



**HAL**  
open science

# Human-computer interface evaluation in industrial complex systems: a review of usable techniques

Martial Grislin, Christophe Kolski, Jean-Claude Angué

► **To cite this version:**

Martial Grislin, Christophe Kolski, Jean-Claude Angué. Human-computer interface evaluation in industrial complex systems: a review of usable techniques. 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems, Jun 1995, Cambridge, United States. pp. 461-466. hal-03419771

**HAL Id: hal-03419771**

**<https://uphf.hal.science/hal-03419771>**

Submitted on 14 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## HUMAN-COMPUTER INTERFACE EVALUATION IN INDUSTRIAL COMPLEX SYSTEMS: A REVIEW OF USABLE TECHNIQUES

Martial GRISLIN, Christophe KOLSKI and Jean-Claude ANGUE

LAMIH, URA CNRS 1775, University of Valenciennes  
BP 311 - 59304 Valenciennes, FRANCE  
e-mail: grislin@univ-valenciennes.fr

**Abstract:** In industrial complex systems, a human-computer interface can be composed with hundreds of graphic screens and thousand variables. Human tasks are complex; HCI evaluation in this application field is difficult. During many interface developments, evaluation is often neglected by lack of knowledge about the existing techniques. Several techniques exist and may be used. The paper is focused on expert-based and analytical evaluation techniques usable in this field.

**Keywords:** Human-Machine Interface, Evaluation, Industrial complex systems.

### 1. INTRODUCTION

Contemporary literature mentions many techniques and tools for the evaluation phase—a crucial problem for the development team. Indeed many human-computer interface evaluation techniques are available today, and it is difficult to choose those really suitable to a particular situation. Moreover, their use is a long and difficult task requiring expertise and patience. In this context, our objective consists in studying and organizing the techniques that can be used for supervisory control systems with several human, ecological, economical and/or technical constraints. We are interested in graphical human-machine interface used in a control room to resolve complex problems in a dynamic situation. In the first part, this paper describes briefly the evaluation principle and properties, and the usability target data to be collected. In the second part, on the basis of a review of evaluation techniques, a set of usable human-machine interface evaluation techniques in industrial complex systems are identified and commented. This paper is focused on expert-based and analytical approaches.

### 2. EVALUATION PRINCIPLE

Usability evaluation aims at assessing the operator's

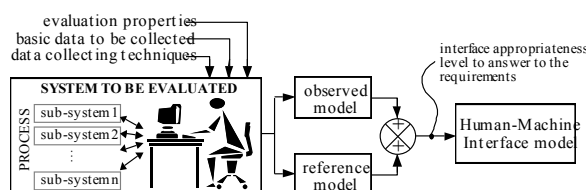


Fig. 1. Evaluation principle

ability to carry out his task. Senach (1990) explains that every evaluation consists in comparing an observed model of the system to be evaluated with a 'reference model' (fig. 1). In interactive software evaluation, the reference model can be based on usability standards or guidelines. During the development of HCI for the supervisory systems this evaluation based on a comparison is not easy. The two models (the observed one and the reference) provide a global model of the human-machine interaction. The *observed model* comes from design methods or from empirical evaluation. The *reference model* comes from task analysis methods, i.e., human oriented design. This evaluation principle helps the development team and enables to take into account the usability guidelines during the cycle. The realization and the validation of used models are the main difficulty in providing good evaluation. These descriptions are formal models of the dialogue between the human supervisor and the machine. Another important difficulty concerns the availability model: formalizing the man-machine interaction is not easy and there is often a gap between theory and reality. The problem is to know whether one can trust an evaluation based on a representation that may not reflect reality. The choice of a HMI model for evaluation is determined by the context. Thus the selection of adequate properties to evaluate quality depends not only on the evaluation objectives but also on the operators' characteristics and on the task requirements. The range of functions and the user-friendliness of an interface are significant properties, defined in the terms *utility* and *usability*. *Utility* determines whether the interface enables the

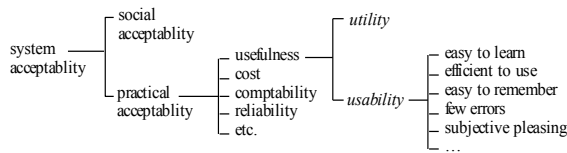


Fig. 2. Principal evaluation data (Nielsen 1993)

user to achieve his goals, in terms of functional abilities, performance and aid quality. *Usability* describes the quality of the interaction between human and machine (ease of learning and use) and reflects the consistency of the design.

An analysis of the interface and the interaction brings to light a set of target variables, i.e., basic data to be collected. There are two families of principal

**Table 1 Evaluation usability criteria, basic criteria in bold and numbered (Bastien and Scapin 1993)**

Criteria	Principle and objective
Guidance C1 C2 C3 C4 C5	Guidance facilitates discovery and usage by helping the user to identify further states and actions possible. <b>Prompting (C1)</b> : provide guidance about what the user can do. <b>Grouping of items by localization (C2)</b> : relationships between displayed information (organizing items into hierarchic lists, distinct areas, etc.), and <b>Grouping of items by format (C3)</b> : distinguishing between functional areas, distinctive label of categories, color coding, etc. <b>Immediate feedback (C4)</b> : results of users' action. <b>Clarity (C5)</b> : facilitate legibility
User workload C6 C7 C8	Workload concerns all interface elements that play a role in the reduction of the users' perceptual or cognitive load, and in the increase of the dialogue efficiency. This criterion includes two first areas. <b>Concision (C6)</b> : individual items (data, commands, etc.) use the shortest possible items to be entered by the user (short items); and <b>Minimal actions (C7)</b> : set of items (procedures) for which it is suggested to allow only the shortest, most necessary action. The third area is the <b>Mental load (C8)</b> or information density. It concerns the users' workload from a perceptual and cognitive point of view regarding the whole set of information presented.
Explicit control C9 C10	Control in user entries favors focusing on the user's actions and therefore limits ambiguities and errors. Two sub criteria are defined. <b>Explicit actions (C9)</b> : unambiguous user specification of an action or set of action; and <b>User control (C10)</b> : allow processing of only what the user is asking for.
Adaptability C11 C12	The adaptability of the system refers to its capability to behave contextual and according to the users' needs and preferences: provide various ways for the user to interact. One general aim is <b>Flexibility (C11)</b> : the capability of the interface to adapt to various user actions. One more specific aim is flexibility applied to users' experience, that is, <b>Users' experience management (C12)</b> .
Error management C13 C14 C15	The aim is to avoid errors as much as possible. When errors occur, identify the error(s), its location, and means to correct the error(s). <b>Error protection (C13)</b> : means available to detect and prevent data entry errors, command errors, or actions with destructive consequences. <b>Quality of error messages (C14)</b> : phrasing and content of error message, their relevance, readability, and specificity about the nature of the errors (syntax, format) and the actions needed to correct them. <b>Error correction (C15)</b> : means available to the users to correct their errors.
<b>Consistency (C16)</b>	Consistency refers to the way interface design choice (codes, naming, formats, procedures, etc.) are maintained in similar contexts, and are different when applied to different contexts. It concerns location (similar window location, message area location), format (similar screen formats), syntax, procedures (similar order in command language syntax, and procedures for menu selection), naming (similar naming of commands or menu option, abbreviations, and labels).
<b>Significance of codes (C17)</b>	It concerns the association between an individual term and/or sign (label, command, etc.) with its reference. Codes and names are significant to the users when there is a strong semantic relationship between such codes and the items or actions they refer to. Not the coding related to the distinction or grouping of several items belonging to the same class.
<b>Compatibility (C18)</b>	Compatibility refers to the match between users' characteristics (memory, perception, customs, skills, age, expectations, etc.) and task characteristic at the one hand, and the organization of the output, input, and dialogue for given application, on the other hand. It's also concerns the coherence between environments and between applications.

variables: a social and a practical acceptability (fig. 2). This second class includes current production, cost, reliability constraints, and usefulness of interface. *Utility* and more particularly *usability* is located in this criterion. Contemporary literature offers numerous usability and ergonomic recommendations: guidelines or standards. They can be used during evaluation as *usability* criteria to observe or measure. Relying on several works, Bastien and Scapin (1993) discussed a set of criteria and sub-criteria that take into account such recommendations (table 1). The whole criteria help the evaluation team to estimate the interface usability quality and to make the necessary decision concerning modification and/or improvement.

### 3. REVIEW OF USABLE EVALUATION TECHNIQUES

HCI evaluation depends on the characteristics of the situation to be evaluated. When the interface to be evaluated exists, user behavior data analysis methods are recommended. Data are collected with observations or measures on real site or on laboratory. Those methods consist in realizing usability tests with real users or selected subjects for experiment. This evaluation type is realized *a posteriori* on an already used interface or a prototype. It provide opinion about their use. In this first case, several methods and techniques exist. They can be classified into three approaches (fig. 3): (1) empirical approaches, (2) observing approaches, (3) expert-based approaches. When the interface to be evaluated does not exist, analytical methods can be used. They allow an *a priori* analysis, i.e. provide an interface valuation before his real use. In this second case, the corresponding techniques can be classed into two approaches (fig. 3): (3) some expert-based approaches (already cited above), (4) analytical approaches. Analytical evaluation allows to describe an interface that responds to the operator requirement in terms of utility and usability. It takes into account the human oriented design recommendation and can be supported by an appropriate architecture of human-machine dialogue. Those methods can be used during interface specification and design, to check *a priori* the proposed interface usability quality. The usable

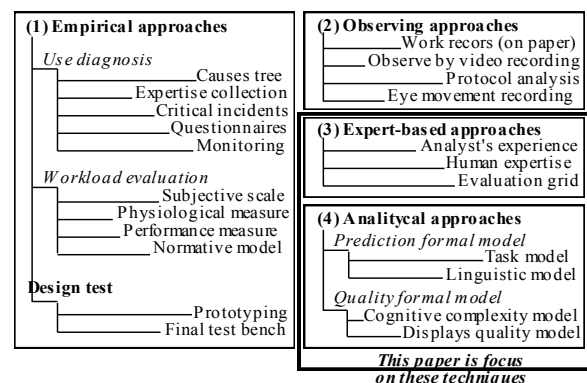


Fig. 3. HCI evaluation techniques classification

techniques or tools for human-computer interface evaluation come from various fields, such as software engineering, knowledge elicitation, ergonomics studies or psychology. Many methods exist. This classification synthesizes some evaluation studies (Maguire and Sweeney, 1989; Senach, 1990; Whitefield *et al.*, 1991; Balbo and Coutaz 1992), with an update. We take into account only those usable for complex systems. We used those techniques currently in real and/or simulated industrial situation (chemical process, nuclear, automobile and rail transport, submarine...). This paper is focused on our last research in evaluation: an expert-based and an analytical approach. The next sub-section gives a short discussion about the techniques of other approaches (for more details, see Kolski, 1993).

### 3.1 Empirical and observing approaches

Empirical approaches consist in measuring representative users' performance with experimental methods and tools. The information from *Use diagnosis* technique facilitates design and improvement decision making. But, they do not take into account the temporal evolution of the system and the identification of the failure sequence is difficult for complex interactive systems. The *Workload evaluation* is based on the user's cognitive work observation. This qualitative or quantitative measuring of the user's activity contributes to the usability evaluation in line of the interactive system. Several methods and techniques are proposed (Wierwille and Eggeneier, 1993). The analysis of those methods leads to the conclusion that workload evaluation may be based on the use of several indicators: (i) a description of the user's work gives information about their different strategies, (ii) a subjective scale give a reference evaluation of the mental workload, and (iii) a physiological indicator informs of the user's capability and his vigilance level. Several recent works, in the LAMIH, devote an important development in a temporal model based method (Berger *et al.* 1989; Riera and Grislin 1995), and an eye scanning movement measure (Simon *et al.* 1993). The *Design test* methods realize an evaluation with an iterative cycle during the whole design process. At every development step, the interface is tested which generate modification, then tested in a new version, and so on until it is satisfactory. This evaluation engineering is very interesting for process control. It permits to take automatically into account the criteria and constraint related to the process control (production, economy...). But, those methods imply

the use of a powerful and realist simulator, that is very expensive. *Observing approaches* group methods only based on the observation of specific aspects concerning the HCI. It consists in collecting representative data of the interaction, and to analyze those user activity traces. The collected data describe the event occurring during the user's activity: failures, orders modifications, stopping, takeoff, errors and correction actions, etc. In the supervisory of an industrial complex system with graphical aids tools, the operator uses mainly his visual sense. In this case, analysis of eye scanning movement is logically integrated in an evaluation process. This technique is usable in parallel of other analysis tools. Eye movement record is used in the LAMIH for the task analysis in the interface specification and/or improvement (Abed, 90).

### 3.2 Expert-based approaches

More often, the HCI evaluation requires an expert analysis. Expert-based approaches consist in comparing the interface with a human-computer interaction informal model (table 2). The **Analyst's experience** consists for the analyst in replacing the operator in his job. This technique let us see the difficulties met by the user. It is recommended for a simple task (office automation), but can rarely used when the task is complex. In a control room, this method raises a huge problem: it is not possible for the evaluator to use the supervisory system in line. The type of evaluation more currently used is the **Human expertise**. It is efficient and simple to organize for a rapid diagnosis of design errors. Every criterion is concerned as human-factor experts have to be able to use them. But this type of method is based on an individual reference model acquired with experience. Every specialist is focused on specific characteristic of the interface and bases his evaluation on his personal model. To get a more exhaustive approach of the HCI problems and to avoid a focusing error, it is necessary to take several different specialists, and therefore it implies an important cost. the **Evaluation grid** is very interesting for the measurement of the interface usability. This technique has to take into account every characteristic of the interface (an exhaustive list of criteria). It obtains an excellent result for a comparative evaluation of interfaces and prototypes. A, evaluation grid is usable by designers, human factor specialists, and sometimes by the final users. But they are difficult to understand for non-expert of human factors: recommendations are subjective and/or can be interpreted differently, and the notation is difficult to analyze. The grid is more adapted for a human-factor expert because they have already a mental model of the user's work.

### 3.3 Analytical approaches

Analytical approaches consist in realizing the

Table 2 Expert-based approaches

Techniques	References	Usability criteria
Analyst's experience	Sperandio, 1991	
Human expertise	Molich <i>et al.</i> 90	every criteria
Evaluation grid	Smith and Mosier, 1986; Ravden and Johnson, 1989	every criteria

evaluation with a model of the interface or of the human-computer interaction. So, it is necessary to identify the pertinent variable, according to the evaluation objective and the task requirement (utility and usability). Then, the collected data analysis require measurement scale to be used during the comparison (a, exploitable and objective notation). Results may be integrated into a global appreciation of HCI. Those approaches use formal models and objective metrics. The models are used to predict some aspects of the interaction: task hierarchy, developing of the user action, time... The metrics allow to measure with objectivity some quality aspect of the interaction: consistency, compatibility with the mental scheme of the user, display quality... Those abstract representations allow to predict the performance of the system, which cannot be obtained with empirical approaches.

*Predictive formal model:* The human task modeling is particularly important if it facilitates the communication between the different agents of the development team: showing the user's assistance requirements, choosing the different aid's function... The modeling can be realized with different type of description. That model is based on a goal, sub-goal description (fig. 4). Table 3 lists the most known ones among those which are usable for interactive applications.

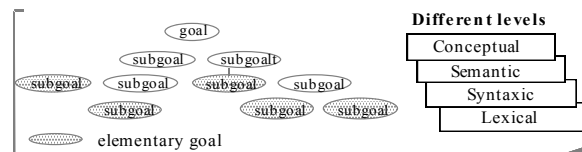


Fig. 4. Task description as a goal sub-goal hierarchic

*Task model.* The task models consist in a hierarchical decomposition of task in order to formalize the cognitive activity of the user. In this type of model, the task is decomposed through a refining process starting from global task or from elementary task (descending and ascending analysis). The analysis provides a formal support for the predictive evaluation of the performance (a measurable description of the user's behavior). **HTA** is a hierarchical model of the user's task. Every task corresponding to a planning, a sequence (or a procedure), can be described for example with a petri net and execution conditions. **GOMS** predicts the length of the task realization, and describes the choice of methods. With its formal approach, GOMS has opened some interesting perspectives and has lead to defining of other models (e.g., KLM). The first version of that model was validated for very simple tasks (office automation). But this model is limited because it takes into account neither the user's errors, nor the resolution of the problem. This limitation does not allow its direct using for the prediction of the user's efficiency when he realizes a complex task. But we can note that a

Table 3 Example of predictive formal model: task and linguistic model

Techniques	References	Usability criteria
HTA	Annet <i>et al.</i> , 1971; Stammer <i>et al.</i> 1990	C1 C7 C8 C11 C12
GOMS KLM	Card <i>et al.</i> , 1983; Irving <i>et al.</i> , 1994	C1 C7 C8 C11 C12
MAD	Scapin and Pierret-Golbreich, 1989	C1 C7 C8 C11 C12
SADT-Petri or TOOD	Abed, 1990; Mahfoudi <i>et al.</i> , 1995	SADT: C1 C7 C8 C11 C12 Petri: C4 C9 C10 C13
ALG	Reisner, 1984	C1 C7 C8 C11 C12
CLG	Moran, 1981	C1 C4 C8 C9 C10 C12 C13 C15 C16 C18
TAG et ETAG	Payne and Green, 1986; Tauber, 1990	C1 C7 C8 C11 C12

recent validation of highly interactive behaviors has been realized with an extended model tell CPM-GOMS based on a critical path method (Gray *et al.* 1993). The INRIA laboratory has developed the **MAD** method to facilitate the user's task description. The obtained task model shows the difference between functional logic and using logic. The user's characteristic is taken into account rather than the application functioning logic and than the computer constraint. This approach is very interesting, but there are some limitations. The interruption of the activity is not visible in the description The human errors and/or procedures are not taken into account in the description. Note that this formalism may be improved in the new version of MAD. Therefore some research in the LAMIH has developed a method based on **SADT and Petri networks**. In this method, SADT formalism realizes the functioning decomposition with a top down and a hierarchical point of view. Every module was structured according to the SADT diagram decomposition. In the end of these functioning decomposition phases, the decision behavior is described using synchronized petri network. The obtained model is next used during the usability evaluation for the user's activity analysis. This method has recently evolved with TOOD: Task Object Oriented Description (Mahfoudhi *et al.* 1995).

*Linguistic model:* The linguistic model can describe the interactive task between the user and his interface, in order to evaluate the command language and its consistency. Those models explain the human-machine interaction structure with a grammar. For instance, **ALG** proposes to describe the interface procedures used with production rules. But with the hierarchical description only, this grammar neither shows the relationship between the display, nor the dynamic of dialogue. **CLG** is a design oriented approach, usable also for the evaluation. The task description is transformed into a set of models realized on six levels: task and semantic for goals and sub-goals, syntax and interaction for the communication between the user and the system, presentation and input-output unit for physical component. This formalism allows to locate where the evaluation decision was critical for the user. For example: the interface efficiency is measured with an

estimation of the action speed or the apprenticeship duration. **TAG** model describes the mental representation of the interaction language. The task is broken up into simple routine sub-tasks (or action) which the user can execute without decision making. But TAG is limited to simple tasks, i.e., the description of specific system function. There is no relationship between the different simple tasks into the upper levels.

*Conclusion.* This hierarchical description of the task in elementary sub-tasks shows the abilities and the information necessary to the user. Tasks are structurally split up to describe the parallelism and the sequential ordering of the procedure. Today, this type of model is in evolution and can be used for design test in the place of empirical studies with prototype. But they limit the evaluation because they predict only the user's performance corresponding to the defined or analyzed model. For a simple task, the linguistic models give a good result. But after our first tests for interface evaluation in a complex industrial system control, those models don't give any more information than task model. The hierarchical aspect of the interaction is to be kept. The description of the human-machine dialogue as a software language (interesting for automation) help to detect some problem: consistency, interaction difficulty, compatibility of the apprenticeship. But some research might be developed to make them valid.

*Quality formal model:* The quality formal models concern the measurement properties of the interface in a more practical view than the above model. They are currently based on a usability and ergonomic knowledge base. Table 4 lists some of them, non-exhaustively. The display evaluation of interactive system is based on the principle that some problems occurring during the use can be avoided *a priori*. The first type of automatic evaluation systems can analyze static information presentation. The **Display analysis program** measures the quality of alphanumeric language exclusively. The prototypes **SYNOP** and **KRI/AG** evaluate, respectively, the industrial mimic displays and the dialogue interfaces in a graphical environment. SYNOP modifies directly the screen, and KRI/AG provides a usability recommendation written like comments in textual reports. These two tools are limited: display assessment one to one without global view, non-

Table 4 Example of quality formal model:  
display quality model and automatic generation

<i>Techniques</i>	<i>Reference</i>	<i>Usability criteria</i>
Display Analysis Program	Tullis, 1986	C2 C3 C5 C8
SYNOP	Kolski, and Millot, 1991	C2 C3 C5 C8 C16
KRI/AG	Lowgren and Nordqvist, 1990	C2 C3 C5 C8 C16
APT	Mackinlay, 1986	C2 C3 C5 C6 C8 C16 C18
Data structure	Perlman, 1987	C2 C3 C5 C6 C8 C16 C18
Ergo-conceptor	Kolski and Moussa, 1991	C2 C3 C5 C6 C8 C16 C18

fulfilment of the knowledge base. KRI/AG system, more recent, interprets the standard UIL format (User Interface Language). This characteristic is very interesting for an automatic evaluation, and can be integrated in a future evaluation tool. This first type of approach shows the possibility to automate the many design guidelines currently proposed in a daunting presentation (manual and guide on paper).

Other automatic evaluation system based on knowledge base are able to generate the interface or specification. From a description of the system, that method aims at obtaining a display respecting basic recommendation. **APT** is too limited for industrial interface design because it can only represent and test static numeric data. The Perlman program can describe the display structure in different abstraction levels. The aspect link to the presentation are formalizes as a network, with relationship between the display structural properties and presentation rules. This type of approach is too limited as a design help of the aspect based on the presentation, without analyzing the use, nor the dynamic. Today, the automatic interface generator might concern the dynamic evaluation of the HCI. **Ergo-conceptor**, developed on the LAMIH, is based on the description of the application (i.e., an industrial system to be supervised). This tool assists the designer with some proposition of representation mode, supposedly adapted to the control task. The designer can choose directly that prescription and generates interactively the different mimic displays. This type of evaluation systems are interesting because they can automate several designs and realization steps. But there are also some problem: non-fulfilment of the knowledge base, difficulty to realize a good description of the application.

*Conclusion:* Nowadays some researches are still carried out about some of the models which can not directly be used for an evaluation in an industrial situation. Their using demands an important and long formalism work and very little information was given about the method to realize the test during design. An other important point is that an evaluation tool of HCI can not be limited to an evaluation of exclusively every usability aspect. It might estimated the advantages and weaknesses. The analysis might concerns the input and output command language, and the screen organization. The different works presented in this section suggest that it is difficult to develop a tool able to evaluate every interface. A really satisfactory tool must be constituted on an evaluation environment.

#### 4. CONCLUSION

Many HCI evaluation methods and techniques exist. Therefore some of them are usable for human-machine interface used in simple task (e.g., office automation).

For several years, we have been testing some of those methods in industrial situations. The empirical, observing and expert-based approaches give relatively good results (of course some adjusting is still needed). The realization of a controlled experiment and the recording, and then the data analyses take a very long time. It is currently non-compatible with the time requirement and delay of the system development. And moreover, the final users are rarely available or accessible, and the designers are not already convinced of the utility of those early-tests. In fact, the evaluation is considered more as a corrective function than as a generative information source for a design solution. Concerning the analytical approaches, some of them raise a problem because of the dynamic of the situations to be evaluated. These are the aim of our present research in complex situations.

## REFERENCES

- Abed M. (1990). *Contribution à la méthodologie de la tâche de conception par des outils de spécification exploitant les mouvements oculaires*. Ph. D. Thesis, Valenciennes.
- Annet J., K.D. Duncan, R.B. Stammers, M.J. Gray (1971). *Task analysis* (HMSO London).
- Balbo S. and J. Coutaz, (1992). *Un pas vers l'évaluation automatique des interfaces homme-machine*. ERGO IA'92, 7-9/10, Biarritz, 367-382.
- Bastien J.M.C. and D.L. Scapin (1993) *Ergonomic criteria for the evaluation of Human-computer Interfaces*. Rapport INRIA n°156, Juin, 79p.
- Berger T., B. Riera, J.C. Angue, P. Millot (1989). *Real time assessment of human operator workload placed in a simulated multitask context of a monitoring and supervising process*. ESA workshop, Human factor engineering: a task-oriented approach, 21-23/11, 13 p.
- Card S., T.P. Moran, A. Newell (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gray W.D., B.E. John, M.E. Atwood (1992). *The precis of project ernestine or an overview of validation of GOMS*. CHI'92, ACM Conference on Human factors in Computing Systems, Bowersfeld et al. (Ed.), Monterey, California, May 3-7, 307-312.
- Irving S., P. Olson, J.E. Irving (1994). *A GOMS analysis of the advanced automated cockpit*. CHI'94, ACM Press, Addison-Wesley.
- Kolski C. (1993). *Ingénierie des interfaces homme-machine, conception et évaluation*. Hermes (Ed.), Paris, 372 p.
- Kolski C. and F. Moussa (1991). Une approche d'intégration de connaissances ergonomiques dans un atelier logiciel de création d'interfaces pour le contrôle de procédé. *Quatrièmes Journées Internationales: le Génie Logiciel et ses applications*, Toulouse, 9-13 décembre.
- Kolski C. and P. Millot (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, 657-674.
- Lowgren J. and T. Nordqvist (1990). *A knowledge-based tool for user interface evaluation and its integration in a UIMS*. In Diaper et al. (Eds.), INTERACT'90, 395-400, North-Holland.
- Mackinlay J. (1986). Automating the design of graphical presentation of relational information. *ACM Transactions on Graphics*, 5 (2), 110-141.
- Maguire M. and M. Sweeney (1989). *System Monitoring: Garbage Generator or Basic for comprehensive Evaluation system?* The fifth conf. of the British Computer Society HCI, Specialist Group, University of Nottingham, september 5-8.
- Mahfoudhi A., M. Abed M., J.C. Angue (1995). *A new methodology for Man-Machine Systems analysis and design Design*. Industrial Ergonomics and Safty Conference, Seattle, Washington USA, June 13-16.
- Molich R., J. Nielsen (1990). Improving A Human Computer Dialogue. *Communications Of The Acm*, 33, 338-348.
- Moran T.P. (1981). The Command Language Grammar: a representation of the user interface of interactive computer systems. *Int. J. of Man-Machine Studies*, 15, 3-50.
- Nielsen J. (1993). *Usability Engineering*. Academic Press, Boston, 358 p.
- Payne S.J., T.R.G. Green (1986). Task-action grammars: A model of the mental representation of task languages. *Human-Computer Interaction*, 2, 93-133.
- Perlman G. (1987). *An axiomatic model of information presentation*. Human Factors Society Meeting, New York, 1219-1223.
- Ravden S. and G. Johnson (1989). *Evaluating Usability Of Human-Computer Interface*. Chichester: Ellis Horwood.
- Reisner P. (1984). Formal grammars as a tool for analysing ease of use: some fundamental concepts. In Thomas and Schneider (Eds). *Human Factors in Computer Systems*. Norwood, N.J.: Ablex.
- Riera B., M. Grislin, P. Millot (1994). *Methodology To Evaluate Man-Car Interfaces*. IFAC/IFORS Transportation Systems: Theory and Application of Advanced Technology, Tianjin, China, August 24-26, 459-469.
- Scapin D.L. and C. Pierret-Golbreich (1989). *MAD : Une méthode analytique des descriptions des tâches*. Colloques sur l'ingenierie des IHM, 131-148, Sophia Antipolis, 24-26 Mai.
- Senach B. (1990). *Évaluation ergonomique des interfaces homme-machine : une revue de la littérature*. Rapport INRIA n°1180, 75 p.
- Simon P., T. Berger, J-C. Angue (1993). *Is quantitative of eyse scanning movement relevant of eye*

- scanning movements/*. International Industrial Ergonomics and Safty Conference'93, Copenhagen, Danemark, june 8-10.
- Smith S.L. and J.N. Mosier (1986). *Guideline for designing user interface software*. Report MTR-10090, ESD-TR-86-278, Bedford, MA: The MITRE Corporation.
- Sperandio J.C. (1991). Les méthodes d'analyse du travail en psychologie ergonomique. In *La recherche en Psychologie (Domaine et méthodes)* de J.P. Rossi, Dunod (Ed.).
- Stammer R.B., M.S. Carey, J.A. Asteley (1990). Task analysis. In Wilson et Corlett (Eds), *Evaluation of human work , a practical ergonomics methodology*. Taylor & Francis Ltd, 135-160.
- Tauber M. (1990). *ETAG: Extended Task Action Grammar-A language for the description of the user's task language*. In Diaper *et al.* (Eds), INTERACT90, 163-168, North-Holland.
- Tullis T.S. (1986). *Display analysis program (Version 4.0)*. Lawrence, Kansas, The Report Store.
- Whitefield A., F. Wilson, J. Dowell (1991). A framework for human factors evaluation. *Behaviour & Information Technology*, vol 10, n°1, january-february, 65-79.
- Wierwille W.W. and F. Eggeneier (1993). Recommendation for mental workload measurement in a test and evaluation environment. *Human Factors*, 35 (2), 263-281.