextual factors. Table IV.8 summarizes all the operands and their properties.

Table IV.8: *Operand properties of the dialog box complexity.*

Name	Type	Switch	Description	Notes
General complexity	Radiobutton group	swDiagno sis	Choice of level from the list	
General complexity for action	Radiobutton group	swAction	Choice of level from the list	
Weight general complexity	Combo box	No	Input Weight	Min 0 Max 1
Precision level of activity	Radiobutton group	No	Choice of level from the list	
Weight precision of activity	Combo box	No	Input Weight	Min 0 Max 1
Mental efforts required	Radiobutton group	No	Choice of level from the list	
Weight mental effort	Combo box	No	Input Weight	Min 0 Max 1
Parallel tasks	Radiobutton group	No	Choice of level from the list	
Weight parallel tasks	Combo box	No	Input Weight	Min 0 Max 1
Weight physical efforts	Combo box	No	Input Weight	Min 0 Max 1
Multipliers complexity	Dialog button	No	Entry to secondary dialog box	
Value highly complex	Text box	No	Input multipliers	Min 2 Max 100
Value moderately complex	Text box	No	Input multipliers	Min 2 Max 100
Value nominal 4	Text box	No	Input multipliers	Require d
Obvious diagnosis	Text box	No	Input multipliers	Min 0 Max 1
Insufficient information 1	Text box	No	Input multipliers	Require d

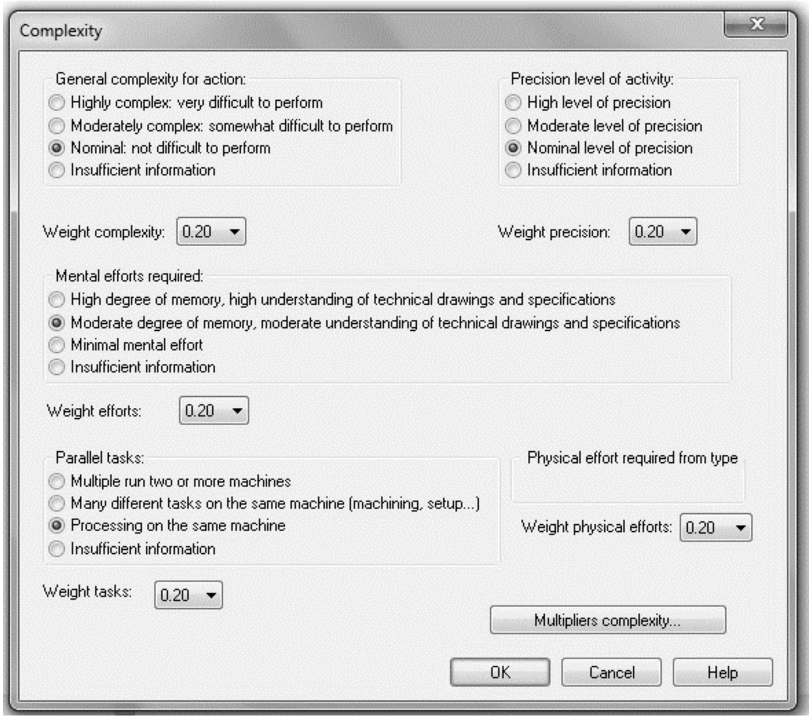

Figure IV.12: *Dialog box complexity data entry.*

The operator's experience and training can be quantified with greater simplicity, as they are defined directly by the operator's number of months of experience (Figure IV.13). All operands are summarized in Table IV.9.

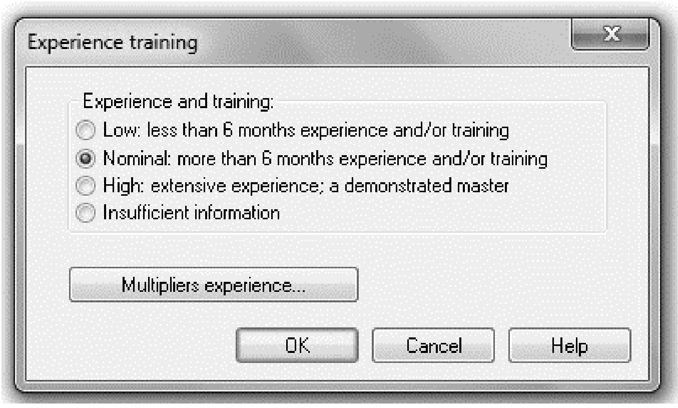

Figure IV.13: *Dialog box experience data entry.*

Table IV.9: *Operand properties of the dialog box experience.*

Name	Type	Description	Notes
Experience and training	Radiobutton group	Choice of level from the list	
Multipliers experience	Dialog button	Entry to secondary dialog box	
Value short for action	Text box	Input multipliers	Min 2 Max 100
Value short for diagnosis	Text box	Input multipliers	Min 2 Max 100
Value nominal 2	Text box	Input multipliers	Required
Value long	Text box	Input multipliers	Min 0 Max 1
Insufficient information 2	Text box	Input multipliers	Required

Procedures PSF refers to the existence and use of formal operational procedures for the activities under consideration. The evaluation of the PSF is not always easy and immediate. In the dialog box two different alternatives are given in the case of action and diagnosis, as shown in Figure IV.14. Table IV.10 shows all the operands of the two windows.

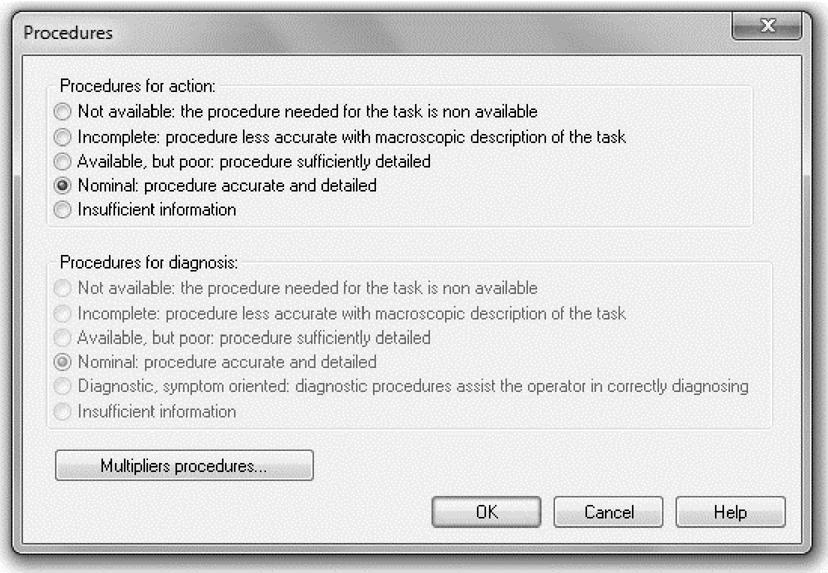

Figure IV.14: *Dialog box procedures data entry.*

Table IV.10: *Operand properties of the dialog box procedures.*

Name	Type	Switch	Description	Notes
Procedures for action	Radiobutton group	swAction	Choice of level from the list	
Procedures for diagnosis	Radiobutton group	Sw Diagnosis	Choice of level from the list	
Multipliers for procedures	Dialog button	No	Entry to secondary dialog box	
Not available	Text box	No	Input multipliers	Min 2 Max 100
Incomplete	Text box	No	Input multipliers	Min 2 Max 100
Available but poor	Text box	No	Input multipliers	Min 2 Max 100
Value nominal 5	Text box	No	Input multipliers	Required
Diagnostic	Text box	No	Input multipliers	Min 0 Max 1
Insufficient information 5	Text box	No	Input multipliers	Required

Fitness for duty considers if worker is physically and mentally fit or not to perform the task at the required time (Figure IV.15). Another important PSF is the cognitive ergonomics. Figure IV.16 shows the dialog box designed for the SHERPA module. The alternatives proposed for the ergonomic level are those defined in the SPAR-H method. Table IV.11 and Table IV.12 summarize the operands with their properties.

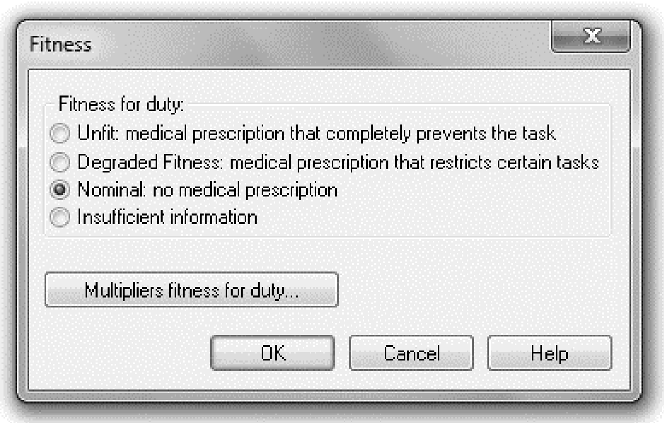

Figure IV.15: *Dialog box fitness for duty data entry.*

Table IV.11: *Operand properties of the dialog box fitness for duty.*

Name	Type	Description	Notes
Fitness for duty	Radiobutton group	Choice of level from the list	
Multipliers for fitness for duty	Dialog button	Entry to secondary dialog box	
Value degraded fitness	Text box	Input multipliers	Min 2 Max 100
Value nominal 3	Text box	Input multipliers	Required
Insufficient information 3	Text box	Input multipliers	Required

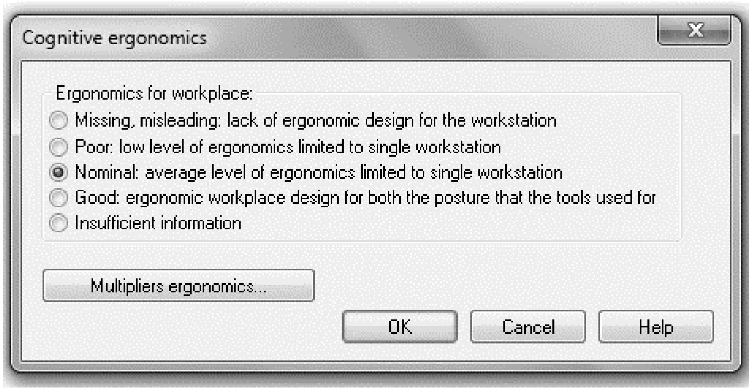

Figure IV.16: *Dialog box cognitive ergonomics data entry.*

Table IV.12. *Operand properties of the dialog box ergonomics.*

Name	Type	Description	Notes
Ergonomics for workplace	Radiobutton group	Choice of level from the list	
Multipliers for ergonomics	Dialog button	Entry to secondary dialog box	
Value missing	Text box	Input multipliers	Min 2 Max 100
Value poor	Text box	Input multipliers	Min 2 Max 100
Value nominal 6	Text box	Input multipliers	Required
Value good	Text box	Input multipliers	Min 0 Max 1
Insufficient information 6	Text box	Input multipliers	Required

The last PSF is work processes, which is divided in two sub-factors: Communication and integration; Work processes (Figure IV.17). All operands are summarized in Table IV.13.

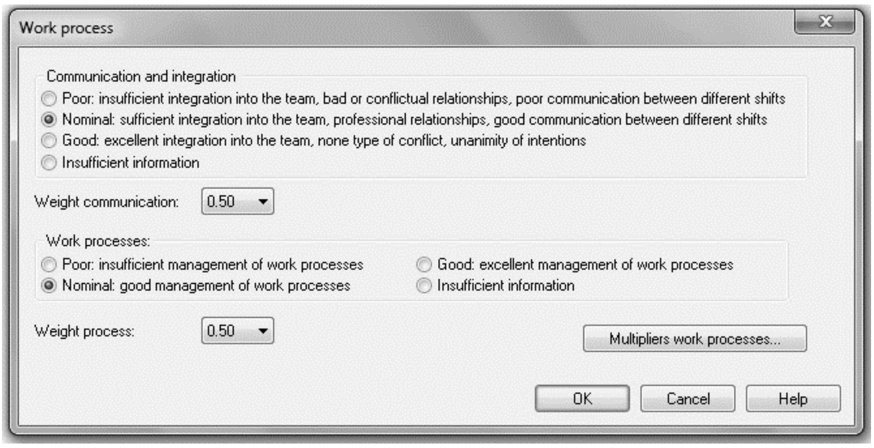

Figure IV.17: *Dialog box work process data entry.*

Table IV.13: *Operand properties of the dialog box work process.*

Name	Type	Description	Notes
Communication and integration	Radiobutton group	Choice of level from the list	
Weight communication	Combo box	Input weight	Min 0 Max 1
Work process	Radiobutton group	Choice of level from the list	
Weight work processes	Combo box	Input weight	Min 0 Max 1
Multipliers work process	Dialog button	Entry to secondary dialog box	
Value poor for action	Text box	Input multipliers	Min 2 Max 100
Value poor for diagnosis	Text box	Input multipliers	Min 2 Max 100
Value nominal 7	Text box	Input multipliers	Required
Value good for action	Text box	Input multipliers	Min 0 Max 1
Value good for diagnosis	Text box	Input multipliers	Min 0 Max 1
Insufficient information 7	Text box	Input multipliers	Required

IV.3.2 Logical implementation

The logic of the template, which represents the heart of the SHERPA module, was implemented in the Logic Window and it is not visible to the user during the simulation runs. The simulation module is structured in seven sections:

1) Input module.
2) Setup module.
3) Human reliability quantification.
4) Simulation without breaks.
5) Simulation with scheduled breaks.
6) Simulation with automatic break scheduling management.
7) Recovery items.

Figure IV.19 shows the logical SHERPA model used to implement the theoretical model in Arena simulation environment. The first section manages the entry of entities into the system, which must respect as constraint the work shift assigned as input. SHERPA has been designed to be used with any type of work shift.

The information related to the beginning of the shift and its length are, in fact, required in the Shift dialog box. Based on this information, *Tstart* and *Tend* variables are calculated, which represent respectively the start and end of the daily shift. *Tstart* and *Tend* are used to manage entry of entities both during working hours and during non-working hours. These variables are defined in the section input module, shown in Figure IV.18.

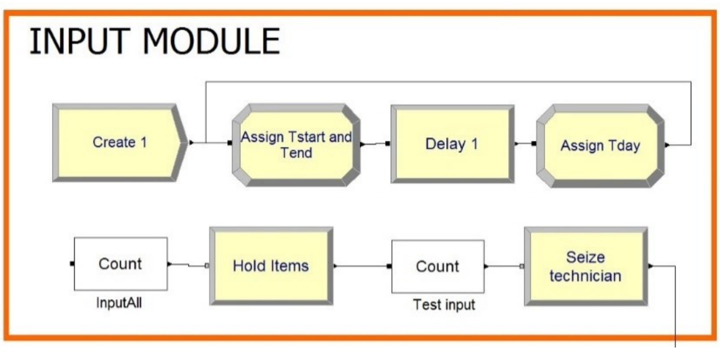

Figure IV.18: *Input Module.*

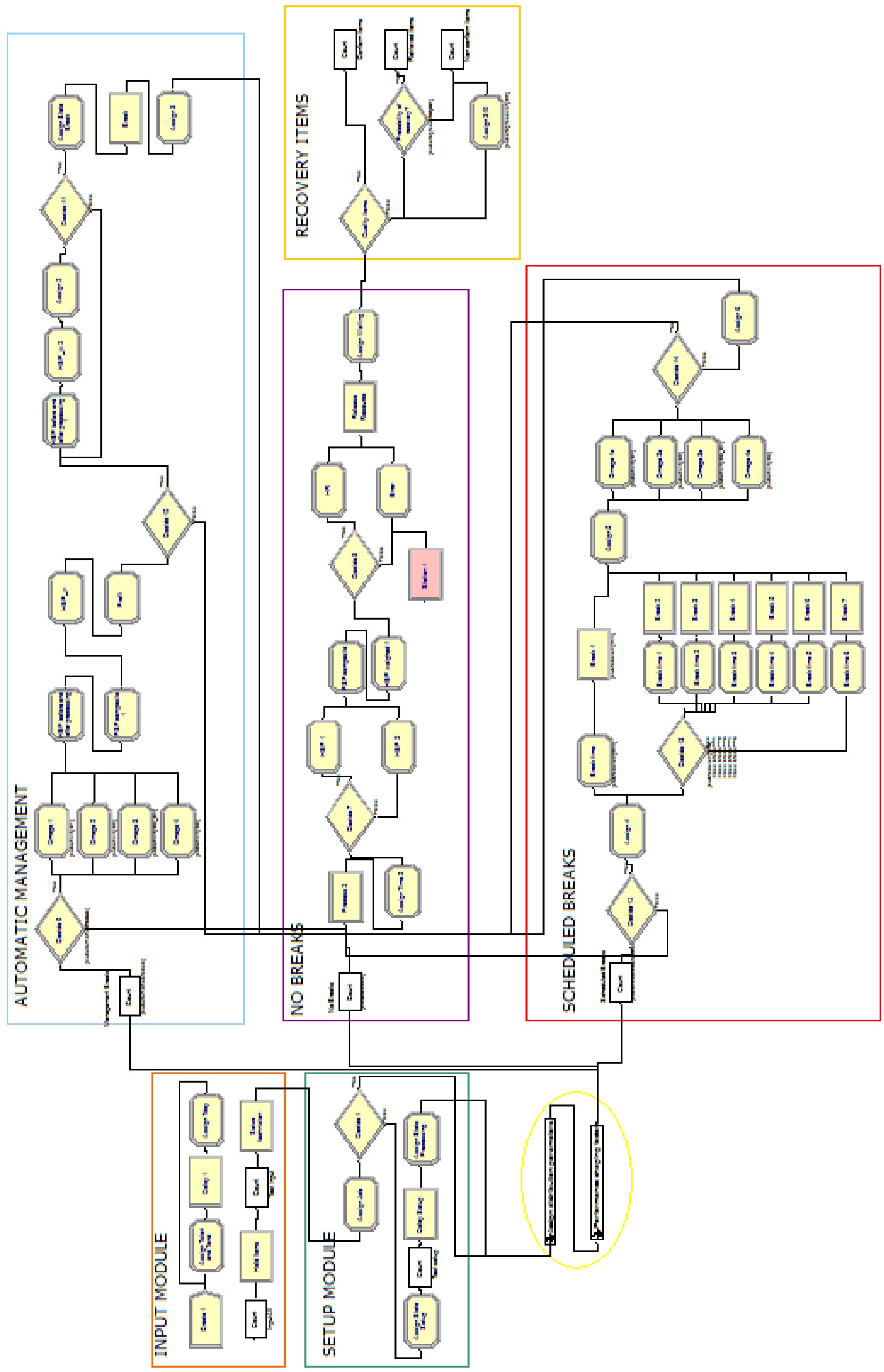

Figure IV.19: *Logical SHERPA model implemented in Arena.*

Once entered through the "*Count: InputAll*" block, the entities are kept in the *Items.Queue* queue, managed via FIFO logic, until the condition is true:

TNOW> Tstart TNOW && <Tend

where *TNOW* represents the current simulation time.

At the beginning of the simulated work shift, the first entity into the system using the *Seize technician module* (Figure IV.18) to engage the operator for the execution of the activity. Once the operator is busy, the entities enter the setup module (Figure IV.20). This SHERPA section allows to manage the setup between machining of different entities, if required.

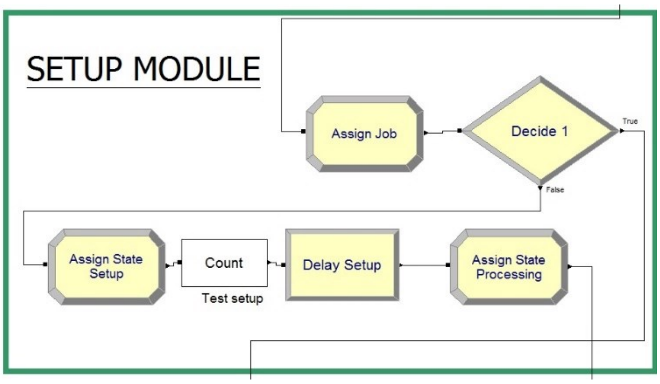

Figure IV.20: *Setup Module.*

IV.3.2.1 Human reliability quantification

The first phase for the evaluation of human reliability is the assignment of the parameters of the Weibull distribution. In the sub-model *Assign distribution parameters* (Figure IV.21), alpha, beta and kappa values are assigned to calculate the nominal HEP distribution, based on the generic task selected by user in dialog box for operation data entry (Figure IV.6).

In this sub-model, the three variables of time, used in the model, are initialized:

- Technician_Relative time = time elapsed since the start of the shift, calculated as the difference between the current time and the start of the work shift.
- Technician_Time = time for calculating operator reliability. It is calculated using the formula:

$$\text{Technician}_{\text{Time}} = \text{Technician}_{\text{Time}} + \text{Technician}_{\text{RelativeTime}} - \text{Technician}_{\text{Last Time}}$$

- Technician_Last time = variable used to calculate the Technician_Time.

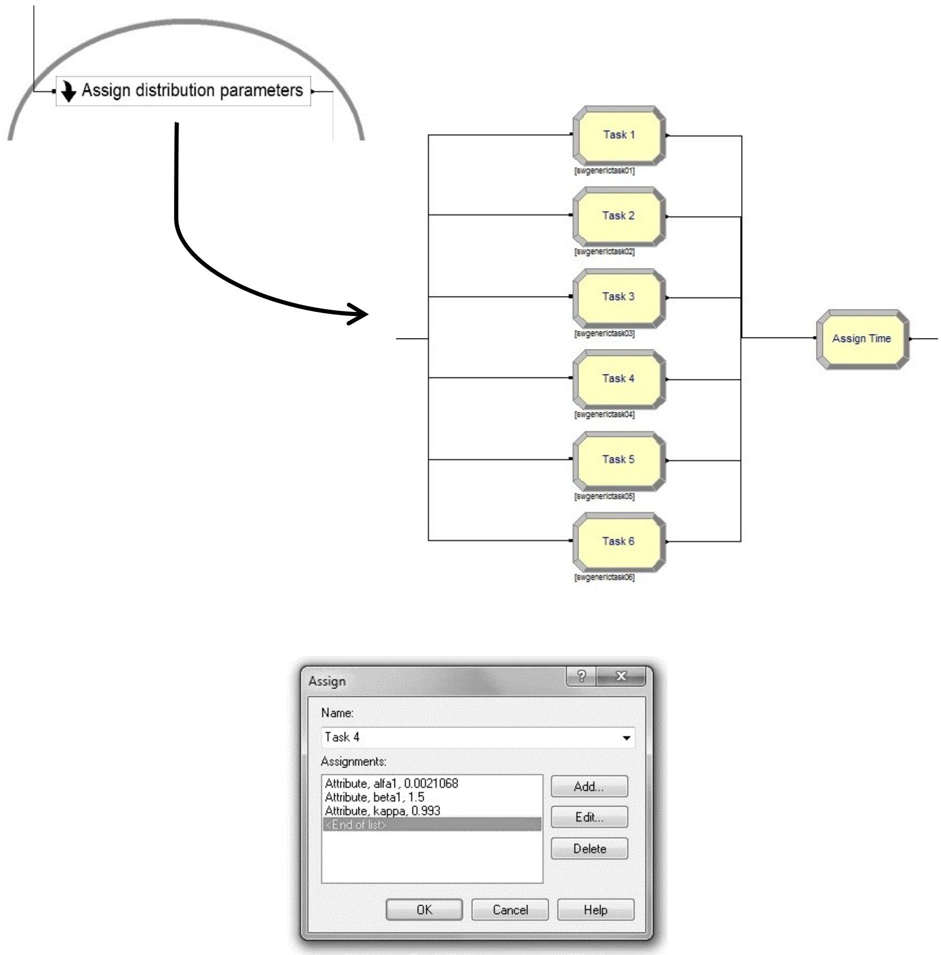

Figure IV.21: *Assign distribution parameters.*

The next phase consists in the evaluation and quantification of the PSF impacts on HEP in the sub-model *Performance shaping factors* (Figure IV.22). Each of the influencing factors is implemented in a sub-model, in which the respective value of the PSF multiplier is assigned, according to the data entered the module through the dialog boxes previously described. The logic, implemented to assign the multiplier of the PSF, is the same for all PSFs whose level is defined and assigned directly by the user through dialogs, such as available time, experience, procedures, ergonomics and fitness for duty.

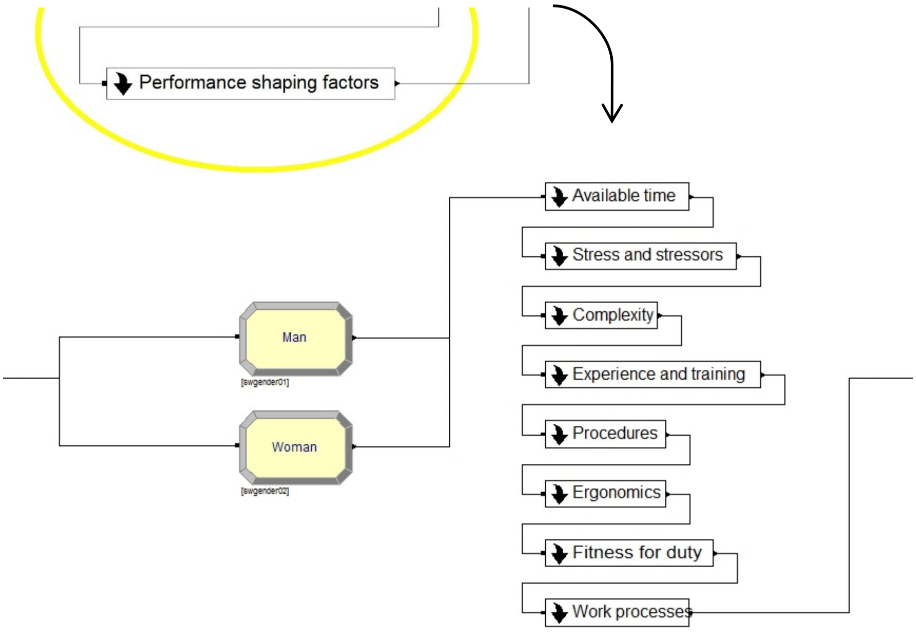

Figure IV.22: *Sub-model performance shaping factors.*

By way of example, the case of available time is shown. The level chosen by the user in the dialog box (Figure IV.9) is used by the specific sub model and it is transferred to the module through the logical part (

Figure IV.**23**). Figure IV.24 shows as *PSF_time* attribute is assigned.

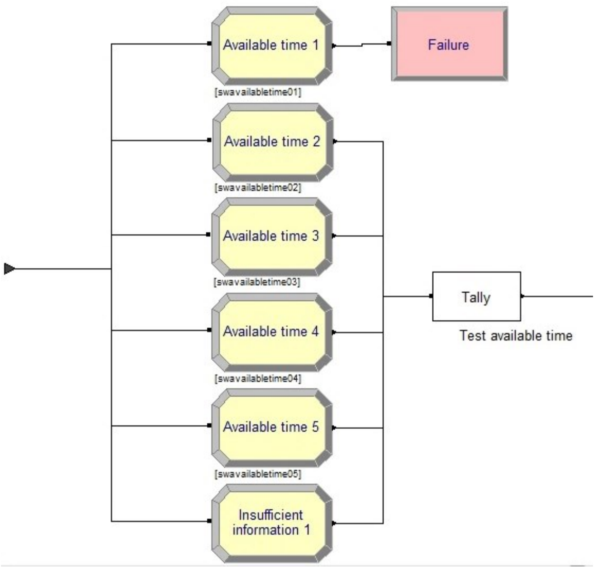

Figure IV.23: Logic assignment of PSF available time.

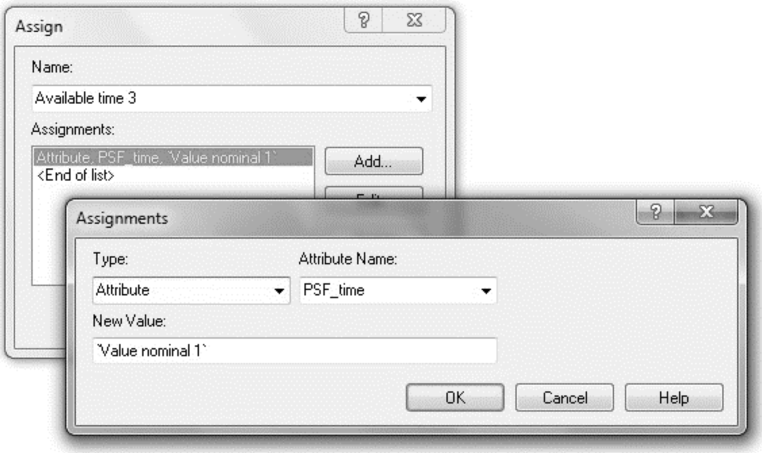

Figure IV.24: *Assignment of the PSF_time attribute.*

The different switches direct the entity to the appropriate *Assign* block in which the *PSF_time* attribute is assigned with the respective value of the multiplier (Table IV.14). Once the value of the *PSF_time* has been assigned, the *Tally* block allows to memorize this value and bring it back to the end of the simulation.

Stress, complexity and work processes, instead, were modelled to take into account their sub-factors. The PSF stress, for example, is divided into nine different sub-factors. For some of them, circadian rhythm, mental stress, pressure time and workplace, the multiplier is assigned, as for the available time, according to the level selected in the user interface (Figure IV.25 and Figure IV.26).

Table IV.14: *Switches used in PSF sub-model.*

Name	Condition
swavailabletime01	\`Time for action\`=="Inadequate Time"||\`Time for diagnosis\`=="Inadequate Time"
swavailabletime02	\`Time for diagnosis\`=="Barely Adequate Time"||\`Time for action\`=="Time available is equal to time required"
swavailabletime03	\`Time for diagnosis\`=="Nominal time"||\`Time for action\`=="Nominal time"
swavailabletime04	\`Time for diagnosis\`=="Extra time"||\`Time for action\`=="Time available is greater than 5 x time required"
swavailabletime05	\`Time for diagnosis\`=="Expansive time"||\`Time for action\`=="Time available is greater than 50 x time required"
swavailabletime06	\`Time for diagnosis\`=="Insufficient information"||\`Time for action\`=="Insufficient information"

The remaining factors were implemented separately because the quantification of the level depends on multiple values and different situations, as reported in Chapter III. For each of them a sub-model has been created, as shown in Figure IV.27, in order to assign each multiplier through a specific attribute. Each attribute assigned is then multiplied by the relative weight and the *Tally* blocks allow keeping track of the values assumed. The same logic was implemented for complexity and work processes.

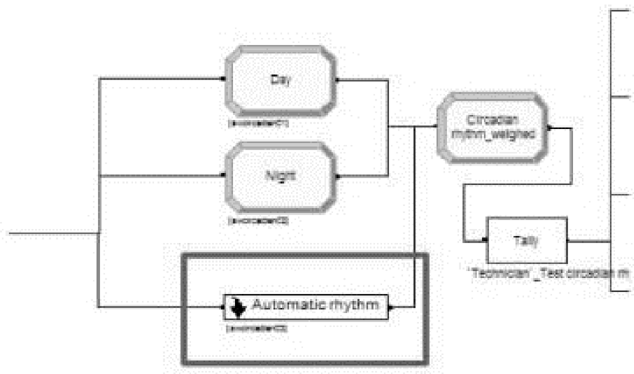

Figure IV.25: *Sub-model of the stress factor (part 1).*

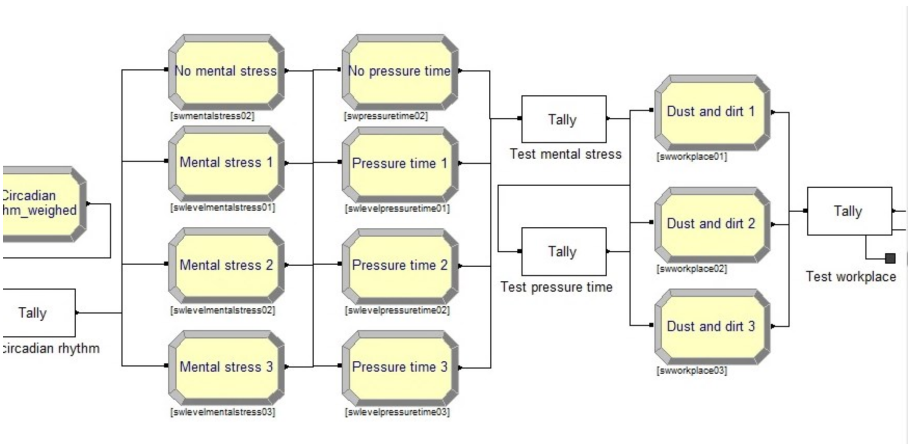

Figure IV.26: *Sub-model of the stress factor (part 2).*

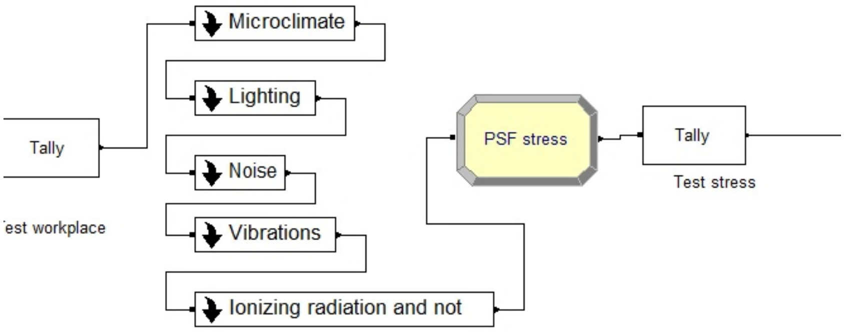

Figure IV.27: *Sub-model of the stress factor (part 3).*

IV.3.2.2 Process simulation for break scheduling management

SHERPA aims to quantify the worker reliability and to use this knowledge for managing the work-rest configurations. Human reliability, or rather the complementary HEP, strongly affects the performance of the operator. The first part of the template, analysed previously, allows obtaining information on the HEP curves, with a careful analysis of PSFS, which represent the most elaborate part of the module. In the second part of the template, instead, a generic work process was implemented, on the basis of which it is possible to simulate an assigned or automatic break scheduling. There are three sections:

1) Work process simulation without breaks: operator works continuously without stopping for a break.
2) Work process simulation with scheduled breaks: the operator stops for a break only when it is scheduled, based on the information inserted in the shift dialog box, previously described.
3) Work process simulation with automatic break scheduling management: the algorithm developed for the model assigns the optimal break scheduling on the basis of the economic and reliability assessment.

The first section is shown in Figure IV.28 and Figure IV.29. The entities are processed according to the FIFO logic and *Count: No Breaks block* allows you to consider the number of entities processed.

Although it is defined as a No Breaks section, all entities cross this section even in the case of scheduled breaks or in automatic management. In these cases, the entities are first sent to the respective section where the break scheduling is managed and then returned to the No break section. This part of the model, in fact, aims to simulate any type of work process and to assess the reliability linked to this activity in terms of:

- worked time;
- performed activity;
- performance shaping factors.

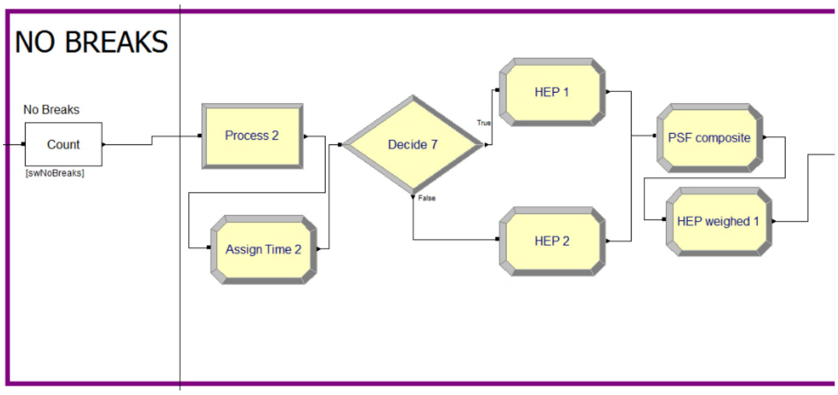

Figure IV.28: *Logic of work process without break (part 1).*

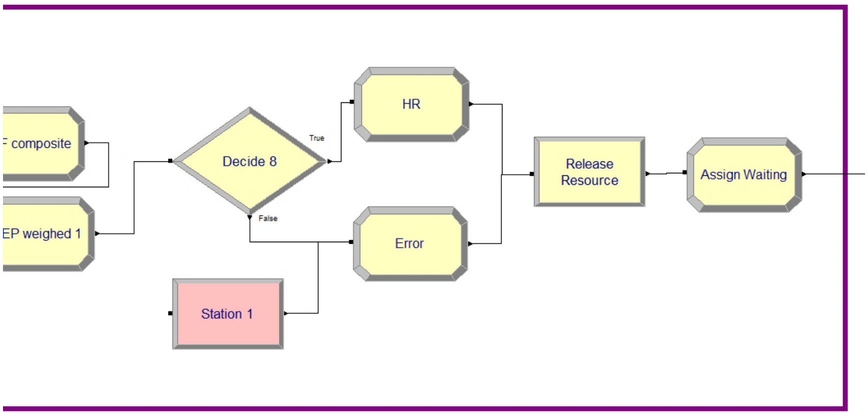

Figure IV.29: *Logic of work process without break (part 2).*

The processing time is variable and is assigned as input, as an attribute of the entities. After processing (*Process 2*), the worker reliability is calculated in the two phases described in Chapter III. First, the probability of nominal error is calculated in the HEP 1 and HEP 2 modules. The two modules calculate the HEP, defined by the *`Technician`_HEP* attribute in the two different conditions managed by *Decide 7*, which divides the entities processed within the first hour of work from those processed subsequently. Alpha, beta and kappa values assigned in *Assign distribution parameters* are used and once the nominal HEP has been determined, PSFcomposite is quantified: $PSF_{composite} = PSF_1 \times \ldots \times PSF_x \times \ldots \times PSF_8$; where PSF_x is the value assumed by the attributes defined in the previous phase, one for each PSF. If the value of the HEPcontext is higher than one, *Decide 8* directs the entities to the *Assign Error* where the attribute *`Technician`_HR*, ie the reliability of the operator, is set equal to one. On the other hand, if HEPcontext is not greater than one, the reliability is calculated as the complement to one of the HEPcontext. Once the process simulation is finished and the reliability calculated, the operator resource is released and then made available for other working activities.

For the management of scheduled breaks, the entities flow through the section shown in Figure IV.30 and Figure IV.31, to coordinate the intervals between the work and the fixed breaks. Entities enter this section thanks to the switch *swscheduledbreaks01: `Scheduled breaks` ==" Yes* "assigned to the *Count: Scheduled Breaks*. The logic of this section provides to transfer the incoming entities to *NoBreaks* section to carry out the processing if entities come when it is not scheduled a break. If the entity arrives in the period in which the pause has to be performed, it is retained to simulate the operator's stop and modify the time variables to identify the new reliability of the operator after the pause.

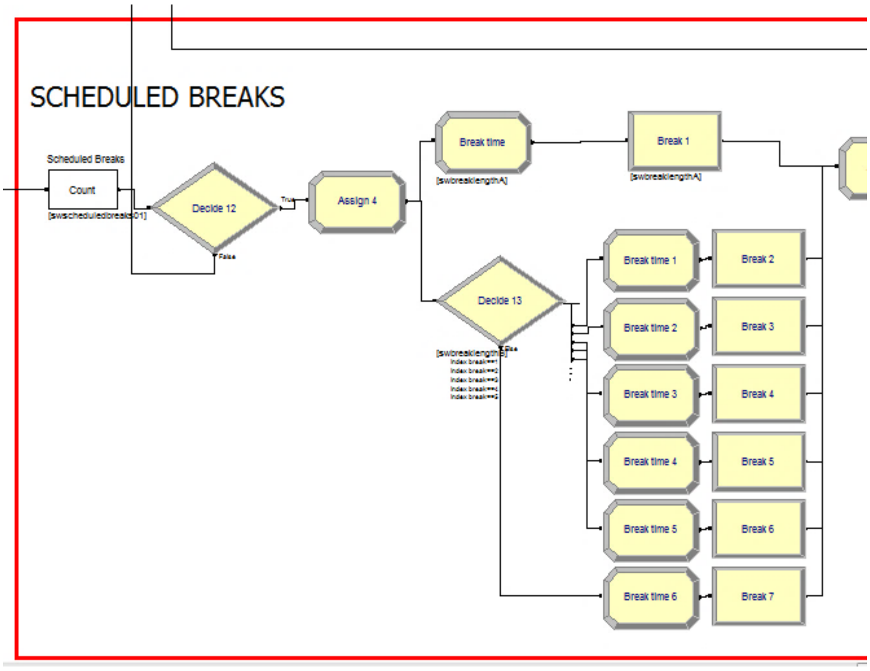

Figure IV.30: *Logic of work process with scheduled break (part 1).*

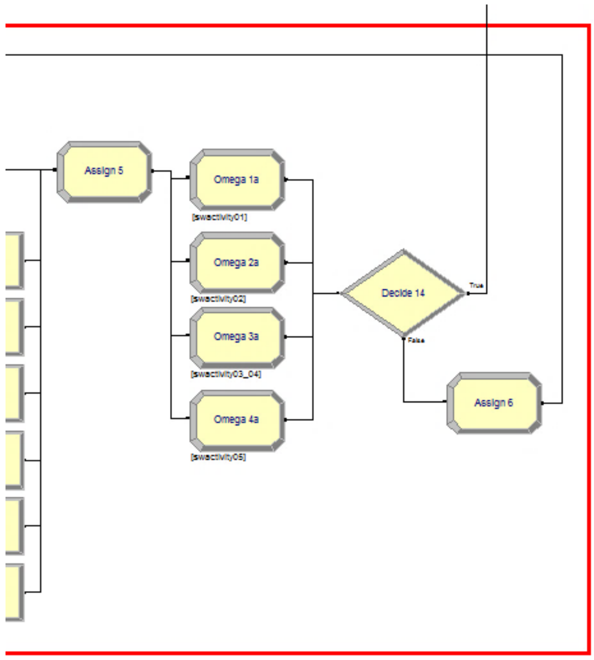

Figure IV.31: *Logic of work process with scheduled break (part 2).*

After the break has been simulated, the *Technician_Last time* variable is modified to insert the break time and the *Technician_time* variable to take into account the recovery due to the pause as follows:

$$Technician\_Last\ time = Technician\_Last\ time + (BreakDuration/60 \tag{4.1}$$

$$Technician\_time = Technician\_time \cdot r_p \tag{4.2}$$

This new time allows to calculate the new reliability of the operator in the *No Breaks* section.

The last type of break scheduling management is the automatic, based on the economic algorithm defined in Chapter III. Entities are referred to this section by the switch *swautomaticbreaks: 'Automatic management of breaks` == "Yes" && Scheduled breaks` == " No."* (Figure IV.32).

A basic assumption of the model is that the lowest probability of error is reached after the first hour of work, considering this hypothesis no breaks are necessary within the first hour. For this reason, the entities in input are sorted according to the *Technician_time* variable in *Decide 9*, if it is less than one processing is carried out directly, otherwise it continues to determine whether to pause. The logic in automatic management is to compare for each incoming entity the profit obtained by directly executing the task or by performing the processing downstream of a break. In the event that it is appropriate to pause, starting from the minimum duration the break is increased until the optimal time is identified, beyond which a break would not be more convenient (Figure IV.33).

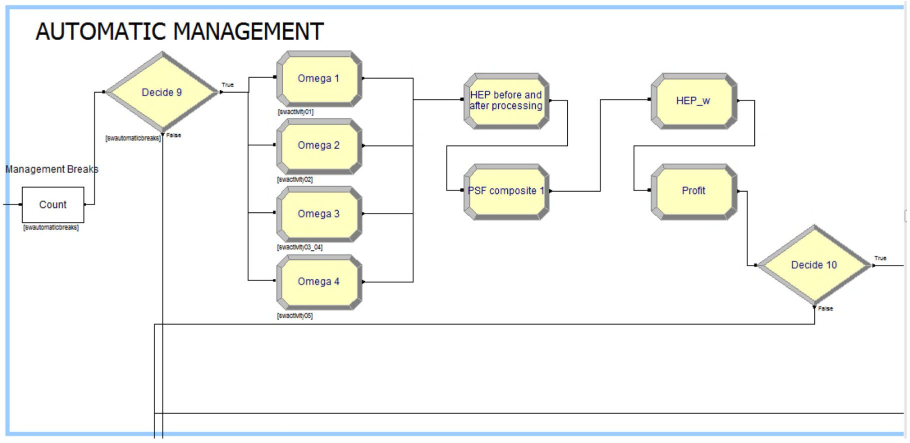

Figure IV.32: *Logic of automatic break scheduling management (part 1).*

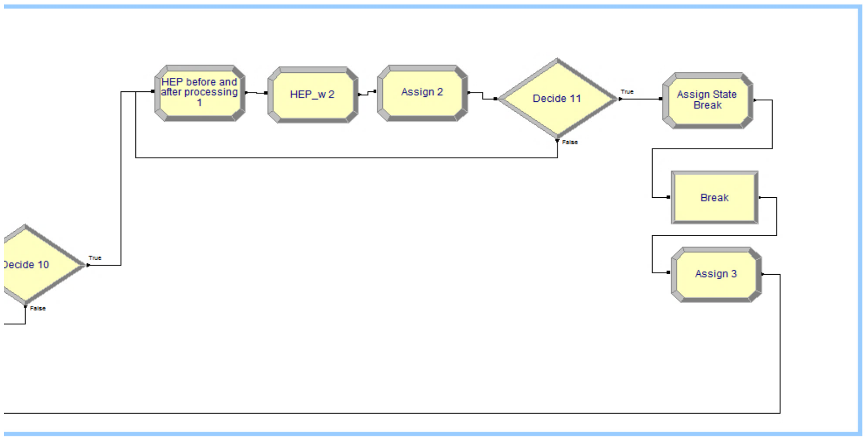

Figure IV.33: *Logic of automatic break scheduling management (part 2).*

To calculate the profit, it is naturally necessary to know the reliability of the operator in these two cases. HEP is calculated in the two alternative conditions for each entitie:

1) Immediate processing without breaks: in this case the reliability is calculated considering the working time reached by the operator at the end of the processing of the entity.

2) Break of minimum length and subsequent processing: it is assumed to make a minimum break (considering the minimum time selected as input) and then the processing. In this case the reliability is evaluated downstream processing, however considering the recovery of the operator for the break performed.

In the second case, the nominal HEP quantification is expressed as follows:

$$HEP_{nominal} = 1 - k \cdot e^{-\alpha \cdot (T \cdot r_p + T_p - 1)^{\beta}} \tag{4.3}$$

where T is the time in which the operator starts the break, r_p the recovery factor and Tp the duration of the subsequent processing. Once the values of the nominal HEP have been obtained with and without the break, the contextual factors are evaluated through the PSF composite. In the *Assign Profit* the Profit_after and Profit_before attributes are defined and calculated:

$$Profit_{before} = \left(HR_{before} + HEP_{before_w} \cdot P_r\right) \cdot Price - Cf - Cv - HEP_{before} \cdot C_r \tag{4.4}$$

$$Profit_{after} = \left(HR_{after} + HEP_{after_w} \cdot P_r\right) \cdot Price - Cf - Cv - HEP_{before} \cdot C_r - Minimum\ break \cdot cb_{after} \tag{4.5}$$

where HR and HEP are calculated in the previous steps and cb_after is the cost of the pause due to the lack of production.

The previous calculation of profit makes it possible to decide whether to pause; in fact, when *Profit_after* is greater than *Profit_before* with the *Decide 10* the entities are addressed towards the calculation of the optimal break, otherwise, since a break is not convenient, the entities enter the *No Breaks* section for process simulation. If a minimum break time is convenient, the model evaluates the possibility of increasing the break until reaching the optimal duration. In the *Assign HEP before and after processing 1* the variable *IncBreak* is defined, which allows to increase the length of the break one minute from time to time. This variable is added to the Minimum Break variable and the recovery factor, the *Profit_before*, recalculated equal to the previous *Profit_after*, and finally the new value of the probability of nominal and contextual error after the pause. In the *Assign 2* the *Profit_after* is then recalculated. On the basis of the new values the profit values are again compared to the minimum pause and increased pause conditions, until all the conditions in *Decide 11* are false, the break is increased by another minute, when at least one is true, the duration has been reached optimal pause. The status of Break is then assigned to the resource and the break is made through the Delay Break. To perform the downstream processing of the break taking into account the recovery of the operator, the variables *Technician_time* and *Technician_last time* are redefined. These variables will be used to calculate the reliability of the *No Breaks* section operator.

IV.3.2.3 Learning and forgetting module in SHERPA

The LFCM module can be easily selected in the main dialog box, as shown in Figure IV.34, and through the LFCM dialog box the algorithm input data, such as the total forgetting time *B,* the learning rate *LR* and the time T_1 can be inserted. The learning rate is fixed according to the working area for which the simulation is performed and, thanks to this value, the levels of learning slope (slow, moderate and fast) are quantified as reported in Table III.19.

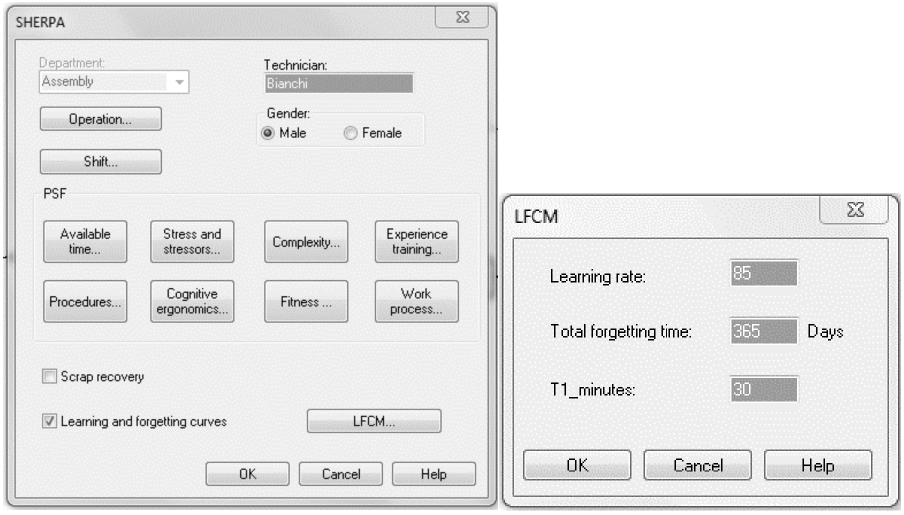

Figure IV.34: *Dialog box learning and forgetting data entry.*

The LFCM logic has been implemented in an Arena sub model that modifies the production time as a result of learning and forgetting processes. When a rest break, whether programmed or automatic, occurs the α-value is quantified, and it is used to decrease the internal q counter taking into account the phenomenon of forgetting during the work shift. At the same way the forgetting is quantified between two successive shifts. Downstream each simulation, the template provides as outputs the learning and forgetting curves and the average production time, in addition to the outputs already given.

The LFCM module may be used to assess how the learning and forgetting impact on productive performances in terms of pieces produced and of human error rates.

IV.3.3 Entities exit

Finally, entities were managed as function of HEP and divided into compliant and non-compliant performed tasks, and, in the case in which the scrap recovery operation is activated, recovered tasks. Figure IV.35 shows the logical flow. Based on the worker reliability in the *Decide Quality Items* the entities are sorted considering the *`Technician`HR* attribute, given in the *No Breaks* section, as a true percentage. The entities recognized as good are transferred to the *Count: Conform Items* while for the non-compliant entities the possibility of performing a rework or not is evaluated. The recovered entities, instead, undergo further processing by the operator and then exit from the template through the *Count: Retrieved items*.

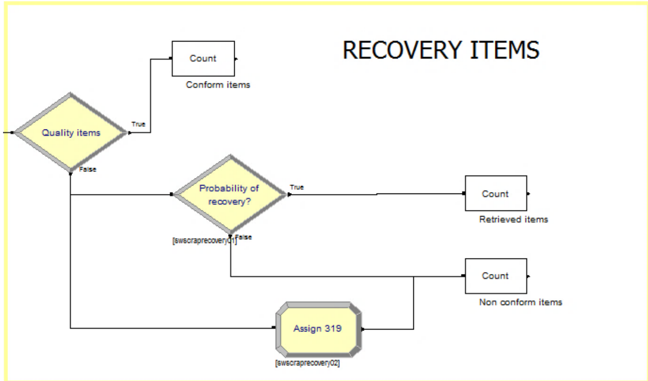

Figure IV.35: *Logic of entities exit.*

IV.4 Anylogic HRA agent

SHERPA has been also implemented in AnyLogic, allows you to build models through a language of nature and extend them with Java code. The agent designed and implemented in Anylogic is shown in Figure IV.36. The basic logic of the model is the same used in the Arena template, with variations due to the different simulation environment. By clicking on this icon, you can access the user interface that allows you to enter the values of the main parameters necessary for operation of SHERPA (Figure IV.37). The user interface is divided into different sections.

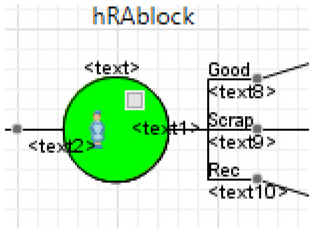

Figure IV.36: *Anylogic SHERPA Agent.*

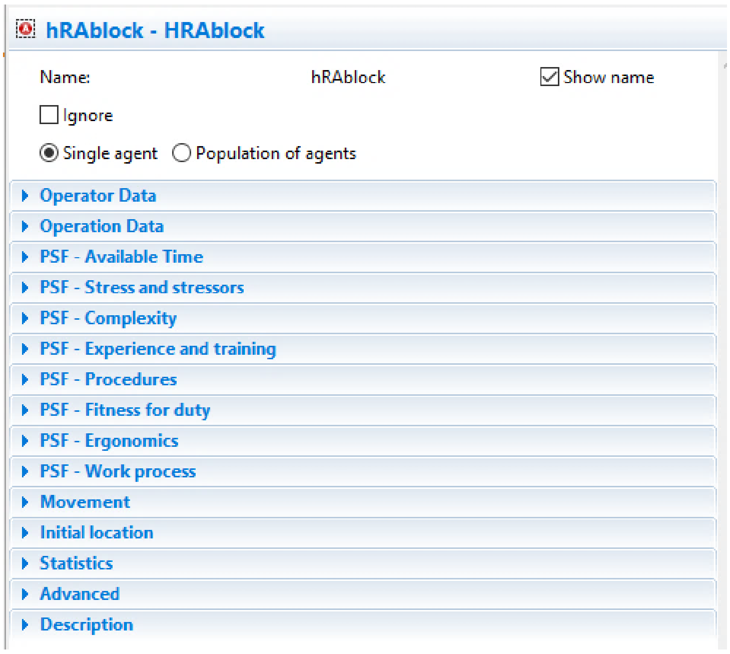

Figure IV.37: *Main SHERPA user interfaces.*

IV.4.1 User interfaces

The first menu allows the identification of data on the operator (Figure IV.38): it is possible specify the name of the operator, gender and age, as well as other data regarding the work context (work shift, end of shift priority, pause).

The Breaks section (Figure IV.38), instead, allows the choice of the psycho-physical recovery speed and the type of break scheduling management. In particular, the choice of automatic management activates the automatic break identification algorithm (it is also necessary to provide the minimum break time); alternatively, it is possible to option to a real break scheduling or to break on a call, when necessary, during the execution of the simulation. In the case of a scheduled break, a *Resource Task* block must be inserted within the main simulation model and selected in the menu. The break is, in fact, seen as a fictitious task that the operator must perform when requested and based on the priority stability. Figure IV.39 shows the data entry screen regarding the activity performed. It is possible to select the type of task, the sector to which it belongs, the priority of the task, the type and the category of activities, as well as the cognitive load required.

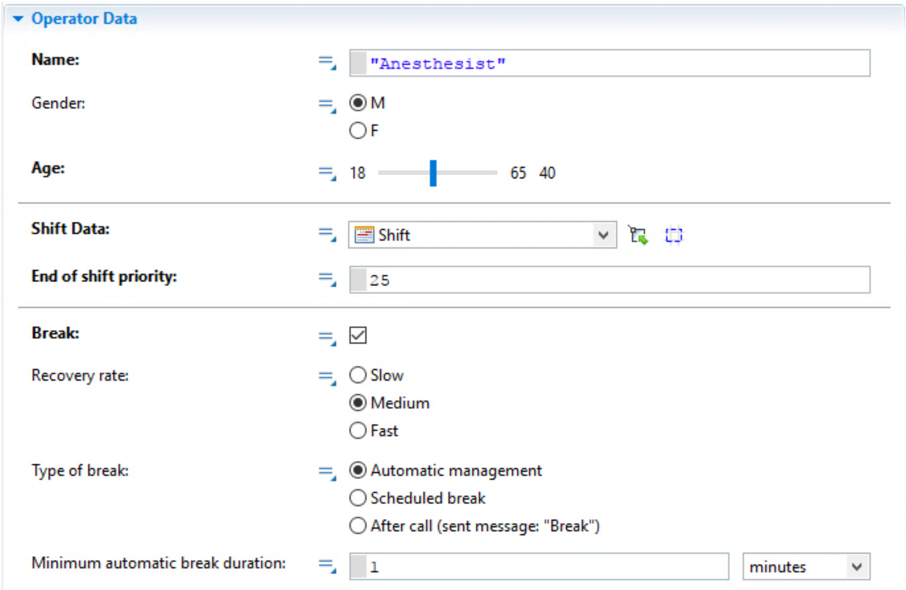

Figure IV.38: *User interface for entering operator data.*

To activate the re-workings, the check on the appropriate entry has to be active: two possible configurations are present in Figure IV.39, depending on whether you choose a fixed rework time (obtained as a percentage of the run time of the main operation) or a variable rework time (in this case the percentage is obtained from a triangular distribution whose parameters are set). The menu also allows the choice of the probability of recovery of the gap and of the unit cost of the recovery itself.

Each PSF, as in the Arena template, has its own menu for the choice of level and the insertion of the information required as input.

Available time. The user interface allows the selection of the PSF level (the menu with the descriptions of the levels depends on the type of activity) and to change the value of the multipliers if the appropriate check box is checked (Figure IV.40).

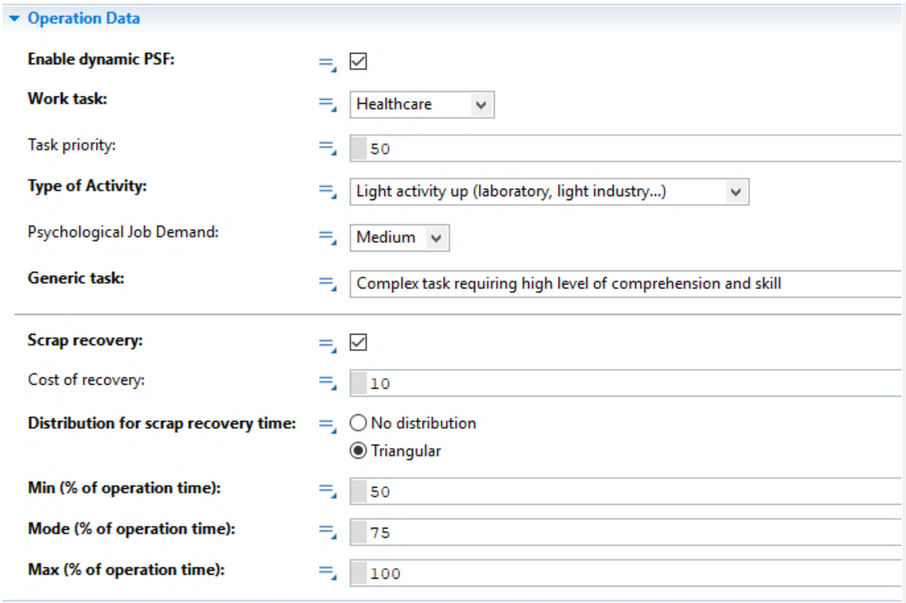

Figure IV.39: *User interface for entering operation data.*

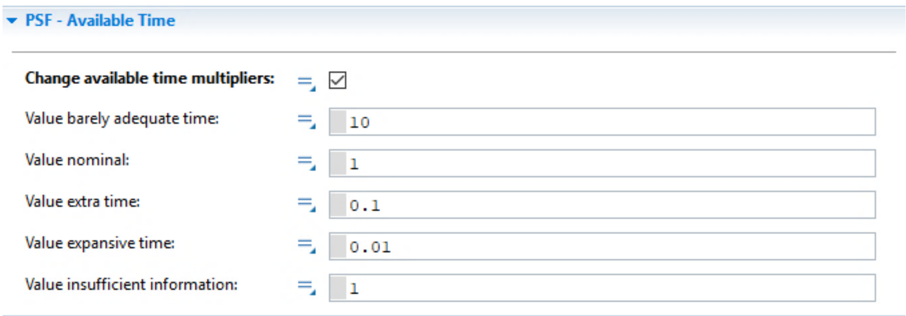

Figure IV.40: *User interface for entering available time data.*

Stress and stressors. The definition of this PSF occurs by specifying the levels (or data, from which the level is derived) of some sub factors:

- Mental Stress: if the activity expects the presence of mental stress, the appropriate section must be activated that allows the selection through a drop-down menu of the level related to this sub factor.
- Pressure Time: the situation is similar to the previous one, by activating the switch it is possible to select the level related to the sub factor.
- Workplace: in this case the level relative to the sub factor must always be selected from the appropriate drop-down menu.

- Microclimate: the choice of level is not immediate, but some data must be provided and then processed within SHERPA. In Figure IV.41 there is a section of the user interface that differs according to whether or not the presence of the temperature control system is selected: if present, the temperature at which it is set is indicated, as well as the relative humidity percentage and the type of clothing used by the operator.
- Lighting: also for this PSF the level is not directly selected, but provided the data necessary to obtain it (Figure IV.41): type of light, visual requirements, type of light source, luminous flux of the illuminated area source, number of light sources, presence of dazzling.
- Noise: the menu connected to this PSF allows first of all the choice on the presence of the factor in the activity performed: only in this case it is possible to enter the noise value expressed in dB from which the level is obtained.
- Vibrations: in this case it is possible to choose if the factor is present or not. In the presence of the factor, the check boxes appear that allow to establish the type of vibrations (vibrations for the hand-arm system only, vibrations for the sole body, vibrations for both): if the corresponding buttons are checked, the boxes appear in which to enter the acceleration values from which the PSF level is then calculated.
- Radiation: the interface and the choices for this PSF are very similar to the previous one. First the presence of the factor is established, therefore the choice of the type of radiation (ionizing, non-ionizing, or both) is allowed. After checking the appropriate boxes, you can enter the values necessary to identify the PSF level. At this point the interface allows to define the weights to be associated with each sub factor (whose corresponding box appears only if the presence of the sub factor has been established, where possible) and, if necessary, to change the values related to each level (by ticking the appropriate check box).

Complexity. This PSF results from the weighted average of several sub factors. The choice of the levels of each sub-factor takes place directly through the user interface, through which it is possible to specify the weights and possibly change the values associated with each level, after the selection of the appropriate box (Figure IV.42).

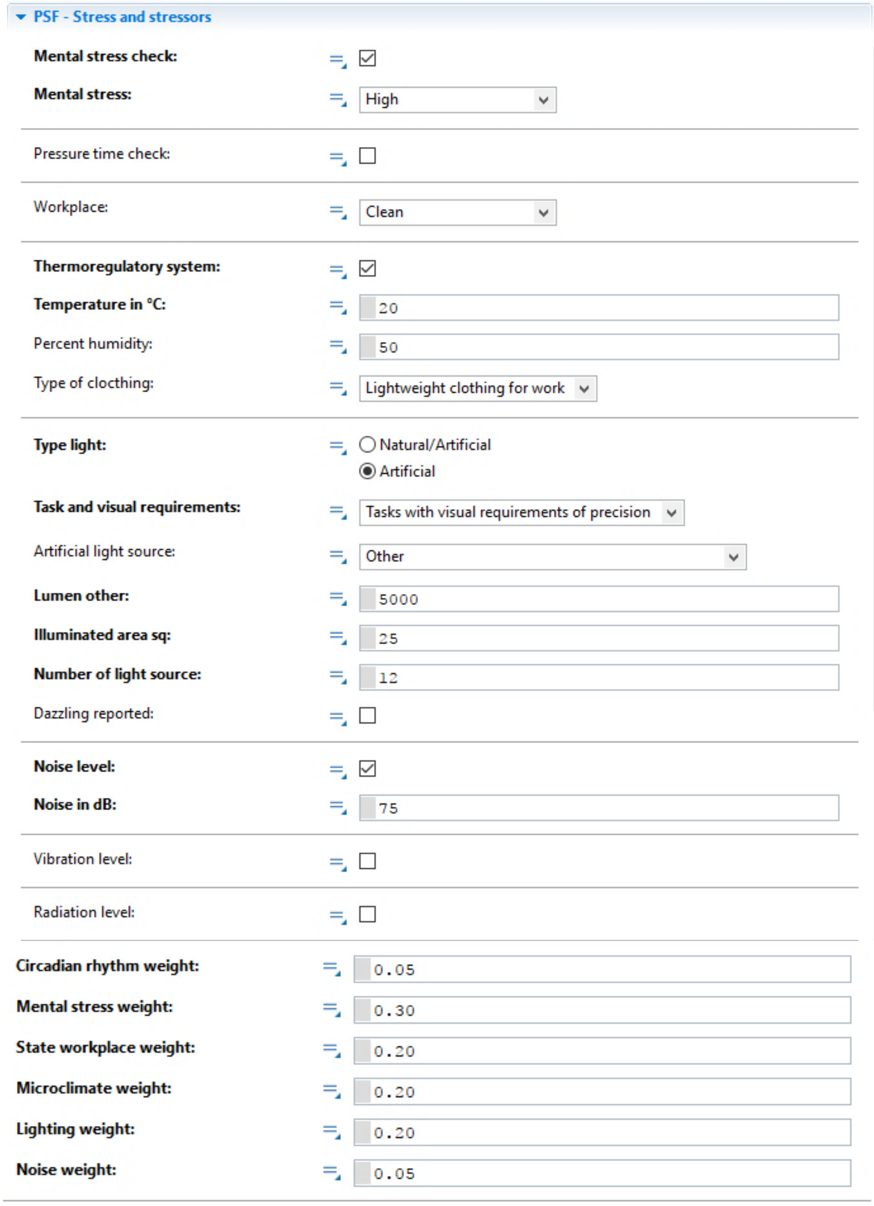

Figure IV.41: *User interface for entering available time data.*

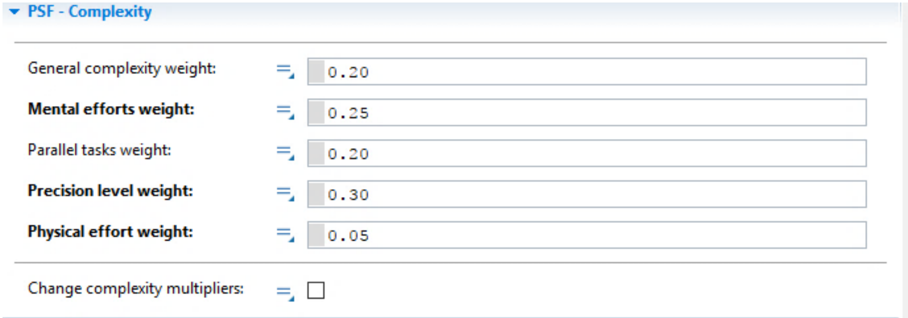

Figure IV.42: *User interface for entering complexity data.*

Experience and training. The user interface allows the assignment directly of the level for the considered PSF and the eventual modification of the values associated to each level through the affixing of the check in the prepared box (Figure IV.43).

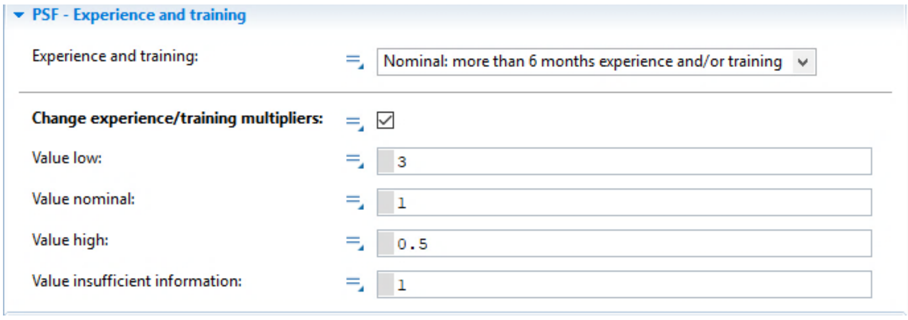

Figure IV.43: *User interface for entering experience and training data.*

Procedures. Through the user interface you can directly choose the PSF level as well as change the values associated with each level after checking the appropriate box. The levels differ from the type of activity.

Fitness for duty. In this case, the user interface allows the direct attribution of the level and the choice of the values associated to each of them.

Ergonomics. The choice of level is direct, just as there is the possibility to change the values associated with the levels after selecting the appropriate box.

Work Process. The user interface returns to be more articulated. For this PSF it is possible to directly select the levels of the sub factors, define the weights to be associated with each sub factor and, after activating the appropriate section after checking the box provided, choose the values linked to each of the levels (Figure IV.44).

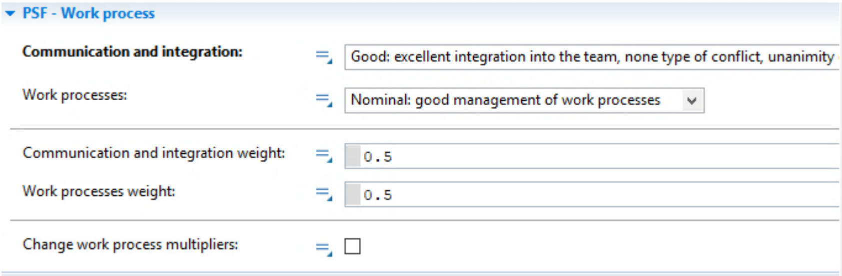

Figure IV.44: *User interface for entering work process data.*

IV.4.2 Logic model

The SHERPA theoretical framework was implemented similar to the previous one seen for the Arena in an agent resource. However, the logic model is visible to the user during the simulation runs. Figure IV.45 shows the implemented SHERPA model. At the beginning of the simulation, the *eventStartSimulation* block is activated, and it allows the system initialization. In particular, based on the choices made through the user interface, the parameters of the Weibull distribution for HEP calculation, the recovery coefficient and, after standardization, the values of the PSF were identified.

Several dedicated functions were designed for applying the theoretical model. In particular, the PSF quantification is immediately available if directly indicated via the user interface, otherwise it is calculated on the basis of the data established through the user interface. The level of the PSF is provided, together with the vector of the standardized values associated with each of the levels of the PSF, as input to the *AssignPSFValue* function, which extracts the desired value. This value is eventually subjected to further reprocessing (weighted average) if the PSF is a combination of several sub-factors.

The second *Hold block* prevent multiple entities in the system, whereas *Seize block* manages the only resource defined within the *ResourceAvailability* block and representative of the operator performing the task. This block is essential for the management of work shifts, breaks (the resource is actually available or there is a break based on what is established through the user interface in the section on operator data) and priorities (it is always carried out the operation with the highest priority).

The temporal attributes indicate the time of the beginning of the shift, the time elapsed since the beginning of the shift (real and correct, where necessary, through the recovery factor r_p) and the time of the last update. These updates take place every time the probability of error is calculated, at the beginning of each shift (*eventShift block*) and at the end of each break (*eventBreak*).

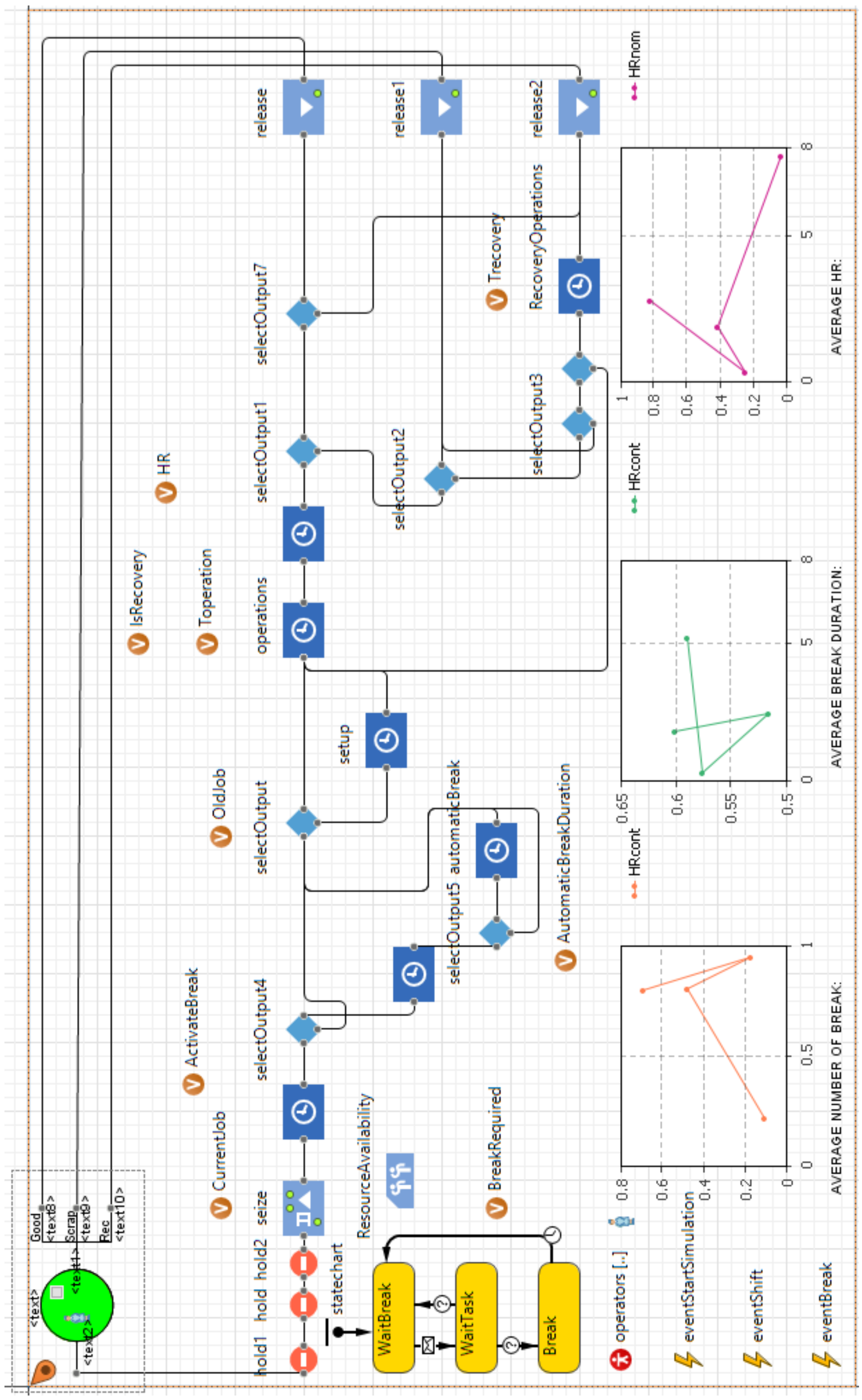

Figure IV.45: *SHERPA Logic.*

The temporal attributes are quantified as follows:

- ✓ Start time of the shift is established (*eventShift* block).
- ✓ Time of the last update is written and coincides precisely with the moment in which the time management routine is launched.
- ✓ Real time, spent since the beginning of the shift, is a simple difference between the current time and the start time of the shift;
- ✓ Correct time elapsed since the beginning of the shift (the one actually used to calculate the probability of error) is, instead, given by its previous value to which must be added the difference between the current time and the time of the last update.

The eventual setup is then managed through a *Select Output* which, reading the parameter of the *Job agent*, allows establishing whether a setup must be carried out. The setup is simulated by stopping the entity within a *Delay block* for the time established by the value given to the attribute of the same agent *Tsetup*.

IV.4.2.1 Human reliability quantification

Human reliability quantification starts with the update of the time variables and the quantification of the nominal HEP. The latter is performed through the *HEPnomCalculation* function, that takes in input the vector containing the parameters of the Weibull distribution and the value of the correct time elapsed since the beginning of the shift. The value of the composite PSF is then calculated and used as input for the *HEPcontCalculation* function, that quantify the contextual HEP. The HEP values are saved within the appropriate attributes of the Operator resource. At the output of the operations block a *Delay* block is used as a computational block in which human reliability is actually calculated as a complement to the contextual HEP and the statistical variables are updated. The reliability value adjusts the operation of the next *Select Output* that directs the entities towards the exit of the good pieces or towards the output of the rejects / reworks.

IV.4.2.2 Break scheduling management

The first step for the break scheduling management is the identification of the recovery coefficient. Optimal break duration and, therefore, recovery factor, based on the parameters indicated in Table III.16, are determined at the beginning of the simulation. The application of the r_p factor takes place at the end of each break both for scheduled and automatic breaks.

The scheduled breaks are managed through the *eventBreak* block. At the beginning of the break, the recovery factor r_p is calculated following the identification of its time as defined through the user interface previously seen. At the same time, the break counting variable is increased and the time variables are updated. At the end of the break, the recovery factor r_p is applied, starting the update routine of the temporal attributes of the Operator resource and modifying the time elapsed since the beginning of the shift through the r_p factor.

The economic algorithm, described in Section III.3.3.2, is implemented in the model section reported in Figure IV.46. Economic variables that regulate the algorithm are read by the entity attributes spanning the SHERPA block, and they allow the quantification of profit with and without break and the choice of the optimal break scheduling.

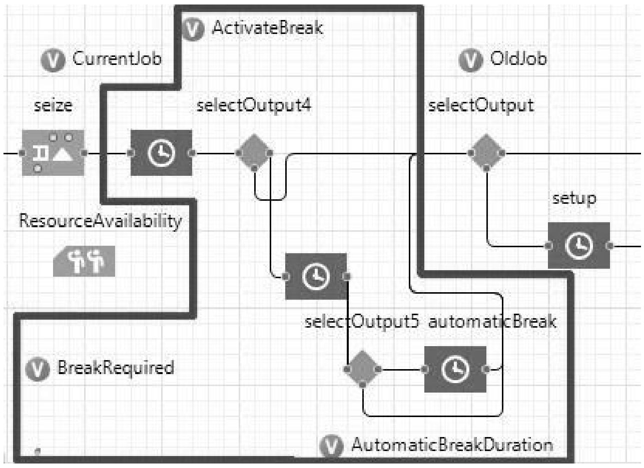

Figure IV.46: *Break scheduling logic.*

IV.5 Methodology for HEP estimation in manufacturing systems

Several studies underline that the use of taxonomies to classify data is a potential solution to produce meaningful information from different types of source using the same framework: i) historical incident data (Moura *et al.*, 2016), ii) incident investigation (Saurin *et al.*, 2008) and iii) prospective analysis (Hollnagel, 1998). Focusing on manufacturing systems, however, this method may not be very suitable and not being able to model the human factor issue.

The estimation of human error probability is a highly complicated task since it involves a huge number of internal and external variables. As previously seen, numerous HRA methods have been proposed over the years, but to date these have several limitations due to validation lacks. Chapter I analysed the consequences of human error in manufacturing systems, proposing a beginning classification for human error data collection. This is the first step of the proposed methodology for HEP estimation in manufacturing systems, starting from experimental data. The methodology, implemented in an Excel tool, allows to use experimental HEP distributions for validation of nominal HEP into SHERPA method, as described below. It is a useful tool for assessing and estimating human error and provides advantageous information about frequency of different human error consequences, length of the transitional phase of human adaption and range of unreliability for every context. In particular, the HEP distributions derived from the realistic and simulated scenarios are compared through two statistical methods in order to verify the goodness of the SHERPA estimation. Figure IV.47 shows the main steps of the proposed methodology.

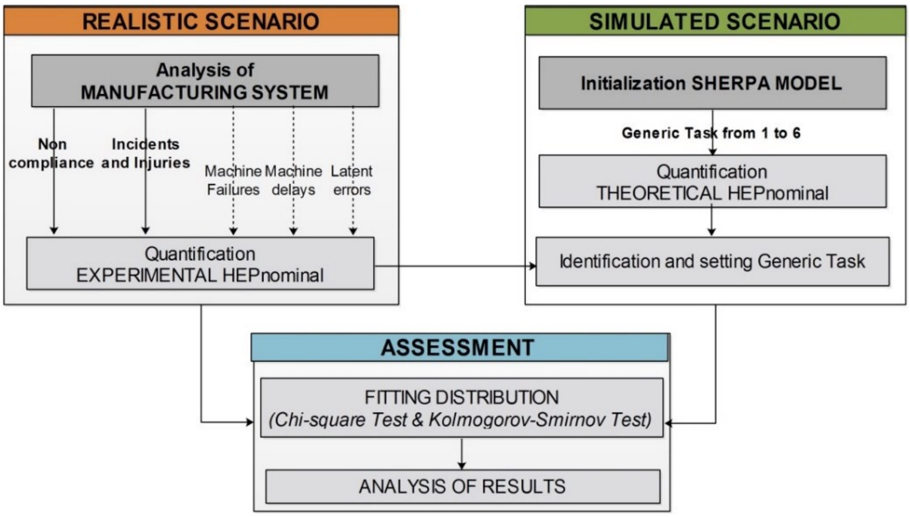

Figure IV.47: *General framework of the proposed methodology.*

IV.5.1 Realistic scenario: data collection and experimental HEP

Data collection is certainly the most complex step. The proposed tool uses historical data on non-compliance (rejected items and reworks offline), accidents and injuries, directly attributable to man, and machine's failures, slowdown and latent errors, when caused by worker, to build experimental HEP curves. Available data are collected in one-hour time slots, covering the entire shift (Table IV.15). They are the number of human errors made during the work.

Considering the formal definition of HEP, the tool calculates the error probability for each hour with respect to all the events, taking into account the average number of errors on the total of activities performed by worker that can be subject to error in every time slot. This allows to build the experimental contextual HEP for the work shift. This distribution is adjusted taking into account the PSFs impact. Table ***IV.16*** reports the obtained experimental HEP points.

IV.5.2 Simulated scenario: theoretical HEP and generic task identification

The theoretical human error distribution on the shift follows a Weibull distribution for six alternative generic tasks, as previously seen. In the proposed methodology, SHERPA simulator is firstly initialized taking into account the impact of contextual and individual factors (Table IV.16), in order to achieve the nominal HEP for the following comparison.

Table IV.15: *Data collection form for experimental HEP*

Non-compliants	**Hours**							
# Shift	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	6	1	4	1	14	13	18	7
2	13	3	4	0	20	12	21	23
….	9	2	4	2	15	24	27	4
250	7	0	6	1	22	7	18	15
Average number of scraps	8	2	4	5	9	11	16	14
Total activities	100	100	100	100	100	100	100	100
HEPcontext	**8,0%**	**2,0%**	**4,0%**	**5,0%**	**9,0%**	**11,0%**	**16,0%**	**14,0%**

Table IV.16: *Experimental HEP distribution.*

	Hours							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Total HEPcontext	18,0%	11,0%	14,0%	17,0%	24,0%	31,0%	35,0%	30,0%
PSFcomposite	5,17	5,17	5,17	5,17	5,17	5,17	5,17	5,17
Experimental HEPnominal	**4,07%**	**2,33%**	**3,05%**	**3,81%**	**5,76%**	**8,00%**	**9,43%**	**7,66%**

The second step, instead, is the identification of generic task that is closer to experimental data, knowing the standard deviations among six theoretical curves and experimental distribution. The method of least squares has been used, identifying the category with the minimum sum of squared residuals, where a residual is the difference between the realistic and simulated HEP value. The selected theoretical curve is then adapted respect to the experimental points, minimizing the sum of squared residuals as function of k, β, α and τ, in order to create an adapted curve that approximates in an optimal way the realistic scenario (Figure IV.48).

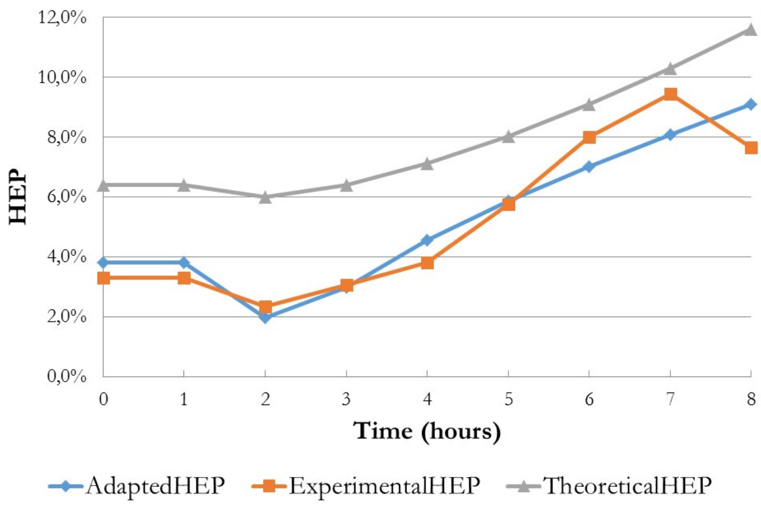

Figure IV.48: *Graphical comparison of the experimental with the theoretical SHERPA curves, standard and adapted.*

IV.5.3 Assessment HEP estimation

In the final stage the experimental and theoretical curves, considering both standard and adapted, are evaluated through two different statistical tests for assessing the goodness of fit between the distributions. The chi-square and Kolmogorov-Smirvon tests establish whether or not an observed frequency distribution differs from a theoretical distribution and in particular whether the hypothesized Weibull distribution is appropriate to describe the phenomenon of human error. Test results provide a lot of information on estimation of error probability into SHERPA model, on statistical significance between theoretical and experimental distribution, on behaviour of human error likelihood in a work shift, on nominal unreliability range for different activities.

Chapter V: Experimental campaigns and case study

V.1 Introduction

The SHERPA model allows to simulate numerous scenarios, considering a plurality of conditions and working activities. The simulator produces in brief time results in terms of compliant and non-compliant performed activities and human error probability that allow to evaluate the impact of context (influencing factors) and break scheduling on system performance.

In this Chapter it has been used to conduct several experimental campaigns to evaluate: the impact of human reliability on system performance; the influence on HR of every PSF level, considered both in singular way and combining with the other factors; the impact of learning and forgetting curves on productivity; and the impact of different work-rest policies on human reliability and system performance. The experimental campaigns simulated a manual assembly processing with the Arena template.

Finally, a case study was conducted in an operating room of Department of Orthopaedics and Traumatology of the University Hospital San Giovanni di Dio-Ruggi d'Aragona of Salerno. The case study aims to study of the human reliability inside an existing operating room, where the consequence of an error is of crucial importance and the spectrum varies from no consequence to serious and fatal.

V.2 Experimental campaigns: simulations of manual assembly processing

The SHERPA template developed in Arena was used to conduct several simulations to evaluate the effect of human reliability as part of manufacturing activity in the prevailing manual content. A manual assembly processing, in which human reliability is critical due to the high contribution of manual tasks, has been simulated as case study in four alternative scenarios:

- Experiment 1: Human Reliability assessment.
- Experiment 2: Simulative analysis of impact of PSFs on human reliability.
- Experiment 3: Simulation of a manual assembly process with LFCM module.
- Experiment 4: Break scheduling management.

The simulated scenarios, even if fictitious, are representative of actual working environments.

V.2.1 Experiment 1: HRA assessment

V.2.1.1 Problem definition

The first experiment was conducted using a simulation model, which reproduces the operator work station involved in manual assembly on an eight-hour shift. The simulation was carried out on an annual basis, considering 235 working days, always with the same work shift. Assembly operation was simulated for three different items with random arrival sequences based on a default mix. For each item, processing times, characterized by a triangular distribution, fixed and variable costs and selling prices, as well as overall production mix, were defined. The data described are shown in Table V.1.

Table V.1: *Features simulated items.*

Features	Item 1	Item 2	Item 3
Processing time (min.)	25	36	45
Setup time (min.)	5	5	5
Price (€)	115	155	200
Fixed cost (€)	52	65	78
Variable cost (€)	18	24	32
SCENARIO 1	20%	5%	75%
SCENARIO 2	15%	65%	20%
SCENARIO 3	50%	30%	20%

The SHERPA template was integrated in a specific Arena model to allow simulation of established scenarios. Figure V.1 shows the Arena model that provides for the entity creation, the assignment of the attributes required for simulation and implementation of the different production mix chosen. Features of items, production mix and SHERPA inputs can easily be modified for subsequent simulations through the following model. Once the required attributes are assigned, entities entered template that allow simulation of the working process and generate the following output: compliant, non-compliant and retrieved items. Of course, the template must be completed with all information concerning the activity and the environmental and behavioural operator conditions, through the dialog boxes discussed in the previous section.

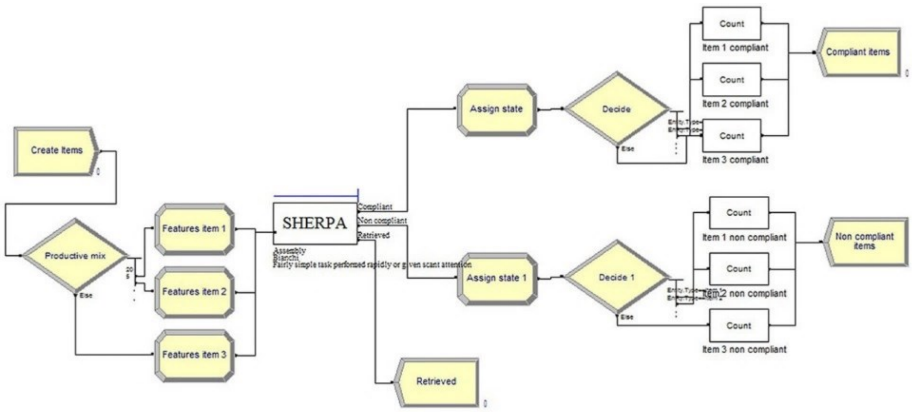

Figure V.1: *Assembly model with SHERPA template.*

V.2.1.2 Results analysis and discussions

The first aspect of SHERPA taken into consideration in this experiment is the impact of several types of modelled generic tasks. Three scenarios were simulated for each of the six categories, keeping the contextual factors at nominal level, with the composite PSF equal to one. Results for every scenario are shown in Table V.2, where the total value of compliant and non-compliant items, their respective percentages and the mean values of HEP nominal and HEP context are reported.

Unlike many HRA methods, SHERPA has been implemented for covering a wide range of working task, for this reason the six modelled categories may represent activities that are more or less reliable. As evident from Figure V.2, the percentage of non-compliant items decreases going from generic task one to task six, due to the increase in the reliability level of each category and the complementary decrease in the human error probability. The generic task one represents the worst activity in terms of reliability; in fact, the average human reliability is approximately equal to 30% in every scenario without taking into account the influence of PSFs. The other categories are higher nominal values of human reliability, reflecting the HEART limitations of unreliability for operation.

Table V.2: *Results of the first step of the simulation.*

	SCENARIO 1					
	GT 1	**GT 2**	**GT 3**	**GT 4**	**GT 5**	**GT 6**
Compliant Items	838	2162	2400	2589	2645	2652
Non-Compliant Items	1814	490	252	63	7	0
Total Items	2652	2652	2652	2652	2652	2652
% Compliant	31.60%	81.52%	90.50%	97.62%	99.74%	100%
% Non-Compliant	68.40%	18.48%	9.50%	2.38%	0.26%	0%
Average HEP Nominal	68.79%	17.90%	8.50%	2.06%	0.39%	0.012%
Average HEP Context	68.79%	17.90%	8.50%	2.06%	0.39%	0.012%

	SCENARIO 2					
	GT 1	**GT 2**	**GT 3**	**GT 4**	**GT 5**	**GT 6**
Compliant Items	917	2380	2643	2845	2896	2911
Non-Compliant Items	1995	532	269	67	16	1
Total Items	2912	2912	2912	2912	2912	2912
% Compliant	31.49%	81.73%	90.76%	97.70%	99.45%	99.96%
% Non-Compliant	68.51%	18.27%	9.24%	2.30%	0.55%	0.04%
Average HEP Nominal	68.73%	17.89%	8.50%	2.06%	0.39%	0.012%
Average HEP Context	68.73%	17.89%	8.50%	2.06%	0.39%	0.012%

	SCENARIO 3					
	GT 1	**GT 2**	**GT 3**	**GT 4**	**GT 5**	**GT 6**
Compliant Items	959	2621	2931	3124	3180	3196
Non-Compliant Items	2238	576	266	73	17	1
Total Items	3197	3197	3197	3197	3197	3197
% Compliant	30%	81.98%	91.68%	97.72%	99.47%	99.97%
% Non-Compliant	70%	18.02%	8.32%	2.28%	0.53%	0.03%
Average HEP Nominal	68.57%	17.84%	8.50%	2.06%	0.39%	0.012%
Average HEP Context	68.57%	17.84%	8.50%	2.06%	0.39%	0.012%

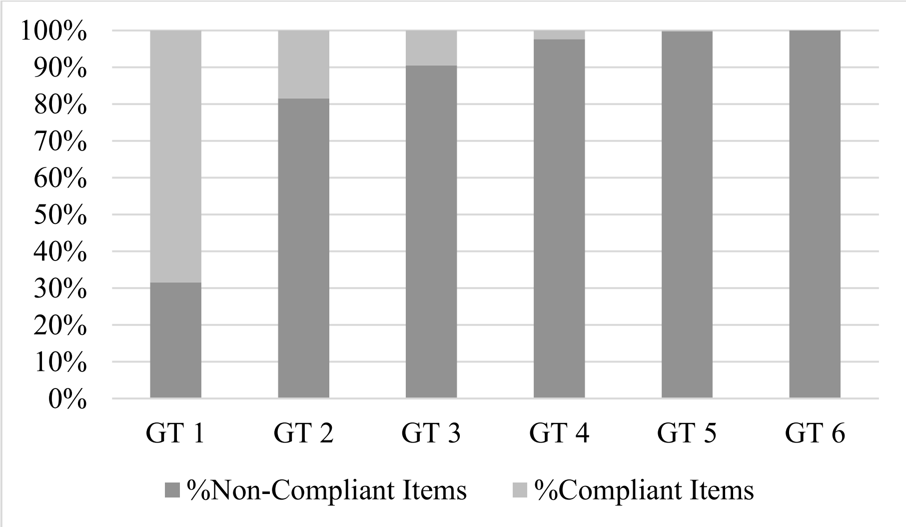

Figure V.2: *Percentage of compliants and non-compliant items for generic categories of activities (Scenario 1).*

The most interesting aspect of the SHERPA model, however, is its ability to simulate several environmental conditions for the same performed activity. The second step of simulation is focused on positive or negative influences of PSFs, keeping constant the type of activity sets equal to generic task three. Table V.3 shows the results of simulations carried out by changing, from time to time, only the complexity and procedures level and keeping all other values at the nominal level.

The performance shaping factors do not always have a negative impact on the reliability, but factors such as experience, ergonomics, time available and work processes may lead to the improvement of the reliability and the consequent decrease in the probability of human error. In the case study, two different conditions were tested where the positive effect of experience was tested and then the positive effect of ergonomics was added. Table V.4 shows a high human reliability improvement due to the decrease in the value of the composite PSF for scenarios two and three.

The simulated scenarios have been used to assess the behaviour of the template when the PSFs levels vary, i.e. with different contextual conditions, for the same performed activity. In the Scenario 1 the simulations highlight the relationship between the composite PSF and the contextual human error probability; in fact, the value of contextual HEP grows with increases in the composite PSF. Starting from the same nominal HEP value, always kept constant, the performance shaping factors increase variably the HEP contextual according to their multiplier (Figure V.3).

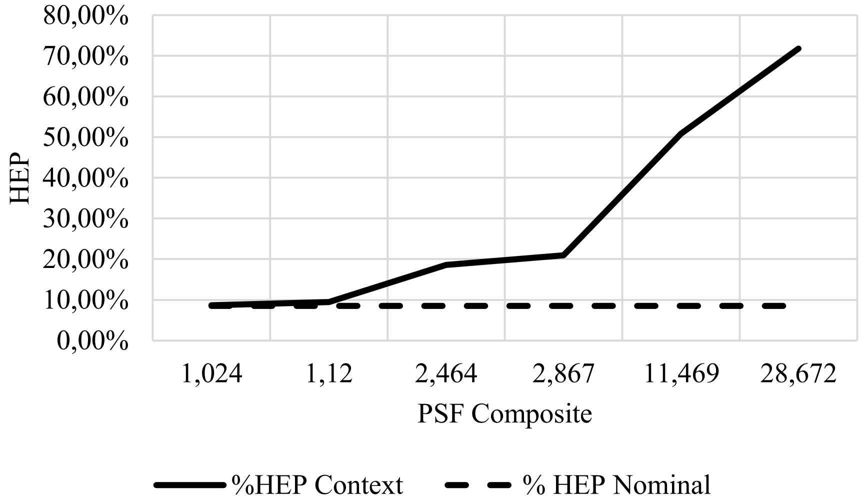

Figure V.3: *HEP as function of increasing composite PSF.*

Table V.3: *Results of simulation for scenario one, while changing complexity and procedures levels.*

SCENARIO 1	**COMPLEXITY**			**PROCEDURES**		
	Nominal	**Moderate**	**Extreme**	**Available but Poor**	**Incomplete**	**Not available**
PSF Composite	1.024	1.12	2.464	2.867	11.4688	28.672
Compliant Items	2396	2372	2150	2106	1311	761
Non-Compliant Items	256	280	500	546	1341	1891
Total Items	2652	2652	2652	2652	2652	2652
% Compliant	90.35%	89.44%	81.15%	79.41%	49.43%	28.70%
% Non-Compliant	9.65%	10.56%	18.85%	20.59%	50.57%	71.30%
Average HEP Nominal	8.52%	8.52%	8.52%	8.52%	8.52%	8.52%
Average HEP Context	8.70%	9.44%	18.55%	20.92%	50.85%	71.80%

Table V.4: *Results of simulation for scenarios two and three, changing experience and ergonomics levels.*

	SCENARIO 2			SCENARIO 3		
	Nominal	Experience	Experience/ Ergonomics	Nominal	Experien ce	Experience/ Ergonomics
PSF	-	0.28	0.28-0.28	-	0.28	0.28-0.28
PSF Composite	1	0.28	0.078	1	0.28	0,078
Compliant Items	2643	2828	2886	2931	3107	3171
Non-Compliant Items	269	84	26	266	90	26
Total Items	2912	2912	2912	3197	3197	3197
% Compliant	90.76%	97.12%	99.11%	91.68%	97.18%	99.19%
% Non-Compliant	9.24%	2.88%	0.89%	8.32%	2.82%	0.81%
Average HEP Nominal	8.52%	8.52%	8.52%	8.52%	8.52%	8.52%
Average HEP Context	8.52%	2.61%	0.75%	8.52%	2.61%	0.74%

In some cases, the variation is limited; for example, when considering a moderate level of complexity there is the increase of approximately 10% compared to the nominal level. The increase grows up to 95%, with an 18.85% of non-compliant items, when the complexity level is extreme. In other cases, the particular environmental or personal conditions can lead to high increases in the probability of error, as in the case of not available procedures where the variations in HEP are larger due to the multiplier theoretically assigned from the SPAR-H method.

A further assessment done in the experiment is relative to the positive performance shaping factors. Positive factors lead to a decrease in the final value of the composite PSF and an improvement in the operator reliability compared to the nominal HEP. As evident in Figure V.4, a multiplier of high experience, amounting to 0.28, improves the human error probability, lowering it to the 2.61%. Where two or more positive PSFs are merged, the improvement is even more evident; for example, in the simulation, high experience level and good ergonomics level allow the values of human error to approach nearly zero, equal to 0.75% in both scenarios.

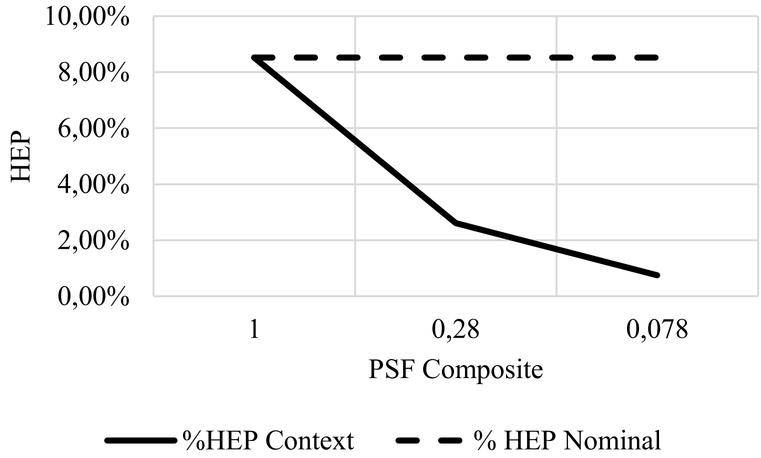

Figure V.4: *HEP as function of decreasing composite PSF.*

In most real cases, positive and negative factors coexist and affect the activity carried out by the operator. This condition was simulated considering factors with positive impact such as high experience, and other negative factors, such as moderate stress, poor procedures and poor working processes. In this last simulation, the same conditions are used for every scenario and the results are shown in Table V.5. Finally, the same kind of task and the same contextual conditions have been used in the three scenarios. The results, in terms of reliability, are very similar to each other, because the difference in production mix translates especially in terms of total units produced, given the different processing times. The error probability, in these cases, is determined more by the type of activity than by performance shaping factors.

Table V.5: *Results of last step of simulation.*

	SCENARIO 1	SCENARIO 2	SCENARIO 3
Composite PSF	1.7208	1.7208	1.7208
Compliant Items	2269	2489	2754
Non-Compliant Items	383	423	443
Total Items	2652	2912	3197
% Compliant	85.56%	85.47%	86.14%
% Non-Compliant	14.44%	14.53%	13.85%
Average HEP Nominal	8.52%	8.52%	8.50%
Average HEP Context	13.77%	13.77%	13.73%

This first experiment underlines the major SHERPA features, described in the theoretical model. In particular, its versatility is useful in revealing the environmental and psycho-physical factors which mainly influence the human reliability and may therefore be subject to improvement in order to reduce errors.

V.2.2 Experiment 2: Simulative analysis of impact of PSFs on human reliability

The influencing factors play a key role in the modelling of human error and many theoretical studies have been carried out to define, to classify and to model these factors as above.

The aim of this experiment is the study of parameters that affect the human performance in workplace, considering how they increase or decrease the human error probability in SHERPA. Influence of every PSF level considered both in singular way and combining with the other factors, was quantified through numerous simulations with the SHERPA template. The following briefly describes the basic steps used in the simulation process:

1) Problem definition: description of the case study.
2) Experiment planning and system definition: identification of the system components to be modelled and the performance measures to be analysed.
3) Results analysis: list of results and discussion of study implications.

V.2.2.1 Problem definition

A manufacturing activity was simulated in an Arena model for the research purpose of this experimental campaign. The construction of the simulation model takes hint from the description of different assembly stations proposed in literature (Falcone *et al.*, 2010, 2011). A 30-minute break after four hours for the shift start is scheduled. The simulations were performed considering the assembly activities belonging to the action category: *Routine, highly-practiced, rapid task involving relatively low level of skill* (Table III.1).

The assembly operation was simulated for three different items with random arrival sequences based on a production mix and with processing times characterized by a triangular distribution, with vertices corresponding to the mean ±10%. For each item, processing times, fixed and variable costs and selling prices, as well as overall production mix, were defined and are shown in Table V.6.

Table V.6: *Features of simulated items in the case study.*

Features	Item 1	Item 2	Item 3
Mean processing time (min.)	25	36	45
Setup time (min.)	5	5	5
Price (€)	115	155	200
Fixed cost (€)	52	65	78
Variable cost (€)	18	24	32
Productive mix	25%	35%	40%

V.2.2.2 Experiment planning and system definition

The SHERPA template allows modelling the context and the psycho-physical condition of the operator through twenty-one PSFs, both main and secondary.

Every PSF impacts in a different way on nominal HEP. SPAR-H method, in fact, uses a nonlinear levels classification and the levels classes are different for every PSF. In the case of available time, for example, there are four levels in addition to the nominal case, while stress or work processes have only two levels. Consider all these factors with a full factorial analysis, would be to make $2^{21} = 2097152$ simulations, taking into account just two levels per factor. For this reason, in the experiment planning a selection of the potentially most significant PSFs was necessary.

Firstly, all factors were classified (Table V.7) compared to the experiment context in:

- ✓ Controllable: you can manage and define values in advance, as the input of the experiment itself.
- ✓ Uncontrollable: are out of hand when they appear; may change during operation of the product or process.
- ✓ Measurable: able to be measured; not subjective, perceptible or significant.
- ✓ Unmeasurable: not able to be measured objectively.

The proposed classification was useful in the subsequent selection of the most relevant factors for the goal of simulative analysis. The choices of the factors took into account this classification, considering at least a factor by category. In the second step a common method of investigating the effects of parameters on a process was applied. The one-factor-at-a-time (OFAT) method allows changing only one factor at a time, to assess the impact of factors considered one at a time instead of all simultaneously and to notice its influence on a given response. Although this method has the advantage of being simple, it requires many trials and does not point out the possible interactions between several factors.

Table V.7: *Factors classification in terms of measurability and controllability.*

	Measurable	**Unmeasurable**
Controllable	Available time Parallel tasks Microclimate Lighting	Workplace Procedures Precision level Physical effort Mental effort Cognitive ergonomics
Uncontrollable	Circadian rhythm Experience Noise Vibrations Radiations	General complexity Work processes Mental stress Fitness for duty Pressure time Communication

The PSF levels were modified one at a time keeping the others at nominal level. The contextual HEP value and the PSF composite were calculated for every simulation. Some factors were set to scenario: microclimate, lighting, circadian rhythm and physical effort, the latter is related to the performed task. Downstream the simulations, the results were analysed and for each factor the ratio between the percentage variation of contextual HEP and PSF composite was considered (Table V.8). It can be clearly seen that the increase of the PSF level determines an increase of the PSF composite and a consequent increase in the probability of error.

The negative changes, such as extra and expansive available time or good cognitive ergonomics, represent the positive effect of the factors on the performance, as seen above. A special factor is the experience. This factor, in fact, determines a very high increase of HEP (ΔHEPc=97.85%) with a modest increase in PSF composite (ΔPSFc=20.63%). Factors with similar changes in their PSF composite, for example the general complexity (ΔPSFc=15.25%) or mental stress (ΔPSFc=26.47%), respectively determine increments of HEP equal to 1.89% and 3.57%. This exception will be thorough better later.

Previous evaluations were used to select the factors for the next step considering the factors with more impact on HEP and mainly representative of a manufacturing context. With the aim of reducing the number of runs the parameters available time, state of workplace, vibrations, radiations, pressure time, precision level, mental efforts and communication and integration in team work were excluded. All these factors were set to the nominal level in the experimental stage, and they have not had their influence on HEP.

Table V.8: *PSF effect on contextual HEP.*

SPAR-H PSFs	PSF Levels	ΔHEPc%	ΔPSFc%	ΔHEPc%/ ΔPSFc%
Available Time	Inadequate	98.73	not available	not available
	Barely adequate	84.09	85.29	0.9859
	Nominal	0	0	-
	Extra	-1355.27	-1370.59	0.9888
	Expansive	-14409.55	-14605.88	0.9865
Mental stress, Pressure time and Noise	Extreme	18.96	70.59	0.2687
	High	3.57	26.47	0.1348
	Nominal	0	0	-
Radiations and Vibrations	Extreme	10.39	54.54	0.1904
	High	1.89	15.25	0.1236
	Nominal	0	0	-
Workplace	Extreme	10.39	70.59	0.1472
	High	1.89	26.47	0.0713
	Nominal	0	0	-
General complexity, Precision level, Mental efforts, Parallel tasks	Highly complex	30.67	70.59	0.4346
	Moderately complex	6.24	26.47	0.2356
	Nominal	0	0	-
Experience/ Training	Low	97.85	20.63	4.7421
	Nominal	0	0	-
	High	-191.72	-1.94	0.9877
Procedures	Not available	95.72	97.06	0.9862
	Incomplete	91.37	92.64	0.9863
	Available, but poor	69.67	70.59	98.69
	Nominal	0	0	-

SPAR-H PSFs	PSF Levels	ΔHEPc%	ΔPSFc%	ΔHEPc%/ ΔPSFc%
Cognitive Ergonomics	Missing	95.72	97.06	0.9862
	Poor	84.09	85.29	0.9859
	Nominal	0	0	-
	Good	-191.72	-194.12	0.9877
Fitness for Duty	Unfit	98.73	not available	not available
	Degraded Fitness	69.66	70.58	0.9869
	Nominal	0	0	-
Work Processes and Communication	Poor	53.95	77.27	0.6981
	Nominal	0	0	-
	Good	-48.71	25.37	-1.92

Then the most significant factors were chosen based on this assessment and taking into account the classification of the factors in terms of measurability and controllability: noise; mental stress; general complexity; parallel tasks; experience; procedures; work processes; fitness for duty; and cognitive ergonomics.

For the chosen factors were considered only two levels from those available for the analysis:

- Noise: Extreme and Nominal levels;
- Mental stress: Extreme and Nominal levels;
- General complexity: High and Nominal levels;
- Parallel tasks: High and Nominal levels;
- Experience: Low and Nominal levels;
- Procedures: Incomplete and Nominal levels;
- Work processes: Poor and Good levels;
- Fitness for Duty: Degraded and Nominal levels;
- Ergonomics: Poor and Good levels;

In the system definition, nine factors were selected with two levels for each one. In this condition we can define the number of simulations to be performed to analyse the scenarios provided by all possible combinations of PSFs and to evaluate their effect on the likelihood of operator error; they are $2^9 = 512$

simulations. The experiment was conducted simultaneously changing the levels of selected factors until you cover the entire experimental plan.

V.2.2.3 Results analysis and discussions

Analysis of variance (ANOVA) was used to examine the effect of significant PSFs on the HEP (Scheffe, 1999; Gelman, 2005). This method, developed by Fisher, is at the basis of many designs of experiments and is used to compare differences of means among more than two groups. It does this by looking at variation in the data and where that variation is found. Specifically, ANOVA compares the amount of variation between groups with the amount of variation within groups. It can be used for both observational and experimental studies.

In performing ANOVA, the experimental factors and the dependent variable or response are identified. The experimental factors are the source of variability whose effect is to be determined based on the results of a dependent variable or response. In the case of study, experimental factors are therefore the PSFs, while the dependent variable is the contextual HEP. The simplest experiment suitable for ANOVA analysis is the experiment with a single factor, used in a first time to assess the impact of each factor on HEP. Table V.9 lists the one-way ANOVA results. The SS stands for Sum of Squares; F-ratio is test statistic used for ANOVA, the p-value is the probability of being greater than the F-ratio. The F is a ratio of the variability between groups compared to the variability within the groups. F-ratio will always be at least 0, meaning that it is always non-negative. The p-values in the last column are the most important information contained in this table. Statistical significance of the effect depends on the p-value, as follows:

- If the p-value is larger than the significance level you selected, the effect is not statistically significant.
- If the p-value is less than or equal to the significance level you selected, then the effect for the term is statistically significant.

Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that an effect exists when there is no actual effect.

Figure V.5 shows the results for all the chosen factors and it underlines graphically the different impacts on error likelihood. Each graph represents the average value of HEP when the factor is set to level one or two. The vertical bars indicate the level of confidence at 95%, that is the probability that the calculated values fall in this range. It is to be noted that when the bars are large the possible values are very different from each other and fluctuate around a mean value.

Table V.9: *One-way ANOVA results.*

FACTORS	SS	F-ratio	p-value
Mental stress	0,046	0,362	0,548
Noise	0,046	0,362	0,548
General complexity	0,128	1,013	0,315
Parallel tasks	0,128	1,013	0,315
Experience	25,196	327,178	0,00
Procedures	11,565	111,289	0,00
Work processes	2,073	16,942	0,000045
Fitness for duty	1,746	14,199	0,000184
Cognitive ergonomics	18,061	198,465	0,00

The most influential factors (experience, procedures and cognitive ergonomics) have a very tight confidence interval, a sign of their strong impact in the calculation of the error probability.

Through this first analysis the greatest difference of average HEP is easily observed for those factors that have the two multipliers more distant from each other, like procedures and ergonomics. The experience is an exception, because it has a strong impact on the probability of error despite its multipliers are comparable to those of stress and complexity. Such behaviour is also clear in the two-way ANOVA.

The two-way analysis of variance is an extension of the previous one-way, which examines the influence of two different factors and it aims at assessing if there is any interaction between factors and how the contemporary presence of two factors affects the variable result. Through this second step of analysis, the interrelationships between multiple PSFs were examined. In this case the p-value is used as an indicator to determine if the two factors have a significant interaction when considered simultaneously. If one factor depends strongly on the other, the F-ratio for the interaction term will have a low p-value. The two-way ANOVA table is structured just like the one-way.

Table V.10, in fact, shows the SS, the F-ratio and the p-value for all factors combinations.

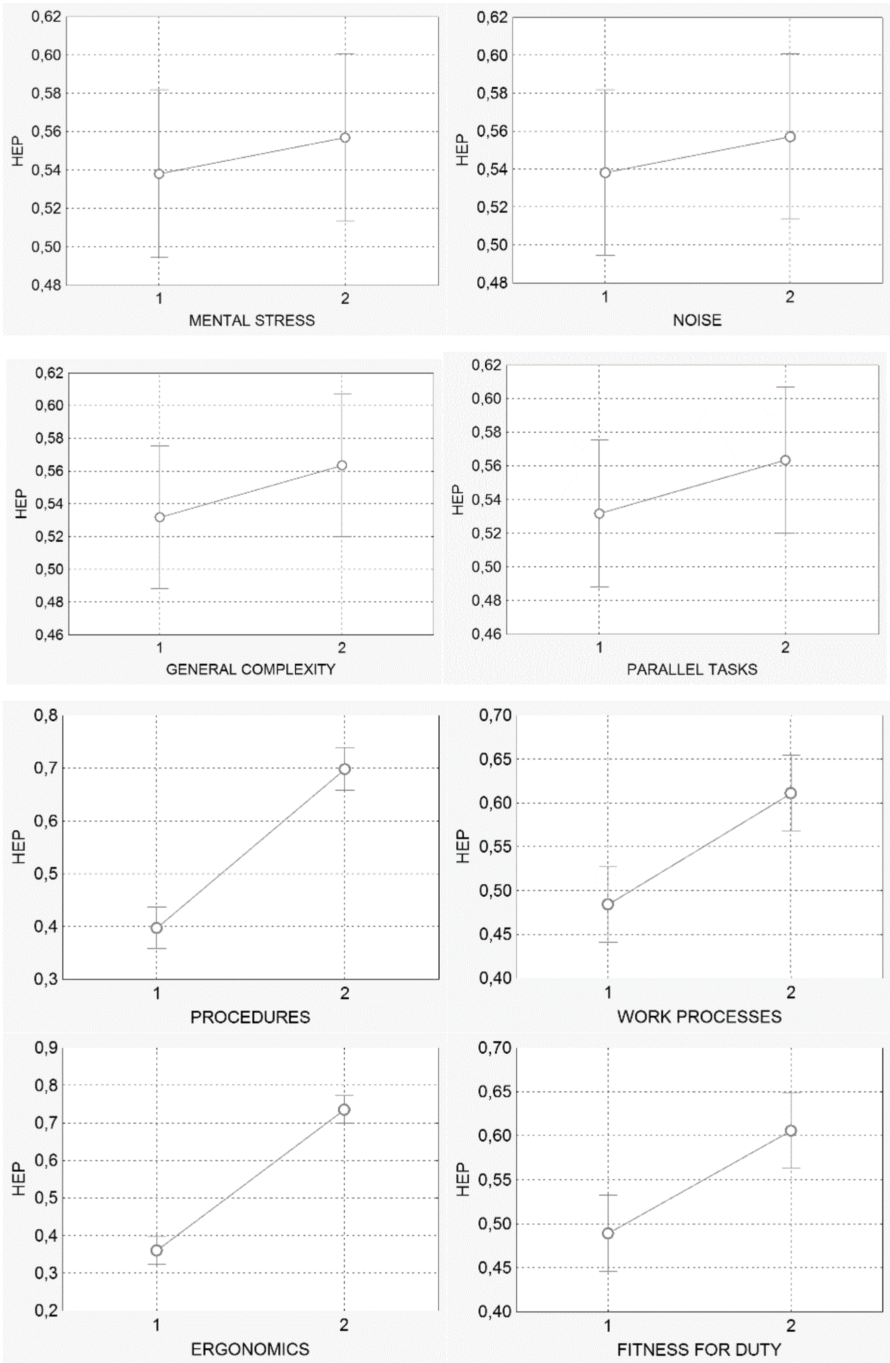

Figure V.5: *Single factor Analysis of Variance.*

Table V.10: *Two-way ANOVA results.*

FACTORS			SS	F ratio	p-value
Mental stress	x	Noise	0,0001	0,001	0,976
Mental stress	x	General complexity	0,0006	0,004	0,946
Mental stress	x	Parallel tasks	0,0006	0,004	0,946
Mental stress	x	Experience	0,0107	0,138	0,710
Mental stress	x	Procedures	0,0007	0,007	0,934
Mental stress	x	Work processes	0,0000	0,000	1,000
Mental stress	x	Fitness for duty	0,0003	0,002	0,963
Mental stress	x	Ergonomics	0,0021	0,022	0,881
Noise	x	General complexity	0,0006	0,004	0,946
Noise	x	Parallel tasks	0,0006	0,004	0,946
Noise	x	Experience	0,0107	0,138	0,710
Noise	x	Procedures	0,0007	0,007	0,934
Noise	x	Work processes	0,0000	0,000	1,000
Noise	x	Fitness for duty	0,0003	0,002	0,963
Noise	x	Ergonomics	0,0021	0,022	0,881
General complexity	x	Parallel tasks	0,0008	0,006	0,937
General complexity	x	Experience	0,0162	0,211	0,646
General complexity	x	Procedures	0,0013	0,012	0,912
General complexity	x	Work processes	0,0001	0,001	0,976
General complexity	x	Fitness for duty	0,0002	0,002	0,965
General complexity	x	Ergonomics	0,0028	0,031	0,860
Parallel tasks	x	Experience	0,0162	0,211	0,646
Parallel tasks	x	Procedures	0,0013	0,012	0,912
Parallel tasks	x	Work processes	0,0001	0,001	0,976
Parallel tasks	x	Fitness for duty	0,0002	0,002	0,965
Parallel tasks	x	Ergonomics	0,0028	0,031	0,860
Experience	x	Procedures	0,2884	5,348	0,021
Experience	x	Work processes	0,0534	0,731	0,393
Experience	x	Fitness for duty	0,1571	2,135	0,145
Experience	x	Ergonomics	0,0926	2,237	0,135
Procedures	x	Work processes	0,0000	0,000	0,992
Procedures	x	Fitness for duty	0,0012	0,012	0,913
Procedures	x	Ergonomics	0,0075	0,109	0,742
Work processes	x	Fitness for duty	0,0001	0,001	0,973
Work processes	x	Ergonomics	0,0021	0,024	0,876
Fitness for duty	x	Ergonomics	0,0002	0,003	0,960

There is a statistically significant interaction between the effects of experience and procedures on HEP (p-value= 0,021<0.05), so the effect on the mean outcome of a change in one factor depends on the level of the other factor (Figure V.6). The significant relationship between these factors depends also on the high impact on HEP of single factors. For all the other combinations there is not statistical dependence. Their p-values, in fact, are included between 0,135 (experience x cognitive ergonomics) and 1,000 (mental stress x work processes). For example, in Figure V.7 the interaction between mental stress and noise (p-value=0,976) shows clearly the statistical independence: the effects of a change in one factor on the outcome do not depend on the value or level of the other factors.

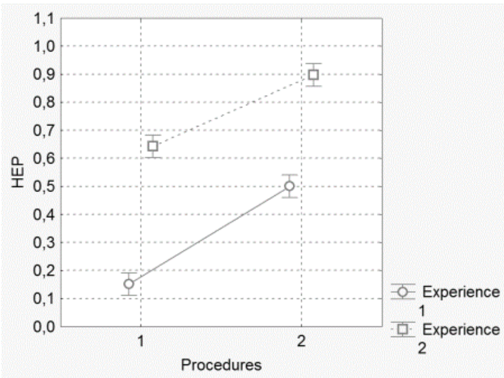

Figure V.6: *Procedures x Experience ANOVA results.*

The experience is the most interesting factors. As already highlighted by the OFAT and one-way ANOVA analysis, the experience has one of the major impact on the error probability. This effect is further confirmed by the interactions between factors (Figure V.8). The level two of experience determines a considerable decrease of human reliability and consequent increase in error probability when it is combined with every factor (i.e. mental stress, general complexity, cognitive ergonomics and procedures). The strong impact does not depend on exclusively from the multiplier, but it derives also by the logic experience evaluation used by the model. Lack of knowledge of the processes, of the machines and of the procedures modifies the nominal HEP, because it impacts on the category of performed task, which can no longer be regarded as routine and highly-practiced.

The performance shaping factors are an integral part of modelling and characterization of errors, and they affect the productivity and the efficiency at work. Their modelling is a problem for each HRA method. Many HRA approaches introduce widespread PSF taxonomies and complex modelling of their mutual influence. Despite the efforts of HRA experts, the PSFs have not explicit role both in error identification and in probability estimation yet. The goal of this experiment was to analyse the PSFs, used in the SHERPA model, and to assess their impact on HEP in order to improve the model and to make it more responsive to working reality.

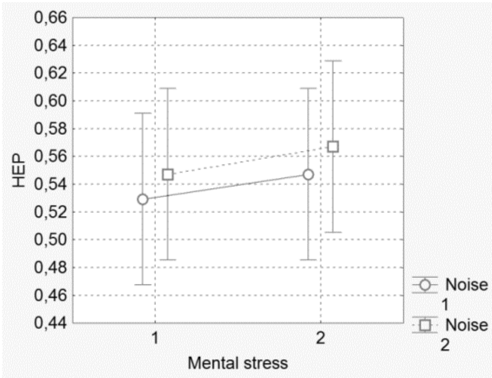

Figure V.7: *Mental stress x Noise ANOVA results.*

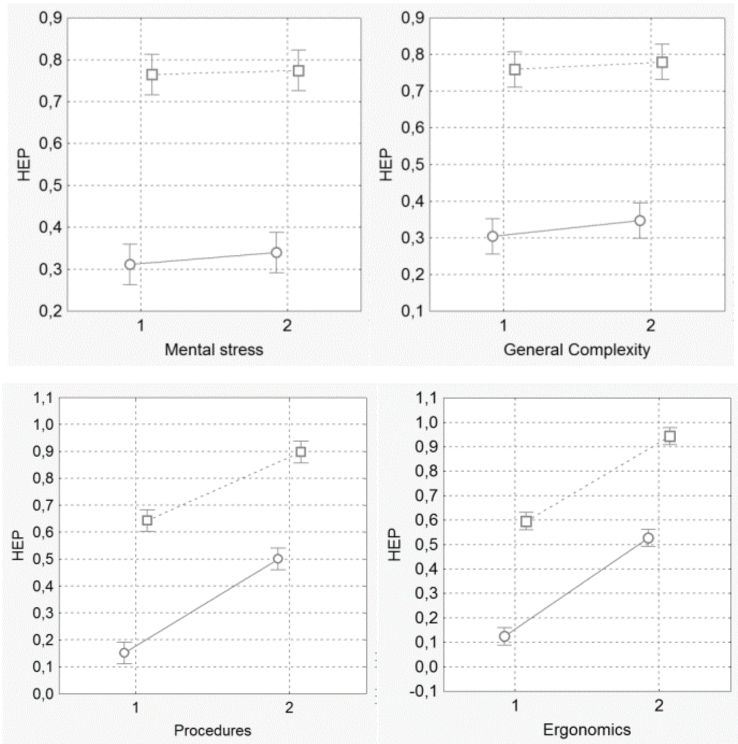

Figure V.8. *Two-way ANOVA results.*

Thanks to the simulative analysis and to the results obtained from one and two-way ANOVA, the influence of every PSF level, considered both in singular way and combining with the other factors, was quantified and evaluated, allowing to realize if there is more or less dependence between them. Several useful considerations can be made downstream of the study. First of all, through the preparatory OFAT analysis the different PSF impacts on HEP in relation to the value of its multiplier is evident. It is certainly useful as a starting point for the improvement of PSF modelling, that is currently under investigation. The one-way ANOVA underlined the higher or lower impact on HEP of individual factors, whereas the results of two-way ANOVA highlight few interactions between factors. There is significance of impact only when the experience is combined with the procedures. As regards the experience, its special behaviour requires further investigations and studies.

V.2.3 Experiment 3: Simulation of a manual assembly process with LFCM module

The LFCM module, validated and tested, was used to simulate a manual assembly process without fixed production rate. The case study reproduces the operator work station involved in manual assembly, considering 235 working days with 5 days of training. PSFs for this scenario were chosen to represent approximately the actual conditions in the assembly plant. Available time, procedures, fitness for duty and work processes were imposed at the nominal level; complexity, stress and ergonomics at the moderate/high level and experience at high level.

The work-break schedule of a Toyota assembly line was used, and it involves two breaks of 10 minutes and one break of 45 minutes as presented in Table V.11 (Givi, et al., 2015). The other parameters of the problem are set according to Table V.11.

Table V.11: *Features of produced item in the case study.*

Cycle time	Break time	Parameters	
125 min	10 min	Price (€)	115
120 min	45 min	Fixed cost (€)	52
120 min	10 min	Variable cost (€)	18
90 min	-	T1 (min)	30
Total		Total forgetting time B (days)	365
455 min	65 min	Learning rate %	85

Case study was simulated for two scenarios with and without the learning and forgetting effects. In the second scenario, the average processing time was fixed to 5 minutes. The learning curve for the first scenario over the entire horizon time (240 days) is presented in Figure V.9. In the graph, the first unit takes 30 minutes to be produced and the processing time per unit decreases as operator gain experience by producing additional units.

During the training phase the production time decreases very fast because the operator is learning a new task. The forgetting phenomenon, instead, is present during the shift and between two consecutive days and it determines an increase in time to produce the first unit in the next production run. The simulation results for both the scenarios are reported in Table V.12.

Table V.12: *Simulation results.*

	Case study with LFCM	**Case study without LFCM**
Number of compliant items	22,433	20,678
Number of non-compliant items	753	716
Average HEP	3.3%	3.3%

The impact of the learning and forgetting processes on the system performance is evident from the comparison between the two scenarios. The number of produced items grows from 20,678 to 22,433; when the learning and forgetting curves are considered in the simulation. It is due to the reduction of the average production time that varies from 5 minutes of the scenario without LFCM to 4.62 minutes of the other scenario.

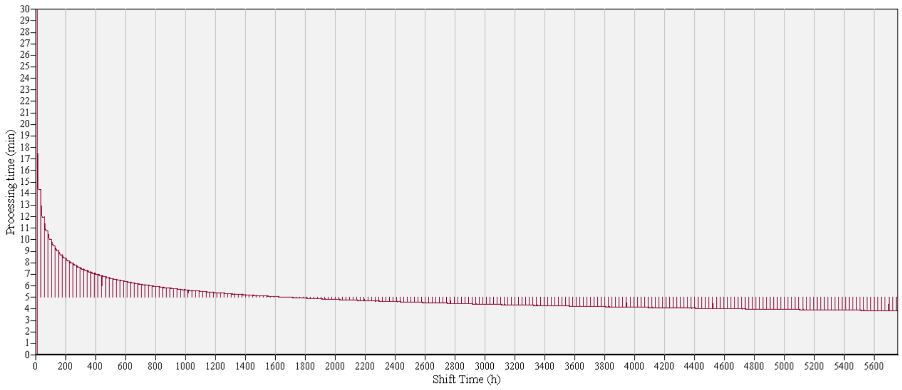

Figure V.9: *The learning and forgetting effect on the processing time.*

The LFCM module allows simulating a large number of scenarios without being resource intensive or time consuming in order to evaluate the human performance under the learning and forgetting phenomenon.

V.2.4 Experiment 4: Break scheduling management

Two different numerical examples were presented for the break scheduling management in this section.

V.2.4.1 Scenario 1: Problem definition

A manual assembly process was simulated as a case study, involving a single 8-h shift for 230 days per year. The simulated assembly task is mixed-model with two different items (P1 and P2) with similar assembly processes and with random arrival sequences based on the fixed production mix (65% P1 and 35% P2). The assembly operation was performed with processing times reported in Table V.13 and characterized by a triangular distribution, with vertices corresponding to the mean ±10%. The economic parameters are set according to what is shown in Table V.13.

Table V.13: *Features of simulated items.*

Features	P1	P2
Productive mix	65%	35%
Mean Processing time (min)	5	7.5
Setup time (min)	0.5	0.5
Price/added value (€)	20	25
Fixed standard cost (€)	3.76	5.64
Variable standard cost (€)	9.84	13.7

SHERPA template, integrated in a specific Arena model, was set to reproduce an operator with high experience (PSF experience = high level) and in good physical fitness (PSF fitness for duty = nominal level) involved in moderately complex tasks. The PSFs for the context were chosen to represent approximately the actual conditions in the assembly plant: available time and work processes were imposed at the nominal level while stress, procedures and ergonomics at the moderate/high level.

The template, as implemented, investigates the performance of different work break configurations and it was applied in different scenarios changing the simulation parameters in Table V.14. Every break configuration was simulated for the three different recovery rates (slow, medium, and fast), for two different reworking probabilities (30% and 60% of the non-compliant items) and for two reworking times, which involve an increase of processing time equal to 15% and 30%. Without considering the four reworking classes, Table V.15 shows the list of the simulated scenarios in order to have a clearer and more immediate understanding of them.

Table V.14: *Simulation parameters.*

Parameters	Levels	Values
Recovery rate	3	Slow, Medium, Fast
Break Length (min)	3	20, 25, 30
Number of breaks	5	0, 1, 2, 3, 4
Reworking time (Tr)	2	+15%, +30%
Reworking probability (Pr)	2	30%, 60%

Table V.15: *Simulated scenarios.*

Tr= x% Pr= x%		**RECOVERY RATE**								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	S-20-1	S-25-1	S-30-1	M-20-1	M-25-1	M-30-1	F-20-1	F-25-1	F-30-1
	2	S-20-2	S-25-2	S-30-2	M-20-2	M-25-2	M-30-2	F-20-2	F-25-2	F-30-2
	3	S-20-3	S-25-3	S-30-3	M-20-3	M-25-3	M-30-3	F-20-3	F-25-3	F-30-3
	4	S-20-4	S-25-4	S-30-4	M-20-4	M-25-4	M-30-4	F-20-4	F-25-4	F-30-4
No breaks		(S-M-F) 0-0								

Two types of work-rest schedule were introduced: a single break in half shift or more breaks distributed at different times on the entire work shift. For each break configuration, the overall rest period length was considered respectively equal to 20/25/30 min. In the case of distributed breaks, the following distributions were hypothesized:

- Scenarios with 20 min of break:

Number of breaks	Length (min.)	Interval (min.)
2	10-10	180-120-120
3	6-8-6	150-90-90-90
4	5-5-5-5	132-72-72-72-72

- Scenarios with 25 min of break:

Number of breaks	Length (min.)	Interval (min.)
2	12.5-12.5	180-120-120
3	8-9-8	150-90-90-90
4	6-6-7-6	132-72-72-72-72

- Scenarios with 30 min of break:

Number of breaks	Length (min.)	Interval (min.)
2	15-15	180-120-120
3	10-10-10	150-90-90-90
4	7-8-7-8	132-72-72-72-72

In summary, the experiment was applied in 148 different scenarios in order to show how effective solutions for the break scheduling problem can be found with the proposed simulator.

V.2.4.2 Scenario 1: Simulation results

Results for every scenario consist of total value of compliant and non-compliant items, their respective percentages, mean values of the HEP context, as well as the economic results in terms of profit, revenue, scraps costs, rework costs, and breaks costs.

Table V.16 shows the average HEP for every scenario. These values reflect the chosen experiment, and they are a function of the performed assembly task as well as of the supposed individual and contextual factors. In addition to the scenarios defined, additional scenarios were simulated in the absence of breaks for every reworking class. The human error probabilities, reported in Table V.16, were significantly lower than those to the reference case in the absence of breaks because of the presence of operator's psychophysical recovery.

As described in Chapter III, the profits, related to the correct execution of each task, depend on the revenues of the compliant items, the scraps costs (fixed and variable unit costs), the costs of the reworking items (rework costs), and finally the breaks costs that stand for lack of production.

Table V.16: *Average HEPs for the simulated scenarios.*

Tr = +15% Pr = 30%		**RECOVERY RATE**								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	12.88%	12.20%	11.65%	11.83%	11.28%	11.19%	11.35%	11.28%	11.19%
	2	***12.85%***	***12.06%***	***11.40%***	11.57%	10.78%	10.18%	9.66%	9.61%	9.55%
	3	12.88%	12.09%	***11.40%***	***11.55%***	***10.72%***	10.09%	9.36%	8.88%	8.82%
	4	12.97%	12.13%	11.45%	11.61%	10.74%	***10.08%***	***9.28%***	***8.67%***	***8.43%***
No breaks		17.86%								

Tr = +15% Pr = 60%		**RECOVERY RATE**								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	12.89%	12.19%	11.64%	11.82%	11.28%	11.19%	11.35%	11.28%	11.19%
	2	***12.85%***	***12.06%***	***11.39%***	11.56%	10.78%	10.18%	9.67%	9.61%	9.55%
	3	12.88%	12.10%	11.40%	***11.54%***	***10.72%***	10.08%	9.36%	8.87%	8.83%
	4	12.96%	12.13%	11.45%	11.60%	10.74%	***10.07%***	***9.29%***	***8.68%***	***8.43%***
No breaks		17.84%								

Tr = +30% Pr = 30%		**RECOVERY RATE**								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	12.90%	12.19%	11.64%	11.82%	11.27%	11.20%	11.36%	11.27%	11.20%
	2	***12.85%***	***12.05%***	***11.39%***	11.56%	10.79%	10.18%	9.67%	9.61%	9.55%
	3	12.88%	12.08%	11.40%	***11.53%***	***10.72%***	***10.07%***	9.35%	8.86%	8.83%
	4	12.97%	12.12%	11.44%	11.59%	10.74%	10.08%	***9.28%***	***8.68%***	***8.43%***
No breaks		17.88%								

Tr = +30% Pr = 60%		**RECOVERY RATE**								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**20**	**20**	**20**	**20**	**20**	**20**	**20**	**20**
# breaks	1	12.86%	12.19%	11.62%	11.80%	11.26%	11.19%	11.35%	11.26%	11.19%
	2	***12.84%***	***12.04%***	***11.38%***	11.54%	10.78%	10.16%	9.66%	9.60%	9.54%
	3	12.87%	12.07%	11.39%	***11.52%***	***10.71%***	10.08%	9.36%	8.88%	8.82%
	4	12.93%	12.11%	11.44%	11.58%	10.73%	***10.07%***	***9.23%***	***8.67%***	***8.44%***
No breaks		17.80%								

Table V.17 reports the profits for the all the configurations which will be analysed in detail hereinafter.

Table V.17:*Profit in euros for the simulated scenarios.*

Tr = +15% Pr = 30%		RECOVERY RATE								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	64.497	64.367	63.794	67.674	66.825	64.971	68.763	66.825	64.971
	2	65.210	***64.494***	62.992	***68.212***	67.794	67.229	73.051	71.293	67.946
	3	***65.399***	64.163	***64.060***	67.620	***67.908***	66.888	72.730	72.719	70.521
	4	63.846	64.103	63.378	67.621	67.457	***67.929***	***74.235***	***72.942***	***70.857***
No breaks		61.339								

Tr = +15% Pr = 60%		RECOVERY RATE								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	76.861	75.207	73.646	79.005	76.699	74.131	79.754	76.699	74.131
	2	***77.272***	75.181	***74.399***	79.378	77.650	76.257	81.344	79.439	77.052
	3	77.249	75.163	74.159	***79.415***	77.363	76.055	***82.513***	80.214	78.101
	4	76.597	***75.755***	74.047	79.228	***77.680***	***76.459***	82.447	***81.186***	***78.943***
No breaks		79.615								

Tr = +30% Pr = 30%		RECOVERY RATE								
		Slow			**Medium**			**Fast**		
		Total Breaks Length			**Total Breaks Length**			**Total Breaks Length**		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	64.191	63.365	63.221	67.420	66.371	63.871	68.158	66.371	63.871
	2	63.489	63.358	63.648	67.234	66.677	66.009	72.251	70.499	68.158
	3	64.041	***64.226***	63.253	***67.713***	***67.638***	66.444	73.294	72.664	69.257
	4	***64.197***	63.872	***64.087***	67.375	67.571	***67.598***	***73.904***	***74.328***	***70.941***
No breaks		59.703								

Tr = +30% Pr = 60%		RECOVERY RATE								
		Slow			Medium			Fast		
		Total Breaks Length			Total Breaks Length			Total Breaks Length		
		20	**25**	**30**	**20**	**25**	**30**	**20**	**25**	**30**
# breaks	1	75.569	74.906	72.800	77.602	75.861	74.707	78.366	75.861	74.707
	2	75.964	75.010	***73.690***	78.072	***77.110***	75.005	81.562	79.010	76.054
	3	***76.105***	75.666	72.772	78.363	77.026	***75.168***	***82.347***	80.257	77.552
	4	75.984	***75.977***	73.265	***78.595***	76.784	75.077	82.328	***80.867***	***77.606***
No breaks		78.347								

V.2.4.3 Scenario 1: Discussion

The purpose of this experiment is the evaluation of impact of different work-rest policies on human reliability and system performance for an assembly process. Reliability evaluation involves three significant aspects associated with the impact of recovery rate, breaks time, and configurations.

The decrease of worker error probability in the simulated scenarios, in fact, derives from the break time and the recovery rate and it is underlined graphically in Figure V.10. There is a statistically significant interaction between the effects of recovery rate and break time on the HEP; therefore, the effect on the mean outcome of a change in one factor depends on the level of the other factor. The vertical bars in the graphs indicate the level of confidence at 95%.

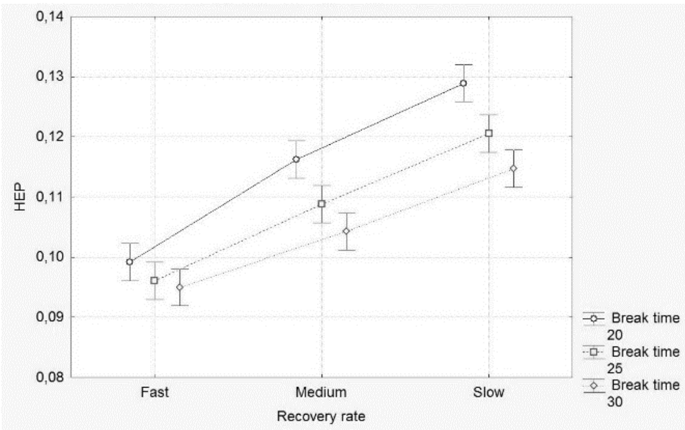

Figure V.10: *Human error probability value as a function of recovery rate and break total time without distinction of rework class.*

The human unreliability is a function of the operator's recovery rate through equation (3.16) and (3.17), as previously explained in Chapter III. With equal break lengths, in fact, a slower recovery rate leads to higher values of HEP. The recovery rate is an inherent feature of the worker that depends on several elements, such as the age (Mohren, Jansen and Kant, 2010), and even if in a limited manner, it can be influenced by the regenerating activities during the same break, e.g., specific physical exercises (van de Heuvel, de Looze and Hildebrandt, 2003; Balci and Aghazadeh, 2004). The increase of total break time in the shift, instead, improves human reliability because the worker has more time to rest and receive a greater psychophysical recovery. This increase is naturally stronger in the case of slow recovery rate than that fast, because in this last case, a shorter time for an adequate recovery is enough. The last assessment is linked to the effect of several work-break configuration in the shift (Figure V.11).

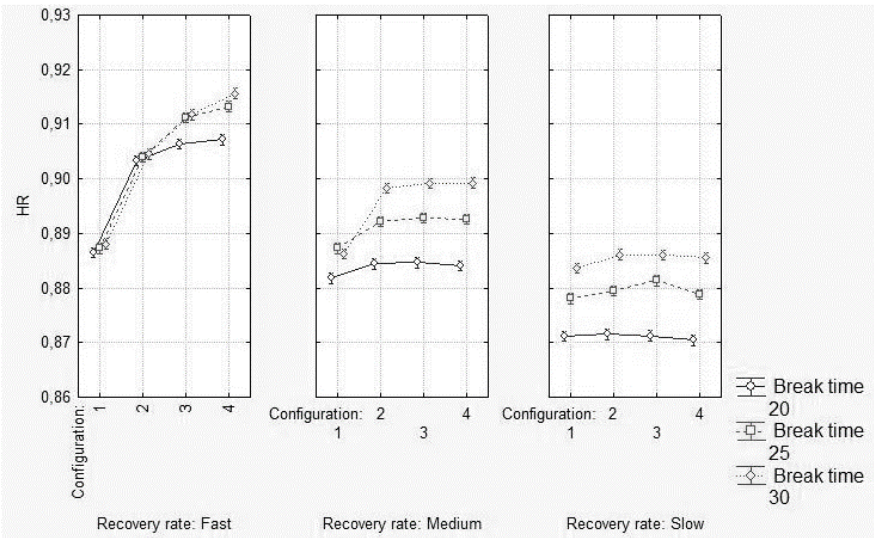

Figure V.11: *The impact of work–rest configurations in terms of HR.*

It is evident that the four work-rest policies impact differently on the worker reliability according to the rate of recovery. The HR improvement is much more stringent for the fast recovery rate, where the single break in half shift is less effective and significantly worse compared to the other three configurations, which exploit the distribution of shorter breaks over 8-h to their advantage. The work-rest policies with three or four breaks in the shift allow the worker more rest moments, increasing its average reliability. These benefits are less marked in the case of medium recovery rate and almost insignificant when the recovery rate becomes slow. This can easily be justified with a propensity to longer pauses that allow a greater recovery for the operator.

The previous evaluations were carried out without discrimination on the reworking class (reworking time and reworking probability) since this has no impact on the HEP. In the economic evaluations, however, the reworking class must be taken into account due to its significant effect on the profits. Comparing Table V.16 and Table V.17, it is evident that when reworking class changes, the HEP value remains unchanged while the profits vary greatly. This effect can be easily justified, considering that the rework has no impact on the human reliability distribution, but it influences the number of compliant items of the system that generate higher profits. For this reason, the scenarios were evaluated separately considering the different reworking classes and recovery rates in order to assess the economic impact of different break configurations.

Figure V.12 reports the profits for the scenarios with rework probability equal to 30% recovery and rework time equal to 30% for the three recovery rates. It is evident that the one-break configuration is always less advantageous compared to three or four distributed breaks in terms of economic performance. Unlike the HEP trend, the increase of the total length does not always have an improving effect on profit and the economic results only partially reflect the previous HR assessments.

z

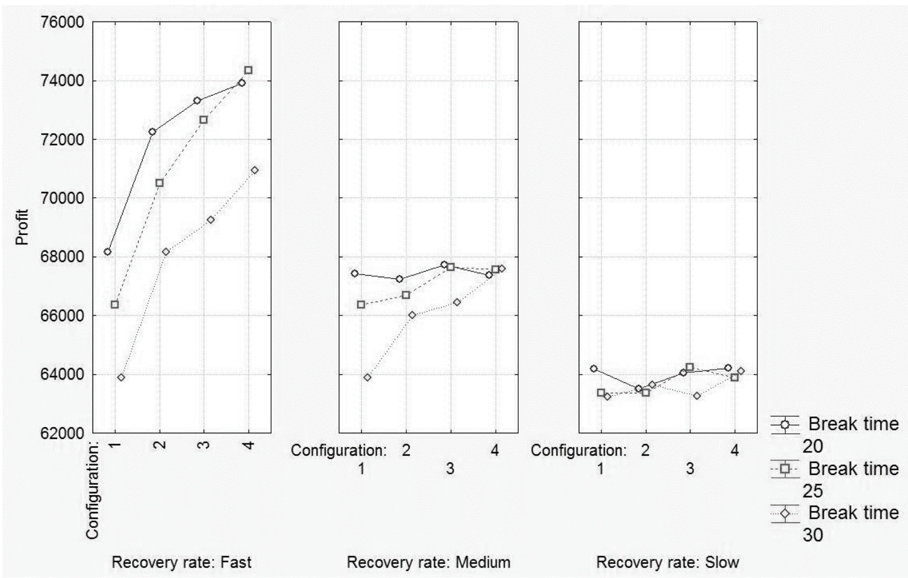

Figure V.12: *Economic performance (profits in euros) to changing work–rests policies with fixed rework class (30% reworking probability and 30% reworking time).*

Table V.18 lists the economic results in the case of reworking time and probability equal to 30% in average conditions of recovery. The best results for each performance parameter are underlined. It is evident that the best economic performances do not correspond in the same order with the best reliabilities. This result derives from the combination of the economic impact of break times and break configurations.

The positive HR variations related to the increase of the number and the length of breaks involves an improvement in the rate of quality of the processing and consequently a lower cost of scraps, while the increase of the break time determines a clear rise of the breaks costs. As previously described, the break costs represent the costs of lost production time; naturally, the transition from 20 to 30 min increases the number of products not manufactured and the break costs, and this is reflected in a reduction of global revenues, which do not result from a deterioration in the quality of work, but only by the reduction of the total worked hours.

Table V.18: *Details of the economic performance.*

SCENARIO		**PERFORMANCES**					
Break time	**# breaks**	**Profits (€)**	**Revenues (€)**	**Scraps costs (€)**	**Rework costs (€)**	**Breaks costs (€)**	**HEP %**
20	1	67,420	96,936	22,121	633	6,762	11.82%
20	2	67,234	96,749	22,086	634	6,795	11.56%
20	3	67,713	96,997	21,855	630	6,799	11.53%
20	4	67,375	96,850	22,060	624	6,791	11.59%
25	1	66,371	96,131	20,597	624	8,540	11.27%
25	2	66,677	96,194	20,317	584	8,615	10.79%
25	3	67,638	96,458	19,604	589	8,626	10.72%
25	4	67,571	96,520	19,742	583	8,623	10.74%
30	1	63,871	95,306	20,560	615	10,261	11.20%
30	2	66,009	95,705	18,670	571	10,454	10.18%
30	3	66,444	95,849	18,352	578	10,475	10.07%
30	4	67,598	96,336	17,731	533	10,473	10.08%

Being the obtained profits strongly dependent on the economic parameters, a further analysis was carried out with the following changes on the reference case:

- Price/added value: ±20%;
- Fixed standard cost: ±50%;
- Variable standard cost: ±20%.

Figure V.13 shows the profits for the new six scenarios obtained modifying such parameters according to the OFAT (One Factor At a Time) analysis technique.

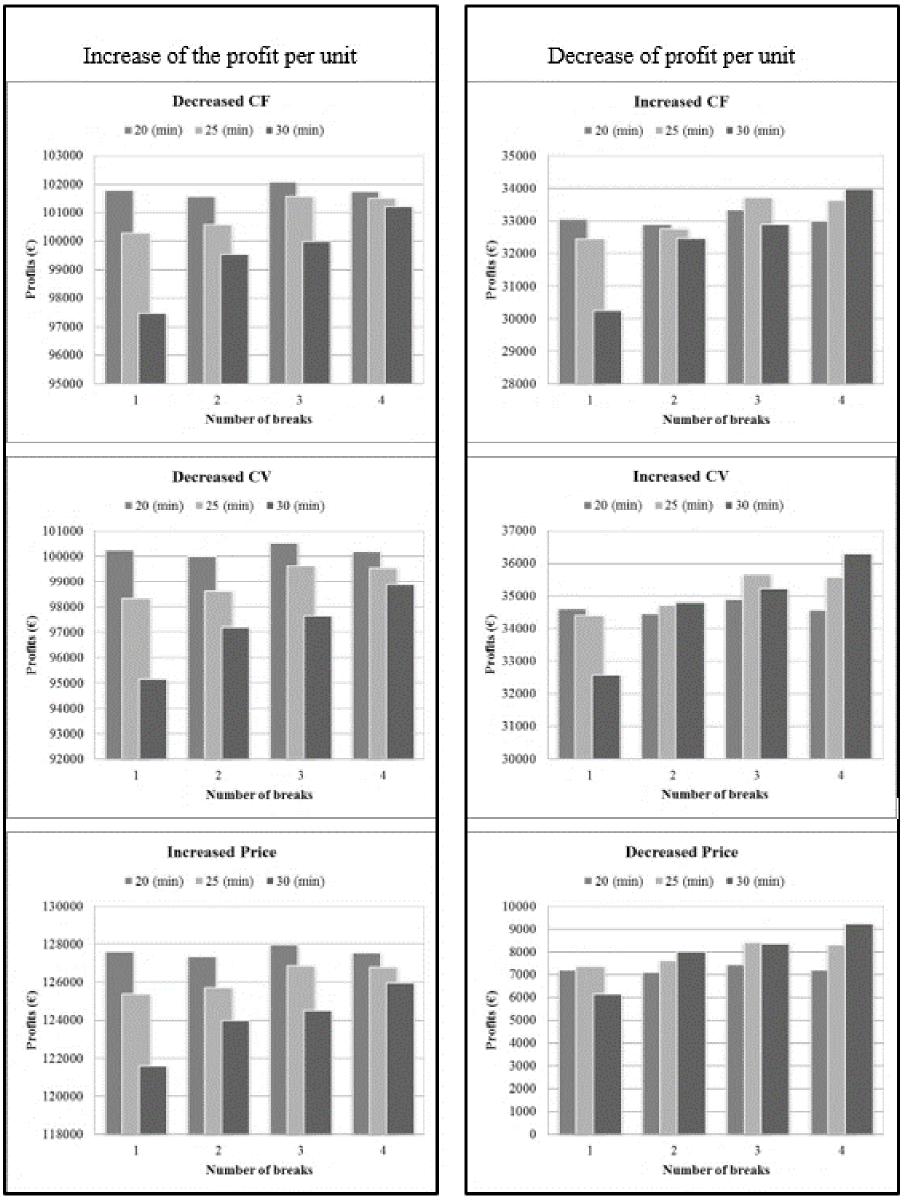

Figure V.13: *Economic performances with changes in the economic baseline.*

The effects of these changes belong to two distinct classes of result: growth and the reduction of the profit per unit. On one side, the reduction of the costs, both fixed and variable, and the increase of the price increase the profit per unit and this leads to an overall increase in profits. In this condition, shorter breaks are preferable in all possible configurations since they reduce the non-working time and increase the total production; in fact, the pauses of 20 min are always the best ones. Furthermore, the advantages associated to a greater number of breaks are more evident in the case of breaks by 30 min, which is noted in the net increase of profits in the passage from one to four break pauses. Otherwise, when the costs rise or the price drops, the reduction of the profit per unit greatly lowers the economic performance of the system and it entails the convenience of longer breaks than the previous case. In this situation, the scenarios with four breaks amounting 30 min always represent the best choice, and in general 25 and 30 min of breaks are economically preferable especially with distributions of 3 and 4 breaks. Such variations are caused to the different impact of scrap costs, break costs, and revenues when the economic parameters of the examined case study change.

V.2.4.4 Scenario 2: Problem definition

The SHERPA simulator was also used to conduct several simulations in order to evaluate the effect of rest break management in a manufacturing industry. The simulation model reproduces the operator work station involved in manual assembly, on an 8-hour shift, considering 235 working days. The assembly operation was simulated for three different items with random arrival sequences based on a production mix and with processing times characterized by a triangular distribution, with vertices corresponding to the mean ±10%. The item input data are shown in Table V.19.

PSFs for this scenario were chosen to represent approximately the actual conditions in the assembly plant. Available time, procedures, fitness for duty and work processes were imposed at the nominal level; complexity, stress and ergonomics at the moderate/high level and experience at high level. Of course, the choices made do not reflect accurately the reality, because this is highly variable from context to context. For this scenario the following breaks scheduling were simulated:

- absence of breaks during the work shift;
- a long pause (30 min.) in mid-turn (fixed 1);
- two breaks of 15 minutes, the first after 2.5 hours and the second after other 3 hours (fixed 2);
- three short breaks (10 min.) every two hours (fixed 3);

- seven breaks of 5 minutes each hour (fixed 4);
- automatic determination of breaks for each operator (minimum break time 1,2,3,5 minutes).

Table V.19: *Features of produced items in the experiment.*

Features	Item 1	Item 2	Item 3
Mean processing time (min.)	25	36	45
Setup time (min.)	5	5	5
Price (€)	115	155	200
Fixed cost (€)	52	65	78
Variable cost (€)	18	24	32
Productive mix	20%	30%	50%
Production target (units/year)	2300		

V.2.4.5 Scenario 2: Simulation results and conclusions

In Table V.20 the simulation results are shown; for the automatic management, the number and the break length assigned by the evaluation of the economic convenience are reported. In first analysis, the percentage of errors, when breaks are absent in the shift, is higher than in the case of automatic management and programmed break (from 23% to 12-14% of non-compliance).

Table V.20: *Results of the case study.*

Breaks scheduling	Time breaks (min)	N. Breaks	%Comple-tion	% Err	Annual breaks cost (€)	Annual Profit (€)
No breaks	0	0	91%	23%	0	96 802
Automatic 1	20.8	6.64	98%	14%	9 694	123 703
Automatic 2	20.85	5.9	97%	14%	9 767	119 553
Automatic 3	20.5	4.8	97%	15%	9 650	117 129
Automatic 5	20.3	3.35	97%	14%	9 715	120 522
Fixed 1	30	1	94%	15%	14 805	108 449
Fixed 2	30	2	100%	13%	14 921	120 000
Fixed 3	30	3	97%	12%	14 878	119 497
Fixed 4	35	7	96%	13%	16 418	115 514

The decrease of scrap has to be attributed to the physical and mental recovery operator following a break, is also clear that as the length of the pause, decreases the total number of units produced, because the break is a period of non-production.

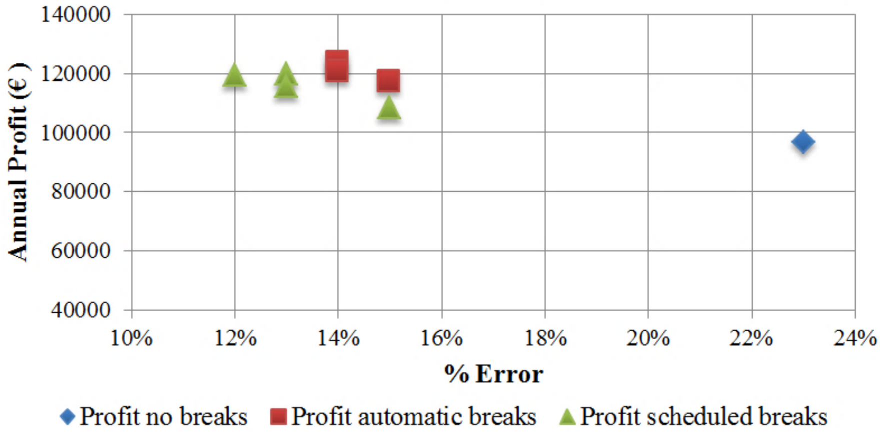

Figure V.14: *Annual profit vs % Error for the simulated scenarios.*

Automatic management in all scenarios determines many breaks but with short time, such as to ensure the recovery of the operator at times of high probability of error. It is clear the advantage of distributing the breaks throughout the work shift and not concentrate them in few interruptions of long duration. From the economic point of view the automatic management appears to be the best not only in terms of percentage of errors but especially considering the total profit, depending on the number of good products but also on the break time and on when the break is carried out (Figure V.14).

V.3 Case study: Orthopaedics Surgery

The SHERPA method was applied through the AnyLogic simulation template for the study of the human reliability inside an existing operating room. The objective was to conduct a scenario analysis for the evaluation of the HR effects on the result of an orthopaedic surgery, both in terms of errors committed and in terms of the time taken to perform it.

V.3.1 Problem definition

The case study was conducted in an operating room of Department of Orthopaedics and Traumatology of the University Hospital San Giovanni di Dio-Ruggi d'Aragona of Salerno. The Department of Orthopaedics and Traumatology deals with both elective and urgent interventions, occupying the operating room on Monday, Wednesday and Friday from 8 am to 2pm.

For the study carried out it was chosen to refer to a particular surgery: the reconstruction of the Achilles tendon following rupture. This intervention is carried out using a particular innovative technique described below (Maffulli *et al.*, 2008). Task analysis was performed to determine the chronological steps involved in the Table V.21.

Table V.21: *Operating activities.*

TASK	DESCRIPTION	RESOURCES
Anaesthesia room	In this phase the identity of the patient and of the site to be operated is verified. The anaesthesiologist administers the anaesthesia and it waits until the patient is sedated.	Nurse Anaesthesiologist
Pre-intervention in the operating room	A nurse and a theatre nurse proceed with the preparation of the material and equipment used during the operation.	Nurse Theatre nurse
Patient positioning	The patient is moved to the operating bed. The patient is positioned prone with a thigh tourniquet.	Nurse
Operative field preparation	Skin preparation is performed, and sterile drapes are applied.	Surgeon Nurse Anaesthesiologist
Operating activity 1	Pre-operative anatomical markings include the palpable tendon defect and both malleoli.	Surgeon
Operating activity 2	The first incision is a 5 cm longitudinal incision, made 2 cm proximal and just medial to the palpable end of the residual tendon.	Surgeon
Operating activity 3	The second incision is 3 cm long and is also longitudinal but is 2 cm distal and in the midline over the distal end of the tendon rupture.	Surgeon
Operating activity 4	The tendon of the semitendinosus is harvested through a vertical, 2.5–3 cm longitudinal incision over the pes anserinus	Surgeon

TASK	DESCRIPTION	RESOURCES
Operating activity 5	An osteotomy of the postero-superior angle of the calcaneus is performed.	Surgeon
Operating activity 6	The calcaneus is pierced and reamed in order to obtain a bone tunnel to pass the semitendinosus to be transplanted.	Surgeon
Operating activity 7	The semitendinosus muscle tendon is passed through an incision of the proximal abutment and secured to the entry and exit points of the incision.	Surgeon
Operating activity 8	The tendin of the semitendinosus muscle is lodged through the bone tunnel first made in the calcaneus, stretched with the foot in the position of complete plantar flexion, fixed with a heel screw and tied to the distal stump of the Achilles tendon.	Surgeon
Operating activity 9	The incisions are sutured, and the limb plastering is performed.	Surgeon

V.3.2 Problem modelling

As a basic assumption, it was decided to study the process that includes the activities performed in the operating room starting from the patient entry time to the anaesthesia room to the ends of operation. Moreover, it was hypothesized (even considering the intervention that was chosen to simulate) that the operating team is composed of two nurses, a theatre nurse, an anaesthesiologist and two surgeons, with individual features reported in Table V.22.

Table V.22: *Resources features.*

Resource	Gender	Age
Nurse	F	40
Anaesthesiologist	M	40
Theatre nurse	M	40
Surgeon	M	50

Possible errors and hazards related to the procedure were identified (Maffulli *et al.*, 2008):

- Inadequate exposure and traction of the proximal stump of the Achilles tendon (operating activity 3).
  - ✓ Solution - Complete the exposure.
- Incorrect positioning, breaking, loss or loosening of the screw (operating activity 8).
  - ✓ Solution - Remove the screw.
- Inadequate tendon tension (operating activity 8).
  - ✓ Solution - Repetition of the intervention.
- Calcaneus fracture (operating activities 5 and 6).
  - ✓ Solution - Fracture reduction and internal fixation.
- Damage to the sural nerve.
- Infection.
  - ✓ Solution - Antibiotics.
- Reduced hip mobility.
  - ✓ Solution - Physiotherapy.
- New rupture.
  - ✓ Solution - Repetition of the intervention.

The possible errors on the operative activities that lead to possible post-intervention complications are to be mainly found in the operating activity 8: the others can be easily recovered, leading to a lengthening of the times of the operation itself.

Figure V.15 shows graphically how the model is organized on a conceptual level. The arrival of the patient determines the beginning of the whole process: the operating team, composed of the various resources, marks the succession of tasks to be performed and it determines the progress of the operation within the model. The SHERPA blocks, one for each resource involved, allow the collection of indications on the reliability of resources and the relative intervention performance, in terms of errors committed and times taken to complete the activities. Each resource (nurse, anaesthesiologist, theatre nurse and surgeon) corresponds to an agent within which a SHERPA block is placed. In this case, the SHERPA block simulates the completion of an elementary operation: each processed entity represents an activity that can end with a success, a rework or an error, being the object of an estimate of the reliability of the individual who performs it. In particular, the rework entails an increase of the operating activity time.

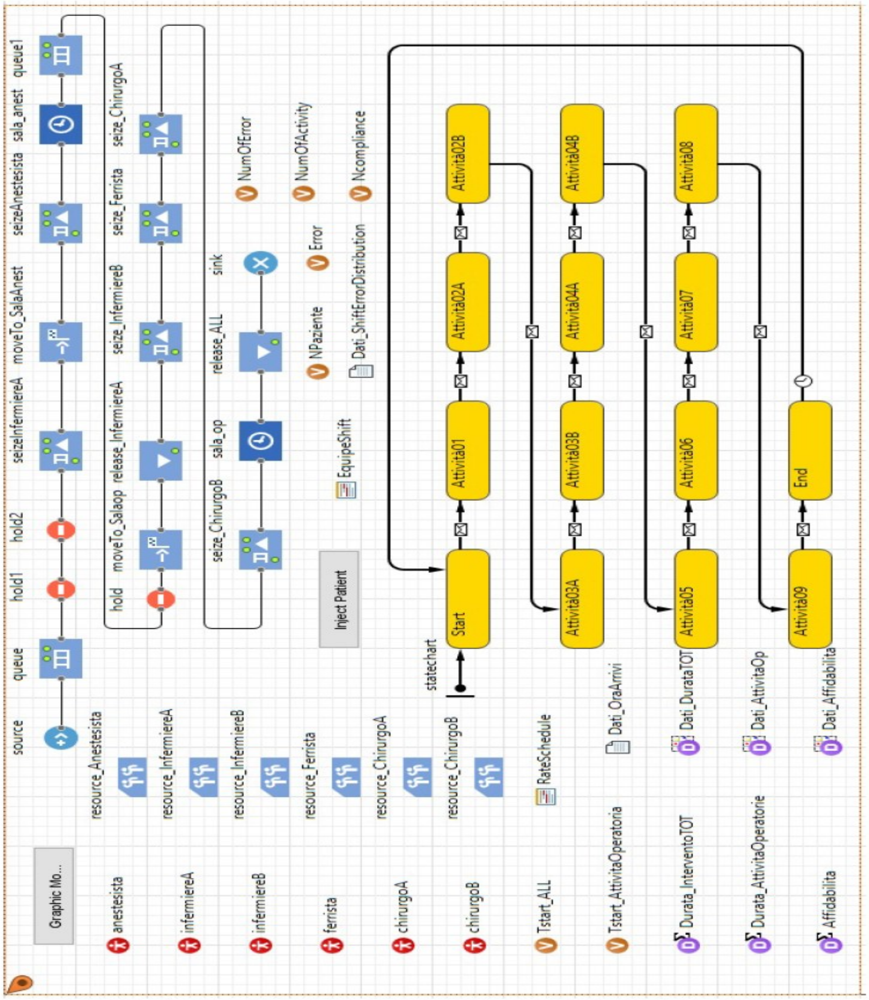

Figure V.15: *Operating room model.*

The activities performed in the pre-operative anaesthesia room, and operating theatre by each resource was modelled in detail. For example, the activities performed by Nurse, Anaesthesiologist and Surgeon were modelled as shown in Figure V.16, Figure V.17 and Figure V.18. The respective *State charts* allow managing and process all the activities performed by the individual operator in the sequence reported in Table V.21, and to calculate in real time the operator's performance and the possible errors committed with the relative consequences. At the end of the operation, the patient operated leaves the model, which is ready to receive another one.

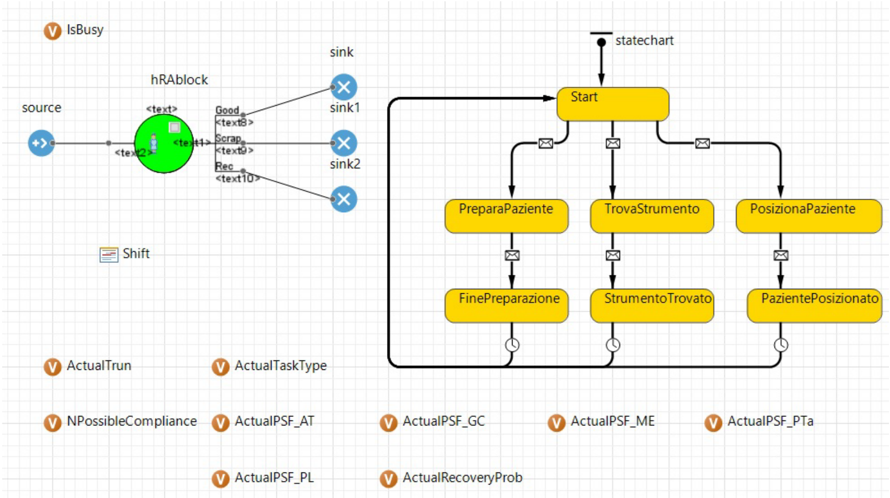

Figure V.16: *Nurse Agent.*

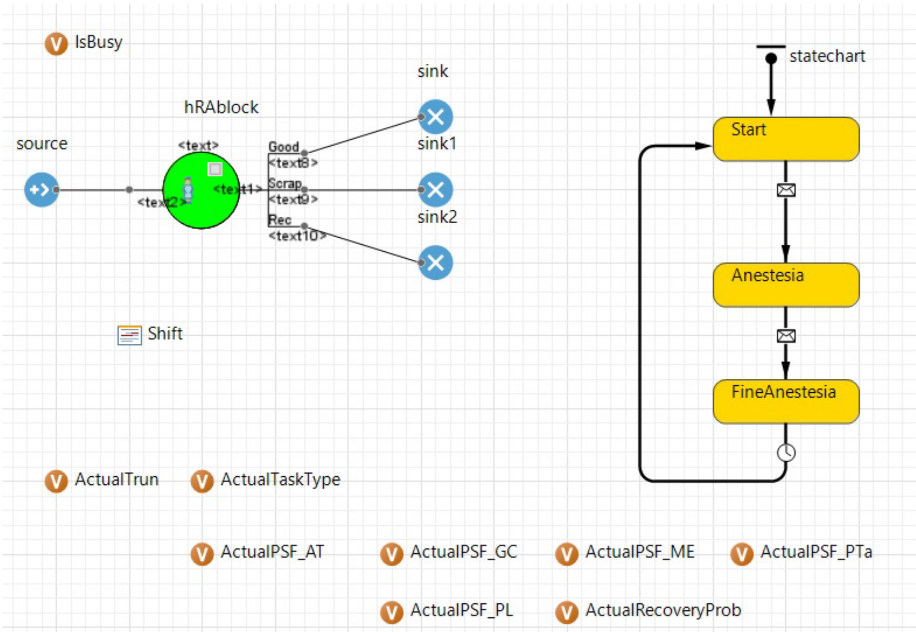

Figure V.17: *Anaesthesiologist Agent.*

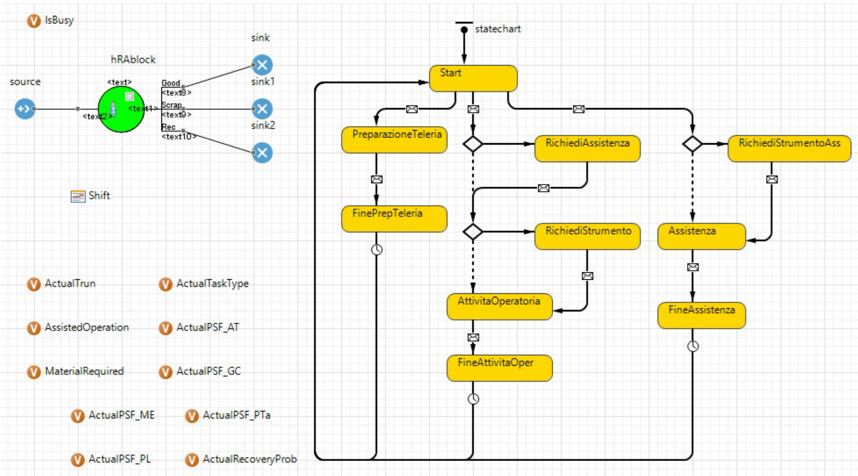

Figure V.18: *Surgeon Agent.*

Table V.23 shows the times of the activities planned for the intervention. These times were obtained through a direct comparison with the personnel taking part in the intervention and were re-elaborated to be implemented in the simulator.

Table V.23: *Time of operating activities.*

Activity	**Minimum time**	**Modal time**	**Maximum time**
Anesthesia room	15 min	20 min	30 min
Pre-intervention in the operating room	5 min	10 min	15 min
Patient positioning	—	—	2 min
Operative field preparation	5 min	7-8 min	10 min
Operating activity 1	—	—	1 min
Operating activity 2	1 min	2 min	3 min
Operating activity 3	—	—	1 min
Operating activity 4	—	—	1 min
Operating activity 5	—	—	2 min
Operating activity 6	—	—	2 min
Operating activity 7	3 min	4 min	5 min
Operating activity 8	—	—	1 min
Operating activity 9	5 min	7-8 min	10 min

For each entity entering SHERPA the attribute corresponding to the duration of the activity must be assigned. The starting point for the choice of the distributions from which to derive this duration is Table V.23. In cases where all the values were known, a triangular distribution was defined. In other cases, not having enough information available, it was decided to use a uniform distribution. Some activities were divided because they are composed of more sub-activities that require the use of different tools (for example the operating activity 2, 3 and 4) or the collaboration among more resources (in cases where there is a request for an instrument or a request for assistance). In Table V.24, there is a collection of all the distributions used to the resource involved and to the state in which the duration is assigned.

Each resource was modelled from relevant field data and the PSF levels were defined as shown in Figure V.19 and Figure V.20.

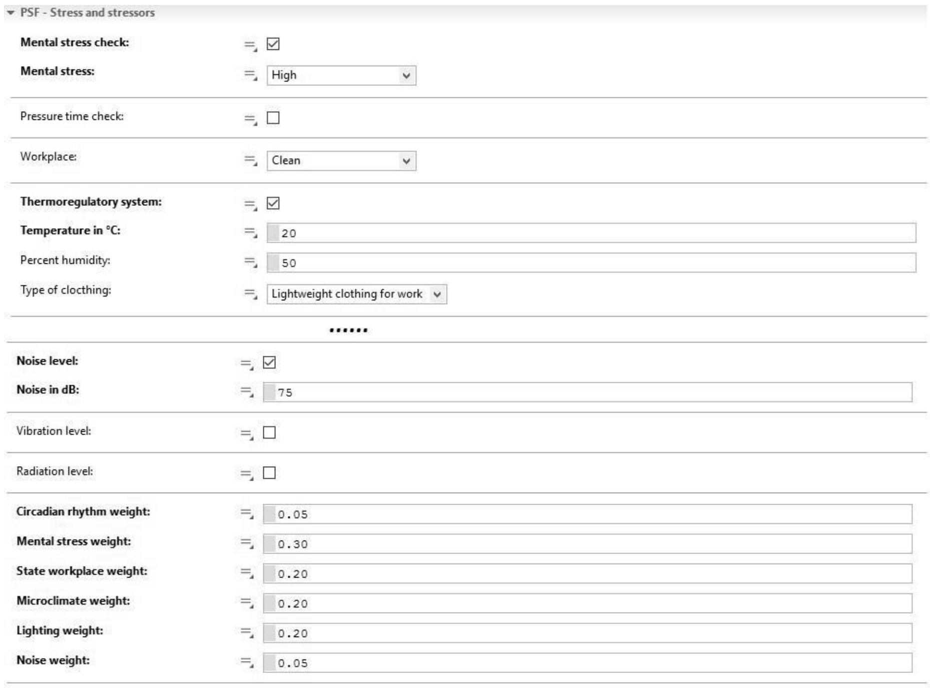

Figure V.19: *PSFs data (part 1).*

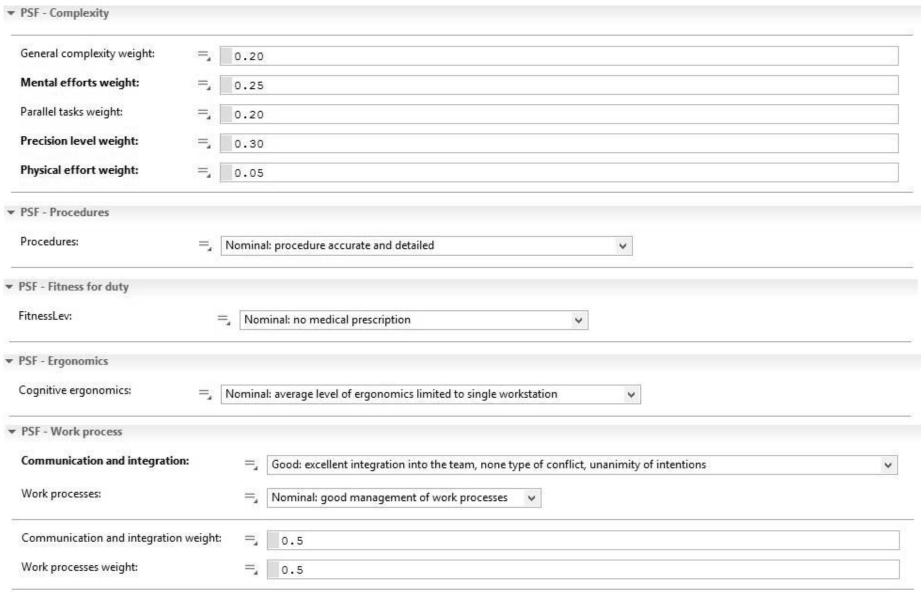

Figure V.20: *PSFs data (part 2).*

Table V.24: *Distribution used in the model.*

Activity	State	Agent Resource	Distrbution (min)
Anesthesia room	PreparaPaziente	NurseA	uniform(0.5, 2)
Anesthesia room	Anestesia	Anesthesiologist	triangular(15,30,20)
Pre-intervention in the operating room	TrovaStrumento	NurseB	triangular(0.15, 1.50, 0.25)
Pre-intervention in the operating room	PosizionaStrumento	Theatre nurse	uniform(0, 0.15)
Patient positioning	PosizionaPaziente	NurseB	uniform(0.25, 2)
Operative field preparation	PreparazioneCute	Theatre nurse	triangular(3, 6, 4.5)
Operative field preparation	PreparazioneTeleria	SurgeonB	triangular(2, 4, 3)
Operating activity	PrendiStrumento	Theatre nurse	uniform(0.05, 0.15)
Operating activity 1	Attività01	SurgeonA	uniform(0.25, 1)
Operating activity 2	Attività02A	SurgeonA	uniform(0.15, 0.25)
Operating activity 2	Attività02B	SurgeonA	triangular(1, 3, 2)

Activity	State	Agent Resource	Distrbution (min)
Operating activity 2	Attività02B	SurgeonB	uniform(0.25, 0.5)
Operating activity 3	Attività03A	SurgeonA	uniform(0.15, 0.25)
Operating activity 3	Attività03B	SurgeonA	uniform(0.50, 1)
Operating activity 4	Attività04A	SurgeonA	uniform(0.15, 0.25)
Operating activity 4	Attività04B	SurgeonA	uniform(0.50, 1)
Operating activity 4	Attività04B	SurgeonB	uniform(0.25, 0.5)
Operating activity 5	Attività05	SurgeonA	triangular(1, 2, 1.5)
Operating activity 6	Attività06	SurgeonA	triangular(1, 2, 1.5)
Operating activity 6	Attività06	SurgeonB	uniform(0.25, 0.5)
Operating activity 7	Attività07	SurgeonA	triangular(3, 5, 4)
Operating activity 8	Attività08	SurgeonA	uniform(0.50, 1)
Operating activity 9	Attività09	SurgeonA	triangular(5, 10, 7.5)

V.3.3 Model validation

Verification and validation of simulation model is usually part of the model development process. In order to enhance the assurance of the results and evaluate the accuracy of models, it is necessary to verify and validate the simulation model. The model validation is defined as proving that the conceptual model is an accurate representation of real system which deals with forming the correct model.

The model presented was validated and verified by statistical test considering the total time of the intervention (from the beginning of the administration of the anaesthesia to the end of the operating activities) and the time of the operating activity with data collected within the operating registry from November 2015 to October 2016 and shown in Table V.25.

Real durations were subjected to a boxplot analysis that led to the exclusion of the real patient number three. Table V.26 contains the durations obtained by running a model run on the simulator. The tests necessary for the validation of the model (Table V.27) were applied both for the total duration of the intervention and for the individual operating activities.

Table V.25: *Historical data from operating registry.*

Patient	Date	Room Entry	Start of anaes-thesia	Start of Surgery	End of Surgery	End anaes-thesia	Room Exit
01	26/01/2016	14:30	14:30	14:50	15:10	—	15.20
02	15/03/2016	8:50	9:00	9:10	9:25	9:30	10:00
03	17/03/2016	8:00	8:00	8:30	9:40	—	9:45
04	29/03/2016	13:00	13:00	13:20	14:00	—	14:10
05	05/04/2016	10:50	11:00	11:30	12:10	12:20	12:30
06	28/04/2016	12:00	12:00	12:15	12:40	—	12:45
07	16/06/2016	8:15	8:30	9:00	9:40	—	9:50
08	23/06/2016	8:00	8:00	8:45	9:15	—	9:20
09	14/07/2016	8:00	8:30	9:00	9:15	—	9:30
10	26/07/2016	8:00	8:10	9:00	9:50	—	10:00
11	06/09/2016	8:00	8:25	9:00	9:30	—	9:40

Table V.26: *Data derived from real system and simulator.*

Real system			Model to the simulator		
Patient	**Total activity duration (min)**	**Duration of surgery (min)**	**Patient**	**Total activity duration (min)**	**Duration of surgery (min)**
1	40	20	**1**	59.131	23.076
2	25	15	**2**	70.701	24.547
3	100	70	**3**	70.879	21.539
4	60	40	**4**	53.956	22.036
5	70	40	**5**	51.302	22.165
6	40	25	**6**	51.991	21.867
7	70	40	**7**	57.687	21.323
8	75	30	**8**	60.466	26.046
9	45	15	**9**	53.025	23.362
10	100	50	**10**	71.152	24.565
11	65	30	**11**	54.003	22.692
			12	60.909	23.279
			13	59.987	28.120
			14	51.008	22.323
			15	60.146	26.596

Table V.27: *Validation results with indication of the p-values obtained.*

	Shapiro-Wilk's test	Fisher 's test	Wilcoxon's test	Smi.-Sat.'s test	Res. Val.
Total real time	0.8090	—	1	—	OK
Total model time.	0.0297				
Real Operating activity time	0.4652	10^{-8}	—	0.0999	OK
Model Operating activity time	0.0698				

V.3.4 Design of experiment

The validated model was used as a decision support system in the operating room to evaluate HR impacts on performance. A scenario analysis was performed, by changing one or more parameters that regulate the functioning of the simulation model, and choosing, after appropriate considerations on a statistical basis, the best solution. The Key Performance Indicators (KPI), taken as a reference, are the overall reliability of the system, the time the patient has passed from entering the anaesthesia room to leaving the operating room and the effective duration of operating activities. In particular, the overall reliability is calculated at the end of each intervention through the following formula:

$$Reliability = \frac{Number\ of\ activities\ performed\ without\ error}{Total\ number\ of\ activities\ performed} \tag{5.1}$$

Errors were considered both if the error is recovered during the interventions and if the error is not recovered and it is, therefore, a source of possible future complications. The total number of activities for the operation under examination (from anaesthesia to actual intervention) was fifty-three, divided as follows: the anaesthesiologist performs only one activity, the nurseA performs only one activity, the nurseB performs eleven activities, the theatre nurse performs twenty-four activities, the surgeonA performs twelve activities and the surgeonB performs four activities.

Two experimental campaigns have been chosen: the first one focuses on factors (PSF) that directly influence reliability calculations using the SHERPA method, and it is used as a tool for verifying the correct modelling of the process. The second focuses mainly on organizational factors such as the break scheduling management, and it is used to determine if changes in this direction lead to improvements compared to the As-Is model.

V.3.4.1 Experimental campaign 1: PSF impacts

Factors chosen among the available PSFs are those that most influence a medical activity (Dollarhide *et al.*, 2014): Mental Stress, Fitness for duty, Communication. In particular, the first two are related to the psycho-physical well-being of the surgery staff involved in the intervention, the last one regards the harmony of the medical team. The changes of the PSF concern all the resources involved. The full list of alternatives is reported in Table V.28, where the alternative As-Is, that is the model of the real system, corresponds to the number five. This experimental campaign allows to verify the hypothesis of decrease of reliability and increase of the operating times with negative value assumed by the PSFs.

Table V.28: *List of alternatives of the first experimental campaign.*

Alternatives	Levels		
	Mental Stress	**Fitness for duty**	**Communication**
1	Nominal	Nominal	Good
2	Nominal	Nominal	Poor
3	Nominal	Degraded	Good
4	Nominal	Degraded	Poor
5 (As-Is)	High	Nominal	Good
6	High	Nominal	Poor
7	High	Degraded	Good
8	High	Degraded	Poor
9	Extreme	Nominal	Good
10	Extreme	Nominal	Poor
11	Extreme	Degraded	Good
12	Extreme	Degraded	Poor

V.3.4.2 Experimental campaign 2: Break scheduling management

Organizational factors chosen for this experimental campaign are:

- the scheduling of break before each intervention, if this occurs after the first hour with respect to the beginning of the shift;
- the reduction of operating times, as if a limit of duration to be respected was imposed, in order to reduce its variability. This effect is obtained by decreasing the maximum time associated with the distributions previously seen. This condition activates the PSF Pressure Time. The effect of this activation is a decrease in reliability with a greater possibility of error and, therefore, an increase in times (in a situation opposite to that desired).

The nine alternatives are shown in Table V.29, where the As-Is alternative is the number one. The goal is to find the alternative that ensures the best balance between the reliability achieved and the durations of the interventions.

Table V.29: *List of alternatives of the second experimental campaign.*

Alternatives	Levels		
	Break time	**Maximum time reduction**	**Pressure Time**
1 (As-Is)	0 min	0%	Absent
2	0 min	5%	High
3	0 min	15%	Extreme
4	10 min	0%	Absent
5	10 min	5%	High
6	10 min	15%	Extreme
7	15 min	0%	Absent
8	15 min	5%	High
9	15 min	15%	Extreme

V.3.5 Analysis of results and discussions

The analysis of the results is divided into four phases: identification of the number of replicas; execution of the simulation runs; statistical analysis of the simulation runs and choice of the best solution.

V.3.5.1 Analysis of the results for the first experimental campaign

Table V.30, Table V.31, and Table V.32 show the results obtained in terms of average KPI of interest in the first experimental campaign. All the relative standard errors are lower than the pre-established threshold of 0.1, and for this reason the 10 replicas made for each alternative are sufficient to proceed with the subsequent analyses.

Table V.30: *KPI total flow time.*

	Average total flow time (min)											
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**
1	60.52	73.12	68.62	74.80	59.51	64.67	62.48	75.87	58.75	66.88	66.18	75.81
2	58.98	62.70	65.15	79.54	63.24	64.59	69.37	76.23	65.19	68.71	67.61	76.78
3	56.92	62.78	61.90	72.97	61.06	65.19	63.80	73.82	59.61	71.74	70.55	81.06
4	54.64	67.09	66.75	73.75	57.70	66.74	61.53	79.36	59.93	69.49	64.06	72.09
5	58.38	66.22	62.41	71.84	61.25	61.01	64.48	78.08	64.39	72.54	67.45	79.53
6	57.26	62.63	66.68	78.92	62.41	68.05	67.35	72.15	61.37	68.20	65.21	71.69
7	59.16	63.62	62.50	74.91	59.75	67.14	66.17	74.67	62.61	63.89	66.89	80.44
8	59.22	71.05	66.61	70.89	62.10	68.46	66.11	71.93	56.33	71.99	65.63	74.05
9	56.95	64.71	59.24	71.75	58.68	67.22	70.08	74.51	64.25	68.31	63.97	78.55
10	57.19	68.30	60.66	70.65	59.88	65.12	67.02	67.06	62.17	66.93	68.75	81.29
μ	***57.92***	***66.22***	***64.05***	***74.00***	***60.56***	***65.82***	***65.84***	***74.37***	***61.46***	***68.87***	***66.63***	***77.13***
σ	1.66	3.68	3.11	3.13	1.75	2.20	2.79	3.48	2.82	2.69	2.07	3.62
e	0.02	0.04	0.03	0.03	0.02	0.02	0.03	0.03	0.03	0.03	0.02	0.03

Table V.31: *KPI duration of the operating activities.*

	Average duration of the operating activities (min)											
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**
1	23.45	26.85	25.17	27.56	24.10	24.76	24.72	30.68	23.58	26.54	25.15	28.84
2	23.12	25.16	24.81	28.78	23.24	23.94	26.40	27.96	24.25	26.95	26.57	29.77
3	22.92	25.91	25.89	27.37	23.84	25.84	24.94	29.35	24.25	27.16	26.81	30.73
4	22.15	25.93	26.13	28.08	23.68	26.75	24.74	32.08	23.31	25.95	26.04	31.13
5	23.33	26.23	26.84	28.66	23.20	28.84	26.83	28.85	25.55	26.31	25.18	30.87
6	23.13	24.65	27.96	28.22	23.56	26.70	25.59	31.36	24.89	25.85	25.36	27.08
7	25.25	25.20	25.70	28.79	23.84	25.54	26.80	28.64	24.21	25.85	27.34	29.42
8	23.85	24.96	25.73	26.42	23.82	26.05	26.06	30.60	23.19	26.87	24.33	29.33
9	23.30	25.25	24.66	26.83	25.30	27.03	27.67	29.55	23.75	25.64	26.71	30.95
10	24.40	26.60	25.59	30.11	23.03	24.61	26.31	27.72	24.58	26.16	26.63	31.06
μ	***23.49***	***25.68***	***25.85***	***28.08***	***23.76***	***26.01***	***26.01***	***29.68***	***24.16***	***26.33***	***26.01***	***29.92***
σ	0.85	0.74	0.98	1.08	0.64	1.42	1.00	1.46	0.73	0.53	0.96	1.30
e	0.03	0.02	0.03	0.03	0.02	0.04	0.03	0.04	0.02	0.01	0.03	0.03

Table V.32: *KPI Human reliability.*

	Average human reliability											
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**
1	0.903	0.764	0.786	0.582	0.913	0.765	0.801	0.553	0.893	0.739	0.795	0.512
2	0.929	0.792	0.814	0.577	0.921	0.775	0.802	0.601	0.882	0.744	0.765	0.567
3	0.931	0.810	0.827	0.615	0.904	0.728	0.823	0.570	0.884	0.743	0.736	0.572
4	0.899	0.765	0.822	0.615	0.911	0.761	0.774	0.537	0.877	0.763	0.761	0.514
5	0.927	0.780	0.818	0.557	0.925	0.744	0.783	0.547	0.896	0.726	0.756	0.526
6	0.934	0.788	0.808	0.620	0.906	0.760	0.770	0.543	0.907	0.742	0.781	0.585
7	0.896	0.799	0.804	0.580	0.898	0.808	0.761	0.580	0.855	0.766	0.756	0.520
8	0.935	0.796	0.805	0.644	0.877	0.740	0.758	0.593	0.898	0.745	0.779	0.578
9	0.919	0.787	0.840	0.633	0.921	0.696	0.750	0.554	0.907	0.755	0.767	0.533
10	0.914	0.769	0.798	0.600	0.916	0.791	0.791	0.613	0.866	0.768	0.755	0.537
μ	***0.919***	***0.785***	***0.812***	***0.602***	***0.909***	***0.757***	***0.781***	***0.569***	***0.887***	***0.749***	***0.765***	***0.544***
σ	0.015	0.015	0.015	0.028	0.014	0.032	0.023	0.027	0.017	0.014	0.017	0.028
e	0.01	0.01	0.01	0.03	0.01	0.03	0.02	0.03	0.01	0.01	0.02	0.04

The statistical analysis of the runs is performed using the software R. The ANOVA results are collected in Figure V.21, Figure V.22 and Figure V.23.

```
                                  Df  Sum Sq Mean Sq  F value    Pr(>F)
Livello_MS                         2  180.72   90.36  11.2712 3.583e-05 ***
Livello_FD                         1 1412.57 1412.57 176.1972 < 2.2e-16 ***
Livello_CO                         1 2078.94 2078.94 259.3165 < 2.2e-16 ***
Livello_MS:Livello_FD              2    0.33    0.16   0.0204   0.97981
Livello_MS:Livello_CO              2   30.75   15.38   1.9180   0.15187
Livello_FD:Livello_CO              1   53.44   53.44   6.6662   0.01117 *
Livello_MS:Livello_FD:Livello_CO   2    3.94    1.97   0.2460   0.78239
Residuals                        108  865.83    8.02
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.21: *ANOVA Results (KPI flow time).*

```
                                  Df  Sum Sq Mean Sq  F value  Pr(>F)
Livello_MS                         2  14.563   7.281   7.0713 0.00130 **
Livello_FD                         1 216.790 216.790 210.5355 < 2e-16 ***
Livello_CO                         1 224.678 224.678 218.1956 < 2e-16 ***
Livello_MS:Livello_FD              2   1.679   0.839   0.8151 0.44531
Livello_MS:Livello_CO              2   4.178   2.089   2.0290 0.13645
Livello_FD:Livello_CO              1   8.603   8.603   8.3551 0.00465 **
Livello_MS:Livello_FD:Livello_CO   2   4.043   2.021   1.9632 0.14539
Residuals                        108 111.208   1.030
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.22: *ANOVA Results (KPI operating time).*

```
                                 Df  Sum Sq Mean Sq   F value    Pr(>F)
Livello_MS                        2 0.03791 0.01895   41.6445 3.923e-14 ***
Livello_FD                        1 0.72237 0.72237 1587.1589 < 2.2e-16 ***
Livello_CO                        1 0.94722 0.94722 2081.2001 < 2.2e-16 ***
Livello_MS:Livello_FD             2 0.00179 0.00090    1.9690    0.1446
Livello_MS:Livello_CO             2 0.00059 0.00029    0.6450    0.5267
Livello_FD:Livello_CO             1 0.04013 0.04013   88.1804 1.134e-15 ***
Livello_MS:Livello_FD:Livello_CO  2 0.00073 0.00036    0.8006    0.4517
Residuals                       108 0.04915 0.00046
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.23: *ANOVA Results (KPI human reliability).*

The asterisks indicate the factors that strongly influence the three KPIs. Furthermore, important interferences between the factors are detected (the effects often combine, overlap and amplify), especially between the levels of the PSF Fitness for duty and Communication. The Duncan test (Table V.33) allows verifying with greater precision what are the effects of PSF on KPIs.

Table V.33: *Results of the Duncan test for the first campaign.*

KPI total flow time (min).

Groups	Alternatives in increasing order of performance											
	A01	**A05**	**A09**	**A03**	**A06**	**A07**	**A02**	**A11**	**A10**	**A04**	**A08**	**A12**
μ	57.92	60.56	61.46	64.05	65.82	65.84	66.22	66.63	68.87	74	74.37	77.13
G1												
G2												
G3												
G4												
G5												
G6												

KPI operating time (min).

Groups	Alternatives in increasing order of performance											
	A01	**A05**	**A09**	**A02**	**A03**	**A06**	**A07**	**A11**	**A10**	**A04**	**A08**	**A12**
μ	23.49	23.76	24.16	25.68	25.85	26.01	26.01	26.01	26.33	28.08	29.68	29.92
G1												
G2												
G3												
G4												

KPI Reliability.

Groups	Alternatives in increasing order of performance											
	A01	**A05**	**A09**	**A03**	**A02**	**A07**	**A11**	**A06**	**A10**	**A04**	**A08**	**A12**
μ	0.919	0.909	0.887	0.812	0.785	0.781	0.765	0.757	0.749	0.602	0.569	0.544
G1												
G2												
G3												
G4												
G5												
G6												
G7												
G8												
G9												

The alternative A01 is always the most performing with respect to each of the KPIs, while the alternative A12 is always the worst. The hypothesis for which the experimental campaign was carried out is verified. In fact, the alternative A01 is the one that owns all PSFs with the best level from the reliability point of view, on the contrary for the alternative A12 these factors assume the worst level.

The first case allows to obtain the highest reliability value and, consequently, the lower times (both of total flow time and duration of the operating activities) given the lower number of errors committed; vice versa in the second case, the high number of errors committed (low reliability) is reflected in an increase in the times of each activity. The results of the Duncan test are also confirmed by other data collected, in particular on the average percentage of errors that could generate complications on the total interventions for each single run. This information was obtained by recording for each run of each of the alternatives the ratio between the number of interventions with possible future complications and the number of total interventions and averaging them (Table V.34).

Table V.34: *Average incidence of interventions with possible complications.*

A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	A11	A12
35.7%	56.3%	60.0%	76.5%	33.6%	61.7%	52.3%	83.8%	27.1%	63.0%	56.2%	77.7%

V.3.5.2 Analysis of the results for the second experimental campaign

Table V.35, Table V.36 and Table V.37 show the results obtained in terms of average KPI of interest in the second experimental campaign. All the relative standard errors are lower than the pre-established threshold of 0.1, and for this reason the 10 replicas made for each alternative are sufficient to proceed with the subsequent analyses.

The ANOVA results (Table V.35, Table V.36, Table V.37) shows that the factors interfere individually with a certain significance on the chosen KPIs, while the combined effect is negligible. The Duncan test (Table V.38) allows us to understand, although there is a strong overlap between the groups, that there is a significant difference from the reliability point of view between the As-Is system (A01) and the set of alternatives A04, A05, A07, A08. The reliability values for the latter models are higher than the alternative A01: this is also evident by observing the duration of the only operating activities for which the alternative A01 has the worst performance, given the highest number of errors that generate waste of time for their recovery. Within the group of more reliable alternatives than the As-Is, the discriminant cannot be the duration of the operating activities, so there are no significant differences according to the Duncan test (the four alternatives all belong to the G4 group).

Table V.35: *KPI total flow time.*

	Average total flow time (min)								
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**
1	60.88	56.12	55.96	64.04	68.47	64.08	70.96	69.08	68.74
2	57.99	59.45	57.04	65.28	66.48	61.59	74.26	74.67	69.00
3	58.14	56.56	56.23	68.11	64.26	59.06	71.07	68.46	67.01
4	63.21	61.60	54.63	65.61	68.50	66.61	71.33	68.91	68.14
5	57.18	58.47	54.31	64.45	65.73	61.67	70.60	66.17	65.13
6	60.48	57.34	59.14	62.51	68.58	65.96	75.05	68.50	61.11
7	62.12	56.32	60.33	66.97	67.70	62.86	73.26	72.91	64.65
8	58.02	62.55	55.46	63.70	64.86	65.87	72.16	72.07	59.73
9	58.92	56.04	59.11	64.59	64.92	63.27	71.06	63.44	68.64
10	60.15	63.64	51.02	67.23	71.34	58.59	72.95	71.31	70.86
μ	***59.71***	***58.81***	***56.32***	***65.25***	***67.08***	***62.69***	***72.72***	***69.55***	***66.30***
σ	1.99	2.87	2.75	1.75	2.22	2.79	1.55	3.31	3.61
e	0.02	0.03	0.03	0.02	0.02	0.03	0.02	0.03	0.04

Table V.36: *KPI duration of the operating activities.*

	Average duration of the operating activities (min)								
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**
1	24.51	22.54	21.84	22.48	23.33	21.36	22.15	22.79	20.79
2	23.57	23.03	23.03	23.38	21.76	21.93	23.74	21.85	21.91
3	22.69	22.16	23.45	23.09	22.42	21.42	22.19	22.25	21.43
4	23.77	22.52	22.46	22.08	22.31	23.07	22.33	22.59	22.20
5	24.43	22.64	22.06	22.83	22.53	21.22	23.68	23.06	20.55
6	24.70	22.87	22.21	22.85	21.94	22.50	23.82	22.42	21.60
7	23.49	23.17	22.29	23.71	22.33	21.07	23.40	23.18	21.45
8	23.67	23.02	22.55	23.66	22.86	22.45	22.98	24.23	21.04
9	23.07	23.21	21.94	22.43	23.84	21.65	24.39	22.01	22.12
10	24.81	23.58	21.65	23.46	24.52	22.16	24.81	22.64	22.40
μ	***23.87***	***22.87***	***22.35***	***23.00***	***22.78***	***21.88***	***23.35***	***22.70***	***21.55***
σ	0.71	0.42	.55	0.56	0.87	0.65	0.92	0.68	0.62
e	0.02	0.01	0.02	0.02	0.03	0.02	0.03	0.02	0.02

Table V.37: *KPI Human reliability.*

	Average human reliability								
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**
1	0.891	0.919	0.909	0.942	0.917	0.929	0.945	0.926	0.938
2	0.916	0.907	0.858	0.934	0.942	0.920	0.926	0.926	0.902
3	0.926	0.928	0.897	0.928	0.927	0.911	0.923	0.942	0.917
4	0.926	0.911	0.868	0.950	0.927	0.900	0.928	0.929	0.902
5	0.890	0.893	0.889	0.933	0.933	0.923	0.935	0.917	0.949
6	0.902	0.910	0.889	0.929	0.982	0.915	0.936	0.937	0.941
7	0.919	0.912	0.892	0.925	0.926	0.898	0.919	0.928	0.919
8	0.907	0.897	0.869	0.914	0.933	0.932	0.916	0.953	0.897
9	0.918	0.911	0.879	0.937	0.922	0.919	0.934	0.943	0.928
10	0.915	0.898	0.889	0.922	0.949	0.913	0.923	0.942	0.900
μ	***0.911***	***0.909***	***0.884***	***0.931***	***0.936***	***0.916***	***0.929***	***0.934***	***0.919***
σ	0.013	0.011	0.015	0.010	0.019	0.011	0.009	0.011	0.019
e	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

The worst performance for A07 and A08 make it possible to understand that the 15-minute pause, while increasing reliability, turns out to be excessive. Observing the levels assumed by the factors for the alternatives

A01, A04 and A05 it is possible to state that the 10 minute pause is recovered partly thanks to the increase in reliability, without the inclusion of a maximum time going to affect too much on the result: pausing allows offering a better service, in terms of patient safety, also gaining in terms of duration of operative activities and without losing excessively as regards the total crossing time with respect to the As-Is system.

```
                        Df  Sum Sq Mean Sq  F value    Pr(>F)
Dur_Pause                2 1878.59  939.30 136.5154 < 2.2e-16 ***
Livello_PT               2  262.28  131.14  19.0599 1.645e-07 ***
Dur_Pause:Livello_PT     4   63.24   15.81   2.2979   0.06597 .
Residuals               81  557.32    6.88
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.24: *ANOVA Results (KPI flow time).*

```
                        Df Sum Sq Mean Sq F value    Pr(>F)
Dur_Pause                2  4.756  2.3781  5.1092  0.008134 **
Livello_PT               2 33.126 16.5628 35.5838 8.121e-12 ***
Dur_Pause:Livello_PT     4  2.476  0.6190  1.3298  0.265955
Residuals               81 37.702  0.4655
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.25: *ANOVA Results (operating activities time).*

```
                        Df    Sum Sq   Mean Sq F value    Pr(>F)
Dur_Pause                2 0.0139236 0.0069618 37.8410 2.485e-12 ***
Livello_PT               2 0.0069711 0.0034855 18.9457 1.778e-07 ***
Dur_Pause:Livello_PT     4 0.0008359 0.0002090  1.1359    0.3456
Residuals               81 0.0149020 0.0001840
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure V.26: *ANOVA Results (KPI human reliability).*

Table V.38: *Results of the Duncan test for the second campaign.*

KPI total flow time (min).

Groups	Alternatives in increasing order of performance								
	A03	**A02**	**A01**	**A06**	**A04**	**A09**	**A05**	**A081**	**A07**
μ	56.32	58.81	59.71	62.96	65.25	66.30	67.08	69.55	72.27
G1									
G2									
G3									
G4									

G5									
G6									

KPI operating time (min).

Groups	**Alternatives in increasing order of performance**								
	A09	**A06**	**A03**	**A08**	**A05**	**A02**	**A04**	**A07**	**A01**
μ	21.55	21.88	22.35	22.70	22.78	22.87	23.00	23.35	23.87
G1									
G2									
G3									
G4									
G5									

KPI Reliability.

Groups	**Alternatives in increasing order of performance**								
	A05	**A08**	**A04**	**A07**	**A09**	**A06**	**A01**	**A02**	**A03**
μ	0.936	0.934	0.931	0.929	0.919	0.916	0.911	0.901	0.884
G1									
G2									
G3									
G4									
G5									

V.4 Results discussion

Experimental campaigns and case study underline the major SHERPA features, described in the theoretical model. In particular, its versatility is useful in revealing the environmental and psycho-physical factors which mainly influence the human reliability and may therefore be subject to improvement in order to reduce errors.

Unlike many HRA methods, SHERPA has been implemented for covering a wide range of working task, for this reason the six modelled Generic Task may represent activities that are more or less reliable. However, the most interesting aspect of the SHERPA model, however, is its ability to simulate several environmental conditions for the same performed activity. The most influential factors (experience, procedures and cognitive ergonomics) have a very tight confidence interval, a sign of their strong impact in the calculation of the error probability.

The analysed results in experimental campaign 4 and in the case study provide a wide overview of the potentialities of SHERPA simulator in the analysis and evaluation of the work-rest policies, which are influenced by several intrinsic system factors such as environmental and individual factors as well as the economic value of the process carried out by the worker. Despite the best breaks configuration for the system varies as a function of several factors and it cannot always be generalized in advance to different working environments, the experiment provides the following results:

- The increase of the total length of the breaks always improves the operator reliability while the economic performance is the result of a trade-off between cost classes with opposed trends. In fact, HR higher values are reflected on the machining quality with lower costs for scrap and reworking, but longer breaks reduce the time worked with an increase of the break costs and a possible decrease in revenues.
- The increase of the number of breaks maintaining fixed the total length has a positive impact on the worker reliability and involves higher profits because the worked hours do not change. This is true especially when the recovery rates are medium or fast, while for the slow rate an excessive fractioning of the rest time limits the reaching of a satisfactory recovery.

These results cannot be easily compared with the literature due to the presence of several criteria of human performance modelling, as seen previously. However, it is evident that one or more breaks in the shift provide a higher level of human performance, especially with short breaks. The proposed recovery modelling, based on an HRA approach, is therefore in agreement with the existing literature, even if it analyses this issue from a different point of view. The obtained results highlight the importance of measuring and evaluating both human reliability and work rate in the break scheduling problems, because of their significant economic and qualitative impact on the system performance. As well as the choices of the optimal work-rest policy cannot be separated from economic evaluations in terms of profits, considering the cost of lost production due to break and the quality costs related to operator errors.

Furthermore, LFCM module allows simulating a large number of scenarios without being resource intensive or time consuming in order to evaluate the human performance under the learning and forgetting phenomenon.

Conclusions

Human reliability is a highly relevant factor with a considerable impact on the overall performance of human-intensive working systems. Human error in the workplace, in fact, can have more or less serious consequences, such as accidents, malfunctions and defects in the quality of the performed task. The evidence that human actions are a source of vulnerability for industrial systems has led to the birth of many Human Reliability Analysis (HRA) methods, which aim at further examination of the human factor, but they have not always been especially useful for this purpose.

The SHERPA model proposed in this thesis has as its main objective the development of a model for quantifying human error probability in any work situation and in every context – quantification that today is hardly possible given the lack of tools similar to that achieved in this work. SHERPA model can be effectively used to evaluate changes in human error probability when changes occur in type of activity, contextual conditions, time spent at work and breaks assigned during the shift. The main advantage of the model lies in its being generic – it is suitable for any environment and working conditions, without limitations related to a particular sector or activity. The ability to change the values of the multipliers makes it easy to modify the weight of each factor PSF, regardless of the values assigned by the SPAR-H method. Through SHERPA, the concept of human reliability, often dealt with only in theory, is taken up in terms of production capacity or quality index (compliant and non-compliant items or retrieved items), and useful information about human reliability can be obtained for every kind of working context.

The break scheduling problems emerge in human-intensive working contexts where rest periods are indispensable due to features of the tasks to be performed, but despite the impact and the importance of breaks, these are not taken into proper consideration and there are ongoing efforts to develop models for optimal shift scheduling with multiple rest breaks. The SHERPA model efficiently evaluates the impacts of the work-rest policies on the HEP and on the economic system performance. It represents a decision support system for the break scheduling problem that quickly compares different break configurations, changing number, duration, and placement of rest breaks over the work shift, according to whatever sets of constraints imposed by legislation or by internal union agreements for the system under consideration. The management of breaks provided by the module allows simulation of all distributions of unintended breaks and evaluation of its effect on both the percentage of non-compliance and the economic return. In this way, different scheduling of breaks can be tested and compared, rapidly and with limited costs, in order to choose the best solution for the particular domain of work. This management can be difficult to apply in some cases but allows you to have an idea of the best condition to be used, which can be tested as fixed breaks with SHERPA. As evident from the experimental campaigns and case study, there are many factors that impact the system performance and the results are heavily influenced by the selected work-rest policies. The results obtained have led to the first considerations about the impact of different work-rest policies on the HR levels, but do not reach a univocal economic generalization because of the strong dependence between the value of the process performed by the operator and profit obtained. In any case, SHERPA results represent many different scenarios and discriminate between the different solutions identifying which ones are the more promising.

The proposed model was not created for a particular industry or application and therefore can be easily applied to contexts that vary widely. For example, the module can equally represent manual maintenance activity, manual assembly tasks, medical task in a surgery room etc., by varying the input variables such as performed task, level of contextual factors, or physical and mental employee condition and by modelling the specific system considering all the working context features. Simulators and tools similar to the one proposed do not exist today, either from the theoretical point of view or from the point of view of the analysis carried out.

The limitations of the current research underscore several issues worthy of additional studies. Many constraints on break scheduling management were relaxed in this first version, given the simulative nature of SHERPA model, as for example the maximal working time without breaks or the minimum and maximum possible break time. Future research should address the integration of these constraints in the simulation model. Furthermore, SHERPA requires additional tests for the validation and the calibration of HRA coefficients, as for example the impact of contextual and individual factors on human performance.

References

Aalipour, M., Ayele, Y. Z. and Barabadi, A. (2016) ‘Human reliability assessment (HRA) in maintenance of production process: a case study’, *Interntional Journal of System Assurance Engineering and Management*, 7(2), pp. 229–238.

Abbassi, R. *et al.* (2015) ‘An integrated method for human error probability assessment during the maintenance of offshore facilities’, *Process Safety and Environmental Protection*, 94(C), pp. 172–179.

Achebo, J. I. and Oghoore, O. (2010) ‘A nonparametric analysis of asymptomatic hazard rates in a brewing plant using the probability failure functions’, in *WCE 2010 - World Congress on Engineering 2010*, pp. 2381–2384.

Åkerstedt, T. (2003) ‘Shift work and disturbed sleep/wakefulness’, *Occupational Medicine*, 53(2), pp. 89–94. doi: 10.1093/occmed/kqg046.

De Ambroggi, M. and Trucco, P. (2011) ‘Modelling and assessment of dependent performance shaping factors through Analytic Network Process’, *Reliability Engineering and System Safety*. Elsevier, 96(7), pp. 849–860. doi: 10.1016/j.ress.2011.03.004.

Anderson, G. F. and Hussey, P. S. (2000) ‘Population aging: A comparison among industrialized countries’, *Health Affairs*, 19(3), pp. 191–203. doi: 10.1377/hlthaff.19.3.191.

Andriulo, S. and Gnoni, M. G. (2014) ‘Measuring the effectiveness of a near-miss management system: An application in an automotive firm supplier’, *Reliability Engineering and System Safety*, 132, pp. 154–162. doi: 10.1016/j.ress.2014.07.022.

Anzanello, M. J. and Fogliatto, F. S. (2011) ‘Learning curve models and applications: Literature review and research directions’, *International Journal of Industrial Ergonomics*. Elsevier Ltd, 41(5), pp. 573–583. doi: 10.1016/j.ergon.2011.05.001.

Asadzadeh, S. M. and Azadeh, A. (2014) ‘An integrated systemic model for optimization of condition-based maintenance with human error’, *Reliability Engineering and System Safety*, 124, pp. 117–131.

Aykin, T. (1996) ‘Optimal Multiple Shift Scheduling Break with Windows’, *Management Science*, 42(4), pp. 591–602. doi: 10.1287/mnsc.42.4.591.

Azadeh, A., Asadzadeh, S. M. and Seif, J. (2014) ‘An integrated simulation-analysis of variance methodology for effective analysis of CBM alternatives’, *International Journal of Computer Integrated Manufacturing*, 27(7), pp. 624–637. doi: 10.1080/0951192X.2013.834466.

Azizi, N., Zolfaghari, S. and Liang, M. (2010) ‘Modeling job rotation in manufacturing systems: The study of employee’s boredom and skill variations’, *International Journal of Production Economics*. Elsevier, 123(1), pp. 69–85. doi: 10.1016/j.ijpe.2009.07.010.

Backes-Gellner, U., Schneider, M. R. and Veen, S. (2011) ‘Effect of workforce age on quantitative and qualitative organizational performance: Conceptual framework and case study evidence’, *Organization Studies*, 32(8), pp. 1103–1121. doi: 10.1177/0170840611416746.

Baez, Y. A. *et al.* (2014) ‘Model of human reliability for manual workers in assembly lines’, in *IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 1448–1452. doi: 10.1109/IEEM.2014.7058878.

Balci, R. and Aghazadeh, F. (2004) ‘Effects of exercise breaks on performance, muscular load, and perceived discomfort in data entry and cognitive tasks’, *Computers and Industrial Engineering*, 46(3), pp. 399–411. doi: 10.1016/j.cie.2004.01.003.

Bao, M. M. and Ding, S. S. (2014) ‘Individual-related factors and Management-related factors in Aviation Maintenance’, in Gizardin, M and Zhang, S and Ding, S. (ed.) *3RD International Symposium on Aircraft Airworthiness (ISAA 2013)*. (Procedia Engineering), pp. 293–302.

Bao, Y. Y. *et al.* (2015) ‘Impact of Human Error on Electrical Equipment Preventive Maintenance Policy’, in *2015 IEEE Power & Energy Society General Meeting*. (IEEE Power and Energy Society General Meeting PESGM).

Battini, D. *et al.* (2017) ‘Ergonomics and human factors in waste collection : analysis and suggestions for the door-to-door method’.

Bechtold, S. E., Janaro, R. E. and Sumners, D. W. L. (1984) ‘Maximization of Labor Productivity Through Optimal Rest-Break Schedules’, *Management Science*, 30(12), pp. 1442–1458. doi: 10.1287/mnsc.30.12.1442.

Bechtold, S. E. and Thompson, G. M. (1993) 'Optimal scheduling of a flexible-duration rest period for a work group.', *Operations Research*, 41(6), p. 1046. doi: 10.1287/opre.41.6.1046.

Bell, J. and Holroyd, J. (2009) 'Review of human reliability assessment methods', *Health & Safety Laboratory*.

Bellamy, L. J. (1991) 'The quantification of human fallibility', *Journal of Health and Safety*, 6, pp. 13–22.

Bieder, C. and Le Bot, P. (1998) 'MERMOS: EDF's new advanced HRA method.', in Springer-Verlag (ed.) *Probabilistic safety assessment and management. PSAM 4*. New York City, NY, United states.

Blackman, H. S., Gertman, D. I. and Boring, R. L. (2008) 'Human Error Quantification Using Performance Shaping Factors in the SPAR-H Method', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52, pp. 1733–1737. doi: 10.1177/154193120805202109.

Boenzi, F., Digiesi, S., *et al.* (2015) 'Modelling workforce aging in job rotation problems', *IFAC-PapersOnLine*, 28(3), pp. 604–609. doi: 10.1016/j.ifacol.2015.06.148.

Boenzi, F., Mossa, G., *et al.* (2015) 'Workforce aging in production systems: Modeling and performance evaluation', in *Procedia Engineering*. Elsevier B.V., pp. 1108–1115. doi: 10.1016/j.proeng.2015.01.473.

Boring, R. L. (2006) *Modeling Human Reliability Analysis Using MIDAS International Workshop on Future Control Station Designs and Human*.

Boring, R. L. (2007) 'Dynamic human reliability analysis: Benefits and challenges of simulating human performance', *European Safety and Reliability Conference*, pp. 1043–1050.

Boring, R. L. (2010) 'How Many Performance Shaping Factors are Necessary for Human Reliability Analysis ?', in *Proceedings of the 10th International Probabilistic Safety Assessment and Management Conference.*

Boring, R. L. (2015) 'Adapting Human Reliability Analysis from nuclear power to oil and gas applications', (October), pp. 2853–2860.

Boring, R. L. and Blackman, H. S. (2007) 'The origins of the SPAR-H method's performance shaping factor multipliers', *IEEE Conference on Human Factors and Power Plants*, pp. 177–184. doi: 10.1109/HFPP.2007.4413202.

Boring, R. L., Griffith, C. D. and Joe, J. C. (2007) 'The measure of human error: Direct and indirect performance shaping factors', *IEEE Conference on Human Factors and Power Plants*, pp. 170–176. doi: 10.1109/HFPP.2007.4413201.

Börsch-Supan, A. and Weiss, M. (2016) 'Productivity and age: Evidence from work teams at the assembly line', *Journal of the Economics of Ageing*, 7, pp. 30–42. doi: 10.1016/j.jeoa.2015.12.001.

Bozkurt, Y. and Kavsaoglu, M. S. (2010) 'MEDA case study for an MRO', in *10th AIAA Aviation Technology, Integration and Operations Conference 2010, ATIO 2010*.

Brewer, S. *et al.* (2006) 'Workplace interventions to prevent musculoskeletal and visual symptoms and disorders among computer users: A systematic review', *Journal of Occupational Rehabilitation*, 16(3), pp. 325–358. doi: 10.1007/s10926-006-9031-6.

Bubb, H. (2005) 'Human reliability: A key to improved quality in manufacturing', *Human Factors and Ergonomics In Manufacturing*, 15(4), pp. 353–368. doi: 10.1002/hfm.20032.

Bye, A. *et al.* (2006) 'Methodology for improving HRA by simulator studies (PSAM-0391).', in *Proceedings of the eighth international conference on probabilistic safety assessment and management (PSAM).* ASME Press 2006.

Cacciabue, P. C. *et al.* (1992) 'COSIMO: A Cognitive Simulation Model of Human Decision Making and Behavior in Accident Management of Complex Plants', *IEEE Transactions on Systems, Man and Cybernetics*, 22(5), pp. 1058–1074. doi: 10.1109/21.179844.

Cacciabue, P. C. (1998) 'Modelling and Simulation of Human Behaviour for Safety Analysis and Control of Complex Systems', *Pergamon Safety Science*, 28(2), pp. 97–110. doi: 10.1016/S0925-7535(97)00079-9.

Cacciabue, P. C. (2004) 'Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training', *Reliability Engineering & System Safety*, 83(2), pp. 229–240. doi: 10.1016/j.ress.2003.09.013.

Calixto, E., Lima, G. B. A. and Firmino, P. R. A. (2013) 'Comparing SLIM, SPAR-H and Bayesian Network Methodologies', *Open Journal of Safety Science and Technology*, 3(2), pp. 31–41. doi: 10.4236/ojsst.2013.32004.

Carr, M. J. and Christer, A. H. (2003) 'Incorporating the potential for human error in maintenance models', *Journal of the Operational Research Society*, 54(12), pp. 1249–1253.

Carrier, J. and Monk, T. H. (2000) 'Circadian rhythms of performance: new trends', *Chronobiol Int*, 17(6), pp. 719–732. doi: 10.1081/CBI-100102108.

Caruso, C. C. (2014) ‘Negative impacts of shiftwork and long work hours’, *Rehabilitation Nursing*, 39(1), pp. 16–25. doi: 10.1002/rnj.107.

Castiglia, F. and Giardina, M. (2013) ‘Analysis of operator human errors in hydrogen refuelling stations: Comparison between human rate assessment techniques’, *International Journal of Hydrogen Energy*, 38(2), pp. 1166–1176.

Chan, G., Tan, V. and Koh, D. (2000) ‘Ageing and fitness to work’, *Occupational Medicine*, 50(7), pp. 483–491. doi: 10.1093/occmed/50.7.483.

Chang, Y. H. J. and Mosleh, A. (2007a) ‘Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 1: Overview of the IDAC model’, *Reliability Engineering and System Safety*, 92(8), pp. 997–1013. doi: 10.1016/j.ress.2006.05.014.

Chang, Y. H. J. and Mosleh, A. (2007b) ‘Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model’, *Reliability Engineering and System Safety*, 92(8), pp. 1014–1040. doi: 10.1016/j.ress.2006.05.010.

Chang, Y. H. J. and Mosleh, A. (2007c) ‘Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 3: IDAC operator response model’, *Reliability Engineering and System Safety*, 92(8), pp. 1041–1060. doi: 10.1016/j.ress.2006.05.013.

Chang, Y. H. J. and Mosleh, A. (2007d) ‘Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 4: IDAC causal model of operator problem-solving response’, *Reliability Engineering and System Safety*, 92(8), pp. 1061–1075. doi: 10.1016/j.ress.2006.05.011.

Chang, Y. H. J. and Mosleh, A. (2007e) ‘Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 5: Dynamic probabilistic simulation of the IDAC model’, *Reliability Engineering and System Safety*, 92(8), pp. 1076–1101. doi: 10.1016/j.ress.2006.05.012.

Chen, W. and Huang, S. (2014) ‘Human Reliability Analysis for Visual Inspection in Aviation Maintenance by a Bayesian Network Approach’, *Transportation Research Record*, 2449(2449), pp. 105–113.

Chen, W. and Huang, S. P. (2013) ‘Human reliability analysis in aviation maintenance by a bayesian network approach’, in *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures - Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013*, pp. 2091–2096.

Chiodo, E., Gagliardi, F. and Pagano, M. (2004) ‘Human reliability analyses by random hazard rate approach’, *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 23(1), pp. 65–78.

Chiu, M. C. and Hsieh, M. C. (2016) ‘Latent human error analysis and efficient improvement strategies by fuzzy TOPSIS in aviation maintenance tasks’, *Applied Ergonomics*, 54, pp. 136–147.

Chu, T., Musicki, Z. and Others (1994) ‘Evaluation of Potential Severe Accidents During Low Power and Shutdown Operations at Surry, Unit 1’, *Nureg/Cr-6144*, 1.

Claeys, A. *et al.* (2015) ‘Framework for evaluating cognitive support in mixed model assembly systems’, *IFAC-PapersOnLine*. Elsevier Ltd., 28(3), pp. 924–929. doi: 10.1016/j.ifacol.2015.06.201.

Cooper, S. E. *et al.* (1996) *A Technique for Human Error Analysis (ATHEANA) – Technical Basis and Methodology Description.*

Cox, A., Dolan, L. and Macewen, C. J. (2008) ‘Human reliability analysis: a new method to quantify errors in cataract surgery.’, *Eye (London, England)*, 22(3), pp. 394–397. doi: 10.1038/sj.eye.6702648.

Crawford, J. O. *et al.* (2010) ‘The health safety and health promotion needs of older workers’, *Occupational Medicine*, 60(3), pp. 184–192. doi: 10.1093/occmed/kqq028.

Cucumo, M. *et al.* (1996) ‘Stato dell’arte nel calcolo della luminanza del cielo e dell’illuminameno naturale di superfici esterne’, in *51° Congresso Associazione Termotecnica Italiana*.

Cuschieri, A. (2000) ‘Human reliability assessment in surgery - a new approach for improving surgical performance and clinical outcome’, *The Royal College of Surgeons of England*, 82, pp. 83–87.

Czaja, S., Nair, S. and Salvendy, G. (2012) ‘Human factors engineering and system design’, in *Handbook of Human Factors and Ergonomics, Fourth Edition*. John Wiley and Sons, pp. 38–45.

Dababneh, A., Swanson, N. and Shell, R. (2001) ‘Impact of added rest breaks on the productivity and well being of workers’, *Ergonomics*, 44(2), pp. 164–174.

Davidson, O. *et al.* (2010) ‘Sabbatical leave: who gains and how much?’, *Journal of Applied Psychology*, 95(5), p. 953.

Demerouti, E. *et al.* (2009) *Research in Occupational Stress and Well-being*, *Research in Occupational Stress and Well Being*. Elsevier. doi: 10.1108/S1479-3555(2009)0000007006.

Demerouti, E. *et al.* (2012) 'Work-related flow and energy at work and at home: A study on the role of daily recovery', *Journal of Organizational Behavior*, 33(2), pp. 276–295. doi: 10.1002/job.760.

Dhillon, B. S. (2009) *Human reliability and error in maintenance*, *Handbook of Maintenance Management and Engineering*.

Dhillon, B. S. (2014) 'Human error in maintenance: An investigative study for the factories of the future', in *IOP Conference Series: Materials Science and Engineering*.

Dhillon, B. S. and Kirmizi, F. (2003) 'Probabilistic safety analysis of maintainable systems', *Journal of Quality in Maintenance Engineering*, 9(3), pp. 303–320.

Dhillon, B. S. and Liu, Y. (2006) 'Human error in maintenance: A review', *Journal of Quality in Maintenance Engineering*, 12(1), pp. 21–36.

Dhillon, B. S. and Shah, A. S. (2007) 'Availability analysis of a generalized maintainable three-state device parallel system with human error and common-cause failures', *Journal of Quality in Maintenance Engineering*, 13(4), pp. 411–432.

Dollarhide, A. W. *et al.* (2014) 'A real-time assessment of factors influencing medication events', *Journal for healthcare quality : official publication of the National Association for Healthcare Quality*, 36(5), pp. 5–12.

Droguett, E. L. and Menêzes, R. (2007) 'Human reliability analysis through Bayesian networks: An application in maintenance of transmission lines [Análise da confiabilidade humana via redes Bayesianas: Uma aplicação à manutenção de linhas de transmissão]', *Producao*, 17(1), pp. 162–185.

Edwards, E. (1998) 'Introductory overview', in Wiener, E.L., Nagel, D. C. (Eds. . (ed.) *Human Factors in Aviation*. San Diego, CA, USA: Academic Press.

Eilon, S. (1964) 'On a mechanistic approach to fatigue and rest periods', *International Journal of Production Research*, 3(4), pp. 327–332.

Elmaraghy, W. H., Nada, O. A. and Elmaraghy, H. A. (2008) 'Quality prediction for reconfigurable manufacturing systems via human error modelling', *International Journal of Computer Integrated Manufacturing*, 21(5), pp. 584–598. doi: 10.1080/09511920701233464.

Emami-Mehrgani, B. B. *et al.* (2016) 'Considering human error in optimizing production and corrective and preventive maintenance policies for manufacturing systems', *Applied Mathematical Modelling*, 40(3), pp. 2056–2074.

Embrey, D. (1986) 'SLIM-MAUD : A Computer-Based Technique for Human R e l i a b i l i t y Assessment *', *International Journal of Quality and Reliability Management*, pp. 4–13.

Falcone, D. *et al.* (2010) 'Study and modelling of very flexible lines through simulation', in *Proceedings of the 2010 Spring Simulation Multiconference*. 11 – 15 April 2010, Orlando, Florida (USA), pp. 11–16.

Falcone, D. *et al.* (2011) 'Modeling and simulation of an assembly line: a new approach for assignment and optimization of activities of operators', in *MAS (The International Conference on Modeling and Applied Simulation)*. 12 –14 September 2011, Rome (Italy), pp. 12–14.

Folkard, S. and Lombardi, D. A. (2006) 'Modeling the impact of the components of long work hours on injuries and "accidents"', *American Journal of Industrial Medicine*, 49(11), pp. 953–963. doi: 10.1002/ajim.20307.

Folkard, S., Lombardi, D. a and Spencer, M. B. (2006) 'Estimating the circadian rhythm in the risk of occupational injuries and accidents.', *Chronobiology international*, 23(6), pp. 1181–92. doi: 10.1080/07420520601096443.

Folkard, S. and Monk, T. H. (1980) 'Circadian rhythms in human memory.', *British Journal of Psychology*, 71, pp. 295–307.

Folkard, S. and Rosen, S. D. (1990) 'Circadian performance rhythms: some practical and theoretical implications.', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 327(1241), pp. 543–533.

Folkard, S. and Tucker, P. (2003) 'Shift work, safety and productivity', *Occupational Medicine*, 53(2), pp. 95–101. doi: 10.1093/occmed/kqg047.

Fritz, C. and Sonnentag, S. (2005) 'Recovery, health, and job performance: Effects of weekend experiences.', *Journal of occupational health psychology*, 10(3), pp. 187–199. doi: 10.1037/1076-8998.10.3.187.

Fritz, C. and Sonnentag, S. (2006) 'Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences.', *The Journal of applied psychology*, 91(4), pp. 936–45. doi: 10.1037/0021-9010.91.4.936.

Fritzsche, L. *et al.* (2014) 'Good ergonomics and team diversity reduce absenteeism and errors in car manufacturing', *Ergonomics*, 57(2), pp. 148–161. doi: 10.1080/00140139.2013.875597.

Galinsky, T. L. *et al.* (2000) 'A field study of supplementary rest breaks for data-entry operators.', *Ergonomics*, 43(5), pp. 622–38. doi: 10.1080/001401300184297.

Di Gaspero, L. *et al.* (2010) 'A hybrid LS-CP solver for the shifts and breaks design problem', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6373 LNCS, pp. 46–61. doi: 10.1007/978-3-642-16054-7_4.

Geibel, W. D., Von Thaden, T. L. and Suzuki, T. (2008) 'Issues that precipitate errors in airline maintenance', in *Proceedings of the Human Factors and Ergonomics Society*, pp. 94–98.

Gelman, A. (2005) 'Analysis of variance—why it is more important than ever', *The annals of statistics*, 33(1), pp. 1–53.

Gentzler, G., Khalil, T. and Sivazlian, B. (1977) 'Quantitative models for optimal rest period scheduling', *Omega*, 5(2), pp. 215–220. doi: 10.1016/0305-0483(77)90104-9.

Gerdes, G. (1997) *Identification and analysis of cognitive errors: application to control room operators*.

Gertman, D. *et al.* (2005) 'The SPAR-H human reliability analysis method', in *American Nuclear Society 4th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technology*. Division of risk analysis and applications: Office of nuclear regulatory research, US Nuclear Regulatory Commission., pp. 17–24.

Gertman, D. I. *et al.* (1992) 'INTENT: a method for estimating human error probabilities for decisionbased errors', *Reliability Engineering & System Safety*, 35(2), pp. 127–136. doi: 10.1016/0951-8320(92)90032-G.

Gibson, G. P. (2000) 'Improving maintenance through reducing human error', in *Foresight and Precaution, Vols 1 and 2*, pp. 399–404.

Giuntini, R. E. (2000) 'Mathematical characterization of human reliability for multi-task system operations', *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, 2, pp. 1325–1329. doi: 10.1109/ICSMC.2000.886037.

Givi, Z. S. and Jaber, M. Y. (2014) 'Human error due to learning and fatigue', in *IIE Annual Conference and Expo 2014*, pp. 1933–1942.

Givi, Z. S., Jaber, M. Y. and Neumann, W. P. (2015) 'Modelling worker reliability with learning and fatigue', *Applied Mathematical Modelling*. Elsevier Inc., 39(17), pp. 5186–5199.

Givi, Z. S., Jaber, M. Y. and Neumann, W. P. (2015) 'Production planning in DRC systems considering worker performance', *Computers and Industrial Engineering*, 87, pp. 317–327. doi: 10.1016/j.cie.2015.05.005.

Grandjean, E. (1968) 'Fatigue: its physiological and psychological significance', *Ergonomics*, 11(5), pp. 427–436.

Griffith, C. D. and Mahadevan, S. (2011) 'Inclusion of fatigue effects in human reliability analysis', *Reliability Engineering and System Safety*. Elsevier, 96(11), pp. 1437–1447. doi: 10.1016/j.ress.2011.06.005.

Grosse, E. H., Glock, C. H. and Müller, S. (2015) 'Production economics and the learning curve: A meta-analysis', *International Journal of Production Economics*. Elsevier, 170, pp. 401–412. doi: 10.1016/j.ijpe.2015.06.021.

Groth, K. M. and Swiler, L. P. (2013) 'Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H', *Reliability Engineering and System Safety*. Elsevier, 115, pp. 33–42. doi: 10.1016/j.ress.2013.02.015.

Haji Hosseini, A. R. *et al.* (2012) 'Factors influencing human errors during work permit issuance by the electric power transmission network operators', *Indian Journal of Science and Technology*, 5(8), pp. 3169–3173.

Hameed, A., Khan, F. and Ahmed, S. (2016) 'A risk-based shutdown inspection and maintenance interval estimation considering human error', *Process Safety and Environmental Protection*, 100, pp. 9–21.

Hannaman, G., Spurgin, A. and Lukic, Y. (1984) *Human cognitive reliability model for PRA analysis.*

Harper, S. and Marcus, S. (2006) 'Age-Related Capacity Decline : A Review of some Workplace Implications', *Ageing Horizons*, (5), pp. 20–30.

Hayama, K. *et al.* (2011) 'Method to support risk management of human error in track maintenance work', *Quarterly Report of RTRI (Railway Technical Research Institute) (Japan)*, 52(3), pp. 168–173.

Henning, R. *et al.* (1997) 'Frequent short rest breaks from computer work: effects on productivity and well-being at two field sites', *Ergonomics*, 40(1), pp. 78–91.

Heo, G. and Park, J. (2010) 'A framework for evaluating the effects of maintenance-related human errors in nuclear power plants', *Reliability Engineering and System Safety*, 95(7), pp. 797–805.

van de Heuvel, S. G., de Looze, M. P. and Hildebrandt, V. H. (2003) 'Effects of software breaks and exercises on stimulating regular programs neck disorders and upper-limb', *Scandinavian Journal of Work, Environment and Health*, 29(2), pp. 106–116.

Hobbs, A. and Williamson, A. (2002) 'Skills, rules and knowledge in aircraft maintenance: Errors in context', *Ergonomics*, 45(4), pp. 290–308.

Hobbs, A. and Williamson, A. (2003) 'Associations between errors and contributing factors in aircraft maintenance', *Human Factors*, 45(2), pp. 186–201.

Hobbs, A., Williamson, A. and Van Dongen, H. P. A. (2010) 'A circadian rhythm in skill-based errors in aviation maintenance', *Chronobiology International*, 27(6), pp. 1304–1316.

Hochrein, S. and Glock, C. H. (2012) 'Systematic literature reviews in purchasing and supply management research: A tertiary study', *International Journal of Integrated Supply Management*, 7(4), pp. 215–245. doi: 10.1504/IJISM.2012.052773.

Hollnagel, E. (1993) *Human Reliability Analysis: Context and Control*. London, Engl: Academic.

Hollnagel, E. (1996) 'Reliability analysis and operator modelling', *Reliability Engineering and System Safety*, 52(3), pp. 327–337. doi: 10.1016/0951-8320(95)00143-3.

Hollnagel, E. (1998) 'Cognitive Reliability and Error Analysis Method (CREAM)', *Cognitive Reliability and Error Analysis Method (CREAM)*, pp. 1–21. doi: 10.1016/B978-008042848-2/50001-4.

Ilmarinen, J. (2001) 'Aging workers', *Occupational and environmental medicine*, 58(8), pp. 546–552. doi: 10.1136/oem.58.8.546.

Ilmarinen, J. (2006) 'The ageing workforce - Challenges for occupational health', *Occupational Medicine*, 56(6), pp. 362–364. doi: 10.1093/occmed/kql046.

Islam, R. *et al.* (2016) 'Determination of human error probabilities for the maintenance operations of marine engines', *Journal of Ship Production and Design*, 32(4), pp. 226–234.

Islam, R. *et al.* (2017) 'Development of a monograph for human error likelihood assessment in marine operations', *Safety Science*, 91, pp. 33–39.

Jaber, M. Y. and Bonney, M. (1996) 'Production breaks and the learning curve: the forgetting phenomenon', *Applied Mathematical Modelling*, 20(2), pp. 162–169. doi: 10.1016/0307-904X(95)00157-F.

Jaber, M. Y. and Bonney, M. (1997) 'A comparative study of learning curves with forgetting', *Applied Mathematical Modelling*, 21(8), pp. 523–531. doi: 10.1016/S0307-904X(97)00055-3.

Jaber, M. Y. and Glock, C. H. (2013) 'A learning curve for tasks with cognitive and motor elements', *Computers and Industrial Engineering*, 64(3), pp. 866–871. doi: 10.1016/j.cie.2012.12.005.

Jaber, M. Y. and Sikström, S. (2004) 'A numerical comparison of three potential learning and forgetting models', *International Journal of Production Economics*, 92(3), pp. 281–294. doi: 10.1016/j.ijpe.2003.10.019.

Jansen, N. W. H., Kant, I. and van den Brandt, P. A. (2002) 'Need for Recovery in the Working Population: Description and Associations With Fatigue and Psychological Distress', *International Journal of Behavioral Medicine*, 52(6), pp. 322–340.

Jett, Q. R. and George, J. M. (2003) 'Work interrupted: a Closer Look At the Role of Interruptions in Organizational Life', *Academy of Management Review*, 28(3), pp. 494–507. doi: 10.5465/AMR.2003.10196791.

Julius, J. *et al.* (1995) 'A Procedure for the Analysis of Errors of Commission in a Probabilistic Safety Assessment of a Nuclear Power Plant at Full Power', *Reliability Engineering & System Safety*, 50(2), pp. 189–201. doi: 10.1016/0951-8320(95)00075-D.

Kelton, W. D., Sadowsky, R. P. and Sadowsy, D. A. (2006) *Simulation with Arena*. 5th edn. McGtraw-Hill Professional.

Kennedy, C. A. *et al.* (2010) *Systematic review of the role of occupational health and safety interventions in the prevention of upper extremity musculoskeletal symptoms, signs, disorders, injuries, claims and lost time*, *Journal of Occupational Rehabilitation*. doi: 10.1007/s10926-009-9211-2.

Kenny, G. P. *et al.* (2016) 'Age, human performance, and physical employment standards', *Applied Physiology Nutrition and Metabolism*, 41(6, 2), pp. S92–S107. doi: 10.1139/apnm-2015-0483.

Kern, C. and Refflinghaus, R. (2011) 'Analysis of human reliability in manual assembly operations', in *21st International Conference on Production Research: Innovation in Product and Production, ICPR 2011 - Conference Proceedings*.

Kern, C. and Refflinghaus, R. (2013) 'Cross-disciplinary method for predicting and reducing human error probabilities in manual assembly operations', *Total Quality Management and Business Excellence*, 24(7–8), pp. 847–858.

Kern, C. and Refflinghaus, R. (2017) 'Modification of the HRA-method ESAT for improving human reliability in manual assembly', in *Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016*, p. 148.

Khalaquzzaman, M. *et al.* (2010) 'A model for estimation of reactor spurious shutdown rate considering maintenance human errors in reactor protection system of nuclear power plants', *Nuclear Engineering and Design*, 240(10), pp. 2963–2971.

Khalaquzzaman, M. *et al.* (2011) 'Optimization of periodic testing frequency of a reactor protection system based on a risk-cost model and public risk perception', *Nuclear Engineering and Design*, 241(5), pp. 1538–1547.

Khalaquzzaman, M. . *et al.* (2010) 'Quantification of unavailability caused by random failures and maintenance human errors in nuclear power plants', *Nuclear Engineering and Design*, 240(6), pp. 1606–1613. doi: 10.1016/j.nucengdes.2010.03.001.

Kim, J. N. (1997) 'The development of K-HPES: a Korean-version Human Performance Enhancement System', pp. 16–20.

Kim, J. and Park, J. (2008) 'Task types and error types involved in the human-related unplanned reactor trip events', *Nuclear Engineering and Technology*, 40(7), pp. 615–624.

Kim, J. and Park, J. (2009) 'Characteristics of test and maintenance human errors leading to unplanned reactor trips in nuclear power plants', *Nuclear Engineering and Design*, 239(11), pp. 2530–2536.

Kim, J. and Park, J. (2012) 'Reduction of test and maintenance human errors by analyzing task characteristics and work conditions', *Progress in Nuclear Energy*, 58, pp. 89–99.

Kim, J. W. and Jung, W. (2003) 'A taxonomy of performance influencing factors for human reliability analysis of emergency tasks', *Journal of Loss Prevention in the Process Industries*, 16(6), pp. 479–495. doi: 10.1016/S0950-4230(03)00075-5.

Kim, M. C., Seong, P. H. and Hollnagel, E. (2006) 'A probabilistic approach for determining the control mode in CREAM', *Reliability Engineering and System Safety*, 91(2), pp. 191–199. doi: 10.1016/j.ress.2004.12.003.

Kirwan, B. (1994) *A guide to practical human reliability assessment*. CRC Press.

Kirwan, B. (1996) 'The validation of three human reliability quantification techniques—THERP, HEART and JHEDI: Part I—technique descriptions and validation issues.', *Applied Ergonomics*, 27(6), pp. 359–373. doi: 10.1016/S0003-6870(96)00044-0.

Kirwan, B. (1997) 'The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI', *Reliability Engineering & System Safety*, 56(2), pp. 107–133. doi: 10.1016/S0951-8320(97)00006-9.

Kirwan, B. (1997) 'The validation of three human reliability quantification techniques--THERP, HEART and JHEDI: Part III--Practical aspects of the usage of the techniques.', *Applied ergonomics*, 28(1), pp. 27–39. doi: 10.1016/S0003-6870(96)00046-4.

Kirwan, B. *et al.* (2004) 'Nuclear Action Reliability Assessment (NARA):

A Data-Based HRA Tool', *Probabilistic Safety Assessment and Management*, 7353(September), pp. 1206–1211. doi: 10.1007/978-0-85729-410-4_195.

Kirwan, B., Basra, G. and Taylor-Adams, S. E. (1997) 'CORE-DATA: a computerised human error database for human\nreliability support', *Proceedings of the 1997 IEEE Sixth Conference on Human Factors and Power Plants, 1997. 'Global Perspectives of Human Factors in Power Generation'*, pp. 7–12. doi: 10.1109/HFPP.1997.624858.

Knutsson, A. (2003) 'Health disorders of shift workers', *Occupational Medicine*, 53(2), pp. 103–108. doi: 10.1093/occmed/kqg048.

Kolaczkowski, A. *et al.* (2005) *Good Practices for Implementing Human Reliability Analysis (HRA)*.

Kontogiannis, T. (1997) 'A framework for the analysis of cognitive reliability in complex systems: a recovery centred approach', *Reliability Engineering and System Safety*. UK, 58(3), pp. 233–48. doi: 10.1016/S0951-8320(97)00078-1.

Kovacevic, S. *et al.* (2016) 'The analysis of human error as causes in the maintenance of machines: a case study in mining companies', *South African Journal of Industrial Engineering*, 27(4), pp. 193–202.

Kumar, V. N. A., Gandhi, M. S. and Gandhi, O. P. (2015) 'Identification and Assessment of Factors Influencing Human Reliability in Maintenance Using Fuzzy Cognitive Maps', *Quality and Reliability Engineering International*, 31(2), pp. 169–181.

Kumar, V. N. A. and Gandhi, O. P. (2011) 'Quantification of Human Error in Maintenance Using Graph Theory and Matrix Approach', *Quality and Reliability Engineering International*, 27(8), pp. 1145–1172.

Latorella, K. A. and Prabhu, P. V (2000) 'A review of human error in aviation maintenance and inspection', *International Journal of Industrial Ergonomics*, 26(2), pp. 133–161.

Lawrence, P. and Gill, S. (2007) 'Human hazard analysis: A prototype method for human hazard analysis developed for the large commercial aircraft industry', *Disaster Prevention and Management*, 16(5), pp. 718–739.

Lee, S. W. *et al.* (2011) 'Development of a qualitative evaluation framework for performance shaping factors (PSFs) in advanced MCR HRA', *Annals of Nuclear Energy*. Elsevier Ltd, 38(8), pp. 1751–1759. doi: 10.1016/j.anucene.2011.04.006.

Leva, M. C. *et al.* (2009) 'Quantitative analysis of ATM safety issues using retrospective accident data: The dynamic risk modelling project', *Safety Science*. Elsevier Ltd, 47(2), pp. 250–264. doi: 10.1016/j.ssci.2008.04.002.

Liang, G. F. *et al.* (2010) 'Preventing human errors in aviation maintenance using an on-line maintenance assistance platform', *International Journal of Industrial Ergonomics*, 40(3).

Lind, S. (2008) 'Types and sources of fatal and severe non-fatal accidents in industrial maintenance', *International Journal of Industrial Ergonomics*, 38(11–12), pp. 927–933.

Liu, P. and Li, Z. (2014) 'Human Error Data Collection and Comparison with Predictions by SPAR-H', *Risk Analysis*, 34(9), pp. 1706–1719. doi: 10.1111/risa.12199.

Lyons, M. *et al.* (2004) 'Human reliability analysis in healthcare: A review of techniques', *International Journal of Risk and Safety in Medicine*, 16(4), pp. 223–237.

Macwan, A. and Mosleh, A. (1994) 'A methodology for modeling operator errors of commission in probabilistic risk assessment', *Reliability Engineering and System Safety*, 45(1–2), pp. 139–157. doi: 10.1016/0951-8320(94)90082-5.

Madonna, M. *et al.* (2009) 'Il Fattore Umano Nella Valutazione Dei Rischi: Confronto Metodologico Fra Le Tecniche Per L'Analisi Dell'Affidabilità Umana', *Prevenzione Oggi*, 52(1), pp. 67–83.

Maffulli, N. *et al.* (2008) 'Ipsilateral free semitendinosus tendon graft transfer for reconstruction of chronic tears of the Achilles tendon.', *BMC musculoskeletal disorders*, 9(1), p. 100. doi: 10.1186/1471-2474-9-100.

Malik, R., White, P. S. and Macewen, C. J. (2003) 'Using human reliability analysis to detect surgical error in endoscopic DCR surgery', *Clinical Otolaryngology and Allied Sciences*, 28(5), pp. 456–460. doi: 10.1046/j.1365-2273.2003.00745.x.

Mathiassen, S. E. *et al.* (2014) 'Can cognitive activities during breaks in repetitive manual work accelerate recovery from fatigue? A controlled experiment', *PLoS ONE*, 9(11), pp. 1–13. doi: 10.1371/journal.pone.0112090.

Mc Leod, J. E. N. and Rivera, S. S. (2009) 'Human Error Management Optimization in CAREM NPP', in *World Congress on Engineering 2009, Vols I and II*. (Lecture Notes in Engineering and Computer Science), pp. 618–622.

Mc Leod, J. E. N. and Rivera, S. S. (2011) 'Human error hotspots in nuclear power plants', in *Proceedings of the World Congress on Engineering 2011, WCE 2011*, pp. 890–894.

Mc Leod, J. E. N. and Rivera, S. S. (2013) 'Human Error on the Risk Reduction Actions during the NPP Maintenance', in *World Congress on*

Engineering - WCE 2013, Vol I. (Lecture Notes in Engineering and Computer Science), pp. 659–663.

McDonnell, D. *et al.* (2015) 'Reducing Uncertainty in PHM by Accounting for Human Factors - A Case Study in the Biopharmaceutical Industry', in *Proceedings of the 4th International Conference on through-life engineering services*. (Procedia CIRP), pp. 84–89.

Mclean, L. *et al.* (2001) 'Computer terminal work and the bene"t of microbreaks', *Applied Ergonomics*, 32(3), pp. 225–237. doi: http://dx.doi.org/10.1016/S0003-6870(00)00071-5.

Mehta, R. K. and Agnew, M. J. (2010) 'Analysis of individual and occupational risk factors on task performance and biomechanical demands for a simulated drilling task', *International Journal of Industrial Ergonomics*, 40(5), pp. 584–591. doi: 10.1016/j.ergon.2010.06.003.

Michalos, G., Makris, S. and Chryssolouris, G. (2013) 'The effect of job rotation during assembly on the quality of final product', *CIRP Journal of Manufacturing Science and Technology*, 6(3), pp. 187–197. doi: 10.1016/j.cirpj.2013.03.001.

Mohren, D. C. L., Jansen, N. W. H. and Kant, I. J. (2010) 'Need for recovery from work in relation to age: A prospective cohort study', *International Archives of Occupational and Environmental Health*, 83(5), pp. 553–561. doi: 10.1007/s00420-009-0491-9.

Monk, T. H. *et al.* (1997) 'Circadian rhythms in human performance and mood under constant conditions.', *Journal of sleep research*, 6(1), pp. 9–18. doi: 10.1046/j.1365-2869.1997.00023.x.

Mosleh, A. and Chang, Y. H. (2004) 'Model-based human reliability analysis: Prospects and requirements', *Reliability Engineering and System Safety*, 83(2), pp. 241–253. doi: 10.1016/j.ress.2003.09.014.

Moura, R. *et al.* (2016) 'Learning from major accidents to improve system design', *Safety Science*. Elsevier Ltd, 84, pp. 37–45. doi: 10.1016/j.ssci.2015.11.022.

Mura, M. D., Dini, G. and Failli, F. (2016) 'An Integrated Environment Based on Augmented Reality and Sensing Device for Manual Assembly Workstations', in *Procedia CIRP*, pp. 340–345.

Nembhard, D. A. and Osothsilp, N. (2001) 'An empirical comparison of forgetting', *IEEE Transactions on Engineering Management*, 48(3), pp. 283–291. doi: 10.1109/17.946527.

Neumann, W. P., Kolus, A. and Wells, R. P. W. (2016) 'Human Factors in Production System Design and Quality Performance – A Systematic Review', *IFAC-PapersOnLine*. Elsevier B.V., 49(12), pp. 1721–1724. doi: 10.1016/j.ifacol.2016.07.830.

Neumayr, T. . *et al.* (2015) 'Analyzing Interaction for Automated Adaptation - First Steps in the IAAA Project', *International Journal of Electronics and Telecommunications*, 61(2), pp. 159–164. doi: 10.1515/eletel-2015-0020.

Nicholas, C. (2009) *Human error in maintenance - A design perspective*, *Handbook of Maintenance Management and Engineering*.

Norman, D. A. (1981) 'Categorization of action slips.', *Psychological Review*, 88(1), pp. 1–15. doi: 10.1037//0033-295X.88.1.1.

Noroozi, A. *et al.* (2013) 'The role of human error in risk analysis: Application to pre- and post-maintenance procedures of process facilities', *Reliability Engineering and System Safety*, 119, pp. 251–258.

Noroozi, A., Khan, F., *et al.* (2014) 'Determination of human error probabilities in maintenance procedures of a pump', *Process Safety and Environmental Protection*, 92(2).

Noroozi, A., Abbassi, R., *et al.* (2014) 'Effects of cold environments on human reliability assessment in offshore oil and gas facilities', *Human Factors*, 56(5).

Okoh, P. (2015) 'Maintenance grouping optimization for the management of risk in offshore riser system', *Process Safety and Environmental Protection*, 98.

Onofrio, R., Trucco, P. and Torchio, A. (2015) 'Towards a Taxonomy of Influencing Factors for Human Reliability Analysis (HRA) Applications in Surgery', *Procedia Manufacturing*, 3, pp. 144–151. doi: 10.1016/j.promfg.2015.07.119.

Papic, L. and Kovacevic, S. (2016) 'Human factor in mining machines maintenance operations', in *Proceedings - 2nd International Symposium on Stochastic Models in Reliability Engineering, Life Science, and Operations Management, SMRLO 2016*.

Parent-Thirion, A. *et al.* (2007) *Fourth European working conditions survey*.

Park, K. S. and Lee, J. (2008) 'A new method for estimating human error probabilities: AHP-SLIM', *Reliability Engineering and System Safety*, 93(4), pp. 578–587. doi: 10.1016/j.ress.2007.02.003.

Pennathur, A. *et al.* (2003) 'Manual dexterity of older Mexican American adults: A cross-sectional pilot experimental investigation', *International Journal of Industrial Ergonomics*, 32(6), pp. 419–431. doi:

10.1016/j.ergon.2003.07.001.

Perez, J. *et al.* (2014) 'Discrete event simulation as an ergonomic tool to predict workload exposures during systems design', *International Journal of Industrial Ergonomics*. Elsevier Ltd, 44(2), pp. 298–306. doi: 10.1016/j.ergon.2013.04.007.

Philips, L. *et al.* (1990) 'A socio-technical approach to assessing human reliability.', in Wiley series in probability & statistics. John Wiley & Sons Ltd., West Sussex, UK (ed.) *Influence Diagrams, Belief Nets and Decision Analysis.*, pp. 253–276.

Pimenta, A. *et al.* (2014) 'Analysis of Human Performance as a Measure of Mental Fatigue', *Hybrid Artificial Intelligence Systems SE - 35*, 8480, pp. 389–401. doi: 10.1007/978-3-319-07617-1_35.

Pires, S. . *et al.* (2015) 'Industrial maintenance for sustainable performance: A systematic literature review', *23rd International Conference for Production Research, ICPR 2015*.

Presley, A. and Liles, D. H. (1995) 'The Use of IDEF0 for the Design and Specification of Methodologies', in *Proceedings of the 4th Industrial Engineering Research Conference.*

Ragsdale, J. M. *et al.* (2011) 'An integrated model of weekday stress and weekend recovery of students.', *International Journal of Stress Management*, 18(2), pp. 153–180. doi: 10.1037/a0023190.

Rankin, W. *et al.* (2000) 'Development and evaluation of the Maintenance Error Decision Aid (MEDA) process', *International Journal of Industrial Ergonomics*, 26(2).

Rashid, H. S. J., Place, C. S. and Braithwaite, G. R. (2013) 'Investigating the investigations: A retrospective study in the aviation maintenance error causation', *Cognition, Technology and Work*, 15(2).

Rashid, H. S. J., Place, C. S. and Braithwaite, G. R. (2014) 'Eradicating root causes of aviation maintenance errors: Introducing the AMMP', *Cognition, Technology and Work*, 16(1).

Rasmussen, J. *et al.* (1981) *Classification system for reporting events involving human malfunctions*.

Rasmussen, J. (1982) 'Human Errors. A Taxonomy for describing Human Malfunction in Industrial Installations', *Journal of Occupational Accidents*, 4, pp. 311–333.

Razak, I. H. A., Kamaruddin, S. and Azid, I. A. (2008) 'Development of Human Reliability Model for evaluating maintenance workforce reliability: A case study in electronic packaging industry', in *Proceedings of the*

IEEE/CPMT International Electronics Manufacturing Technology (IEMT) Symposium.

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

Rekik, M., Cordeau, J. F. and Soumis, F. (2010) 'Implicit shift scheduling with multiple breaks and work stretch duration restrictions', *Journal of Scheduling*, 13(1), pp. 49–75. doi: 10.1007/s10951-009-0114-z.

Rouch, I. *et al.* (2005) 'Shiftwork experience, age and cognitive performance.', *Ergonomics*, 48(10), pp. 1282–1293. doi: 10.1080/00140130500241670.

Rouse, W. B. and Rouse, S. H. (1983) 'Analysis and classification of human error', *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(4), pp. 539–549. doi: 10.1109/TSMC.1983.6313142.

Salthouse, T. (2010) 'Selective review of cognitive aging', *Journal of International Neuropsychology*, 16(5), pp. 754–760. doi: 10.1017/S1355617710000706.

Salthouse, T. (2012) 'Consequences of age-related cognitive declines.', *Annual review of psychology*, 63, pp. 201–26. doi: 10.1146/annurev-psych-120710-100328.

Saptari, A., Leau, J. X. and Mohamad, N. A. (2015) 'The effect of time pressure, working position, component bin position and gender on human error in manual assembly line', in *IEOM 2015 - 5th International Conference on Industrial Engineering and Operations Management, Proceeding*.

Saurin, T. A. *et al.* (2008) 'An algorithm for classifying error types of front-line workers based on the SRK framework', *International Journal of Industrial Ergonomics*, 38(11–12), pp. 1067–1077. doi: 10.1016/j.ergon.2008.02.017.

Schafhauser, W., Musliu, N. and Wild, M. (2009) 'A Memetic Algorithm for a Break Scheduling Problem', *Metaheuristics International Conference*, (January).

Scheffe, H. (1999) *The analysis of variance*. Jonh Wiley & Sons.

Schemeleva, K. *et al.* (2012) 'Human Error Probability Computation for Manufacturing System Simulation Using Cream'.

Serdet, E. and Le Bot, P. (2012) 'MERMOS-A: A new method to analyze pre-initiator for specific HRA', in *11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012*, pp. 3218–3235. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84873176314&partnerID=40&md5=862edf1b1ed7b46e8473df92c3c4085e.

Sharit, J. (2012) ‘Human error and Human Reliability Analysis’, in *Handbook of Human Factors and Ergonomics, Fourth Edition*, pp. 734–800.

Sheikhalishahi, M. *et al.* (2016) ‘Human Factors Effects and Analysis in Maintenance: A Power Plant Case Study’, *Quality and Reliability Engineering International*.

Sheikhalishahi, M. *et al.* (2016) ‘Maintenance Scheduling Optimization in a Multiple Production Line Considering Human Error’, *Human Factors and Ergonomics in Manufacturing & Service Industries*, 26(6), pp. 655–666.

Sheikhalishahi, M., Pintelon, L. and Azadeh, A. (2016) ‘Human factors in maintenance: a review’, *Journal of Quality in Maintenance Engineering*, 22(3), pp. 218–237.

Shephard, R. J. (1999) ‘Age and physical work capacity’, *Experimental Aging Research*, 25(4), pp. 331–343. doi: 10.1080/036107399243788.

Silverstein (2008) ‘Meeting the challenges of an aging workforce’, *Americal Journal of industrial medicine*, 5172(10), pp. 269–280. doi: 10.1002/ajim.20569.

Singh, S. and Kumar, R. (2015) ‘Evaluation of human error probability of disc brake unit assembly and wheel set maintenance of Railway Bogie’, in *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015)*, pp. 3041–3048.

Smith, L. *et al.* (1998) ‘Industrial shift systems’, *Applied Ergonomics*, 29(4), pp. 273–280. doi: 10.1016/S0003-6870(97)00059-8.

Strater, O. (1996) ‘Assessment of cognitive errors and organisational aspects based on evaluation of plant experience.’, in *Proceedings of the International Topical Meeting on PSA 1996*. UTAH, USA, pp. 245–252.

Strater, O. and Reer, B. (1999) ‘A comparison of the applications of the CAHR method to the evaluation of PWR and BWR events and some implications for the methodological development of HRA.’, in *Proceedings of PSA’99 – International Topical Meeting on Probabilistic Safety Assessment. American Nuclear Society,*. La Grange Park, Illinois, USA, pp. 869–875.

Su, K. W., Hwang, S. L. and Liu, T. H. (2000) ‘Knowledge architecture and framework design for preventing human error in maintenance tasks’, *Expert Systems with Applications*, 19(3), pp. 219–228.

Swain, A. D. (1987) *Accident sequence evaluation program: human reliability analysis procedure*.

Swain, A. D. and Guttmann, H. E. (1983) ‘Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR

1278’, (August).

Takano, K., Sasou, K. and Yoshimura, S. (1995) ‘Simulation system for behavior of an operating group.’, in *14 th European Annual Conference on Human Decision Making and manual Control,*. Delft, the Netherlands, pp. 1–26.

Taylor-Adams, S. E. (1995) *An overview of the development of the computerized operator reliability and error database (CORE- DATA)*.

Thun, J.-H., Größler, A. and Miczka, S. (2007) ‘The impact of the demographic transition on manufacturing: Effects of an ageing workforce in German industrial firms’, *Journal of Manufacturing Technology Management*, 18(8), pp. 985–999. doi: http://dx.doi.org/10.1108/MRR-09-2015-0216.

Trucco, P. and Leva, M. C. (2007) ‘A probabilistic cognitive simulator for HRA studies (PROCOS)’, *Reliability Engineering and System Safety*, 92(8), pp. 1117–1130. doi: 10.1016/j.ress.2006.06.003.

Tucker, P. *et al.* (2006) ‘The impact of rest breaks on temporal trends in injury risk.’, *Chronobiology international*, 23(6), pp. 1423–34. doi: 10.1080/07420520601070315.

Tucker, P. and Folkard, S. (2012) *Working time, health and safety: a research synthesis paper*.

Tucker, P., Folkard, S. and Macdonald, I. (2003) ‘Rest breaks and accident risk’, *Lancet*, 361(9358), p. 680. doi: 10.1016/S0140-6736(03)12566-4.

United States General Accunting Office (2003) ‘Older workers: Policies of Other Nations to Increase Labor Force Participation’, Washington(February).

Wang, C.-H. . and Hwang, S.-L. . (2004) ‘A stochastic maintenance management model with recovery factor’, *Journal of Quality in Maintenance Engineering*, 10(2), pp. 154–164.

Wang, Y. B., Zhang, T. and Xue, Q. (2014) ‘The design and realization of HRA service system for the engine assembly’, *Applied Mechanics and Materials*, 635–637, pp. 1900–1905.

Westgaard, R. H. and Winkel, J. (1997) ‘Ergonomic intervention research for improved musculoskeletal health: A critical review’, *International Journal of Industrial Ergonomics*, 20(6), pp. 463–500. doi: 10.1016/S0169-8141(96)00076-5.

Whaley, A. M. *et al.* (2011) ‘SPAR-H Step-by-Step Guidance’, *(No. INL/CON-12-24693). Idaho National Laboratory (INL)*.

Whalley, S. P. (1987) *Factors affecting human reliabiltiy in the chemical* *process industry.* Aston University.

Wild, M. and Musliu, N. (2010) 'An improved memetic algorithm for break scheduling', in *Hybrid heuristics*. Berlin Heidelberg: Springer, pp. 133–147.

Williams, J. C. (1985) 'A Proposed Method for Achieving High Reliability in Process Operation by means of Human Factors Engineering Technology.', in *Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society*. Southport, United kingdom.

Williamson, A. *et al.* (2011) 'The link between fatigue and safety', *Accident Analysis and Prevention*. Elsevier Ltd, 43(2), pp. 498–515. doi: 10.1016/j.aap.2009.11.011.

Woods, D. D., Roth, E. M. and People, H. E. (1987) *Cognitive environment simulation: an artificial intelligence system for human performance assessment.*

Yang, L. *et al.* (2012) 'A novel method of analyzing quality defects due to human errors in engine assembly line', *2012 International Conference on Information Management, Innovation Management and Industrial Engineering*, 3, pp. 154–157.

Zhou, X. *et al.* (2015) 'A Cognitive Reliability Assessment Method For Aviationmaintenance Error', in *Proceedings of the 2015 First International Conference on Reliabiltiy Systems Engineering 2015 ICRSE*.

De Zwart, B. C. H., Frings-Dresen, M. H. W. and Van Dijk, F. J. H. (1995) 'Physical workload and the ageing worker: A review of the literature', *International Archives of Occupational and Environmental Health*, 68(1), pp. 1–12. doi: 10.1007/BF01831627.

Appendix A

Table VIII.1: *Optimal break time (male).*

Physical Activity	Psychological Demands	Recovery Speed	Age Groups (male) 18-25	26-35	36-45	46-55	56-65
Sedentary activities (office, laboratory)	Low	Slow	40	42	45	47	43
		Medium	25	27	28	30	28
		Fast	11	11	12	13	12
	Medium	Slow	34	37	38	41	40
		Medium	23	25	25	27	27
		Fast	11	12	13	14	13
	High	Slow	28	29	30	31	29
		Medium	21	23	23	24	23
		Fast	13	14	14	14	14
Activity light, standing (laboratory, light industry)	Low	Slow	44	46	49	52	48
		Medium	29	31	33	35	32
		Fast	16	17	18	19	18
	Medium	Slow	38	41	42	45	44
		Medium	27	29	29	32	31
		Fast	17	19	19	21	20
	High	Slow	30	32	33	34	32
		Medium	25	27	27	28	27
		Fast	19	21	21	21	21
Medium activity, standing (work machines)	Low	Slow	48	50	54	56	52
		Medium	34	36	38	40	37
		Fast	22	23	24	26	24
	Medium	Slow	41	44	46	49	49
		Medium	30	33	34	36	36
		Fast	23	25	25	27	27
	High	Slow	36	35	36	37	35
		Medium	29	30	31	32	30
		Fast	26	27	28	29	27
Activities heavy (heavy work on machines)	Low	Slow	55	59	63	66	61
		Medium	42	45	47	50	46
		Fast	32	34	37	38	35
	Medium	Slow	48	52	53	57	57
		Medium	38	41	42	46	45
		Fast	34	37	38	41	40
	High	Slow	39	41	42	43	41
		Medium	36	38	39	40	38
		Fast	39	41	42	43	41

Table VIII.2: *Recovery rate (male).*

Physical Activity	Psychological Demands	Recovery Speed	Age Groups (male)				
			18-25	**26-35**	**36-45**	**46-55**	**56-65**
Sedentary activities (office, laboratory)	Low	Slow	3.49	3.28	3.09	2.94	3.19
		Medium	5.48	5.16	4.85	4.62	5.01
		Fast	12.79	12.04	11.32	10.78	11.69
	Medium	Slow	4.04	3.73	3.64	3.37	3.42
		Medium	6.07	5.60	5.46	5.05	5.13
		Fast	12.13	11.19	10.93	10.11	10.25
	High	Slow	5.00	4.71	4.62	4.50	4.70
		Medium	6.43	6.06	5.94	5.79	6.04
		Fast	10.71	10.09	9.89	9.65	10.07
Activity light, standing (laboratory, light industry)	Low	Slow	3.17	2.98	2.81	2.67	2.90
		Medium	4.70	4.42	4.16	3.96	4.29
		Fast	8.52	8.03	7.54	7.19	7.79
	Medium	Slow	3.68	3.39	3.31	3.06	3.11
		Medium	5.20	4.80	4.68	4.33	4.39
		Fast	8.09	7.46	7.29	6.74	6.84
	High	Slow	4.55	4.28	4.20	4.10	4.27
		Medium	5.51	5.19	5.09	4.96	5.18
		Fast	7.14	6.73	6.60	6.44	6.71
Medium activity, standing (work machines)	Low	Slow	2.91	2.74	2.57	2.45	2.66
		Medium	4.11	3.87	3.64	3.47	3.76
		Fast	6.39	6.02	5.66	5.39	5.84
	Medium	Slow	3.37	3.11	3.04	2.81	2.85
		Medium	4.55	4.20	4.10	3.79	3.85
		Fast	6.07	5.60	5.46	5.05	5.13
	High	Slow	4.17	3.93	3.85	3.75	3.92
		Medium	4.82	4.54	4.45	4.34	4.53
		Fast	5.36	5.05	4.95	4.83	5.04
Activities heavy (heavy work on machines)	Low	Slow	2.49	2.35	2.20	2.10	2.28
		Medium	3.29	3.10	2.91	2.77	3.01
		Fast	4.26	4.01	3.77	3.59	3.90
	Medium	Slow	2.89	2.66	2.60	2.41	2.44
		Medium	3.64	3.36	3.28	3.03	3.08
		Fast	4.04	3.73	3.64	3.37	3.42
	High	Slow	3.57	3.36	3.30	3.22	3.36
		Medium	3.86	3.63	3.56	3.48	3.63
		Fast	3.57	3.36	3.30	3.22	3.36

Table VIII.3: *Optimal break time (female).*

Physical Activity	Psychological Demands	Recovery Speed	Age Groups (female) 18-25	26-35	36-45	46-55	56-65
Sedentary activities (office, laboratory)	Low	Slow	40	39	41	44	45
		Medium	26	25	26	28	29
		Fast	11	11	11	12	12
	Medium	Slow	35	35	35	39	42
		Medium	23	23	23	26	28
		Fast	12	12	12	13	14
	High	Slow	28	27	27	29	31
		Medium	22	21	21	22	24
		Fast	13	13	13	13	14
Activity light, standing (laboratory, light industry)	Low	Slow	44	43	45	49	50
		Medium	30	29	30	33	34
		Fast	16	16	17	18	19
	Medium	Slow	38	38	38	42	47
		Medium	27	27	27	30	33
		Fast	17	17	17	19	21
	High	Slow	31	30	30	32	34
		Medium	26	2	25	26	28
		Fast	20	19	19	20	22
Medium activity, standing (work machines)	Low	Slow	48	47	49	53	54
		Medium	34	33	35	37	39
		Fast	22	21	22	24	25
	Medium	Slow	42	42	41	46	51
		Medium	31	31	31	34	38
		Fast	23	23	23	26	28
	High	Slow	34	33	33	35	37
		Medium	29	28	28	30	32
		Fast	26	26	25	27	29
Activities heavy (heavy work on machines)	Low	Slow	56	55	57	62	64
		Medium	43	42	43	47	48
		Fast	33	32	33	36	37
	Medium	Slow	49	49	48	54	59
		Medium	39	39	38	43	47
		Fast	35	35	35	39	42
	High	Slow	39	38	38	40	43
		Medium	36	36	35	37	40
		Fast	39	38	38	40	43

Table VIII.4: *Recovery rate (female).*

Physical Activity	Psychological Demands	Recovery Speed	Age Groups (female)				
			18-25	**26-35**	**36-45**	**46-55**	**56-65**
Sedentary activities (office, laboratory)	Low	Slow	3.43	3.51	3.39	3.13	3.04
		Medium	5.39	5.51	5.32	4.91	4.78
		Fast	12.57	12.86	12.42	11.46	11.16
	Medium	Slow	3.97	3.99	4.00	3.58	3.26
		Medium	5.96	5.98	6.00	5.37	4.89
		Fast	11.92	11.96	12.00	10.75	9.79
	High	Slow	4.92	5.03	5.07	4.79	4.49
		Medium	6.32	6.47	6.52	6.16	5.77
		Fast	10.53	10.78	10.86	10.26	9.61
Activity light, standing (laboratory, light industry)	Low	Slow	3.12	3.19	3.08	2.84	2.77
		Medium	4.62	4.72	4.56	4.21	4.10
		Fast	8.38	8.57	8.28	7.64	7.44
	Medium	Slow	3.61	3.62	3.64	3.26	2.97
		Medium	5.11	5.12	5.14	4.61	4.20
		Fast	7.95	7.97	8.00	7.17	6.53
	High	Slow	4.47	4.57	4.61	4.35	4.08
		Medium	5.42	5.55	5.59	5.28	4.94
		Fast	7.02	7.19	7.24	6.84	6.41
Medium activity, standing (work machines)	Low	Slow	2.86	2.92	2.82	2.61	2.54
		Medium	4.04	4.13	3.99	3.69	3.59
		Fast	6.28	6.43	6.21	5.73	5.58
	Medium	Slow	3.31	3.32	3.33	2.99	2.72
		Medium	4.47	4.48	4.50	4.03	3.67
		Fast	5.96	5.98	6.00	5.37	4.89
	High	Slow	4.10	4.19	4.22	3.99	3.74
		Medium	4.74	4.85	4.89	4.62	4.33
		Fast	5.27	5.39	5.43	5.13	4.81
Activities heavy (heavy work on machines)	Low	Slow	2.45	2.51	2.42	2.23	2.17
		Medium	3.23	3.31	3.19	2.95	2.87
		Fast	4.19	4.29	4.14	3.82	3.72
	Medium	Slow	2.84	2.85	2.86	2.56	2.33
		Medium	3.58	3.59	3.6	3.22	2.94
		Fast	3.97	3.99	4.00	3.58	3.26
	High	Slow	3.51	3.59	3.62	3.42	3.20
		Medium	3.79	3.88	3.91	3.70	3.46
		Fast	3.51	3.59	3.62	3.42	3.20

Appendix B

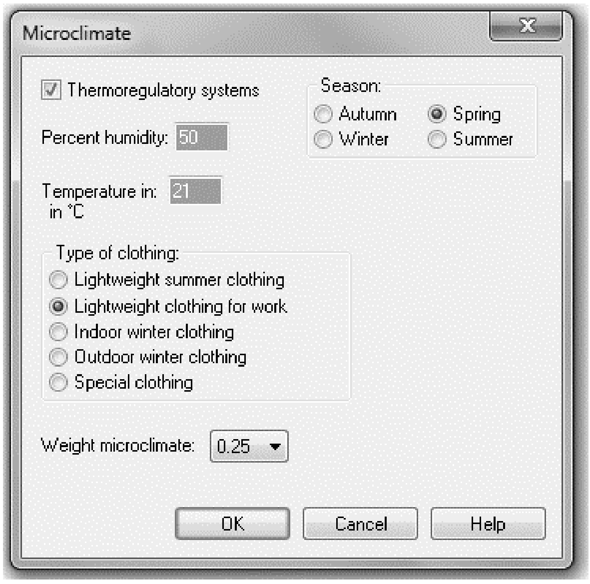

Figure IX.1: *Dialog box microclimate data entry.*

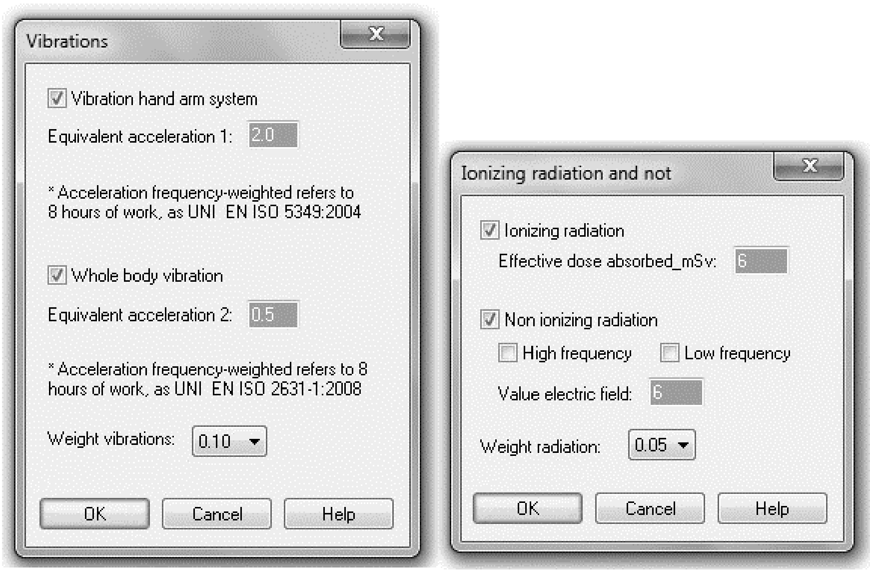

Figure IX.2: *Dialog box vibration and ionizing radiation and not data entry.*

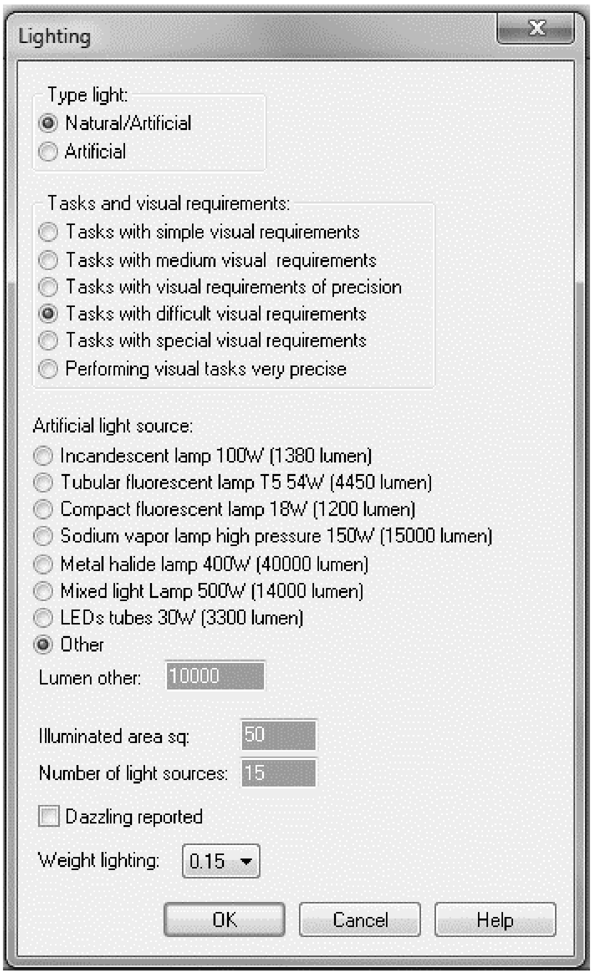

Figure IX.3: *Dialog box lighting data entry.*

Figure IX.4: *Dialog box workplace data entry.*

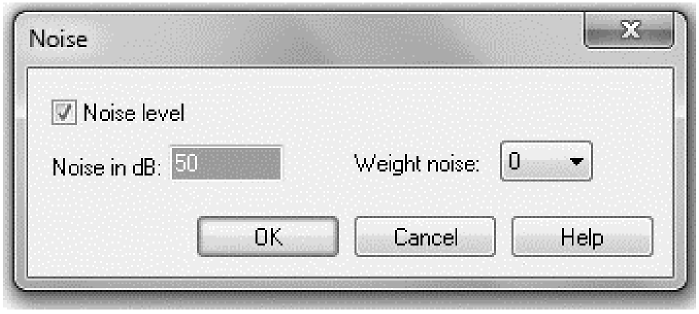

Figure IX.5: *Dialog box noise data entry.*