A version of this paper has been published in:
Cognition Technology & Work, volume 14, issue 2, pp. 107-128, 2012.

Please see: http://www.springerlink.com/content/1435-5558/5/4/

The performance of future designers on the specification of supervisory HCI: case study of a simulated work situation

Christophe Kolski[1,2,3], Pierre Loslever[1,2,3], Mouldi Sagar[1]

[1] *Univ Lille Nord de France, F-59000 Lille, France*
[2] *UVHC, LAMIH, F-59313 Valenciennes, France*
[3] *CNRS, FRE 5503, F-59313 Valenciennes, France*
{christophe.kolski, mouldi.sagar, Pierre.loslever}@univ-valenciennes.fr

Abstract:
This article concerns the performance of future designers on the specification of supervisory human-computer interfaces (HCI). This paper reports the results of a case study of a simulated work situation. Sixty-three groups of 3 to 5 students (Master's program in Electrical Engineering and Industrial Informatics) were involved. In this article, we first describe the study context and explain how we ran it. Next, we describe the industrial process to be supervised, composed of five interconnected stations. We then present our informal analysis of the global study. After our informal analysis, the article focuses on a second level of analysis, based on the 63 HCI specification reports produced by the groups. Each report is analyzed according to eight groups of criteria. The results of this second-level analysis are presented and discussed.

Key words: human-computer interface (HCI), specification, supervisory, design error, simulation, ergonomic design project.

1. Introduction

1.1. The importance of supervisory HCI in industrial situations

The rapid technological evolution of recent years has resulted in significant changes in human activity in the control room. Indeed, human operators increasingly have to perform complex mental problem-solving tasks, using graphical tools in dynamic situations with very specific requirements. Human-machine interaction plays an essential role in the efficiency and reliability of the complete human-machine system (Millot 1988; Kolski 1993 and 1997; Moray 1997; Sheridan 1992 and 2002; Piccini 2002; Lind 2003; Petersen 2004; Liu et al. 2004; Johannsen, 2007). The users of human-computer interfaces in the control room of complex systems are often physically far from the process and must solve problems involving many variables, using tools at their disposal.

Human tasks in complex systems require a high level of knowledge and can be grouped into four broad classes (Rouse 1983; Klein et al. 2005):

- *Transition tasks*: These tasks correspond to system regime changes (e.g., stops, start-ups, operating point changes). The human-computer interfaces allow users to perform pre-established procedures, to assess the impact of these procedures on the system and to monitor the system's evolution constantly.
- *Control and monitoring tasks*: The HCI should facilitate system state monitoring, so as to detect and anticipate the occurrence of abnormal events and to optimize the production by fine-tuning the system.
- *Detection and diagnosis tasks*: Through alarms and/or observation of an abnormal development in certain system variables, the human operators must be able to detect faults and carry out diagnoses. Human-computer interfaces have to allow operators to understand the cause-effect relationships between system variables.
- *Compensation and/or correction tasks*: In order to restore normal system functions, the HCI must help the human operators to determine the actions that should be performed and visualize the effect of these actions on the system.

In addition to these four classes of human tasks, many authors—for example, De Keyser et al. (1987), Van Dale and De Keyser (1990), Artman and Waern (1999), Klein (2006), Kontogiannis (2010), Nonose et al. (2010)—insist specifically on the collective aspect of work, which must be carefully considered by designers. Indeed, in certain situations inherent to the tasks described above, human operators must organize a distribution of actions over time. They then must manage and synchronize sequential and/or parallel actions, possibly made by other stakeholders. These tasks must be based on estimates of time and the duration of the action. To accomplish these complex tasks, which sometimes involve many variables, the user has human-machine interfaces that graphically present a set of information. The interfaces can be coupled with assistance modules, sometimes qualified as "intelligent" if they use artificial intelligence techniques. These modules may offer, for example, diagnosis assistance, defect prediction or alarm filtering.

The HCI must synthesize the state of the complex system and assist the human operators in their activities. The operators have no direct view of the system and the result of their actions, and must therefore form a mental representation of the system and its variables at a distance. They thus work through the mental representation obtained from the HCI (Norman 1986; Rasmussen 1986; Ezzedine and Kolski 2005). Consequently, the HCI must be centralized on one or more screens. It consists of sets of views, sometimes several hundred, depending on the complexity of the system. Each view can display many kinds of information. The development environments currently available on the market allow HCI to be developed relatively easily, perhaps coupled with assistance modules.

In this context, given the high number of variables to be considered and the complexity of the tasks involved, it appears that using a badly designed HCI may cause human errors, sometimes with catastrophic human, economic and/or ecological consequences (e.g., Three Mile Island, Chernobyl, Seveso, Bhopal, Mont Saint Odile). In the literature, there are many human error classifications and studies, for example, the ones by Reason (1990), Nicolet et al. (1990), Senders and Moray (1991), Hollnagel (1998), Gertman and Blackman (2001), Cacciabue et al. (2005), Wallace and Ross (2006) or Dekker (2006). The interested reader can consult these publications.

1.2. Global problem of teaching about Supervisory HCI

Teaching about HCI has become a part of standard engineer training (e.g., software engineering, mechanical engineering) in addition to training in the cognitive sciences (e.g., ergonomics, cognitive psychology). However, unfortunately, the number of classroom hours devoted to HCI is too often limited; HCI is one subject among many others. This article describes how we teach our HCI class in a Master's degree, with a limited number of classroom hours (18 hours).

For teaching HCI classes, teachers often use their personal experience, drawing on their research if they come from the university or their practice if they come from industry. They can rely on many books, journals and conference proceedings about HCI. They can also use the ACM SIGCHI's Curricula for Human-Computer Interaction, whose first version was published in 1992 and the second in 2002 (Hewett et al. 2002). There is also a set of HCI educational resources in HCIBib managed by Gary Perlman (http://hcibib.org/). The documents available for teaching about HCI from a global perspective are rich in potential. In addition, the languages already taught in other teaching modules (e.g., Java, Visual C++, Delphi) can be used to illustrate the HCI module, for example, HCI implementation, programming or mock/prototyping.

However, when the targeted HCI application domain becomes specific, with important security, ecological or economic stakes (e.g., industrial system supervision, computer-assisted surgery), teaching about HCI becomes more problematic, requiring particular concepts, methods, models and architectures. Practical work, under real and/or simulated conditions, also requires special development environments that are often expensive, requiring a long apprenticeship. Moreover, the pedagogical problems related to HCI teaching are often underestimated or completely ignored. For example, there are forms of education focused mainly on norms and standards (e.g., ISO) or style guides (Gilmore et al. 1989), and where they exist, regulation constraints, particularly related to physical ergonomics (e.g., lighting, noise, the thermal environment, vibration, ventilation), can be problematic.

Unfortunately, the trend in the complex industrial systems is still too often confined to applying normative or regulatory methods, relatively rapid in terms of education (or rather requiring a high level of personal learning), regardless of underlying theoretical and real pedagogical problems. Consequently, many potential designers/developers often start work for their companies unprepared or even neutralized in terms of HCI design and, more generally, the direction of ergonomic industrial projects, despite their fairly strong interest and expertise in key aspects related to supervisory HCI.

1.3. Objectives of this article

In this article, we focus on performance, not from the point of view of the system users, but from the point of view of the designers. For this reason, we situate ourselves upstream in the design process, on the level of the designers of supervision systems.

The study context is a situation that simulates the design process for the specifications of a supervisory HCI that is representative of a real industrial application. The subjects (i.e., our students) all come from a Master's program in Electronic Engineering and Industrial Informatics, training control engineers and industrial system designers at the University of Valenciennes and Hainaut-Cambrésis (France). For these future designers, the supervision of industrial systems is one of their most important objectives. In this simulated roleplay situation, the students work in a group, which competes with other groups of students in their class. Each group must provide the teacher with a detailed report corresponding to a supervisory HCI specification.

This study was conducted over nine consecutive years with various groups of students. All the reports were analyzed according to the same criteria. The performance of the student groups, compared to these criteria, was analyzed and is discussed in this paper.

The rest of this paper is structured as follows. The case study is described in section 2. Section 3 provides qualitative results, and section 4 provides quantitative results. The last section offers our conclusions and proposes several prospects for future research.

2. Method

The method followed in this research is described in this section. First the objectives of the study are presented. The study is then described (Study context, organization and population; main phases; data analysis).

2.1. Objectives of the study

This study focuses on the performance of future designers on the specification of supervisory HCI. The reports produced over a 9-year period have been analyzed. [1]

The case to be handled by these future designers is representative of current industrial projects (supervisory HCI for mixing stations). Its complexity is manageable in the short period of time allocated to the considered teaching module (a few weeks).

Starting with a theoretical course about the elements judged important for supervisory HCI, we wanted to discover to what point each work group, made up of future designers, used these elements in their solution. These elements were assessed using 8 criteria groups, described in 3.2.1.

2.2. Description of the study

The study context, organization and population are first described, and then the main phases of the study are presented.

2.2.1. Study context, organization and population

The global principle of the study is the following: within an academic context, students (i.e., the subjects of this study) initially take a formal theoretical course about the design of supervisory HCI. They are then placed in a design situation representing an industrial setting in order to apply the theory to the practice as soon as possible. This global context was originally described by Kolski et al. (2004a, 2004b). In this article, we describe our teaching approach in the form of a case study. We also provide feedback from over 9 successive years of practice, insofar as we analyzed the data resulting from these 9 years.

This study was run in three phases, which are described in the sections below (Figure 1). Each year, the course involved approximately 20 to 25 potential designers (i.e., our students). The course organization and content are constant and thus the years are globally comparable with one another.

[1] It is important to note that the conditions of conduct of the course have not fundamentally changed over these nine years.

The subjects were the students in the first year of a Master's program in Electronic Engineering and Industrial Informatics. They had been registered in the same program for at least the previous two years. The average age was about 23 years old. We consider that these students are representative of potential designers of supervisory HCI, and that they are likely to find themselves very rapidly involved in supervision projects in various industrial companies.

2.2.2. First phase -- theory

Each year, during a preparatory phase, the subjects take a 10-hour course (2 hours/week) on the design of supervisory HCI.

The course is structured as follows (Kolski 2001). In week I, the basic concepts about human operators in control rooms are presented (e.g., control room examples, global tasks, problem-solving models, human workload, human limits, priority estimation in control rooms, stress, control models, process representation, knowledge). In week II, the information presentation methods for control rooms are explained, as well as the traditional methods (e.g., mimic displays, trends, receipt representation, operations and incidents, alarm representation, event representation, specific displays) and the advanced methods (e.g., pattern representation, fluence models, mass-data-display). In week III, the possible reactions of the human operators with respect to the HCI at their disposal are studied (e.g., acceptance, rejection, absence of certain information). In week IV, general ergonomic knowledge is provided (e.g., information relevance, decision-making aids, abstraction levels; information allowing prediction, control of the user inputs, consideration of the human operator's skill and knowledge, HCI response time, logical-state representation, printer use in control rooms, alarm management). In week V, knowledge about how to present information on the screen is explained (e.g., number of screens; commands; information coding, organization and structuring; information homogeneity and standardization; information perception and interpretation; legibility, the specific rules for each representation mode).

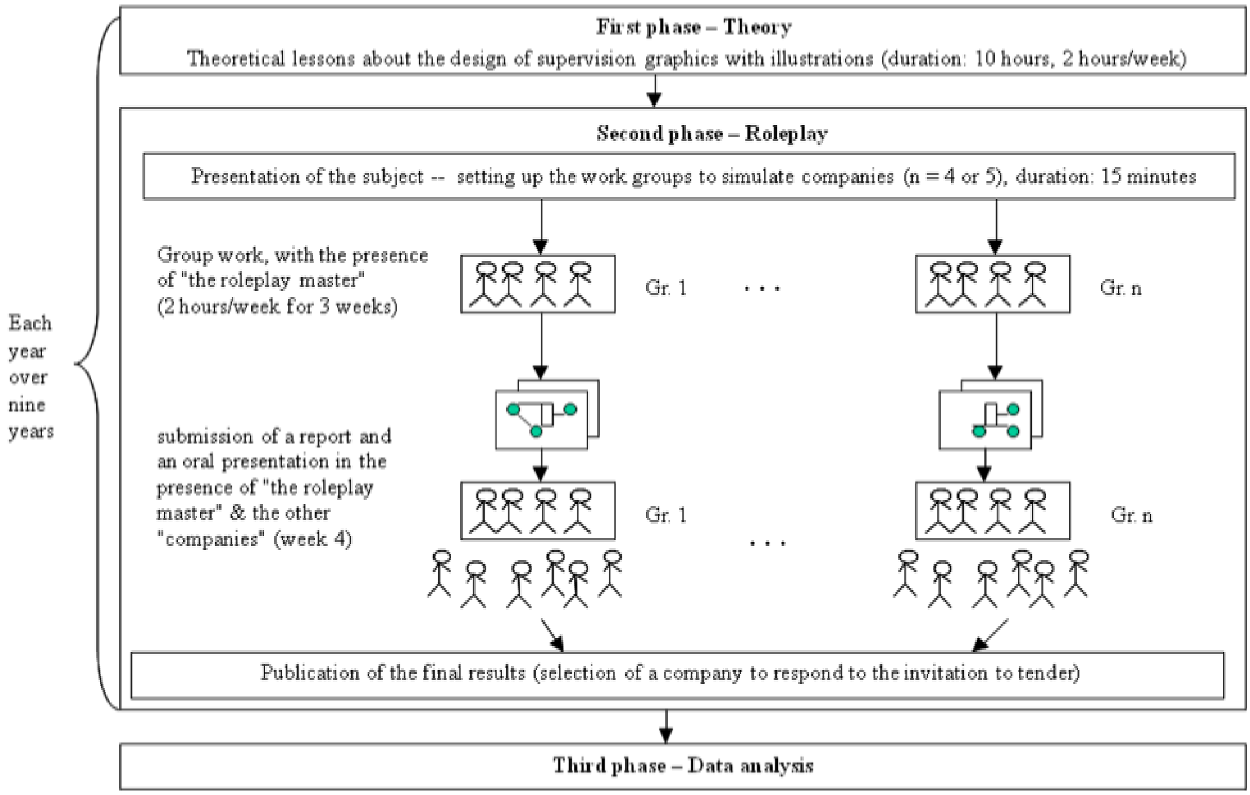


Fig. 1. Phases of the study

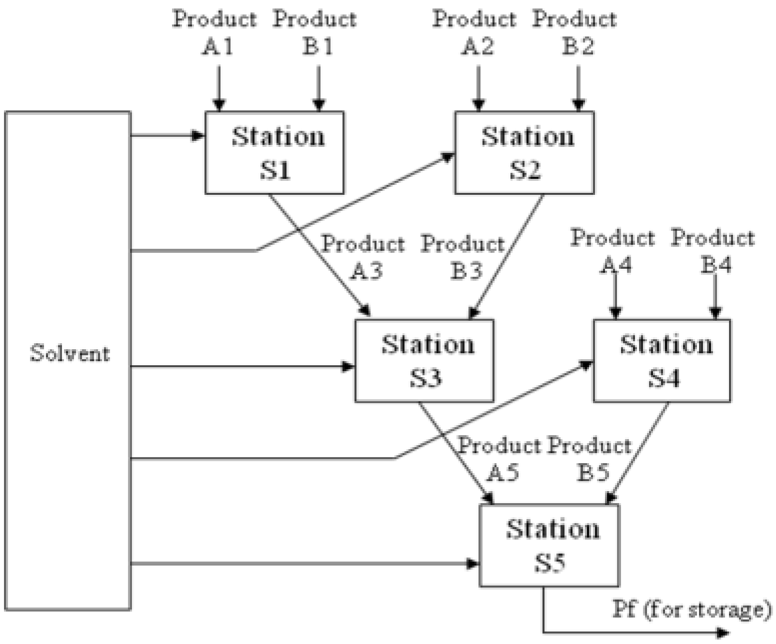
2.2.3. Second phase -- roleplay

The second phase consists of a roleplay simulation that is very close to a real industrial design situation. For 8 hours (2 hours per week, during 4 weeks), the students play the role of the design team in a company. They gather in separate groups of 4 or 5 in a big room. The students in each group choose the name of their company. The groups compete to win an invitation to tender for the specification of a control room and a human-computer interface intended for the supervision of five interconnected product mixing stations. The system stakes are productivity, economy and safety, for the personnel as well as the installations. These aspects are explained in more detail below.

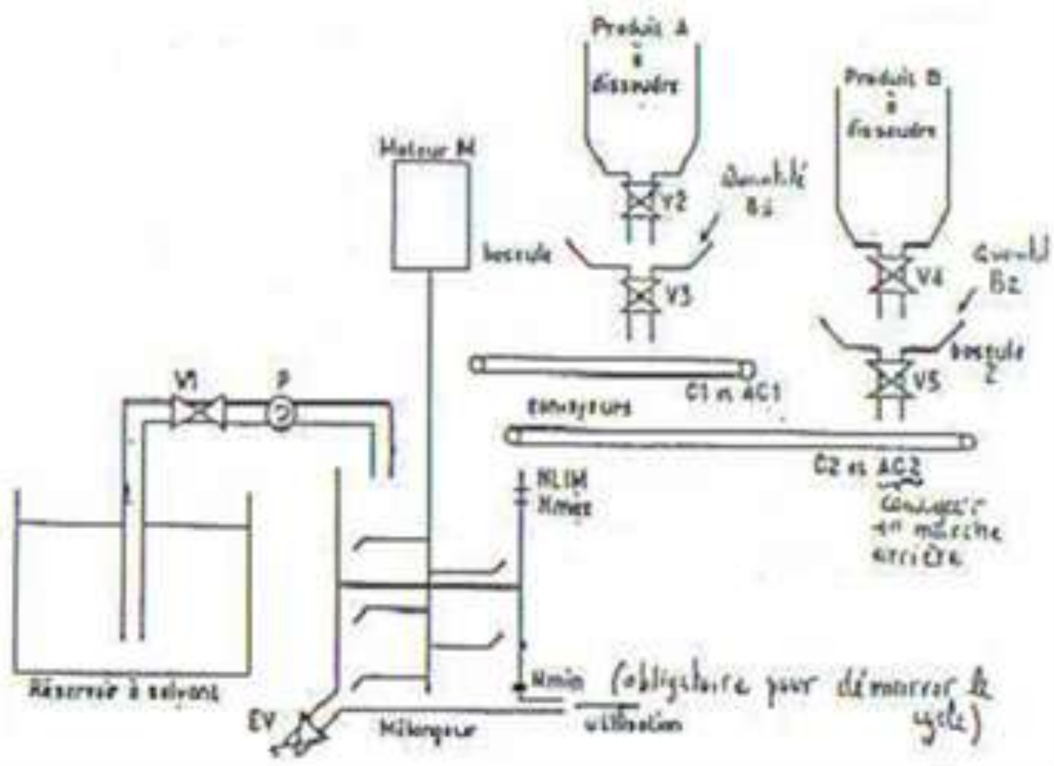
2.2.3.1 Project

Each group was given the following project: propose a control room installation and the specifications for a supervisory HCI for the process shown in Figure 2(a), composed of five mixing stations. This HCI must be displayed on a single screen in a control room occupied by one operator, called the supervisor, working in rotating shifts (3 x 8 hours). This operator uses a walkie-talkie to remain in contact with five experienced roundsmen equipped with mopeds. Orders cannot originate in the control room, but only the supervisor is authorized to decide what actions can be executed on the process. Each action (e.g., start cycle, solvent measure) is carried out by the roundsmen on the instructions of the supervisor.

A roundsman needs about 5 minutes to move from one station to another. The mixing process at each of the five stations, called S1, S2, S3, S4 and S5, lasts 10, 10, 20, 20 and 50 minutes, respectively. Starting from the basic products (A1, B1, A2, B2, A4 and B4), the total process, composed of the mixtures from these five mixing stations, results in the production of a finished product, FP. The students are provided with a complementary description of a mixing station's global operations, represented as a GRAFCET model (similar to Petri nets).



(a)



(b)

Fig. 2. Extract of the documents provided to each group. These documents describe (in French) the industrial facilities for which the groups must propose a control room installation and a supervisory HCI specification
(a) Global diagram of the five interconnected stations (b) Schematic diagram of a station

Several constraints have to be considered:

(1) Each mixer uses less than 100% of its capacity since the supervisor could ask for further solvents to be added.

(2) The process must run uninterrupted 24/7.

(3) The mixing operations for stations S1 and S2 are toxic, and there is a risk of explosion when the temperature of the mixture goes over 18 degrees (however, the stations are situated several hundred meters apart, so an explosion in one mixing station will not have consequences for the others).

(4) When the temperature of one of the 5 mixtures goes over 15 degrees, the V1 valve must be ordered to send additional solvent to the station concerned.

(5) If the temperature is over 13 degrees and under 15 degrees, only a few drops of additional solvent are sufficient. On the other hand, above 15 degrees, 20 liters per degree may be necessary. If the temperature of the mixture is under 13 degrees, it is necessary to purge a little bit of the mixture.

(6) When a mixing problem occurs, it is necessary to purge the station and start again. Two roundsmen are needed to carry out a purge.

(7) Information (e.g., the cycle phase or possible malfunctions) must be exchanged when the shift changes (after 8 hours of work).

(8) The number of revolutions of the mixing engine must be maintained at 200 RPM. If the forward and reverse operation speed is no longer continuous or if a slight rattling is heard, the conveyors are likely to jam.

2.2.3.2 Organization

The roleplay master in fact has three roles: teacher (he teaches the theoretical course and "publishes" the end results), tender caller (he issues the Invitation to Tender and provides the requirements for the HCI specifications), and supervision operator (he is likely to use the control room and HCI, and he is available to the groups to answer any questions). In the rest of this paper, this roleplay master, and his three roles, will be referred to as "the roleplay master".

The "companies" are seated in a large room, a few meters apart from one another. They must not communicate with the other "companies"; in fact, the groups work in relative silence so as not to reveal information to the other groups responding to the Invitation to Tender (Figure 3).

Fig. 3. Competing groups, working in parallel

During the first three 2-hour meetings, the groups have the opportunity to question the roleplay master discreetly in his role as the supervision operator. Except for organizational reasons, the roleplay master no longer plays the role of teacher until he publishes the end results. During these 3 meetings, he therefore never gives advice about the HCI specification: each solution proposed must come only from the members of each "company". During the fourth and final 2-hour meeting, each group hands in a report containing its analysis of the problem in the form of a proposal for a control room installation and a mock-up of the supervisory HCI (examples of these mock-ups are presented further on in this paper.) These reports generally run to about fifteen pages.

After handing in the report, each "company" must make a technical/commercial presentation, describing its proposal. The students can explain only what is in their report; they cannot use the other presentations as a starting point. Following these presentations, a ranking of the companies is posted with a comment on the reports presented by each company. The students' final grades, assigned by the roleplay master in his role as teacher, depend directly on their company's ranking, and thus on the quality of the specification. Feedback about the strong and weak points of each proposal is given orally and in writing; this feedback allows the students to improve their knowledge with respect to all of the reports.

2.2.4. Third phase - data analysis

The idea of study did not come to us initially with a view to annual assessment. Consequently, there was no annual data analysis. It is only after nine years that the idea has matured, leading us to assemble the elements necessary for this study. Thus, at the end of a 9-year period, we analyzed the data from this case study. Sixty-three reports were analyzed.

We started with a global and informal analysis of the educational approach followed. This analysis was based on the impressions of the teacher, gathered throughout the nine years, the observations of students during the simulation sessions, along with feed-back from them in both oral (in informal discussions and formal presentations imposed in the approach), and written form (reports). Some results from this global analysis are described in §3.1.

For purposes of detailed analysis, we then designed a set of criteria (described in §3.2.1, Table 1) to analyze, mark and compare the reports produced. This detailed analysis is described in §3.2.

3. Results

3.1 Preliminary global analysis

The students, as potential designers, all appreciated the approach. Over these nine years, each group was motivated to succeed. Absenteeism in the group meetings was quasi-null, and each group provided a report containing many positive aspects with respect to the evaluation criteria listed in section 4.2 (Table 1). Thanks to the roleplay, these student designers felt that they were working on a project close to the industrial reality; they appreciated the methodology used, especially the spirit of competition between the groups.

3.1.1 Design errors made by each group

Although the roleplay master in his role as teacher took the time necessary during the first theoretical phase to highlight the errors to be avoided, each group without exception made many typical and relatively serious errors, related to information structure, information relevance, the use of the colors, the coherence of the presentation from one slide to another and information overload (e.g., by including useless details), for example.

Figure 4 shows a screen shot that has many errors and shortcomings: The presentation is not ergonomic. Indeed, it is overloaded and has no logical structure according to its use (it should include areas specific to the types of information). The key should be in a separate area (and should not be highlighted so much). The representation does not take the real scales into account. The colors are not necessarily meaningful and can be used in different ways (for example, red is used to represent both excessive heat (which makes sense), but also to represent an instrument (such as a valve). The physicochemical parameters are not represented by numerical values, and do not refer to any curves or histograms detailing the information in question. Furthermore, no information regarding the flow of input and output is specified, both in terms of flow and content. Other graphic details are not relevant (such as the product mixing blades; the conveyors are also poorly positioned and oriented) and load the human operator unnecessarily. The title of the screen page is also poorly positioned (currently at the bottom right). (Section 4 presents our analysis of the errors made, according to eight criteria groups).

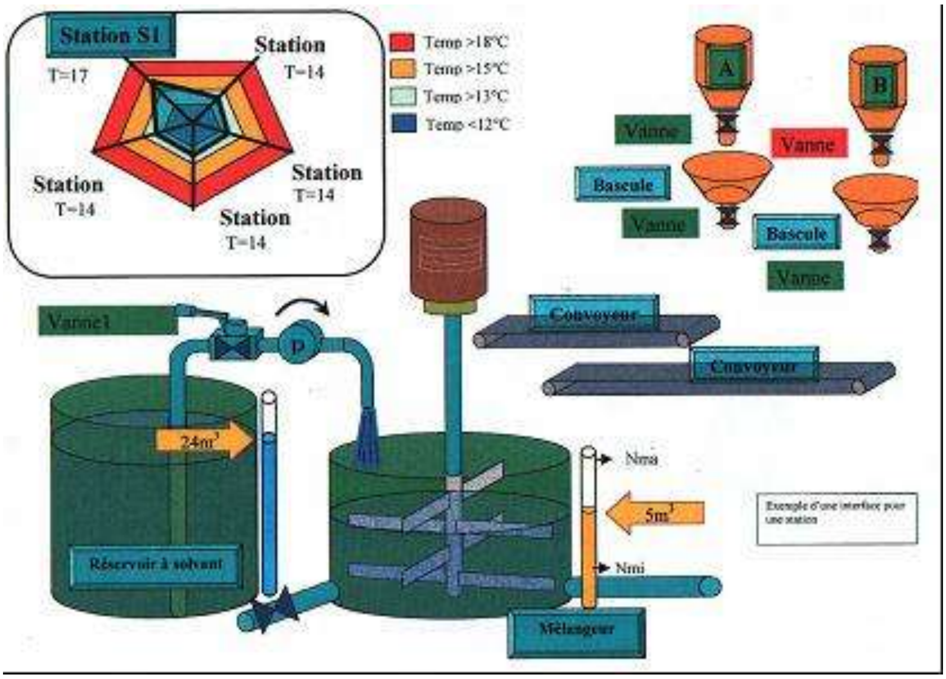


Fig. 4. Example of screen page coming from a report, containing several typical errors [2]

3.1.2 Influence of software prototyping tools

During the first years, when the students seldom had access to software tools (i.e., personal tools or tools provided by the program) that facilitated the creation of mock-ups or prototypes, several reports had solutions that can be seen as too similar to the technical data provided at the beginning of the second phase (see Figure 2). With the availability of office software (e.g., text editors, presentation editors) and development environments for supervision application at lower costs, such unsatisfactory solutions began to disappear. Figure 5 shows 2 screen shots that simply copied technical data given at the beginning of phase 2 to produce a screen shot showing a station: the first mock-up (top) is made by hand by sticking a photocopy of part of the topic given to the students and the second (bottom) uses a software prototyping tool for the structure of the page screen, with a collage of the same part as used in the first mock-up.

It is important here to specify that in our notation of the different reports (see Detailed Analysis, §3.2), we did not take into account the fact of whether they used prototyping tools or not.

---

[2] The non-optimal quality of Figures 4, 5, 6, 7 and 8 is due to the fact that these figures were scanned from the reports provided by the student groups.

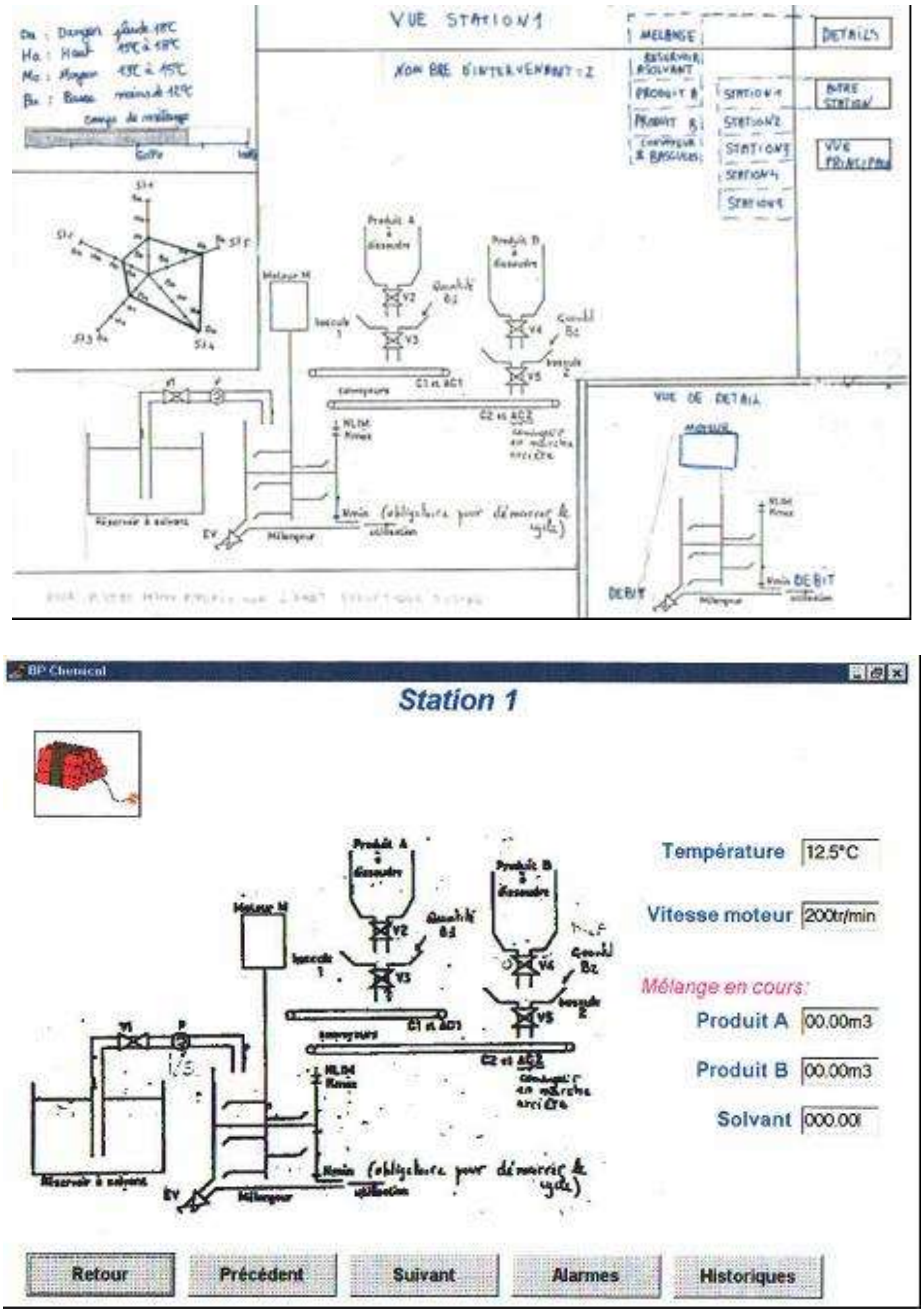


Fig. 5. Two examples of solutions that were too directly inspired by the technical data provid-ed

3.1.3. Generalization of the synthesis view

Although the literature often emphasizes the parcelling out of information on the supervisory HCI (Lejon 1991), the theoretical course emphasized the need for a synthesis view. This seems to have been assimilated quite well by the student because, without exception, a synthesis view was proposed by each group. Figure 6 gives two examples of the synthesis views proposed (Kolski et al. 2004b). The first one represents a global view of the five stations from a functional perspective, while at the same time associating each station to the relevant targeted information. The second example provides on the same screen shot the aggregated information concerning the stations and a table that is supposed to facilitate the management of the roundsmen (knowing where each roundsman is can have effects on the strategy adopted to implement certain actions on the stations). A star display, whose principle was proposed initially by Coekin (1968), also gives a synthesis of a targeted set of variables that represents the global functional state of the whole system.

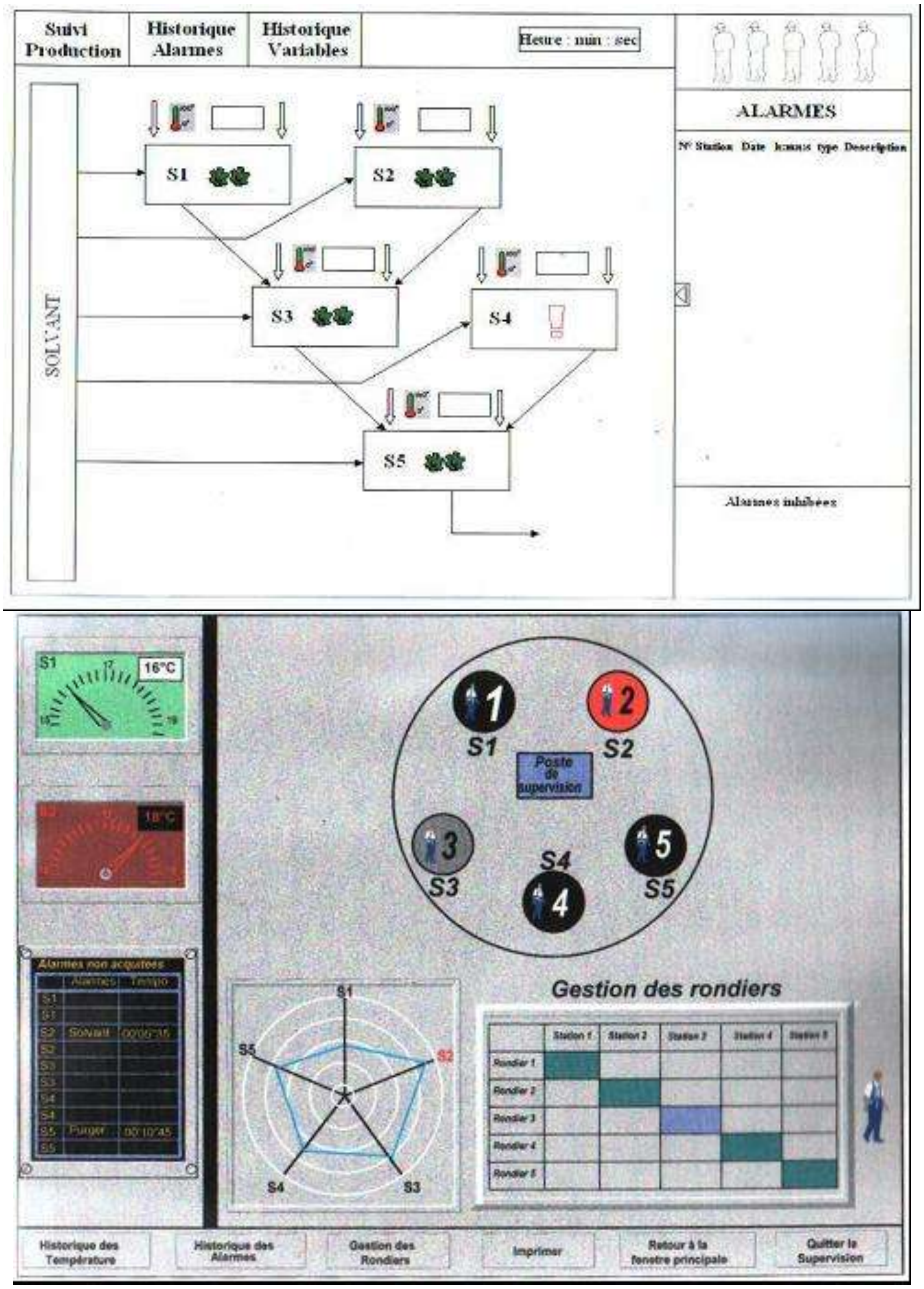


Fig. 6. Two examples of synthesis views proposed in the reports

3.1.4 Creative potential of the designers

By studying the 63 reports produced during these nine years, it is not possible to say that new proposals have emerged, which could become a standard for the future.

Each year, considering that the subjects are novices (no experience in supervisory HCI design in real situations, with real human operators), it can be noted that each solution is generally valid, truly enabling the supervision of stations. This is very positive.

Moreover, several interesting and/or relatively original ideas emerged from the groups, or more specifically from the fertile imagination of one or more of their members. Figure 7 shows an example of an alarm management system, making it easier for the operator to obtain a quick idea of an abnormal situation. It was proposed by one of the groups. The screen shot in Figure 7 shows a frame of reference with a fixed place for each type of alarm that also provides more detailed information, such as the seriousness of the situation. It is not really innovative, but can be considered as relatively interesting and not commonplace; this system has potential and could be improved by complementary studies, for example, by looking at the appearance frequency of the alarm.

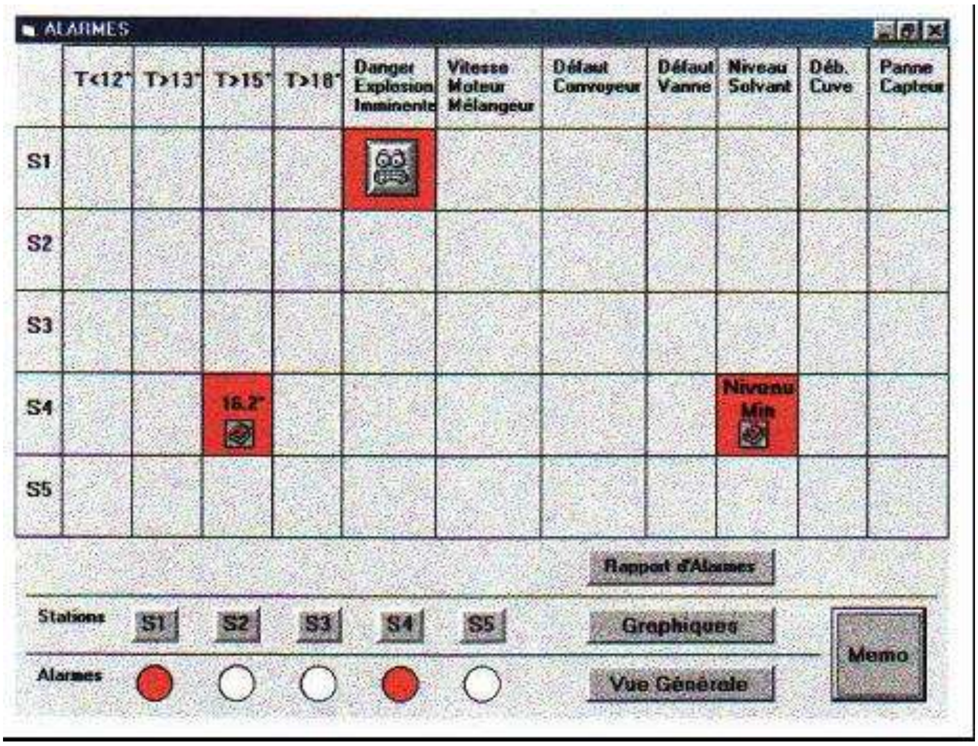


Fig. 7. Example of a relatively original screen shot related to an alarm management system (proposed by one group, scanned from their report)

3.1.5 Diversity of the solutions proposed

While the groups work in the same conditions, have same information and the same documents, it is sometimes surprising to see the diversity of the solutions that the potential designers in the same group propose after analysis and dialogue. This goes in the direction of the diversity of design approaches highlighted by several authors, for example, Theureau and Jeffroy (1992). Figure 8 shows four very different solutions for the control room installation, in terms of the details required by the specification and the level of comfort provided. Figure 8(a) shows a solution represented in three dimensions. Figure 8(b) shows a textual description. Figure 8(c) and Figure 8(d) show schematic views seen from the top and offer more, or fewer, details.

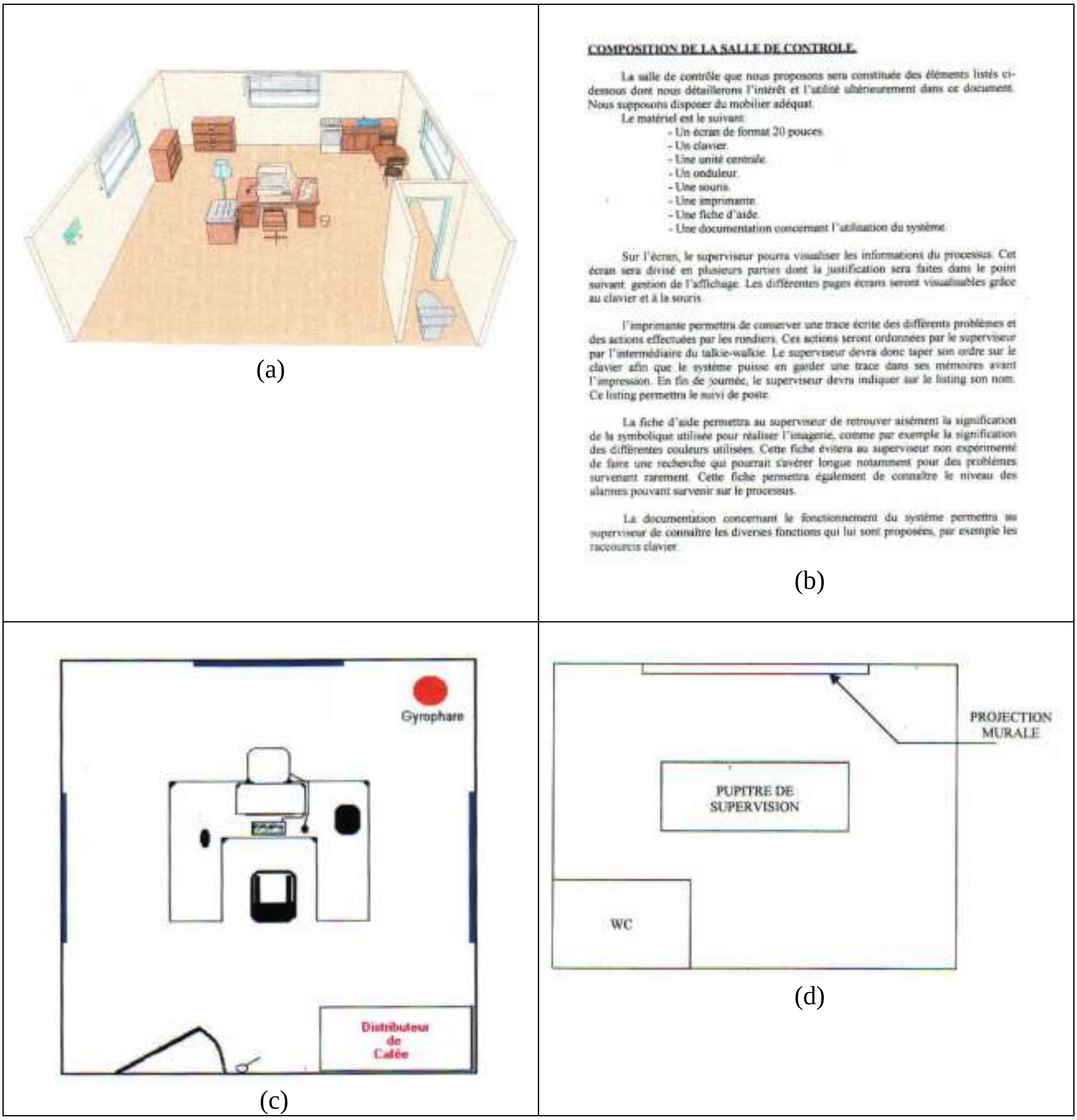


Fig. 8. Four very different proposals for the control room installation

3.2. Detailed analysis

First, we will explain the criteria used for the statistical analysis (§3.2.1). These criteria were applied to the reports of each of the 63 student groups. Then, we will describe the statistical analysis problem (§3.2.2). In the next section (§3.2.3), we will report the analysis results and discuss them.

3.2.1. Student work assessment criteria

The central idea is to analyze the typical errors (e.g., missing information) made by future designers. To achieve this aim, 8 criteria groups have been considered important in this specific application domain: room layout, graphics and representation of the information on screens, the information available on the screens, alarms, communication, time management, linked options and propositions, and suggested work approach. As Table 1 shows, each group is broken up into several criteria for a total of 35 criteria (e.g., 2.1: Is a color code proposed?). Each report is given a grade for each criterion.[3] [4]

The mass of data collected from the reports is extremely large. A first global analysis was carried out; from this analysis, several results can already be highlighted (see section 5).

Table 1. Criteria groups and criteria

Criteria Group	Criteria (questions)
Room layout (Group 1)	1.1. Is the room pleasant to work in? 1.2. Has the lighting system been studied? 1.3. Has the heating system been studied? 1.4. Is the work station suitably situated? 1.5. Does the work station meet ergonomic criteria?
Graphics and representation of the information on screens (Group 2)	2.1. Is a colour code proposed? 2.2. Is the data suitably distributed over the screen? 2.3. Is the process synthesis view correct? 2.4. Is it easy to navigate between screens? 2.5.Is it clear when an alarm appears? 2.6.Is the information suitably represented? 2.7.Does the data make it possible to anticipate future events?
Information available on the screens	3.1.Is all the essential data shown? 3.2.Are all the views represented? 3.3.Is support provided for dynamic use? (e.g., advice on actions, diag-

[3] It is worth noting that the criteria were designed once all the 63 reports had been carefully read (these criteria were not used for a yearly assessment of the reports).

[4] These criteria were defined and proposed based on the experience of the first author (specialized in Human-Computer Interaction) and the third author (specialized in ergonomics) in the field of supervisory HMI design and evaluation. They teach in this area at different levels, and have both been doing so for around twenty years. Their courses are enriched by a set of field studies; thay have published a set of documents in this area, including (Jeffroy and Sagar, 1991, 2003, 2006; Kolski, 1993, 1997; Ezzeddine et al. 2008; Idoughi et al., 2010), while drawing on a rich literature (see references cited in the introduction). These criteria relate mostly to the themes in the module in question (Groups 1 to 7, cf. Table 1), but it is important to note that those in Group 8 (see Table 1) are taught in other courses (eg : courses in ergonomics, safety).

(Group 3)	nostic methods) 3.4.Are support documents provided?
Alarms (Group 4)	4.1.Are the alarms managed or filtered according to their degree of seriousness? 4.2.Are the suggested temperature thresholds appropriate? 4.3.Is an alarm forecast provided? 4.4.Is it possible to cancel the alarms? 4.5.Is the formulation of the alarms understandable?
Communication (Group 5)	5.1.Is the management of the roundsman activity dealt with? 5.2.Is it possible to communicate with the outside? 5.3.Is the data concerning shift changes ergonomic? 5.4.Is the jargon and language used in the company taken into account? 5.5.Is it possible to print data out?
Time management (Group 6)	6.1.Are the date and time shown on all the views? 6.2.Is the time remaining per station shown?
Linked options and propositions (Group 7)	7.1.Has the implantation of new sensors been planned for? 7.2.Is a "recipe" view envisaged? 7.3.Is a "statistical" view envisaged? 7.4.Is it possible to personalize the views on the screens? 7.5.Are back-ups planned for?
Suggested work approach (Group 8)	8.1.Is the approach based on safety and reliability? 8.2.Is the approach ergonomic?

The way that we evaluated these criteria is explained below.

Q1.1. Is the room pleasant to work in?: With this question, we sought to know if the students took into account the fact that the supervision operator spends eight uninterrupted hours in the control room. It is thus necessary to put a minimum of conveniences in the room so that the operator can work correctly. In fact, the concept of user-friendliness is rather vast. For us, this concept encompasses all the elements related to the room installation, particularly the room architecture, the space management, the annexes, the physical ergonomics, the temperature and light in the environment. The expected minimum was thus a WC and a kitchenette. But, on reading various reports, we became aware that our expectations were the strict minimum. Thus, we appreciated the proposal of such things as pleasant colors on the walls, tables, a sufficient and well-arranged space and a comfortable seat.

Q1.2. Has the lighting system been studied?: A good lighting system in work rooms is important, especially when the work is done on a screen. This is particularly the case in a control room where the safety stakes are important. Thus, methods are needed to study the lighting and insure that the light quality is as high as possible for the operator. Therefore, the students had to be capable of proposing a study of the lighting, which in our opinion is a very important subject to insure an effective work environment. Some groups only proposed a very brief outline of the points to be set up; others gave detailed diagrams of the room, with the position of the lights, the windows and the screen compared to various sources of light, for example. The reports were thus assigned a grade according to the depth of the study of the light environment.

Q1.3. Has the heating system been studied?: The study of thermal environment was never dealt with formally in the reports. Nevertheless, some groups thought of the importance of the right temperature in the control room and proposed either a heating system or air-conditioning, and/or an automatic temperature control.

Q1.4. Is the work station suitably situated?: Work station design influences work conditions. To design a work station correctly, the designer needs to think about the activity/ies of those who will use it. To assign a grade for this criterion, we had to determine whether or not the work station location was well positioned in relation to the various light sources. In addition, we wanted to know if the groups had paid attention to the safety of the work station location (e.g., a fire extinguisher in the room, a telephone to call for help) or if they had allowed enough space to move around the room comfortably (e.g., movement limited by the furniture). For example, many people in addition to the supervision operator (e.g., foremen, maintenance workers) need to consult the measurements available in the control room, so it would be a good idea to have enough space for them move comfortably in the control room.

Q1.5. Does the work station meet ergonomic criteria?: There are specific accessories—for example, a desk or table, a comfortable seat, communication devices, a computer and a keyboard—that are intrinsic to a control room. If the control room furniture is not designed to integrate these accessories, they sometimes pile up and are then a source of inconvenience, or the furniture has to be adapted later on. In addition, the work station has to include a command zone, a reading zone, a writing zone and a communication zone. Evaluating this criterion is thus a question of verifying whether or not these dimensional aspects of the work station were taken into account.

Q2.1. Is a color code proposed?: The color code (i.e., color chart) has a major importance in the design of supervisory HCI and must be as coherent as possible with reality. It can be more or less complex depending on the type of system and sometimes on standards. Certain processes require a lot of symbols, and it is important to stay as close as possible to reality to avoid confusion. In the HCI designed by the students, the information that had to be visualized was limited. The groups had to be careful not to overload the displays with color unnecessarily, in order not to overload the short-term memory of the operator. We thus evaluated the similarity of the colors and reality. We also expected the groups to give the meaning of the colors in both the screen page and the reports. In fact, many groups used a color code, but unfortunately no key was given either on the screen page or in the report.

Q2.2. Is the data suitably distributed over the screen?: The goal of this question was to determine if the screen page had been divided into several sections. Given the importance of certain information, it is necessary for the supervisor to have his/her eye on this information constantly, especially for safety reasons. It is appreciated if there is at least one fixed window that constantly displays the significant variables: in the case of the mixing process, these variables are the temperatures S1 and S2, which must not exceed 18°c; at this temperature, there is a great risk of explosion. In addition to these variables, the student reports proposed windows displaying alarms or advice about system control, for example. Some elements of information were more relevant than others.

Q2.3. Is the process synthesis view correct?: The process synthesis view has a central role in the control of the system. With a single glance, it allows the operator to distinguish the most important variables, from both the safety and control perspectives. It is thus important that this synthesis view does not display only the global structure of the system. This view must contain the various temperatures, the remaining processing times and the position of the roundsmen. It is appreciated if the variables are not only displayed in their digital form but, for example, by a histogram that shows the evolution of the variables since such an image can be interpreted more quickly than a number. It is also necessary for the view to show the system's functional state. Most of the time, this was accomplished by using colors.

Q2.4. Is it easy to navigate between screens?: It is of course preferable to click, for example, on one of the stations in the synthesis view to reach its detailed view rather than use keyboard short-cuts. Navigation must be really studied in constant co-operation with the user(s). We expected buttons to reach the various screen pages. In addition, we appreciated the fact that certain groups proposed a diagram presenting the global architecture of the HCI, with the hierarchy of the various screen pages.

Q2.5. Is it clear when an alarm appears?: The most obvious function of an alarm system is to attract the attention of the operator to a modification of the state of the process, even when the operator is not working on the interface. The alarm must "tell" him/her to refer to more precise information so that he/she will be able to diagnose the current problem and then make the appropriate decisions. The student groups practically always proposed a "beep". Is this really a good solution? No, not really. In fact, we find the "beep" solution rather anxiety-producing in the long term. But, this is the most effective solution available at the moment for grabbing the operator's attention when he/she is not in front of the screen. The display resulting from an alarm must also be quickly understandable. To accomplish this, certain groups chose to surround the given station in red. We think that is not enough. Something much more visible on the screen is needed, such as coloring in red of all the symbols representing the station in danger. In addition to the stations' coloring, certain groups proposed a band of text on each screen page, announcing the alarms.

Q2.6. Is the information suitably represented?: For readability, the information on the screen must not only be displayed in its numerical value but also, when it is possible, in the form of small synthetic images, which would better convey the information than a number. Since the system is not very dynamic, we expected such images, which would be easy to integrate into the graphics. We also wanted the system representation to reproduce the process structure in order to avoid any mental overload on the part of the operator (i.e., views built according to same hierarchy).

Q2.7. Does the data make it possible to anticipate future events?: The purpose of this question was to make the students think about evaluating and anticipating the alarms to facilitate the operator's detection of problems. Though the majority of the groups thought of establishing a temperature history for a certain period, others, a more limited number had the excellent idea of representing the temperature change tendency, using the differential calculation of the temperature between the moment (T) and moment (T-1), according to the lapse of time defined by the system's inertia. They simply had to display this tendency on the screen in the form of an arrow whose slope indicates the importance of the evolution. In addition to this, certain groups colored the arrow according to the value of the current temperature. We think that this is a very interesting idea since it is perhaps easier to interpret the temperature in this form, instead of using trends.

Q3.1. Is all the essential data shown?: Among the essential data, we expected to find the temperature of the stations, the time remaining in the process, the position of the roundsmen, the alarms related to the system dysfunction or the dysfunction of one of its components, as well as any irregularity in the speed of the conveyors (i.e., the engine speed). The students proposed this information and other information features as well, for example: the activity of the roundsmen, the state of the valves, the weight of the products on the scales and the process stage at each station.

Q3.2. Are all the views represented?: We expected four main types of views: a synthesis view, detailed views of each station, views of the temperature tendency and a historical view of the alarms. Some groups integrated the station's temperature history directly into the detailed view of the given station. The other views proposed are the subject of the questions in criteria group 7, dealing with linked options and propositions.

Q3.3. Is support provided for dynamic use? (e.g., advice on actions, diagnostic methods): The reports mainly proposed control assistance when the alarms went off: either a specific window appears, mentioning the alarm name, its probable cause and an action to solve the problem, or the same information is displayed in the temperature history view. In addition, some groups proposed a pre-alarm, resulting in advice about better production management or time management, for example.

Q3.4. Are support documents provided?: The support documents provided related to two points. First, it was necessary to provide assistance for using the proposed HCI and not only about the process because some people do not learn to use computer tools easily. Second, certain groups thought of providing help concerning the equipment used by the system. This was in most cases provided in the form of a profile sheet for the equipment, which can both better inform the supervisor about the equipment in order to better guide the roundsmen and help the operator to continue his/her training.

Q4.1. Are the alarms managed or filtered according to their degree of seriousness?: A system can be associated with a large number of alarms. However, some can be much more important than others, especially in relation to safety. Although a small number of groups actually filtered the alarms according to the seriousness of the problem or a set of priorities, it is very important to do this. That is particularly useful when several alarms occur at the same time. An alarm of secondary importance must not hide or interfere with an alarm of primary importance. It is also judicious to highlight the connections between alarms because resolving the problems that lead to the less important alarms may solve the problems that provoke the alarms of greater importance. In our case study, the temperature alarms, the mixing engine speed alarms and the additional solvent alarms could be linked, for example, and some groups did link these alarms. Other groups covered the subject a little differently. Instead of filtering alarms in order of priority, they did it according to their level of importance. Nevertheless, their methods were globally less relevant than showing the connections between the alarms. In addition, they used different colors of text or background to show the seriousness of the alarm in the alarm history display. In our opinion, this is not sufficient for real alarm management.

Q4.2. Are the suggested temperature thresholds appropriate?: With this question, we wanted to see if the subjects paid attention to the various temperature thresholds. Although that seems paramount for control, some groups did not mention anything about this subject. The notation was fine-tuned in different ways. Sometimes, it was possible for the operators themselves to configure the thresholds according to their operational practices. We were looking for threshold accuracy. For example, a pre-alarm at 17°C to prevent approaching 18°C (i.e., the danger zone) would potentially allow the operator to anticipate how the system would evolve.

Q4.3. Is an alarm forecast provided?: Mathematical tools based on statistics, automatic control or artificial intelligence today are able to propose very advanced techniques that make it possible to anticipate future problems. This anticipation function uses database of process knowledge accessed with inference mechanisms. Such tools were not known to all the groups; however, since it is an important aspect for the future of supervised systems, we were pleased to find that some groups proposed such a solution.

Q4.4. Is it possible to cancel the alarms?: Although essential legitimate, the possibility of canceling an alarm is not always recognized by the students as highly important. Not being able to cancel an alarm is practically equivalent to not taking it into account, since it can be forgotten under the pressure of the job and/or when other alarms occur. This question is of major importance for correct process execution. Obviously, it is essential that the alarm cancelling procedure is designed so as not to increase the operator's workload.

Q4.5. Is the formulation of the alarms understandable?: The number of alarms possible in a process can be high, hence the importance of formulating the alarm correctly in order to designate the problem as explicitly as possible. For this question, we graded the reports according to the alarm importance and the need for the precise details given about the formulation of the alarm.

Q5.1. Is the management of the roundsman activity dealt with?: Only five roundsmen are available. In addition, they are not always assigned to a particular station; in fact, they are supposed to move among the five stations. If there is a danger of explosion, a roundsman may need the assistance of another roundsman. Thus, it is necessary to know where the roundsmen are and their availability. For this reason, we expected some information about the roundsmen. The students proposed various possibilities in their reports. The simplest was to post only the number of roundsmen per station. However, with this solution, the supervisor cannot visualize the activity of the roundsmen. Since the supervisor guides the roundsmen, he/she knows the tasks of these roundsmen, but he/she may not remember their activities very long. Accordingly, some groups created pull-down menus containing the various activities that the roundsmen can carry out. Thus, the supervisor just needs to update these menus at each new activity. The supervisor can then visualize the activities of all the roundsmen and can thus choose the roundsman available to help the others when they have difficulties or assign the roundsman differently.

Q5.2. Is it possible to communicate with the outside?: This question complements the question 1.4 (“Is the work station suitably situated”). "Communication with outside" includes several topics. A telephone should obviously be available to call for help (e.g., to call firefighters) if need be. But, communication also includes transmitting information about production to the foremen or the people in charge of production. For this, a fax could be one solution, but another solution would be a computer network that connects the supervision computer directly to the management computers located in other rooms.

Q5.3. Is the data concerning shift changes ergonomic?: We mentioned above that the shift changes three times a day (3 x 8 hours). If a problem begins at the end of a shift, it is important that this information can be transmitted to the supervisor responsible for the next shift. In addition, if a problem occurs during a shift but is not completely resolved or a risk still remains, it is important to be able to transmit this information to the supervisor in the next shift. The various methods used were a memo pad or a post-it on the desk. This is not really a very ergonomic solution because a paper can easily be lost. We expected the “memo pad” to be an integral part of the HCI.

Q5.4. Is the jargon and language used in the company taken into account?: This question relates to the fundamental problems associated to the dialogue between stakeholders at various levels. First of all, there is a language that is specific to each domain and each company, which has to be taken into account. There are also grammatical and spelling rules that it is absolutely essential to respect in each message from the help system. In addition, the messages are addressed to an operator, and this should not be forgotten. One or two groups used some terms in English for a French operator, but the messages were globally understandable.

Q5.5. Is it possible to print data out?: Printing a hard copy remains one of the most common means of keeping the important information for data analysis or for archives. This possibility was found in practically all the reports, even if some groups did not explicitly show it. The groups were evaluated according to the techniques that they used for printing. Among the possibilities was the continuous printing of each alarm or important event, made possible by pressing on a specific button.

Q6.1. Are the date and time shown on all the views?: The date and time are important elements of information in process control. This information makes it possible for operators to orient themselves over time. They can thus note down all the defects that they consider important to archive, even if in theory the supervision software manages these defects locally. We thus expected this information, clearly represented on all the views.

Q6.2. Is the time remaining per station shown?: In time management, it is important to visualize the time remaining in the various processes in order to make the best use of down time. Displaying the remaining time avoids the operator having to calculate time left before reaching the end of the process. We gave good grades to those groups that displayed the remaining time per station but also to groups that gave the duration of the process instead of remaining time. Some groups ignored this subject completely.

Q7.1. Has the implantation of new sensors been planned for?: Since the production system was only partially automated, in order to have access to all the information mentioned above, it was necessary to install new sensors: temperature sensors, tachometers for the various engines, position sensors to determine the position of the valves (i.e., open or closed), flowmeters to measure the flow of the liquid solvent poured into the mixers, the level sensors for the mixers and the solvent tank, smoke detectors. We wanted to test the students' innovation and the imagination on this subject.

Q7.2. Is a "recipe" view envisaged?: Factories like this one are not intended to constantly manufacture the same product. Consequently, the proportions of raw materials are not always the same. As the objective is to control the process, it would be useful to design a “recipe” view making it possible to visualize the raw material proportions for all the products manufactured by the factory. The supervisor, who can visualize the weight of the products poured into the mixer, will thus be able to better guide the roundsmen in order to have the closest possible proportions needed to obtain a quality product. A direct connection to the company's network from the supervisor's computer to a computer in the research department responsible for developing new products could be envisaged. This direct connection would avoid useless movements of the supervisor; the new recipes could be visualized directly in the “recipe” view of the supervision HCI.

Q7.3. Is a "statistical" view envisaged?: A "statistical" view would be mainly used by the maintenance department. This view would make it possible to visualize all alarms and defects that occur during one period, which must be defined. It would be possible to integrate reliability indicators into this view. The statistical results would make it possible to identify the weak points of the process and thus would make it possible to supervise certain parts of the system in particular and to establish action plans for maintenance. These statistics could be coupled to an artificial intelligence module that would determine the moment to act on the process elements. This module would facilitate the physical monitoring of maintenance materials.

Q7.4. Is it possible to personalize the views on the screens?: The personalization of the views on the screen must be considered with attention. In reality, it is not the designer who will use the HCI but the supervisor. What is good for the designer is not necessarily good for the supervisor. The supervisor is probably the person who has the most experience with the system. He/she thus has already his/her preferences, which should not be changed in order to avoid any risk of confusion. For this reason, it is important that the supervisor has the ability to choose the information to be visualized and configure the way of visualizing it. The designer has the technical skills to guide the supervisor to important information and can consequently propose a range of possibilities. Accordingly, the supervisor will formulate his/her choices among the possibilities, in cooperation with the designer. In addition, as stated above, work is carried out in three 8-hour shifts, so there are at least three supervisors likely to use the supervision HCI. This means considering the preferences of three different people, and thus three different visions of the process. All people have their own personal preferences for how things are run, so it is important to be able to change the HCI to coincide with the preference of each supervisor. This can be accomplished, for example, by entering the supervisor's name while connecting to the computer.

<u>Q7.5. Are back-ups planned for?:</u> Data back-ups are necessary to store temperature and alarm information, for example. Since practically all the groups proposed a history of the temperatures in the form of trends and the histories of various alarms that occurred, we deduced that a data back-up function existed implicitly. Some groups envisaged an information storage button for important information that the supervisor considers necessary to access at any moment.

<u>Q8.1. Is the approach based on safety and reliability?:</u> With this question, we sought to verify whether or not the groups thought that safety and reliability were important.

<u>Q8.2. Is the approach ergonomic?:</u> This question was intended to see whether or not ergonomics was the focus of the design approach. For this to be true, it would be necessary to study the real activity of the operators (e.g., human tasks, work conditions). In other words, we wanted to see to what degree the operators had an influence on the design. The answer to this question would tell us whether or not the groups took into account the competences to be analyzed, the characteristics of the activity and the "functioning" of humans in terms of sensory, mental and physical activities.

Given these criteria, the next problem is the way to yield a specific assessment for each student group report. To achieve this aim, the following procedure was used: first it was important to know whether a given topic (e.g. safety and reliability as in Q8.1) was considered, thus involving a no/yes answer; then in the latter case, a rating could be given. Due to the difficulty to assess the student group report, an ordinal scale with few levels was used (instead of a 0-20 continuous scale for instance, the latter being very often used in France). Six levels were used, namely *bad* (0 labeled), *just mentioned but not detailed* (1), *not enough* (2), *average* (3), *good* (4), *very good* (5). Figure 9 shows the assessment procedure and the final scale. Keeping in mind the usual scale models (Stevens, 1974), our scale is rather qualitative, the 1+6 levels yielding 7 ordered modalities.

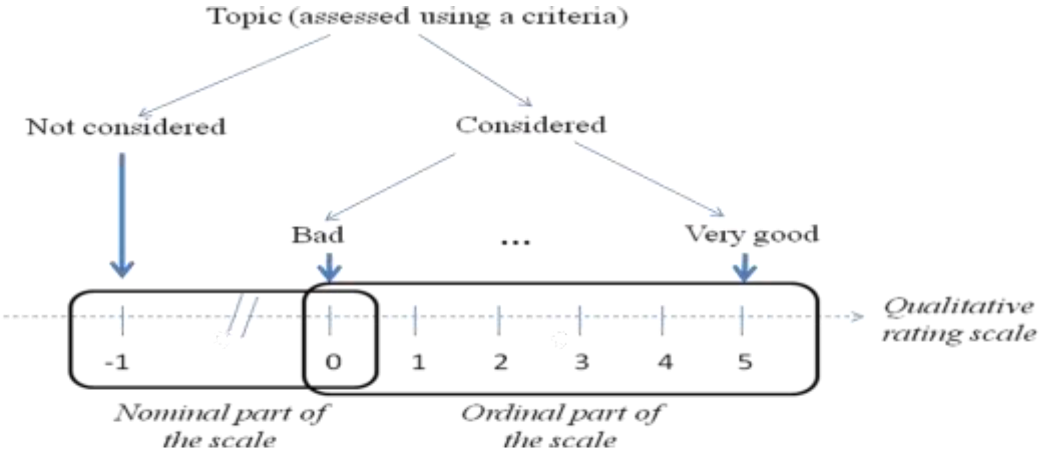


Fig. 9. Assessment scale design

3.2.2. Fundamental problems of statistical analysis

Even though the responses are not so numerous and appear simple, it is not very easy to analyze them. The primary reason is the complexity of rating scale used to grade the reports, figure 9. Another difficulty comes from the high number of criteria (each with its own complex scales) compared with the number of reports (i.e., 35 vs. 63). In addition, the number of reports is quite different from one year to the next (e.g., 11 reports for the first year and 3 for the eighth year). To begin the statistical analysis, let us first evoke various ways, i.e., data analysis paths, (DAP) to solve the stated problems to later select one data analysis path (DAP) that we think provides interesting results.

The first stage of the DAP concerns data coding. Using the usual mathematical scale models (Stevens 1974; Corroyer and Wolff 2003), one way to consider the data could be to simply consider a quantitative scale model, e.g., introducing a 0 value when a topic is not considered and a value from 1 to 6 when a topic is considered. Keeping in mind the properties of the mathematical scale (Stevens 1974), such a coding means that the difference between two near values (0 and 1, 1 and 2, …) has the same meaning. Of course, the individual who grades the reports can make the decision to use this method of coding the data, but any other individual may disagree with this choice. From a statistical perspective, it could be more interesting to consider another method for coding the data.

A second possible method is to choose the ordinal scale model, which is less constraining than the quantitative scale model. With the ordinal scale model, it is only necessary to state that considering a criterion is better than not considering a criterion, even though the way that it is considered is bad. Nevertheless, such a coding method considers the 6 levels of the ordinal scale as having the same importance.

A third option is to first make a distinction between not considered and considered topics and then later include the grade (with its 6 levels, figure 9). This third method for data coding uses a nominal scale model with N modalities, one modality for not considered and N-1 modalities for the grade of the considered case. The maximal value for N is 7. The choice of the data coding method will mainly depend on the histogram of the data (see § 3.2.3).

In addition to this data coding problem, there is the statistical analysis problem. In view of the different ways of looking at data analysis methods (Loslever 2001), there is a choice between performing 35 monovariate analyses vs. one multivariate analysis of the 35 variables and the choice between a descriptive approach, using only the 35*63 grades, vs. an inferential approach, using hypothesis tests. Given the lack of room for presenting the results and the suggestions of many statisticians (Corroyer and Wolff 2003; Saporta 1990), we will focus on a monovariate descriptive analysis, using histograms and variable vs. factor visualizations. Given the rather low number of statistical units (63 reports), the absolute frequencies (from 0 to 63) will be considered instead of relative frequencies (from 0% to 100%). One of the main reasons is that, since we are using a descriptive approach, a relative frequency of 5% is not the same if the total number of occurrences is 63 or 630 (The total number of occurrences considered in our calculation will be always given, which will allow the reader to calculate the percentages if need be.)

In the next section, only statistical results are given. These results will be associated in the discussion to both ergonomic and pedagogical aspects.

3.2.3. Results and discussion

A first result is the rather high frequency of the modality *not considered* (Fig. 10). In terms of the specification, when a point is not considered, this means that this point is not mentioned at all in the student report. For instance, for the "time management" criteria group n°6, criterion 6.1. (Table 1), no HCI showed date and time to the human operator, which is curious since these two dynamic pieces of information have great importance in industrial supervision.

Fig. 10 shows that for topics that have been taken into account (i.e., with a rank from 0 to 5), the overall assessment is rather good: the cumulated frequency for rates from 3 to 5 is much larger than the cumulated frequency for rates from 0 to 2. This constitutes an encouraging result and shows that our future designers were implicated and interested in the project. The overall histogram of Fig. 10 and more local histograms show 1) a high frequency of considered topics (let us remember that our designers are novices, since they are students) and 2) when a topic is treated, the treatment is performed rather in depth.

From an educational point of view, these results show the necessity to focus on topics considered as important (in our case those listed in Table 1) in a domain, during the teaching program. In a project such as this, there is much overall interest from the students in relation to these criteria. It would be interesting to know or to study how this interest is changes (positively or negatively) once they find themselves in a real work situation in companies. It would also be interesting to design systems supporting activities in supervisory HCI design, with an incentive for in-depth consideration of the various criteria relevant to the project in question.

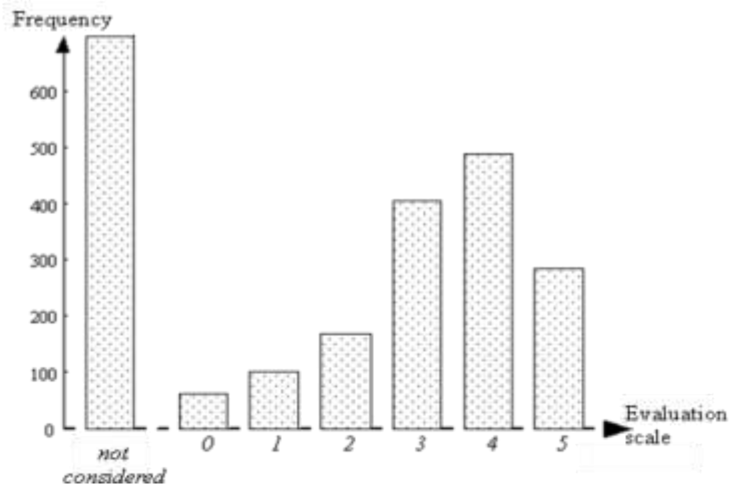
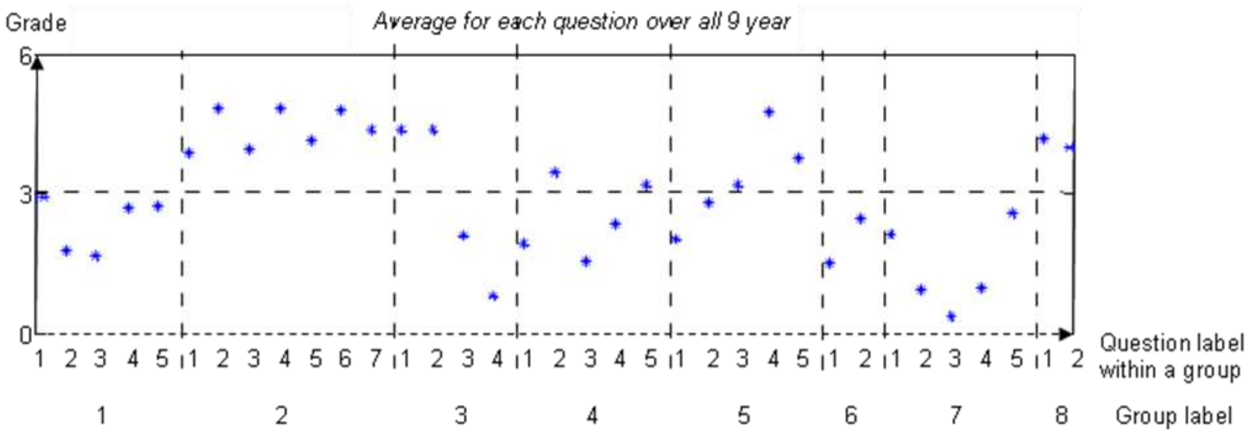


Fig. 10. Histogram computed for all the evaluations (35*63=2205 data)
NB: The horizontal axis is drawn in dashed line to emphasize that this axis represents a qualitative variable.

To have an idea of which criteria give the worst and best ratings, the data corresponding to each of the 35 criteria were averaged. These average values are shown graphically in Figure 11. The main result is that there are large differences between the criteria groups (e.g., group 1 vs. group 2). Furthermore, differences within the criteria groups can be low (e.g., groups 1, 2, 6 and 8) or large (e.g., group 3). If one considers that a well-respected criterion should involve both high values and low differences from one question to another, criteria groups 2 and 8 show the best compromises.

As teachers, we expected that each criterion would be taken into account (or at least considered) and thus yield a rating equal or greater than 3. However, as Figure 11 shows, only criteria groups 2 and 8 were globally rather well considered, and criteria groups 3, 4 and 5 were partially considered. Nonetheless, criteria groups 1, 6 and 7 should have produced better results. The most positive point is that, on average, no criteria group had a really bad score (0 or 1).



(*: Average for the 63 reports)

Fig. 11. Averages for each of the questions of each of the 8 groups

Figure 11 provides interesting results. However, it leaves the *not considered* modality aside, and it is rather biased since the vertical axis corresponds to average values, which could convey the wrong meaning if the *0-5* scale is not quantitative. Figure 12 allows to better maintain the initial structure of the evaluation scale (due to the high number of variables (35), a shades-of-gray graph is used instead of a histogram.) Although this figure is more complex than Figure 11, it shows that the number of occurrences for the *not considered* modality shows large differences from one criteria group to another; however, within one group, the differences are not so large, except for criteria groups 3, 5 and 7. Like Figure 11, this figure also shows that criteria groups 2 and 8 yield the best evaluations.

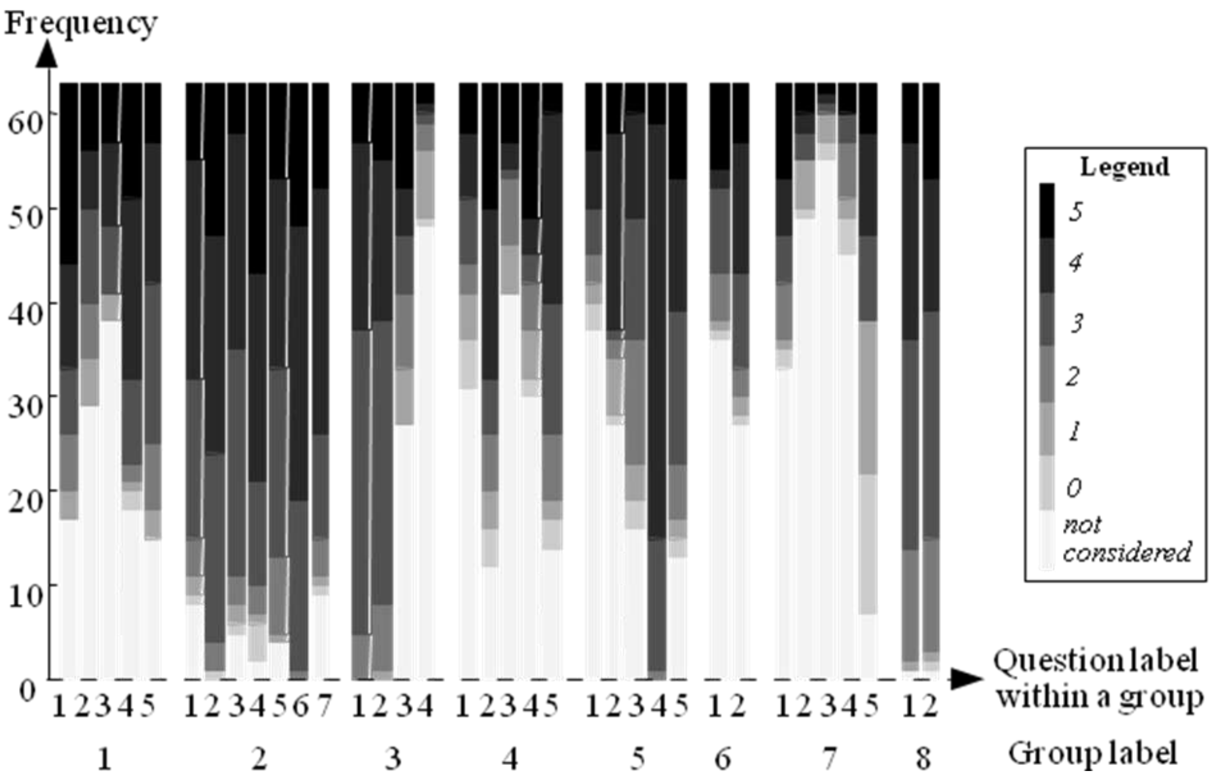


Fig. 12. Frequencies of the modalities, shown using shades-of-gray, for the 35 variables (total effectives for each variable: 63 values)

The results shown in Figures 11 and 12 can be discussed in relation to each of the 8 criteria groups (see 4.1). We start with the criteria groups that yielded globally satisfactory results and finish with the criteria groups that had more important deficiencies:

- Group 2 (Graphics and representation of the information on screens): Figures 11 and 12 show us that the graphics and representation of information on the screens were a key factor for the future designers, and so they obtained excellent results. Indeed, the graphics and screen representations were dealt with in-depth in the course (see 2.2), both from the theoretical point of view and the normative one; many good and bad examples were described. Moreover, in another course (*Ergonomics* module), the teacher had also stressed the importance of graphical aspects of information representation and information in general.

- Group 8 (Suggested work approach): This criteria group was well handled by most groups, which can be explained by the fact that several courses (in addition to the one described in section 2.2; it is the case for instance with the *Software Engineering* course) in the students' program emphasize methodological elements. These elements seemed to interest our future designers very much. If we look closer at the results for both criteria, the same importance was given on one hand to safety and reliability, and partly to the ergonomic approach, which is very positive since the students were free to propose the approach to take, from their point of view.

- Group 3 (Information available on the screens): The essential information, as well as the types of views used to represent these essentials, was for the most part dealt with adequately. On the other hand, both the control assistance and the documentary support were treated too lightly. The results to questions 3.1 and 3.2 show the interest of students in the analysis of important information to be found on the HCI; on the other hand, they have difficulties in moving beyond the supervisory HCI, have not tried to approach different complementary software tools or support systems (bad results to questions 3.3 and 3.4). This lack of interest should be offset by the courses covered in the final year of training (where the concept of the support system in general is presented and discussed).

- Group 1 (Room layout): All the questions were dealt with in roughly the same way, although the theoretical course only mentioned them (see section 2.2). However, the ergonomics course covered them in-depth, which could explain why the students dealt with them like they did. Some student groups went much further in the search for information to deal with the five questions in this criteria group, particularly user-friendliness (1.1), the situation of the work station (1.4) and its ergonomic criteria (1.5). In fact, in Figure 12, we see basically two categories of designers: (1) those who attached little or no importance to the development of the control room (very low marks) and conversely (2) others who, as soon as they started treating this aspect, did so in depth, since they could rely in this case on the *Physical Ergonomics* course (good or very good marks).

- Group 4 (Alarms): The groups just proposed the traditional lists of alarms, with a serious study of only the alarm thresholds (4.2) and their formulation (4.5). On the other hand, crucial questions about process control were often neglected: management and filtering of alarms (4.1), fault prediction (4.3) and even canceling of alarms (4.5), which should have been basic concerns for future designers, were not dealt with seriously, despite the fact that many suggestions had been given about these questions in the theoretical course (see section 2.2). Given the emphasis on the safety and reliability approach (see Group 8, question 8.2), and the fact that the concept of the alarm is fundamental in their training, we can consider this outcome to be disappointing and even ironic, given the dangerousness of the process studied (explosion risk in certain situations).

- Group 5 (Communication): Only the questions related to jargon and language (5.4) were covered in depth. The results for managing the activity of the roundsmen (5.1) and communication with the outside (5.2) were fairly mediocre, although these questions are very important for the specific application in this case study. It is true that these questions were covered only slightly in the theoretical course (see section 2.2). It is equally true that these kinds of information are more visible in real work situations. The results for the problems of shift changes (5.3) (which is in itself a difficult problem, especially in the context of rotating shifts (3 x 8 hours), for example in terms of information to be transmitted to the following team: instructions, situation, steps to follow, probable incidents ...) and printing information (5.5) were much better.

- Group 6 (Time management): it is well known that the time management is an essential factor in process control. Nonetheless, the results obtained for this group were poor, both for the appearance of the date and time on all the views (6.1), which should have been obvious, and the time remaining per station (6.2), which is essential information. Concealing that information is a serious problem of lack of reference with respect to time, while in the process, everything is managed using temporal references (incident management, mixing time, temporal links of the production phases ...); this is also important in relation to the management of the alternating teams (schedules, transmitting of information, instructions ...).

- Group 7 (Linked options and propositions): This group includes complementary aspects useful for control. The results were poor for installation of new sensors (7.1); although these aspects had not been covered in the theoretical course (see section 2.2), they had been covered in another course about sensors. The results were also poor for envisioning a “recipe” view (7.2) and a "statistical" view (7.3) (it gets the worst mark of all the questions from groups 1-

8), which would have been relevant to this case study. As regards 7.2, this was underestimated by the student groups, but was covered in the theoretical course in the first phase of our case study. Regarding 7.3, these future designers underestimated the quantitative aspect of the data, although it is very important for production management, for the quality service and analysis laboratory (aspects that the course had not addressed, or just touched upon very briefly and orally). Another aspect that was covered only slightly in the theoretical course, the personalization of the views (7.4), was also neglected by the future designers (note that it will be addressed in more detail next year in their training). On the other hand, data storage (7.2) was dealt with relatively appropriately.

5. Conclusion

In this article, we focus on the performance of student designers in the specification of supervisory HCI. We describe a simulated work situation in a case study involving an industrial process with five interconnected mixing stations that needs a supervisory HCI. The student designers had to deal with productivity, economic and safety issues for people and installations.

The study data in the form of HCI specifications was very rich. A preliminary global analysis produced several interesting results, for example: many design errors were committed by each of the student groups; the non-exploitation of software prototyping tools influenced the contents of some solutions; synthesis views were proposed systematically by all groups; the designers had great creative potential; and there was great diversity in the solutions proposed.

Following the presentation of this preliminary analysis, we then moved towards a more intensive analysis. We explained the criterion used to evaluate student performance. Sixty-three groups of Master's degree students (3 to 5 students/group, though most groups had 4 members) were involved in this case study. These students can be seen as future designers of supervisory HCI. The results of our second analysis are interesting. They show clearly that, depending on the criterion studied, these potential HCI designers committed many errors; most groups did not address some of the important aspects at all or very little, despite the fact that the importance of these aspects was addressed in the 10-hour theoretical course. Some of these errors are likely to persist in real design situations in companies.

The two conclusions we draw from the study are:

- by bringing together future designers, we can see that from their common work, there are interesting results in terms of overall performance (quality of the reports produced). As stated previously, there are positives and negatives in each report (relating to criteria considered), but an "ideal" report, that is to say, with a score of 4 or 5 for each criterion, has never been produced over the period of 9 years. Thus a report getting a good or very good mark for one criterion or set of criteria, can get average or poor marks for other criteria or criteria groups.

- Different types of solutions can be envisaged to improve the performance of these future designers

  - ✓ Since we can identify shortcomings in relation to a list of criteria, it would be possible to propose software support systems based on such criteria (which may join the research carried out in the 10 or 15 years after 1990 on model-based development approaches (Szekely, 1996) and the "Tools for Working With Guidelines" research way (Vanderdonckt and Farenc, 2000) (including design and evaluation tools dedicated to supervisory HCI (cf. (Kolski and Millot, 1991; Johannsen, 1995; Moussa et al, 2000)), very oriented currently towards Web HCI (Bereikdar et al., 2005). In this case, several categories are possible: they could go from static aids presenting good and bad examples (see also the next proposition), to dynamic aids with critical capabilities, progressively able to produce proposals for designers, to offer suggestions.

- ✓ It would be possible to provide students with reports set in previous years (a set of PDF files or a dedicated web site). These could be discussed broadly with respect to their different strengths and weaknesses (or systematically compared to all the criteria listed, which could increase the teacher's workload, if the number of reports is high every year). Thus, every future designer could learn from mistakes made in previous years. But of course, a problem arises if the cases to study change significantly from year to year.

- ✓ It is also possible to provide the project teams in a learning situation, with real human operators, expressing real needs (this would be possible if new, innovative collaborations between training and production centers,are developed).

- ✓ The ideal situation would be to include this type of approach in the conduct of ergonomic industrial projects with a composition of different members in the team of students, using simulated skills (project manager, developer, ergonomist, graphic designer, production engineer, maintenance manager, roundsman, supervisory operator, etc.). Thus, each team member in a learning situation would have to play his/her role fully in the situation, defending his/her position with solid arguments since they refer both to the role attributed to them and the concern of the service he or she represents in the working group. In this case, we can truly speak of user centered design even though it is clear that the real progression of the project will be more complicated since it must proceed in light of a consensus at decision-making level. That is why a project manager with a strong personality is essential to enforce the decisions with firmness (the learning situation will be different depending on whether the role assigned to the project leader is taken by a teacher or one of the students in the learning process, or even a professional).

Concerning the possible studies, the research perspectives are numerous. It would be possible to conduct several studies of this type, varying the experimental conditions: emphasizing one or more HCI concepts during the theoretical course or changing the subject profile or the complexity of the industrial process, for example. It would be interesting to work with not only groups of designers but also individual designers, and then compare their performance. For the groups, it would be possible to study the aspects related to collaborative work (e.g., communications, contradictory propositions, HCI specification problem-solving). It would be also relevant to compare the results obtained by novice designers like the subjects in this case study with the results obtained by relatively experienced designers.

Concerning the possible evolutions of the university lecture course, it is now possible to consider new developments inherent in new Web technologies, especially Web Services and Service Oriented Architecture (SOA) (Pelz, 2003). It is possible to move towards new concepts of highly distributed HCI, particularly in new supervisory organizations (concept of e-supervision, with strong connections with other services and concepts: e-maintenance, e-logistics ...) involving stakeholders both inside and outside the control rooms, even outside the company (Hou and Su, 2006; Idoughi and Kolski, 2007, Lu et al. 2009; Idoughi et al., 2010). For this reason, the new simulations during the course will involve new organizations (and in consequence the need for new criteria related to these highly distributed organizations), involving adapted HCI (on different interaction supports that can be carried: PDA, smartphone, tablet ...). These new simulations are being studied and implemented.

Acknowledgments

The authors thank Julien Fayolle, Philippe Vellemans and Laurent Wawrzyniak for their contribution in this project. They would like to thank the Nord/Pas-de-Calais Regional Council, the FEDER program, the CISIT project, and the French Ministry of Education, Research and Technology for their financial support of this research. The authors also thank the anonymous reviewers of the journal for their numerous constructive remarks.

REFERENCES

Artman H, Wærn Y (1999) Distributed Cognition in an Emergency Co-ordination Center. *Cognition, Technology & Work,* 1 (4), 237-246

Beirekdar A, Keita M, Noirhomme-Fraiture M, Randolet F, Vanderdonckt J, Mariage C (2005) Flexible Reporting for Automated Usability and Accessibility Evaluation of Web Sites. In MF Costabile, F Paternò (Eds.), *Human-Computer Interaction - INTERACT 2005, IFIP TC13 International Conference*, Rome, Italy, September 12-16, 2005, LNCS 3585, Springer, 2005, 281-294

Cacciabue PC, Amalberti R, Paries J (2005) *Human Error Management in Aviation*. Ashgate Pub Ltd

Coekin JA A (1968) Versatile presentation of parameters for rapid recognition of total state. International *Symposium on Man-Machine Systems, IEEE Conference Record 69 C58-MMS*, September

Corroyer D, Wolff M (2003) *L'analyse statistique des données en psychologie*. Editions Admand Collin, Paris

De Keyser V, Decortis F, Housiaux A, Van Daele V (1987) *Les communications homme-machine dans les systèmes complexes*. Rapport politique scientifique FAST n°8, Université de Liège

Dekker S (2006) *The Field Guide to Understanding Human Error*. Ashgate Publishing

Ezzedine H, Kolski C (2005) Modelling of cognitive activity during normal and abnormal situations using Object Petri Nets, application to a supervision system. *Cognition, Technology & Work*, 7 (3), 167-181

Ezzedine H, Bonte T, Kolski C, Tahon C (2008) Integration of traffic management and traveller information systems: basic principles and case study in intermodal transport system management. *International Journal of Computers, Communications & Control (IJCCC),* 3 (3), 281-294

Gertman DL, Blackman H S (2001) *Human reliability and safety analysis data handbook*. Wiley

Gilmore WE, Gertman DI, Blackman HS (1989) *User-computer interface in process control, a Human Factors Engineering Handbook.* Academic Press

Hewett TT, Baecker R, Card S, Carey T, Gasen J, Mantei M, Perlman G, Strong G, Verplank W (2002) *Curricula for Human-Computer Interaction*. ACM SIGCHI

Hollnagel E (1998) *Cognitive reliability and error analysis method: CREAM*. Elsevier

Hou J, Su D (2006) Integration of web services technology with business models within the total product design process for supplier selection. *Computers in Industry,* 57 (8–9) 797–808

Idoughi D, Kolski C (2007) Towards web services oriented unified supervisory HCI. In: J Jacko (Ed.), *Human–Computer Interaction, Part IV, HCII 2007,* Lecture Notes in Computer Science (LNCS) 4553, Springer-Verlag, 916–925

Idoughi D, Kerkar M, Kolski C (2010) Towards new web services based supervisory systems in complex industrial organizations: basic principles and case study. *Computers in Industry,* 61 (3), 235-249

Jeffroy F, Sagar M (1991). Analysis of work activity and design of intelligent support system for process control. Y. Queinnec, F. Daniellou (Ed.), *Designing for everyone, Proceedings of the 11th Congress of the International Ergonomics Association*, volume 1, Taylor and Francis, London, 826-829

Johannsen G (1995) Knowledge based design of human-machine interfaces. *Control Engineering Practice,* Vol. 3, pp. 267-273

Johannsen G (2007) From control to cognition: Historical views on human engineering. *Studies in Informatics and Control*, Vol. 16 (4), pp. 379 – 392

Klein G (2006) The strengths and limitations of teams for detecting problems. *Cognition, Technology & Work*, 8 (4), 227-236

Klein G, Pliske R, Crandall B,Woods DD (2005) Problem detection. *Cognition, Technology & Work*, 7 (1), 14-28

Kolski C (1993) *Ingénierie des interfaces homme-machine, conception et évaluation*. Hermès, Paris, ISBN 2-86601-377-8

Kolski C (1997) *Interfaces homme-machine, application aux systèmes industriels complexes* (2nd edition). Editions Hermès, Paris

Kolski C (1991) *Conception d'imagerie de supervision* (in french). Tutorial notes, LAMIH, University of Valenciennes and Hainaut-Cambrésis, France, 117 pages

Kolski C, Millot P (1991) A rule-based approach to the ergonomic "static" evaluation of man-machine graphic interface in industrial processes. *International Journal of Man-Machine Studies,* 35 (5), 657-674

Kolski C, Sagar M, Loslever P (2004a) Experiment based on participative ergonomics and performed as part of a supervision engineers' training course. In E.F. Fallon, W. Karwowski (Eds.), *Proceedings HAAMAHA'2004 "Human and Organisational Issues in the Digital Enterprise" (25-27 August), 1,* National University of Ireland, Galway, 336-346

Kolski C, Sagar M, Loslever P (2004b) Spécification d'IHM dans les systèmes critiques : retour d'expérience sur une pratique en enseignement de l'IHM. *Proceedings of IHM 2004, International Conference Proceedings Series,* ACM Press, Namur, 157-164

Kontogiannis T (2010) A contemporary view of organizational safety: variability and interactions of organizational processes. *Cognition, Technology & Work,* 12 (4), 231-249.

Lebart L, Morineau A., Piron M. (1995) *Statistique exploratoire multidimensionnelle,* Editions Dunod, Paris

Lejon JC (1991) *L'évolution de la conduite sur SNCC : l'ergonomie des systèmes numériques de contrôle commande*. Editions ANACT, Paris

Lind M (2003) Making sense of the abstraction hierarchy in the power plant domain. *Cognition, Technology & Work*, 5 (2), 67-81

Liu Q, Nakata K, Furuta K (2004) Making control systems visible. *Cognition, Technology & Work*, 6 (2), 87-106

Loslever P (2001) Obtaining information from data statistical analysis in human component studies. (I). Methods and performances. *Information science*, 132, 133-156

Lu Y, Zhang L, Sun J (2009) Task-activity based access control for process collaboration environments. *Computers in Industry*, 60, 403–415.

Millot P (1998) *Supervision des procédés automatisés et ergonomie*. Editions Hermes, Paris

Moray N (1997) Human factors in process control. In *Handbook of human factors and ergonomics*, G. Salvendy (Ed.), John Wiley & Sons, 1944-1971

Moussa F, Kolski C, Riahi M (2000. A model based approach to semi-automated user interface generation for process control interactive applications. *Interacting with Computers*, 12 (3), pp. 279-292.

Nicolet JL, Carnino A, Wanner JC (1990) Catastrophes ? Non merci ! La prévention des risques technologiques et humains. Masson, Paris

Nonose K, Kanno T, Furuta K (2010) An evaluation method of team situation awareness based on mutual belief. *Cognition, Technology & Work*, 12 (1) 31-40.

Norman D.A (1986) Cognitive engineering. In D. A. Norman et S. W Draper (Eds.), *User centred system design: new perspectives on human computer interaction*, Erlbaum, Hillsdale, NJ, 31-61

Peltz C. (2003) Web services orchestration and choreography. *Computer*, 36 (10), 46–52

Petersen J (2004) Control situations in supervisory control. *Cognition, Technology & Work*, 6 (4), 266-274

Piccini M (2002) Human Factors in the Design of Supervisory Control Systems and Human–Machine Interfaces for Highly Automated Complex Systems. *Cognition, Technology & Work*, 4 (4), 256-271

Rasmussen J (1986) *Information processing and human-machine interaction, an approach to cognitive engineering*. Elsevier Science Publishing

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

Rouse WB (1983) Models of human problem solving: detection, diagnosis and compensation for system failures. *Automatica*, 19 (6), 613-625

Sagar M (2003) Maquettage d'interface homme-machine selon trois points de vue dans un contexte d'imagerie de supervision. K. Zreik, E. Brangier, C. Kolski (Ed.), *Proceedings of IHM 2003, International Conference Proceedings Series*, ACM Press, Caen, 190-197

Saporta G (1990) *Probabilités, analyse de données et statistique*. Editions Technip, Paris

Senders J, Moray N (1991) *Human error: Cause, prediction, and reduction*. Lawrence Erlbaum Associates

Sheridan TB (1992) *Telerobotics, Automation, and Human Supervisory Control*. The MIT Press, Cambridge

Sheridan TB (2002) *Humans and Automation: System Design and Research Issues*. Wiley-Interscience

Stevens SS (1974) *Scaling: a sourcebook for behavioral scientists*. Aldine Publishing Co., Chicago

Szekely P (1996) Retrospective and challenges for model-based interface development. In *Proceedings of Second International Workshop on Computer-Aided interface development CADUI'96*, J. Vanderdonckt (Ed.), Presses Universitaires de Namur, Namur, pp. xxi-xliv

Theureau J, Jeffroy F (1992) *Ergonomie des situations informatisées : la conception centrée sur le cours d'action des utilisateurs*. Toulouse: Octares

Van Daele A, De Keyser V (1990) Distributed decision-making and time in the continuous processes. In J. Rasmussen, B. Brehmer & J. Leplat (Eds), *Distributed decision making and cooperative work*. London: Wiley and Sons

Vanderdonckt J, Farenc C (Eds.) (2000) *Tools for Working With Guidelines*. Springer Verlag, ISBN: 1-852233-355-3

Wallace B, Ross A (2006) *Beyond human error*. CRC Press

Winograd T, Flores T (1986) *Understanding Computers and Cognition: A New Foundation for Design*, Ablex Publishing Corporation, Norwood, NJ