DGLR

CSAPC `01

8th Conference on Cognitive Science Approaches to Process Control

"The Cognitive Work Process: Automation and Interaction"

24 - 26 September 2001

Universität der Bundeswehr, Neubiberg, Germany

R. Onken, The Editor

An EACE Conference Series

European Association of Cognitive Ergonomics

CSAPC `01

8th Conference on Cognitive Science Approaches to Process Control

"The Cognitive Work Process: Automation and Interaction"

24 – 26 September 2001

Universität der Bundeswehr, Neubiberg, Germany

An EACE Conference Series
European Association of Cognitive Ergonomics

DGLR-Bericht 2001-05 (2. Edition)
ISBN 3-932182-21-9

Godesberger Allee 70
D-53175 Bonn
Germany

e-mail: klaus.schubert@dglr.de
phone: +49 (228) 30 80 5-18
fax: +49 (228) 30 80 5-24

Foreword

This publication contains the presentations made at the CSAPC'01 conference, the eighth conference on „Cognitive Science Approaches for Process Control". This conference took place at the Universität der Bundeswehr München in Neubiberg (Germany-Bavaria) from the 24th to the 26th of September 2001.

CSAPC is one of the two conference series organised by EACE (the European Association of Cognitive Ergonomics). The first conference was held in 1987 near Paris and the subsequent conferences have taken place in Denmark, Finland, France, Italy, and United Kingdom. This biennial conference series aims at bringing together well known European experts in the fields of cognitive psychology and ergonomics, human-machine systems, and artificial intelligence, to discuss multidisciplinary research in the design and evaluation of complex, dynamic, and risky human-machine systems. This research has an important impact in a number of application domains such as industrial process control (e.g., nuclear power plant or steel industry), aviation (e.g., cockpit automation and air traffic management), ship navigation, automotive industry (e.g., advanced driver assistance), and medicine (e.g., anaesthesiology in operation room). The CSAPC series of conferences supplement more the specialised meetings within each of these applications, by addressing the common research topics and problems at a level corresponding to solutions from cognitive ergonomics and engineering. This focus on similarity and differences across application domains is one of the main objectives of CSAPC.

The theme of CSAPC'01 was „The Cognitive Work Process: Automation and Interaction". Research relating to human-machine interaction has traditionally been the prevailing theme in this series of conferences. Studies of cognitive automation have evolved from applying artificial intelligence techniques and existing knowledge about human cognition in the field of automation. Both cognitive automation and cognitive interaction have a lot in common and are merging in many aspects.

The single track conference was run with seven sessions of extended paper presentations (half an hour each), two one hour sessions for poster presentations and, in addition, specific discussion sessions each day. The sessions for extended papers were entitled with

- Process Control
- Aviation I and II
- Land-based Transportation
- Modeling I and II and
- User Interface

Evidently, modeling of cognitive behavior and work domains is becoming an increasingly active issue and seems poised to become one of the key areas that will ensure the success of the cognitive science approach.

More than 20 poster papers were presented to supplement and enrich the topics of the beforementioned sessions. The poster papers covered topics like safety-related issues, test and performance assessment of cognitive processes, cognitive tutoring and training as well as additional application domains such as medical systems.

Reiner Onken
Pietro Carlo Cacciabue
Jean-Michel Hoc
Erik Hollnagel

CSAPC'01

Cognitive Science Approaches to Process Control

Neubiberg, Germany, 24-26 Sep 2001

Organizing Committee

R. Onken	UniBw München, Germany (Chairman)
P.C. Cacciabue	JRC, Italy
J.M. Hoc	CNS-UVHC, France
E. Hollnagel	IKP-Liu, Sweden
K. Schubert	DGLR, Germany

Programme Committee

R. Amalberti	IMASSA, France
S. Bagnara	Univ. Siena, Italy
M.C. Beuscart-Zèphir	CERIM, France
G. Boy	EURISCO, France
B. Döring	FGAN, Germany
B. Färber	UniBw München, Germany
D. Harris	Cranfield Univ., UK
G. Johannsen	Kassel, Germany
P. Jones	Univ. Illinois, USA
P. Jorna	NLR, The Netherlands
K.-F. Kraiss	RWTH, Germany
M. Lind	TU Denmark
L. Norros	VTT, Finland
J. Patrick	Univ. Wales, UK
B. Pavard	IRIT U.P. Sabatier, France
J. Reising,	WPAFB, USA
F. Saad	INRETS, France
T. van der Schaaf	Eindhoven Univ., The Netherlands
O. Stock	IRST, Italy
O. Svenson	Univ. Stockholm, Sweden
R. Taylor	DERA, UK
K.P. Timpe	TU Berlin, Germany
G. van der Veer	Univ. Amsterdam, The Netherlands
U. Völckers	DLR, Germany
P. Wieringa	TU Delft, The Netherlands
H. Yoshikawa	Kyoto Univ., Japan
B. Zimolong	Univ. Bochum, Germany

The support extended to this conference by DGLR, EADS and ESG is highly appreciated.

CONTENTS

Session 1 Process Control

Session 2 Aviation I

Session 3 Aviation II

Session 4 Land-Based Transportation

Poster-Session

Session 5 Modeling I

Session 6 Modeling II

Poster-Session

Session 7 User Interface

Process Control

Co-operation on bridge in piloting situations. Analysis of 13 accidents on Finnish fairways

Maaria Nuutinen and Leena Norros
VTT Automation
P.O. Box 1301 VTT, Finland

In the autumn of 1997 several accidents occurred on the Finnish fairways for foreign ship during piloting. The Finnish Accident Investigation Board (AIB) decided to launch a joint investigation for these accidents[*]. The purpose of this paper is to present a method developed for the analysis of the piloting accidents. The target of the method is to enable parallel investigation of several accidents with the aim to create important knowledge of the domain, i.e. of the demands set by the sociotechnical system and of the contextual constrains to fulfil these demands. Based on the emerging cumulative picture of the domain a single accident can be analysed more profoundly. The developed method consists of two phases: a case examination and an organisational examination. The focus in the paper is on the case examination and co-operation on the bridge during the accident voyage. The main result of the investigation was that the demands of piloting activity have changed so that the prevailing practices and the prerequisites no more fully correspond with them.

INTRODUCTION

Piloting is a difficult task, especially on the rocky and narrow fairways of Finland. Pilotage is always executed in navigationally demanding and partly unanticipated conditions. The situational control of the piloting task calls for a preventive preparation. A successful performance of the task requires not only a skilled pilot but also considerable contribution from the bridge crew and effective co-operation among the two partners.

Along with technical development different kinds of navigation equipment are installed in the new ships. The most advanced steering and navigation systems allow planned and highly controlled turns, which is necessary on fairways where the safety marginal is small. However, the technical development does not by itself promote safety or efficiency of marine transport. The prevailing practices must be made compatible with the technical means, and the principles and constraints of equipment should be made known to the users. Moreover, demands set by navigation on open sea seems to be dominating the development of equipment while special demands on navigation on archipelago routes are mostly neglected. In addition, the ships sailing on the oceans represent a wide historical spectrum and different manufacturers. There are modern ships with advanced automation and integrated navigation systems and also ships with simple bridge layout and old head-up radars. This puts extra demands on the pilots'

---

[*] The committee of investigators comprised of Martti Heikkila (chairman) and Risto Repo from the AIB assisted by a group of experts including Pirjo Valkama-Joutsen, Kari Larjo, Leena Norros, Kristiina Hukki, Maaria Nuutinen, Matti Hellevaara and Antti Haapio.

expertise and co-operation on the bridge in piloting situations. Historically, however, piloting has not been considered a co-operative task but, instead, a highly individual and tacit practice the art of which has earlier been mediated only within piloting families, and still today in an apprenticeship relation from expert to novice within a normatively defined and restricted occupation. A major change in this traditional situation is the recent new law that enlarges the right of bridge crews of ships to acquire the qualifications for piloting in the Finnish fairways.

In the autumn of 1997 several accidents occurred on the Finnish fairways for foreign ship during piloting. These did not result in damage to persons or to the environment, and the damage sustained by the vessels remained relatively minor. These accidents were, however, interpreted as indicative of underlying problems. Therefore the Finnish Accident Investigation Board (AIB) decided to launch a joint investigation for these incidents with the aim to prevent similar occurrences in the future. Accident investigators and experts of navigation, psychology, maritime safety training and ship manufacturing technology participated in the investigation. Altogether 13 piloting accidents were included in the joint investigation. A similar comprehensive analysis was conducted for 10 cases. The results of the separately analysed 3 cases were merged with the results of the former during the investigation.

The main purpose of this paper is to present a method developed for the analysis. The reasons for a new method are basically two: First, the presently used methods are mostly case-oriented. They do not therefore provide tools for comparative analysis of several incidents with the aim of revealing generic features. The case-oriented methods can take account organisational or other conditional or causal factors but still they focus on single occurrence. This is due to the type of explanation and modelling typical in them. As Rasmussen [Rasmussen 1996, Rasmussen & Svedung 2000] has claimed a functionally oriented analysis of the boundaries of safe action in a system or domain is necessary for preventive measures. A suggestion of a formative modelling of domains made recently by Vicente [1999] is in concert with these claims. The second reason is based on the conviction that in order to improve learning from incidents and accidents in complex sociotechnical systems, also the adopted conceptions of human action should be modified. Our core-task analysis method, through the introduction of the concept of habit in the analysis, provides a tool for a formatively-oriented analysis of practices that may, in particular situations, lead to violating the boundaries of safe actions [Norros in preparation].

The paradoxes in the functioning of the sociotetechnical systems demonstrate the difficulty of finding relevant measures for preventing accidents. Reason [1998] handles paradoxes of defences, barriers and safeguards by noting that as well as protective measures can cause harm, conversely, small doses of a harmful entity can provide long-lasting protection, as in vaccination or inoculation [Reason 1998, p.42]. He notes further that accident investigators are required not only to establish the causes of an event but also to recommend measures to fix the problems. It may, unfortunately, sometimes turn out that these very same measures and 'fixes' play a major part in causing some subsequent accident [Reason 1998, p. 52].

Reason does not deny the benefits of the defences but emphasises the need among those who manage and operate complex technologies to appreciate both the advantages and the dangers of the multi-layered defences. We would like to add that there is also a need for a complementary approach. The starting point then is to take account *the paradox of human factor* [Nuutinen et al. 2000]. A human agent has a great potential to adapt his/her activity to uncertain and complex environment in order to reach his/her goals, which also is the reason why human part is needed to complete technical systems. But the very same features of the human being that make him/her non-replaceable provide a weak link in the system. This concerns not only cognitive but as well emotional and social aspects of human behaviour.

From this perspective, in addition to making barriers for negative effects of human weakness, also the strengths of the human agents should be supported. The aim is to clarify and improve prerequisites for efficient and safe working. This puts challenges for the methods used in accident investigations. The target of presenting the method developed for the presently discussed investigation is to show how the parallel investigation of several accidents creates important knowledge of the domain, i.e. of the actual demands set by the system and of the contextual prerequisites and constrains to fulfil these demands. Based on the emerging cumulative picture of the domain a single accident can be analysed more profoundly.

The developed method is based on the work carried out in two previous accident investigations by the AIB (summarised in Norros & Nuutinen 1999) and the research on decision-making in dynamic situations made by the VTT Automation Human Factors team (e.g. Hukki & Norros 1998, Norros & Hukki 1998).

THE METHOD

The development process

The investigation started by going through each single case one by one in the investigation group. Then it was decided that the target should not only be to prevent similar occurrences but also to create such knowledge and understanding of the piloting activity that makes possible increasing of safety of the whole system. The general frame of analysis, it's structure and course were constructed, and relevant factors for the analysis were identified. Finally, criteria for analysis and evaluation were formed. As a result, a preliminary method was developed. It is continuously under improvement during the still ongoing investigation.

The method for the parallel accident investigation

The developed method consists of two phases: a case examination and an organisational examination. In the case examination each single accident is first described in the light of the evidence. This includes the description of the vessel, the course of accident events and, rescue activity. Then every case is analysed with the same analysis structure and criteria. The used concepts enable comparisons between single accidents.

In the organisational examination the results of the case analyses are compared and concluded and piloting activity is analysed as an activity system [Engeström 1987, 1999]. Also the background material gathered during the investigation is used. The basic idea is to assess what kind of constraints and possibilities the organisation provides for safe and efficient piloting. Through a further systemic and historical perspective, developmental tensions within the activity system are revealed and used for explaining the nature and problems of prevailing practices. The organisational examination is still in progress and we will therefore concentrate on the case examination in this paper.

The primary aim of the case examination is basically same as in any accident investigation: to analyse the course of events and find out causal and conditional factors behind the particular accident. On this basis conclusions and recommendations are made. The way of doing this is however different at least regarding two aspects. The first relates with the modelling of the domain and the situation. In addition to causal modelling of the course of accident, the domain-related *functional constraints* of the situations are conceptualised in respect to preconditions they offer to taking care of *the core task* of piloting.

Second, the actions in the accident are analysed in causal respect as the course of actions and also in the further sense of what they tell about prevailing practices as expressed in the persons' own *habits*. The habits reflect the actors' personal accounting of the functional constraints of the domain, which they do according to different logics. These are conceptualised by the analysts in references to theoretical assumptions regarding the adaptability of action in uncertain environments. As a result domain specific demands on working practices are constructed. These are then utilised as evaluation bases for *ways of acting*, i.e. task specific concretisation of habits [see further Norros in preparation].

Way of navigating and way of co-operating both refer to a habitual generalised level of action that in the pragmatist tradition [Charles Peirce, Gerorg Herbert Mead and John Dewey] are seen to create continuity to the human-environment relationship. Through repeating a *way* of interacting with the environment a meaning is expressed and control over behaviour in situationally changing conditions achieved. The first term "way" in way of acting or navigating refers to this habitual aspect and the potential to act. The terms "navigating" or "co-operating" refer to concretisation of the concept in empirical analysis towards its particular actualisation in courses of action [Norros in preparation].

The case examination has following steps:
1) The functional specification of the constraints for navigation and piloting on the particular archipelago route. This includes assessment of organisation of piloting in that specific piloting district, practices and procedures of a shipping company in piloting situation, route geometry in comparison with ships' size and navigation marks, bridge layout, steering and navigation equipment and available number and quality of bridge crew.
2) Analysis of the navigation performance on the accident voyage. This includes three main phases: a causal construction of the events leading to accident based on evidence and simulations; a habitual analysis of steering and navigating action during piloting resulting in identification of *ways of navigating*; a habitual analysis of co-operation on the bridge resulting in identification of *ways of co-operating*. The two together form the piloting practice. They are determined through analysing how the actors take the situational weather conditions and the preconditions as resources and demands for navigation co-operation.
3) Specification of rescue activity.
4)Conclusions and recommendations.

The steps one and two are based on the construction of new concepts, methods and assessment criteria. A first version for criteria for assessment of the appropriateness of the bridge layout and equipment was created in an earlier work [Norros & Hukki 1998, Norros et al. 1997]. With diverse domain expertise available in this investigation a more elaborate set of criteria was worked out (to be explicated in further publications). The step two is based on the conceptual core-task analysis of the material of the accident cases.

The core task analysis for piloting work

We use the term core task to indicate the essential content of a particular work. The core task comprises of the functional demands that are to be fulfilled by the operator for the safe and qualified performance of the system. Due to internal and external pressures for change in the dynamic sociotechnical systems the core tasks tend to change. As a result the working practices that fulfil the core task may become invalid. Also training may become disoriented because of the out-dated or deficient conceptions concerning the qualification demands of particular work [Norros & Nuutinen 1999]. The graph in figure 2 illustrates the conceptual logic behind the core task analysis carried out during the investigation.

The object and aim of sea piloting is an efficient navigation and steering of a ship in a sea area where a knowledge and experience about local conditions is necessary for a safe, economic and ecological running of a ship (AB I). There exist historically formed contradictory pressures in the activity system, which create tensions in the interpretation of the aim and the conditions of its fulfilment. These are analysed in the organisational investigation.

The phases of the task are (A1):
- preparation for piloting
- construction of a piloting team
- co-operative navigation and steering during piloting (a hierarchical task-model of this was constructed).

These phases are dependent on each other. This becomes evident when the phases are set in relation to the functionally significant features (B1) that give meaning to the tasks included in the phases. We can make an assumption that, for example, in the preparation phase an explicit route plan improves co-operation possibilities. Further, in the team construction phase a division of the work (a co-operation method) [see Norros et. al 1998] sets different demands for co-operation with regard to the transparency of the behaviour of actors and the ship's route.

The functionally significant features are based on the definition of the aim and the domain specific general constraints as result critical functions of the fulfilment of this aim (B1). They are divided in navigational and co-operative features :

In navigational respect we distinguished the following critical features based on our earlier analysis [Norros et al. 1997].
- uniqueness of the ship-sea area system
- hydrodynamic uncertainty and complexity of the system
- dynamic of a moving system and the delay of its control
- diversity of representational forms of the process information (direct and mediated) and transformations between representations

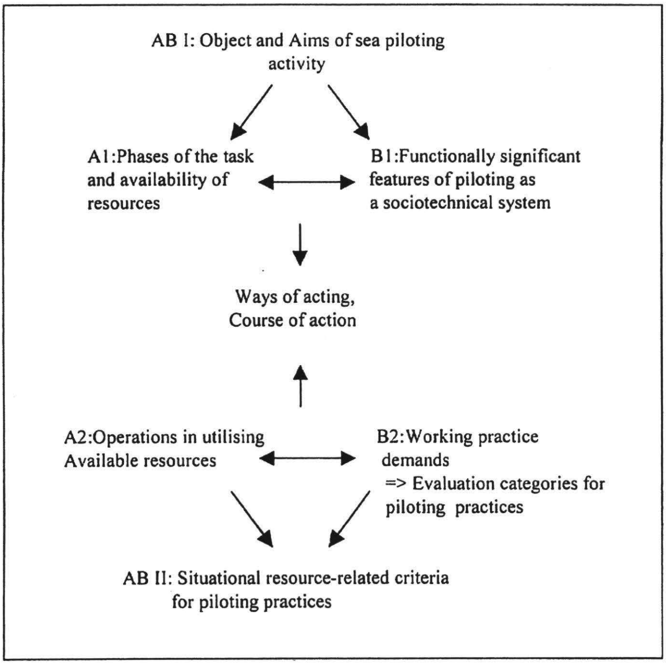


Figure 2. The conceptual construction of the core task of sea piloting. The analysis comprised of many iterative interpretative phases (the scheme adapted from Norros & Klemola 1999)

In co-operative respect there are several critical features that were defined based on earlier analysis of the core task of piloting [Norros et. al. 1997, Hutchins 1995]. The first critical feature is transparency of performance. This means the possibility for a bridge crew and pilot to perceive what is going on. There are at least three aspects that have an effect on transparency. They are transparency of the goal, interaction and tools. For example, a route plan is a way to make the goal transparent and a helmsman repeating course commands makes interaction open. The transparency of tools means how easy it is to perceive what is happening, e.g., an automatic mode a course changing with a little button is not very easy to notice. The second critical feature is a common knowledge and experience base. A bridge crew and a pilot should have at least the partly common knowledge and experience base in order to be able to interpret each others' behaviour and to adapt one's own behaviour. The third critical feature is common norms and procedures and a commitment to them. There is a need for norms and regulations in situations with lots of uncertainty and several independent actors. Every traffic system provides an example of the importance of the norms.

The above analysis provides a basis for the analysis and evaluation of the empirical observations of the course of action during the voyage. This can be more or less comprehensive depending on the material of the accidental voyage.

The ways of utilising the available resources (A2) are defined as operations in the different phases of the task. The operations describe a use of the material resources and a consideration of situational demands. The central material resources are:
-maps, plans and other information of the route
-information about a progress of the voyage (direct observations of the environment, information from the navigation equipment, communications)
-professional concepts.

In the texture of navigational and co-operative operations the analysts urge to identify meaning structures (habits) which express the actors accounts of the functional constraints. These are interpreted by the analysts as expressions of ways of navigating and co-operating. These are evaluated on the bases of the working practice demands derived by the analysts (B2). The derivation of the evaluation categories is based on an idea that the situational solution in the balancing between the various result critical demands (B1) promotes an appropriate human-environment interaction.

The following set of evaluation categories was used in our analysis. It was derived from the available material, and changes are assumed to be necessary with experience from further studies. The criteria refer either to the navigating or to the co-operative task demands, which is indicated in brackets (N for navigation and C for co-operation).

- integrating expertise regarding the environment, routes and the vessel (C)
- constructing and maintenance an ad-hoc team (C)
- formation of a cumulative interpretation of the situation (N)
- formation and maintenance of a shared interpretation of the situation (C)
- checking attitude in operations (N)
- exercising control of performance through monitoring (C)
- taking account of the whole traffic situation (N)
- creating an adequate orientation (N)
- maintaining orientation in the moving vessel (N)
- integrating information from different representations (N)
- shifting from one representation to another (N)
- proper timing of changes in division of work (C)
- planning of operations and their timely execution (N)
- communication of intentions (C)
- up-dating of shared plans (C)
- commitment to norms and common practices (C)
- practical and emotional preparedness for possible difficulties and high demands (N).

The situational resource related criteria for ways of navigating and co-operating refer always to particular piloting situations (ABII). The consideration of the same process control demand showed itself in different ways in different situations.

As a result of the analysis the ways of acting can be defined and used for explaining actual courses of action (in the centre of figure 1). Due to its functional significance for the safe and qualified performance of the system the core task has to be fulfilled in all circumstances.

THE RESULTS

The results of the investigation are of double character, conceptual and empirical. In the next the main results in both respects are summarised.

The piloting practices

Resulting from the work during the investigation a definition of piloting practice was derived and its major concrete forms could be identified. We defined piloting practice in the following way: *Piloting practice is a way a pilot interacts with the object of his/her activity, ship-sea area by navigating and steering in co-operation with a bridge crew in a piloting situation. Piloting practice is a habit and is based on experience and training.*

In the case examination piloting practices were decomposed in two parts: way of navigating (and steering) and way of co-operating on the bridge.

Way of navigating
The way navigating refers to how information from the environment, from the navigation equipment and socially communicated information is utilised in steering and navigation. The way of navigating must be considered both by analysing situated action and in its historical perspective. As a result of conceptual analysis on the investigation material two basic ways of navigating were identified, the traditional and technical way of navigating. The traditional way is based on 'inside-out view' to environment. It has been a natural way to navigate when there were no technical tools. This resembles the characterisation made by Edwin Hutchins [1995] regarding the navigation method used in Micronesia. The historically later technical way of navigating is based on the real movements of the ship on the plane, which is locked to north. This way of navigating is based on 'outside-in" or "bird's eyes" view to environment.

Way of co-operating
The way of co-operating refers to co-operation practises on the piloting situation to fulfil the navigational tasks. In the analysis of co-operation communication on the bridge and with the VTS (Vessel traffic service), the use of steering and navigation equipment, the number of the crew and division of work, and the taking account of the norms and procedures are essential. The way of co-operating is in a close connection with that of navigating, because in both the way of using available tools is central. In the investigation we observed almost perfectly individually performed navigation on the side of both the pilot and the bridge crew. The observed co-operative actions did not fulfil the demands of the defined evaluation categories

and the prevailing way of co-operating could be named reactive in the sense that communication took place post hoc and co-operation was organised merely on external demands and, even then, in an implicit manner.

Piloting practices as explanations of the course of actions in the accidents

The main results of the case investigation are, first, that the current piloting practices do not answer to the defined demands of piloting, the core task of pilotage. There were two main discrepancies that expressed themselves as problems in piloting performance during the accidental voyages. It was found out that the prevailing way of navigating among the pilots is traditional. Within this practice the pilots have difficulties to fulfil the defined piloting practice demands. This was due mainly to the incompatibility between the "inside out" view of the navigating practice with the "bird's eyes" view build in and required by the navigation equipment, on the one hand, and with the small safety marginal due the narrowness of routes in relation to the size of the ships, on the other hand. The tacit nature of the traditional navigating practice is combined with individual performance and scarce communication of the basis of action. The new demands on co-operative practices are clearly not fulfilled and co-operation is not conceived a part of piloting practice.

Also the prerequisites for piloting are inadequate against the demands. The deficiencies were recognised in the organisation of piloting in the piloting districts, the practices and the procedures of a shipping company in piloting situation, the route geometry, the bridge layout, the steering and the navigation equipment and the available number and quality of bridge crew. Furthermore, the present norms do not give support for the co-operation. According to the norms a pilot is a captains' advisor and a pilot is responsible for piloting. The norms and a few procedures concerning piloting are quite different from the practice, where the pilot actually takes the command. The general and quite superficial nature of the norms and procedures and the fact that they are drifted apart from the practice have led the situation where a pilot has not supported by the norms.

These inadequate prerequisites are mostly compensated with a pilots' individual efforts, his/hers experience and skills and only rarely with the effective co-operation and the division of work between a pilot and a bridge crew. This 'pilot-centred' way of doing seems to become even stronger when the situation becomes more demanding. The dominating way of co-operating is based on a centralisation of the tasks to the pilot, who has command and the control of the process. Captains' role is in a quite superficial monitoring and ensuring of the piloting.

The above mentioned problems can be interpreted as an evidence of the need for the redefinition of the piloting task. The emerging new concept of piloting emphasises co-operation as the central mean to cope with the complexity, uncertainty and reliability demands of the navigation task on archipelago routes. In our earlier study [Norros et al. 1997] the piloting task was defined as teamwork, where a navigation team locates the ship in relation to the environment and other ships, and manoeuvres it to the destination safely and economically. The navigation team was defined to consist of a pilot, master (or watch keeping officer) and possible other members of the bridge crew. It became evident in that study, that co-operation has not become a central aspect in the pilots' professional expertise or in piloting practice. These findings get further support from the incident analysis.

DISCUSSION

In summary, the main result of the investigation was that the demands of piloting were changed, while the practices and the prerequisites were not. The recommendations made in the investigation concern mainly a necessity to redefine piloting task and to create better prerequisites on the basis of this new definition.

Parallel investigation is informative as such. This was increased thanks to the developed method. It's focus on finding out habitual features in the co-operation on the bridge supported the development of a deeper understanding of the reasons behind the single case. It also opened up a way of connecting the situated actions with their systemic and historical context. As a consequence we were able to reveal plenty of potential safety relevant aspects in the current navigation practices. We also claim that the achieved results support formulation of recommendations, that do not appear as situated "fixes" but through their connection to the critical functional features of the activity promote genuine development of the system. The other side of the coin is however, that such recommendations are not easy to fulfil. A prerequisite for the difficult task of developing new practices is to understand their development in historical perspective. We anticipate the results of the organisational analysis to provide essential new insight in this respect.

Cacciabue et. al. (2000) have listed three conditions that must be met in order for an accident or a system failure to be imaginable. It should be possible to understand what the problem is, to envisage the consequences and to differentiate between large and small risk, and to conceive means by which the risk can be reduced or eliminated" (Cacciabue et. al., 2000, 180). This list can be used to reflect the investigation and the developed method. The profound analysis of incidents and minor accidents is an important way to understand what the problem is and sometimes also to differentiate between large and small risk. Then we also have time to figure out means be which the risk can be reduced or eliminated.

It is generally agreed that the accidents are only the top of the iceberg and the potential to learn from the incidents is multiple. Also the modern risk assessment methods requires statistical extensive and reliable material, which is in principle available by incident reporting systems. However, the focus of the incident reporting systems is more on the quantitative reporting and analysing than the qualitative and comprehensive analysis of the system. Then the real understanding of the nature and the historical origin of the problem, which are prerequisites for effective treatments, tend to be under-emphasised. Picking up several incident or minor accidents and analysing them more profoundly, as illustrated in this paper, gives the frame of reference, not only to see where the problems are but also to formulate effective protective means. Perhaps we then could create less dangerous defences (see Reason 1998), which are based on deeper understanding of the whole system and also the human strengths

As a safety critical domain shipping and piloting is still somehow in its infancy if it is compared with nuclear industry or aviation. Therefore making use of the experiences of the latter domains in safety control is a commonly expressed advice. While taking account of the previous work it should also be kept in mind that even in these domains, despite of the great advances in technology or the efforts of human factors science, improvement is not seen in the global accident rate (O'Leary 1999). This indicates that there is need for critical reflection regarding the main stream human factors science and its methodical approaches. We feel also that in creating new approaches the uniqueness and the nature of development history of each industry should be the basis for the future development strategies.

References

Cacciabue, P. C. , Fujita, Y.,Furuta, K. & Hollnagel, E. The rational choice of "error". Cognition, Technology & Work, 2: 179-181.

Engeström, Y. 1987. Learning by Expanding. An activity-theoretical approach to developmental research. Helsinki: Orienta.

Engeström, Y. 1999. Activity theory and individual and social transformation. In: Engeström, Y., Miettinen, R. & Punamäki, R-L. (eds) Perspectives on Activity Theory. Cambridge, UK: Cambridge University Press, 19-38.

Hukki, K. & Norros, L. 1998. Subject-centered and systemic conceptualization as a tool of simulator training. Le Travail Humain, 4, 313-331.

Hutchins, E. 1995. Cognition in the Wild. Cambridge, MA: MIT Press.

Norros, L. In preparation. Coping with Dynamic, Complex and Uncertain Situations. A habitual conception of action in analysis of work and expertise.

Norros, L. & Hukki, K. 1998. Utilization of information technology in navigational decision-making. In: Waern, Y. (ed) Co-operative process management. London: Taylor & Francis 77-88.

Norros, L. & Klemola, U.-M. 1999. Methodical considerations in analysing anaesthetists' habits of action in clinical situations. Ergonomics, 42: 1521-1530.

Norros, L. & Nuutinen, M. 1999. Development of an approach to analysis of air traffic controllers' working practices. Human Error, Safety and System Development. Liege, BE, 7 - 8 June 1999 (1999).

Norros, L. , Hukki, K. , Haapio, A. & Hellevaara, M. 1997. Decision making on bridge in pilotage situations. Research report in Finnish. Summary in English.

Nuutinen, M., Norros, L. & Pyy, P. 2000. Inhimilliset tekijät voimalaitoksen käyttövarmuudessa: Integroitu käyttövarmuuden arviointi- ja kehittämismenetelmä
VTT Symposium 206. Käyttövarmuussuunnittelu ja diagnostiikka. Espoo, 21.11.2000. Holmberg, Kenneth (toim.). VTT Valmistustekniikka. Espoo 2000, 91 – 104

O'Leary, M. 1999. The British airways Human factor Programme. Pre-Proceedings in Human error, safety, and system development Conference. Liege, June 7-8, 1999.

Reason, J. 1998. Managing the risk of organizational accidents. Ashgate, UK: .

Rasmussen, J. 1996. Risk management in a dynamic society: A modelling problem. Key note address: Conference on Human Interaction with Complex Systems, Dayton Ohio.

Rasmussen , J. & Svedung, I. 2000. Proactive Risk Management in a Dynamic Society. Karlstad, Sweden: Swedish Rescue Services Agency.

Vicente, K.J. 1999. Cognitive work analysis. Toward safe, productive, and healthy computer-based work. Mahwah, New Jersey: Lawrence Erlbaum Associates.

AUTOMATED DIAGNOSTIC AIDS: A STUDY OF OPERATORS' UTILIZATION STRATEGIES

Douglas A. Wiegmann, Ph.D.
Institute of Aviation, Aviation Human Factors Division
University of Illinois at Urbana-Champaign
#1 Airport Road
Savoy, IL 61874

SUMMARY

This experiment examined users' tendencies to either agree or disagree with automated diagnostic aids under conditions in which (1) the aids were less than perfectly reliable but aided-diagnosis was still more accurate that unaided diagnosis, and (2) the system was completely opaque, affording users no additional information upon which to base a diagnosis. The results revealed that some users adopted a strategy of always agreeing with the aids, thereby maximizing the number of correct diagnoses made over several trials. Other users, however, adopted a probability matching strategy in which agreement and disagreement rates matched the rate of correct and incorrect diagnoses of the aids. The probability-matching strategy, therefore, resulted in diagnostic accuracy scores that were lower than was maximally possible. These findings underscore the need for both interface and training solutions that facilitate the adoption of the most effective concurrence strategies by users of automated diagnostic aids.

INTRODUCTION

Human performance in complex systems, such as nuclear power and aviation operations, is marred by a history of operator error. This fact is due, at least in part, to the ever-increasing amount and intricacy of information that system operators must process. This growth in complexity, in turn, has increasingly removed the operator from the ultimate processes under his or her control [Wickens and Hollands, 2000]. Furthermore, the lack of appropriate or even possible interface solutions has often severely reduced the transparency or "visibility" of these systems. As a result, operators of such "opaque" systems are generally incapable of visually assessing the state of the system or alternatives for action. This increased distance and decreased visibility can often deter from the successful control and diagnosis during system failures. As a result, diagnostic and decisions errors on the part of operators are often cited as the primary cause of accidents and fatalities within such systems [Wiegmann and Shappell, 1997].

Hence, a great deal of effort has been expended to reduce both the occurrence and consequences of human error in complex systems through the design of decision support systems (DSS), such as computer decision-making and automated diagnostic aids. Unfortunately, however, such systems have generally failed to realize their full potential in improving operators' performance because operators frequently underutilize (disuse) automated aids when such aids are not perfectly reliable (i.e., less than 100% accurate)

[Parasuraman and Riley, 1997]. In the case of automated diagnostic and decision aids, this disuse is often manifested in the form of operator disagreements with an aid, even under conditions where aided diagnosis (albeit imperfect) is still, on average, statistically more accurate than unaided diagnosis [Wiegmann and Cristina, 2000]. Consequently, the goal of human factors research is to develop a better understanding of the strategies that operators adopt when interacting with less than perfectly reliable automation so that system interfaces can be designed to help users appropriately calibrate their trust in automated aids.

One utilization strategy that operators may adopt when interacting with automated diagnostic aids is probability matching, which is a strategy that has been repeatedly observed in a variety of human-learning experiments [Walker, 1996]. Within the context of automation utilization, the probability matching strategy would involve operators concurring with an automated aid's recommendation at a rate or frequency equal to that of the aid's true reliability. For example, if a certain diagnostic aid had a hit and correct rejection rate of .8 and a miss and false alarm rate of .2, then operators might concur with the aid 80% of the time and disagree 20% of the time.

Estes' (1982) Stimulus Sampling Theory (SST) accounts for probability matching by assuming that human-automation interactions involve a finite population of stimulus elements, of which only a small portion is randomly sampled by the operator on any given interaction. For example, if an operator agrees with an automated aid's diagnosis and is subsequently correct, then the stimulus elements in the sample will become connected to the "concurrence" response. If the operator's response is incorrect, however, the elements sampled will be connected to the alternative "disagreement" response. A similar process presumably applies when the operator initially disagrees with an aid's diagnosis, which is then either confirmed or disconfirmed. The frequency of agreement and disagreement responses is therefore determined simply by the probability in which each response is, on average, correct or incorrect (i.e., reinforced or not reinforced).

More formally, SST can be stated in the following equation

$$p(A_{1,n}) = \pi - (\pi - p_1)(1 - \theta)^{n-1} \qquad (1)$$

where $p(A_{1,n})$ is the probability that an operator agrees with the automated diagnostic aid on trial n and π reflects the probability that an agreement response will be reinforced, which is generally contingent upon the response being correct. The value of π is therefore a direct function of the reliability of the aid. The value p_1 is the probability of an agreement response occurring by chance on the fist trial, assuming an random distribution of agreement and disagreement responses across users, or no bias for one type of response over the other on the first trial. Finally, θ is a theoretical variable that represents the proportion of stimulus elements sampled on any given interaction with the aid.

If an operator must either always agree or disagree with an automated diagnostic aid (i.e., abstaining or reserving judgment is not allowed), then it follows from Equation 1 that the probability that an operator disagrees with an aid $p(A_{2,n})$ is:

$$p(A_{2,n}) = 1 - p(A_{1,n}). \qquad (2)$$

The curves depicted in Figure 1 were plotted using Equations 1 and 2, letting $\pi = .8$, $p_1 = .5$, and $\theta = .05$. These curves show the expected course of both agreement and disagreement probabilities across 100 interactions with an automated aid. An examination of this figure reveals that agreement probabilities exhibit a negatively accelerating learning-curve function. Specifically, the probability of the agreement response rises from .5 on the first trial and approaches .8 as an asymptote. Conversely, the probability of the disagreement response decreases from .5 on the first trial and then gradually plateaus at approximately .2.

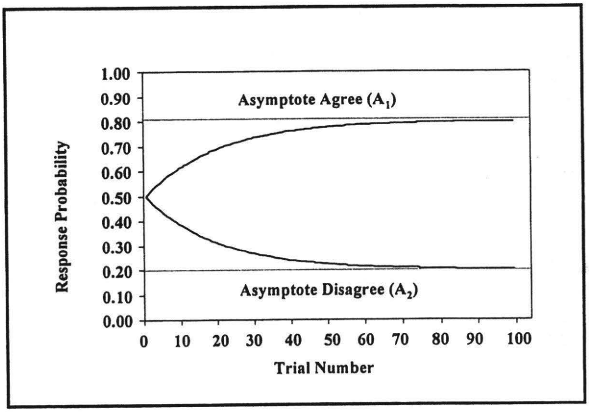


Figure 1. Theoretical curves for probability learning where probability of an agreement response will be correct, π, is .80. Adapted from J. T. Walker (1996) The psychology of learning. Upper Saddle River, NJ: Prentice Hall.

Whether automation users adopt a matching strategy when interacting with a less than perfectly reliable diagnostic aids has yet to be fully examined. Indeed, it is not at all obvious that probability matching should even occur, because probability matching often yields a lower level of accuracy than always agreeing with a highly (albeit not perfectly) reliable automated aid. For example, if an aid is 80% reliable and an operator agrees with the aid 80 times across 100 interactions, the yield would be 64 correct diagnoses (.8 x 80 = 64). The remaining 20 disagreement trials would yield only 4 correct diagnoses (.2 x 20 = 4), producing a total accuracy level of 68%. However, if the operator always agrees with the aid, 100 interactions would yield .8 x 100 = 80 correct diagnoses. Therefore, always agreeing with the automated aid would be the better strategy that would maximize the number of expected correct diagnoses. Even so, several researchers have found that unless pains are taken to explain the nature of event probabilities to participants in learning experiments, probability matching rather than maximizing is the rule [Walker, 1996].

Still, probability matching starkly conflicts with the general assumption that operators typically trust automated aids until such aids prove unreliable. Indeed, Estes' model suggests that automation users originally assume that such systems are generally unreliable. An operator's trust and agreement frequency then gradually increase over repeated interactions with the aid to a point where it matches the true reliability of the system. However, such does not appear to be the case within the context of human-automation interaction, where the general findings suggest that users initially trust automated aids and that trust declines only after faults or failures are introduced. Mistrust and disuse of the aid then occurs, although trust may be gradually regained, albeit not always to its initial levels (i.e., "under trust") [Lee and Moray, 1994].

A variety of factors, however, could be responsible for the general lack of empirical evidence for probability matching in the automation literature. One major factor is that most researchers in this area generally have not employed experimental procedures nor analyzed their automation utilization data in such a way that would reveal the potential use of a probability matching strategy by users. Another possibility is that there may be considerable individual differences in the types of strategies that user's adopt when interacting with automation. For example, Lee and Moray (1994) found that operators' decision to utilize either automated or manual control depended on both their trust in the system and their self-confidence in their own abilities to control the system. Consequently, individual differences in self-confidence also need to be considered when analyzing automation utilization strategies.

The purpose of the present experiment, therefore, was to address these issues by directly examining the type of utilization strategies adopted by users of automated diagnostic aids under conditions in which (1) the aids are less than perfectly reliable but aided-diagnosis is still more accurate that unaided diagnosis, and (2) the system is completely opaque, affording users no additional information upon which to base a diagnosis. The potential relationship between strategy selection and individual differences in self-confidence was also examined.

METHOD

Participants

A total of 50 undergraduate students from the University of Illinois at Urbana-Champaign participated in this study. Participation was voluntary and all participants were compensated for their time at the rate of $6 per hour. The total participation time did not exceed 1½ hours.

Task and Procedures

Participants began the experiment by completing a pre-experimental questionnaire that asked them to rate their problem-solving and decision-making abilities compare to an average college student using a scale that ranged from – 4 (worst than most) to 0 (about the same) to 4 (better than most). Participants also rated their risk-taking tendencies compared to other college students using a scale that ranged from – 4 (less risky than most) to 0 (about the same) to 4 (riskier than most). Upon completion of this questionnaire, participants performed a computer simulation task that required them to repeatedly diagnosis the validity of a pump failure within a waste processing facility. The simulation was developed using Visual Basic for MS-DOS and presented using a 486 MHz PC equipped with a 15-inch color monitor and standard keyboard.

The control panel in the computer simulation contained two radial-dial pressure gauges and an alarm indicator. At the onset of every trial, the gauge bars appeared to move over a range of readings, and the alarm light was shown as green. After a set delay of five seconds, the system appeared to fail, as the alarm light changed from a steady green to a flashing red. The gauges stopped moving, and the text "SYSTEM FAILURE" was displayed above the alarm light. A true system failure occurred in half of all trials (i.e., 50% true-failure rate). The program then presented two diagnostic aids to help the user determine if the failure was true or false. Only one aid could be chosen on any given trial, but both aids were equally accurate and had an equal mean delay in returning a diagnosis. The likelihood of either aid presenting the correct diagnosis (hit or correct rejection) was .8. The probability of an incorrect diagnosis (miss or false alarm) was .2.

After an aid was selected, an interim screen appeared with the message "Please wait." This message flickered on the screen for an average of 15 seconds, followed by the diagnostic aid's conclusion as to whether or not the pump had really failed. The participant then chose either to accept the aid's proposed diagnosis or to ignore it. The user was prompted to input their answer using the Y key, indicating that "yes", the system had really failed or the N key, indicating that "no", the system did not really fail. Feedback was provided as to whether the correct diagnosis was given, and participants' scores were updated by either awarding 10 points for a correct diagnosis or subtracting 10 points for an incorrect diagnosis. Testing then continued until the participant had earned 600 points (an average of 80% correct over 100 trials) or when a maximum of 120 trials was completed.

RESULTS

Preliminary Analyses

The percentage of trials on which participants agreed with the diagnostic aids ranged from a high of 100% to a low of 38%. However, only four (8%) of the 50 participants in this study concurred with the aids on every trial. Exactly half of the participants (50%) first disagreed with an aid before either aid had provided a wrong diagnosis. Yet, only a small number of participants (7 or 14%) disagreed with an automated diagnostic aid on the very first testing trial. A more in depth analysis of agreement-score distributions using a

Kolmogorov-Smirnov test revealed that participants' agreement scores were not normally distributed about the mean ($D = .156, p < .01$). Follow-up analyses indicated that the distribution of agreement scores was both highly kurtotic (9.73, *SE* = .662) and negatively skewed (-2.51, *SE* = .337). Therefore, the distribution of agreement scores was split, dividing participants into high- ($n = 26$) and low-concurrence groups ($n = 24$). Utilization strategies were analyzed separately for participants in each concurrence group.

Strategy Analysis

The distribution of agreement scores of participants in the high-concurrence group was consistent with a maximization strategy. Specifically, average agreement scores (*M* = 95.93%, *SD* = 2.86) were relatively high and stable across testing trials. As a result, the frequency of correct diagnoses of system failures (*M* = 78.06%, *SD* = 4.09) approached the maximum accuracy score obtainable given the 80% reliability of the diagnostic aid and the lack of any additional information upon which to base a diagnosis. In contrast, agreement scores in the low-concurrence group generally reflected the use of a probability matching strategy. Specifically, the average agreement scores (*M* = 81.65%, *SD* = .11) in the low-concurrence group were similar to, and did not differ significantly from, the 80% reliability level of the diagnostic aids. As a result, participants in the low-concurrence group had lower accuracy scores (*M* = 69.21, *SD* = 9.00) than participants in the high-concurrence group, $t(48) = 4.54, p < .01$, and subsequently took more trials (*M* = 115.67, *SD* = 8.31vs. *M* = 107.0, *SD* = 12.54) to achieve the designated goal of 600 points and thus complete testing, $t(48) = 2.86, p < .01$.

For those participants in the low-concurrence group who adopted the probability matching strategy, Estes' SST predicted a negatively accelerating agreement curve that started at a probability of agreement, $p(A_{1,n})$, of .50 and approached an asymptote of .80 as the number of testing trials increased. To explore this possibility, testing trials were grouped into blocks of 10 trials and the average agreement probabilities for participants in the low-concurrence group were computed for each block. An attempt was then made to fit Estes' model to the observed agreement probabilities.

Both the average observed and modeled agreement probabilities across testing blocks are depicted in Figure 2, assuming $\theta = .129$. These values are based on a sample size of $n = 24$, except in Blocks 10 and 11 ($n = 23$) and Block 12 ($n = 19$) where values are based on a slightly smaller samples due to some participants achieving the designated performance criteria (i.e., 600 points) and thus completing testing prior to the termination of the experiment (i.e., trial 120). An examination of Figure 2 reveals that the observed agreement probabilities in the low concurrence group did exhibit a negatively accelerated learning-curve function as predicted by the SST. In addition, the SST did a relatively good job at modeling the rate at which observed agreement probabilities would reach the 80% reliability value of the diagnostic aids. However, given that the values generated by the SST were limited by an asymptotic value set by the 80% reliability of the diagnostics aids, the agreement probabilities generated by the SST never exceeded .80. As is evident from the data plots, this resulted in the SST underestimating the observed agreement probabilities during the later testing blocks. The observed agreement probabilities were generally higher than .80 during the latter part of testing, averaging .84 during the last six testing blocks.

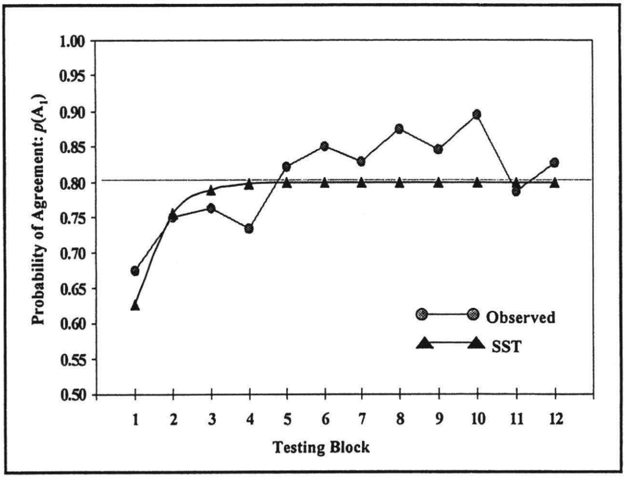


Figure 2. Observed and predicted agreement responses across testing blocks for participants in the low-concurrence condition.

Individual Difference Ratings

Participants made self-ratings of problem-solving skills, decision-making abilities, and risk taking tendencies on the pre-experimental questionnaire. Analysis of these ratings using a Mann-Whitney test revealed that participants in the high-concurrence group had higher self-ratings of problem-solving skills ($M = 1.96$, $SD = 1.22$) than participants in the low-concurrence group ($M = .96$, $SD = 1.12$), $z = 2.78$, $p < .01$. Participants in the high-concurrence group also had higher self-ratings of decision making abilities ($M = 1.81$, $SD = 1.36$ vs. $M = .71$, $SD = 1.30$), $z = 2.87$, $p < .01$. However, there was no difference in self-ratings of risk taking tendencies between the high- ($M = .92$, $SD = 2.02$) and low-concurrence groups ($M = .33$, $SD = 1.34$).

DISCUSSION

Participants in the present study periodically disagreed with the diagnostic aids, even though aided-diagnosis was more accurate than unaided diagnosis and no other additional information was available upon which to base a diagnosis. Furthermore, half the participants in this study disagreed with the aids prior to the aids making an error. One explanation of these findings is that participants were simply exploring the behavior of the aid to gain experience with it. Another possibility, however, could be that participants fell prey to the "gambler's fallacy." In other words, participants may have had the conviction that, since the aids had been correct for several trials in a row, an incorrect diagnosis by the aids was becoming increasingly more probable. This reasoning may have predisposed participants to "prematurely" disagree with the aids before such an automation failure had actually occurred.

Still, only seven participants in this study disagreed with an aid on the very first trial, suggesting that participants tended to trust the automation, at least initially. However, participants did differ in the type of utilization strategy that they employed when interacting with the aids. Evidence was found for both maximization and probability matching strategies. Those participants who adopted the maximization strategy generally agreed with the aids across most of the trials, which optimized their number of correct diagnoses. In contrast, participants who adopted the probability matching strategy only agreed with the aids on roughly 65% of the trials during the first block of testing and agreed with the aids on only 80% of the trials across all testing blocks. The probability matching strategy, therefore, resulted in lower accuracy scores than was maximally possible. Apparently, participants who adopted the maximization strategy initially trusted the aids and were less affected by aid failures. In contrast, those who adopted the matching strategy may have had lower initial levels of trust, which they then adjusted to match actual aid reliabilities.

Significant differences in individual self-confidence ratings were identified between participants who adopted the maximization and probability matching strategies. Specifically, those who adopted the maximization strategy had higher self-ratings of both problem-solving skills and decision-making abilities than those who adopted the probability matching strategy. At first, this difference may seem counterintuitive given the results of previous research showing that higher self-confidence tends to reduce automation utilization [Lee and Moray, 1994]. However, in the present study, participants who adopted the maximization strategy did in fact adopt the more optimal strategy, compared to the probability matching strategy, thereby maximizing their performance accuracy. In this regard, the higher self-ratings of problem solving and decision-making abilities by those participants who adopted the maximization strategy were indeed correct.

For those participants who adopted the probability matching strategy, the SST predicted a negatively accelerating agreement curve that started at 50% agreement and approached an asymptote of 80% as the number of interactions with the aids increased. This theory was partially supported in that the observed agreement probabilities did exhibit a negatively accelerated learning-curve function. However, counter to the theory's predictions, agreement frequencies exceeded 80% during the latter part of testing. Perhaps, during the latter blocks of testing, participants in probability matching group were in process of gradually shifting to the more optimal maximization strategy of always agreeing with the aid. Another possible explanation, however, is simply that SST focuses solely on the role of reinforcement in increasing response frequencies; it does not take into consideration the negative consequences of incorrect responses. Yet, in the present study, points were subtracted when agreements and disagreements with the aids result in incorrect diagnoses. Possibly, the "punishment" of being incorrect, which was more likely to occur on disagreement trials, further served to reduce disagreements, thereby increasing the rate of agreement responses to a point higher than what the SST would predict.

CONCLUSION

The results of the present study suggest that there are considerable individual differences in the type of strategy that operators adopted when interacting with automated diagnostic aids that are less than perfectly reliable. Furthermore, these different utilization strategies do not always result in optimal performance in terms of diagnostic accuracy. However, more research is clearly needed to determine the exact process by which automation trust and utilization strategies are acquired in real world situations. A better understanding of these strategies will potentially allow system designers to develop interfaces that help users appropriately calibrate their trust in automated aids, thus enhancing overall system performance.

REFERENCES

Estes, W. K. (1982). *Models of learning, memory, and choice: Selected papers*. New York: Praeger.

Lee, J. D., and Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40*, 153-184.

Parasuraman, R., and Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230-253.

Walker, J. T. (1996). *The psychology of learning*. Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., and Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wiegmann, D. A., and Cristina, F. J., Jr. (2000, in press). Effects of feedback lag variability on the choice of an automated diagnostic aid: A preliminary predictive model. *Theoretical Issues in Ergonomic Science.*

Wiegmann, D. A., and Shappell, S. A. (1997). Human factors analysis of postaccident data: Applying theoretical taxonomies of human error. *The International Journal of Aviation Psychology, 7*(1), 67-81.

OPERATIONAL READINESS VERIFICATION (ORV) A STUDY OF SAFETY DURING OUTAGE AND RESTART OF NUCLEAR POWER PLANTS

Vincent Gauthereau
Quality and Human-Systems Engineering
Dept. of Mechanical Eng.
Linköping University, SE-581 83 Linköping
vinga@ikp.liu.se

Erik Hollnagel
CSELAB
Dept. of Computer and Information Science
Linköping University, SE-581 83 Linköping
eriho@ida.liu.se

Abstract. Between July 1995 and October 1998, 9 incidents related to Operational Readiness Verification (ORV) have been reported in Sweden. These incidents pointed out the need for Swedish Nuclear Power Plants (NPP) to improve their related routines. This article describes of a survey of most of the Swedish NPPs in relation to ORV: its aim was to take stock of the work conducted under that past years. After presenting the different solutions developed, an analysis of ORV as a barrier system is proposed. Finally conclusions are drawn and suggestions for further research briefly exposed.

1. INTRODUCTION

Nuclear Power Plants (NPP) regularly have periods of outage and maintenance, during which many systems are disconnected and disassembled. Throughout these periods the NPP is effectively off-line and the reactor is shut down. Before the NPP can be restarted and brought on-line, it is necessary to ensure that repairs and modifications have been completed and have achieved the intended effects, that the systems have been assembled correctly, that they can be operated or activated according to their functional specifications (design criteria), and that the NPP as a whole – as a complex assembly of subsystems – is able to function. Operational Readiness Verification (ORV) refers to the test and verification activities that are necessary to ensure that a plant system is able to provide its required function at the required time. In basic language, it means ensuring that systems are able to work as they should and as they have been designed to do.

Although this paper focuses on ORV for NPP, the situation is not unique for NPPs, as all complex technological systems require that some kind of maintenance be carried out on a regular basis. For some systems maintenance can be done while the overall system is still running, for instance by having redundant subsystems (as in the electrical grid and most communication systems). For other systems, predominantly in the transportation domain (airplanes, trains, and ships) periods of maintenance and operation are interspersed in a regular fashion, often because the systems by their very nature or purpose are incapable of functioning reliably for very long periods of time. In all cases regular or scheduled maintenance is essential, because unscheduled maintenance usually is associated with costly interruptions of normal functioning and/or accidents with potentially significant negative consequences.

1.1 The incidents

The motivation for this study basis was a set of events that took place in Sweden between 1995 and 1998 and which involved ORV issues. In each case safety components / systems were left inoperable when the plants were restarted after outages (Table 1). Each event has been analysed for root causes, which were distributed among four classes: (1) weaknesses of administration processes, (2) weaknesses in management, (3) weaknesses in human performance and (4) weaknesses in the control room layout.

Table 1: Operational Readiness Verification Incidents in Swedish NPP

Date	**Plant**	**Type**	**Power (MW)**	**Start of operation**	**Incident**	**Time before discovery**
JUL 1995	F-2	BWR	1006	1981	Valves in the Containment Pressure Relief systems erroneously closed	2 days after start-up
JUN 1996	B-2	BWR	615	1977	Erroneously left open valve caused degraded containment pressure suppression function	Unknown
JUL 1996	F-1	BWR	1006	1980	Valves in the Containment Pressure Relief systems erroneously closed	Power ascension
NOV 1996	O-2	BWR	630	1975	Erroneously left open Disconnecting Switch caused inoperability in the Low Pressure Core Spray System	8 days after start-up
AUG 1997	R-2	PWR	917	1975	Steady State Protection System erroneously left inoperable	16 hours
SEP 1997	R-4	PWR	960	1983	Valves in the Containment Spray System erroneously closed	Unknown
OCT 1997	O-1	BWR	465	1972	Valves in the Containment Pressure Relief systems erroneously closed	7 months
AUG 1998	O-2	BWR	630	1975	Valves in the Residual Heat Removal System erroneously left inoperable	1 year (previous outage)
OCT 1998	O-1	BWR	465	1972	Valves in the Standby Liquid Control system erroneously left inoperable	12 days

1.2 The Practice of ORV

ORV problems occur mainly in relation to outage periods, where re-qualification is considered as a task for operations personnel. Control room operators are responsible for making specific systems available for maintenance as well as for re-qualifying these systems when maintenance has been completed. This obviously presumes that the maintenance has been carried out in an appropriate manner. The maintenance activities themselves may, however, require that some kind of checking is carried out by the maintenance personnel, for instance to ensure that a repaired object is not leaking. Since the first step of re-qualification is to check whether the maintenance task has been done, it is important for operations personnel to get the correct information. This communication is implemented via the work-orders management group and does not seem to have shown weaknesses.

There are however a few exceptions to this strict division of responsibility. Thus, for systems that are quite "*far*" from reactor safety, re-qualification is usually part of the maintenance task. This raises the question of how clearly this separation is defined: in other words, when is a system "*far enough*" from the reactor so that the re-qualification can be left in the hand of maintenance personnel?

Moreover, while it is necessary to re-qualify a system after maintenance, it can be done in a more or less formal manner. While a high level of formality is required by Swedish regulations for specific safety systems, the requirements are less strict for specific production systems. Yet for operators the latter may be just as important: while there is a need to be safe, there is also the need to produce electricity. This often leads to contrasting strategies for safety systems and production systems and raises the question of where the line should be drawn – or even whether such a line should be drawn at all?

Another problem underlying the incidents has been the organisation's incapacity to manage the unexpected. Whatever weight an organisation lays on planning, improvisation always occurs (Weick, 1998). Interactive complexity and tight-coupling are properties of NPP that render them subject to unforeseeable accidents (e.g., Perrow, 1984; Reason, 1988); it is impossible to plan everything, and it is therefore necessary to be able to improvise.

2. THE STUDY

After an initial survey of the literature on ORV, a study was undertaken which included visits to most of the Swedish NPPs and interviews with the technical staff responsible for ORV. The visits and interviews took place from December 2000 to March 2001.

2.1 Findings

One purpose of the study was to find which solutions the various NPPs had developed to cope with the problem, and which steps had been taken specifically to improve the efficiency of ORV. However, it soon became clear that it was problematic to separate ORV from the rest of the work done in a NPP during outages. In practice, many of the proposed solutions have a broader scope than the ORV problems they were meant to address. For instance, solutions such as a *better* organisation of the outage period, *better* processes in order to learn from experience and a *better* "safety culture" are features which benefit ORV as well as other aspects of work in the plant. In this paper we will, however, focus on specific ORV remedies.

2.1.1 Technological Solutions

First, a number of technical solutions were developed. One system is a *Blocked Safety Function*, which also is called "Red Lamps" because of its appearance in the control room. It consists of a logical grouping of position indicators for an important number of valves, which visualises the availability of different subsystems, such as reactivity, activity-barriers, hard-emergency cooling, cooling system and electricity system. In addition to provide the operators with an overview of the availability of the plants' safety systems, the link to a computer system enables operators further to investigate the reason behind an indicated unavailability.

Another device, called *Computerised Position Check* is based on a computerised system, which makes it possible to compare the actual status of systems with their expected status – relative to the condition of the plant. The computer shows the deviations from the "normal situation", but is used on demand rather than being on-line. A further, simpler device is an *Overall Re-qualification Schema*, which provides operators in the control room with a general view of the plant situation regarding operational readiness. The schema consists of a hierarchically organised representation of the plant systems.

Most of these technical solutions are part of the design of newer NPPs. To implement them in older plants requires important changes both in the station and in the control room (apart from the overall re-qualification schema). Their implementation thus usually goes together with general reconstruction plans.

2.1.2 Organisational Solutions

Second, remedies were developed which relate to organisational issues. Parallel to the maintenance plan, an operational readiness plan is nowadays available to control room operators. It describes the OR status of the different systems in parallel to the maintenance plan, and specifies the expected state of the different systems along the whole outage period.

Following the observation that the logistic of operability tests was inadequate or missing, a systematic way of working has been defined which should apply to any system, in a more or less formal manner. This logistic defines four steps to achieve operational readiness:

1. **Reinstating Control**: Operations personnel controls that the maintenance work has been done. This is an administrative control but it provides an opportunity for operations personnel to "sit back and think" whether it is actually the right time to restore the system, whether all the maintenance work is done on that system, etc. The

aim of this phase to prevent any further work once the system has been declared ready for operation.

2. **Resetting basic configuration**: All valves are repositioned.

3. **Activation**: The components (pumps, etc.) are tested functionally one by one.

4. **Testing:** Finally, the whole system is tested, and then if the test is OK, the system is declared operationally ready! However, this test may introduce new risks since some systems have to be shut down in order to run the test. Thus new ORV tasks must be performed consecutive to this testing phase.

New instructions / procedures for ORV have also been written for specific safety systems. Their structure follows the principles of the four steps of the work logistic described above.

Before changing from one level of operation to another, an independent control is performed on all the valves at the very end of the outage period. This is a physical control, which often is performed from a different point of view: for instance while the first control might have been done by systems, the second may be done by functions or even topographically. In addition, the people who perform the control come from the outside. These people may, of course, have been involved in the outage to some extent but have usually not been so active. This redundant control is performed during a time-out at the end of the outage, but the plant operators do not see the time-pressure on this time-out period as high. Yet even though none of the interviewees ever needed more time than initially planned, they all felt they could use more time whenever necessary!

3. ORV IN A BARRIER SYSTEM

The literature survey found different models for ORV relating to human and organisational aspects of NPP safety, but recommended that a systematic description in terms of barrier systems and barrier functions with specific characteristics would provide a useful way to characterise important ORV issues. This should be combined with some way of accounting for the dependencies between activities as well as barriers, for instance using a kind of functional modelling as suggested by Rasmussen & Petersen (1999).

In general, ORV takes place as part of the following sequence of events: (1) the maintenance work itself, (2) ORV as post-condition for completed work, (3) ORV as pre-condition for beginning start-up, (4) the start-up sequence itself, (5) power operation, and finally (6) the functioning of the automatic safety systems. Putting all these together, leads to a view of a set of barrier systems as shown in Figure 1. Each level represents a specific type of barrier system as defined by Hollnagel (1999). As the arrows illustrate, the barriers are intended to capture and "reject" errors, so that in the end nothing is left that can lead to incidents or accidents.

Practical experience shows that these principles do not always work as planned, although the findings of Pyy et al. (1997) hopefully overstates the problem.

We can use this representation of the various barriers to discuss the findings from the study.

- **Maintenance work** is carried out subject to procedures and job descriptions of various types, which are followed to a larger or smaller extent by the operators. The procedures are in principle a kind of barrier, since they among other things serve to prevent incorrect actions from taking place. One finding from the study was that not much work has been done to involve maintenance personnel in ORV issues.

- **Work approval – ORV as post-condition**. These are the tests carried out when the maintenance work on a specific subsystem or component has been completed. Such tests are often prescribed by the procedures, but are in this context considered to be conceptually separate. Over the years different barriers have been added and/or improved. In particular, the four phases described in section 2.1.2 represent an important symbolic barrier when used together with a specific procedure, while they are an immaterial barrier when no specific procedure is required for a specific system.

- **Preventive actions – ORV as pre-condition**. The use of ORV as a pre-condition test aims to ensure that the reconnected system is in a ready state before the start-up is begun. The main barrier is the so-called integrated testing, during which the functions (from different systems) are tested. These tests however affect the status of some systems. In order to reduce such side effects, the position of some valves is now indicated in the control room, reducing the number of tests to be performed.

- The **start-up** is regulated by procedures, which include various checks and tests. It is assumed that some of the checks are carried out automatically, i.e., by the technological systems themselves. Such tests are normally defined as operational pre-conditions and are part of the I&C logic. Right before start-up the redundant / independent control constitutes an important barrier as well. The multiplicity of the different views used to perform the test is seen as important in order to reduce the interrelations between the different barriers.

- **Power operation** is an extended phase of steady-state operation, but regular checks are a part of that. Technological devices such as *Red Lamps* or *Computerised Position Check* constitute symbolic barriers as well.

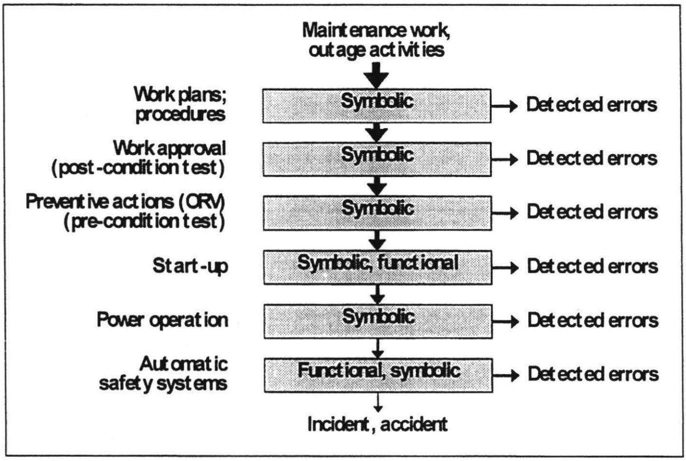


Figure 1: Layers of barriers related to ORV.

3.1 Additional Aspects

The above rendering of ORV as a barrier system does not tell the whole story. In fact a major weakness of ORV, considered as a system of barriers, is that their order cannot be guaranteed. In other words, operators and maintenance workers may with the best of intentions deviate from the prescribed order, for instance to accommodate an immediate demand or to avoid slowing down the work of others. The importance of safety culture and of learning from experience has been argued extensively in the literature and no further discussion will take place in this paper. However, it seems important to focus on the ability of the organisation to plan the outage in a proper manner, as well as on its ability to improvise in case unexpected events occur. Concerning the former work has been done in different directions, and it seems that work planning has been improved. Moreover, tools such as the operational readiness plan and the overall re-qualification schema were developed to support operators in following the plan. With regard to the organisation's ability to improvise in order to cope with unplanned events, not much work seems to have been conducted.

4. DISCUSSION – CONCLUSION

Since the incidents took place much work has been conducted in Swedish NPPs in order to improve ORV and plant safety during restart. Much as well happened to the electricity market and to the context of nuclear power generation. As a combined result, most of the interviewees acknowledged a certain "improvement" in the way outages are planned. Many different factors could lead to this increased confidence. Of course the main factor could actually be increased organisations' abilities, but this is still to be determined.

During this study we realised the ambiguity behind ORV issues. In this paper we highlighted solutions which have been specifically developed after the incidents (both technical and organisational). However, the organisation of the outage period is an important factor of success and should not be denigrated.

We also observed important distinctions between plants, since each utility has its own favoured solution. Yet, these peculiarities could barely reflect differences in plant safety, since the social and cultural anchoring of each solution showed to have much importance.

Research will go on in two major directions: first, a closer look will be given to testing tasks: from object tests up to safety function tests. There is also an interest of looking at the plants' ability to plan / re-plan and improvise during outage.

5. ACKNOWLEDGEMENTS

The work reported here was carried out as the first task in the study on "Operational Readiness Verification: A study on safety during outage and restart of nuclear power plants" funded by the Swedish Nuclear Regulatory body, SKI. The authors would like to thank Gerd Svensson, Ola Svenson, and Martin Fridleifer, for valuable and inspiring discussions throughout the project, and numerous staff members at the Swedish NPPs for so willingly to have given their time and effort to the study.

6. REFERENCES

Hollnagel, E. (1999). Accidents and barriers. In J.-M. Hoc, P. Millot, E. Hollnagel & P. C. Cacciabue (Eds.), Proceedings of. *Lez Valenciennes, 28*, 175-182. (Presses Universitaires de Valenciennes.)

Perrow, C. (1984). *Normal Accidents: Living With High-Risk Technologies*. New-York, USA, Basic Books, Inc.

Pyy, P. Laakso, K. & Reiman, L. (1997). A study on human errors related to NPP maintenance activities. *IEEE Sixth Annual Human Factors Meeting*, Orlando, Florida, p. 12-23 – 12-28.

Rasmussen, B. & Petersen, K. E. (1999). Plant functional modelling as a basis for assessing the impact of management on plant safety. *Reliability Engineering and System Safety, 64*, 201-207.

Reason, J. T. (1988). Cognitive aids in process environments: prostheses or tools? *Cognitive engineering in complex dynamic worlds*. E. Hollnagel, G. Mancini and D. D. Woods. London, Academic Press.

Weick, K. E. (1998). Improvisation as a Mindset for Organizational Analysis. *Organization Science* 9(5): 543-555.

Aviation I

COGNITIVE AUTOMATION FOR TACTICAL MISSION MANAGEMENT: CONCEPT AND PROTOTYPE EVALUATION IN FLIGHT SIMULATOR TRIALS

A. Schulte, P. Stütz and W. Klöckner

ESG – Elektroniksystem- und Logistik-GmbH
Einsteinstr. 174, 81675 München, Germany
phone: +49-(0)89-9216-2623, fax: +49-(0)89-9216-2632
e-mail: aschulte@esg-gmbh.de

SUMMARY

This paper describes an approach to flight-deck automation in the field of tactical combat mission management. A concept for a functional prototype system, the so-called *Tactical Mission Management System (TMM)*, will be given. The TMM has been implemented as a functional prototype in the Mission Avionics Experimental Cockpit (MAXC), a development flight simulator at ESG and evaluated with German Air Force pilots as subjects in simulator trials. Therefore, the TMM has been compared with a reference cockpit avionics configuration in terms of task performance, workload, situation awareness and operator acceptance. After giving an overview on the system concepts his paper reports on the experimental design and results of the simulator trial campaign.

INTRODUCTION

Performing military combat missions in an uncertain dynamic tactical environment, presents a potentially intolerable workload for the crew. Therefore, research and development activities are conducted in order to assist flight crews while performing the tasks with respect to a safe and successful mission completion. The functions required to provide intelligible interaction and efficient use of such a system are derived from a concept taking into consideration the process of human information processing. Such cognitive automation concepts have already been proved successful in other crew assistant programmes for transport aircraft [Stütz & Schulte, 2000]. They are incorporated in the concept and extended towards an application for the air-to-ground attack role. According to task analyses of the air combat role, the demands put on the crew in terms of decision-making are even more complex than in air transport. The crews possess more freedom in the decision how to proceed, instead of following rather fixed procedures (e.g. IFR). The approach aims at the provision of crew assistant functions which focus upon the monitoring and if necessary the retrieval of the integrity of superior goals such as safety, combat survival and mission accomplishment rather than merely supporting the compliance with procedures. In order to realise this approach, a way has to be found how to deal with human goal knowledge in machine systems.

SYSTEM DESIGN CONCEPT

Dealing with general goals, an approach has to be found to how to design a machine system to process this kind of information. Today's cockpit avionics completely miss out this level. Of course, the treatment of goals is implicit (e.g. GCAS is designed to avoid ground collision per definition). But, goals, goal violations and the possible interference of different goal domains (e.g. combat survival vs. mission accomplishment) are not yet processed explicitly.

The following sections discuss the approach to cognitive flight-deck automation on the basis of a co-operative automation philosophy. A simple model of the human problem solving strategy based upon Rasmussen's work [1986] is used to identify shortfalls in automation. A model of superior goals and the deduction of resulting tasks is given.

Co-operative automation

The variety of tasks to be performed by the pilot during a tactical flight mission result in a workload on all work process levels (see Figure 1), ranging from skill-based manipulatory control (bottom) through rule-based system interaction (middle) up to general knowledge-based problem solving tasks (top). Conventional automation traditionally focuses on relieving the crews from exhausting routine actions, thereby being granted full autonomy in certain well defined areas (Figure 1, left). Expanding this strategy of automation into task domains primarily subjected to rule- and knowledge-based crew action, leads to severe problems in the area of man-machine interaction (Figure 1, middle). Significant for this kind of development are very complex avionics structures and functions taking over full autonomy for comprehensive parts of the flight, while reducing the pilot to a mere solver of abnormal situations [Billings, 1997]. Therefore, several authors, e.g. [Onken, 2000] demand new progressive methods when it comes to expand automation into all aspects of flight and mission management. One promising way to proceed is the concept of an automated system acting in a co-operative relation rather than separating the crew from the basic aircraft systems (Figure 1, right). Being well aware of the complex task to perform, the crew interprets the output of the automation system as the recommendation of an additional electronic crewmember. The decision-making to accept or reject the machine's advice is allocated to the crew. Proceeding this way, the crew is kept continuously in the decision loop and is able to employ the full strength of human performance. At the same time the crew takes advantage of the particular strengths and abilities of the system.

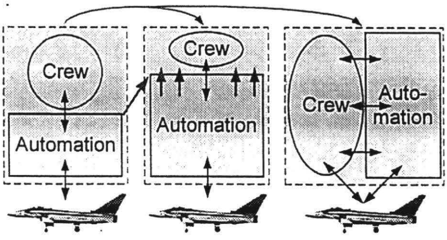


Figure 1: Conventional / co-operative automation

Human problem-solving strategy

Besides an effective means of communication, a prerequisite to establish this partnership relation between man and machine is the implementation of transparent functional behaviour within the automation system. In order to make the machine output easy to comprehend and evaluate, and to establish a close-partner work relationship, both crew and machine have to reason from the same principles. Thus, the analogue problem-solving strategies and mechanisms have to be implemented in the automated system in a similar way to that which can be found in the human counterpart. This concept is the core element of cognitive automation.

Figure 2 depicts the elementary steps of human problem solving [Rasmussen, 1986], which are to be transferred equally into the machine system. On the lowest level a state-oriented acquisition of environmental signals is conducted and direct manipulatory output is generated. Problems demanding a certain amount of data abstraction and knowledge transfer typically cannot be solved on this level. Therefore, further data interpretation is necessary, taking into account superior context knowledge. A task-related aggregation and fusion of information can be derived as a result. On the structure-oriented level the data so-gained is further processed using additional rule and knowledge bases in order to reach a more profound problem diagnosis and a spectrum of possible solutions. Again, back on the context-level, a decision on how to proceed is found by the use of planning and forward-simulation results. Then, the derived solution is passed to the state-level for execution, thus ensuring successful problem solving under consideration of all relevant circumstances and all available information.

Simple automation implementations within clearly-cut task domains such as auto-pilot or flight director systems usually can be seen as immediate instantiations of the processing steps on the state-level, which are directly connected through functional relations. Autonomous planning functions are state-of-the-art in today's Flight Management Systems. Assisting the crew also on the higher levels of their problem solving tasks, some abstract and therefore more versatile task knowledge has to be made available to the machine.

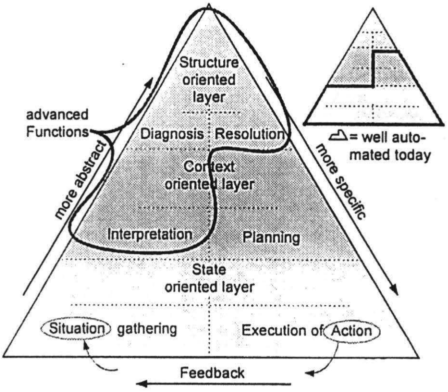


Figure 2: Human cognitive problem solving behaviour

Functional breakdown

In order to derive a functional breakdown of the TMM the generic processing steps are translated into specific functions. On the state-oriented layer situational parameters such as aircraft sensor and system signals, data-link information such as tactical elements and mission order, and on-board database entries e.g. terrain elevation data are to be considered. These data are analysed in order to identify context-specific features concerning pilot behaviour, aircraft movement and the external tactical situation. The monitoring of the threat accumulation along the planned trajectory is a typical example for a context-spanning analysis.

The determination of the current phase of flight is essential for the model-based prediction of expected crew actions required to cope with the mission constraints. It considers all relevant task domains such as flight guidance along the pre-planned track, systems operation, weapon deployment, as the human operator does. It computes the current tasks and task parameters relevant for the crew in the present situational context.

The *conflict detection and resolution* function builds up a hierarchy of general goals to be followed throughout the mission, such as flight safety, combat survival and mission accomplishment. Utilising the results of the tactical situation interpretation and the monitoring of the flight situation-dependent tasks, the system figures out violations of these goals. After negotiation with the pilot, a proposal how to resolve the conflict will be passed to appropriate machine agents for conflict resolution. Implementing this human-like goal-task-model ensures machine problem solving strategies which are easy to anticipate for the pilot.

Planning is the most important conflict solving agent activity. The tactical mission management system offers a fully autonomous mission and route planning capability to the crew, including terminal operations planning, transit flight planning, tactical low-level flight trajectory planning and the use of attack procedure templates. Feedback is obtained in two ways; externally, due to manipulatory action of the crew and, thereby, alteration of the situation; and internally by the continuous re-consideration of the planning result in terms of goal integrity.

Finally, the Tactical Mission Management System provides an appropriate *man-machine interface* on the flight deck, in order to manage the information flow and the crew interactions. The main components are an advanced primary flight display utilising a perspective 3D synthetic vision symbology, a tactical mission management and navigation interface, and speech synthesis.

Goal and task model

As the structure of functions stated above makes obvious, a machine-immanent representation and processing of goals and tasks is compelling, thus establishing a common ground of understanding between the avionics system and the crew. Reasoning from first principles, the machine system is capable of inferring the same solutions a human pilot is most likely to reach in this situation. This goal model considers the principal kinds of pilot motivation such as flight safety, combat survival and mission accomplishment (see Figure 3).

The primary aim of the introduction of the goal/task model is the detection and prioritisation of conflicts in the mission progress. A conflict is defined as the violation of a goal. Starting from the abstract goal classes, application-specific sub-goals are derived (e.g. meet a given time-over-target, TOT). The parameterisation of the sub-goals is performed using the mission order and the current situation. This might include the application of scheduled values or tolerances and even the de-activation of single sub-goals according to the current flight phase (e.g. goal 'meet TOT' is no longer relevant in the flight phase Ergress). For conflict detection the specified sub-goals (i.e. current tasks) are compared with the actual situational parameters. In the case of an intolerable deviation a conflict is detected and can be passed on for further processing.

This approach heavily relies on its capability to analyse all situational elements and their influence on the given goal structure. Several sub-functions using expert knowledge are necessary to perform the needed interpretation of flight progress, environmental, tactical and aircraft data. Only when this overall situational picture is available to the machine system it is possible to determine and prioritise goai violations reliable by the centralised goal conflict analysis.

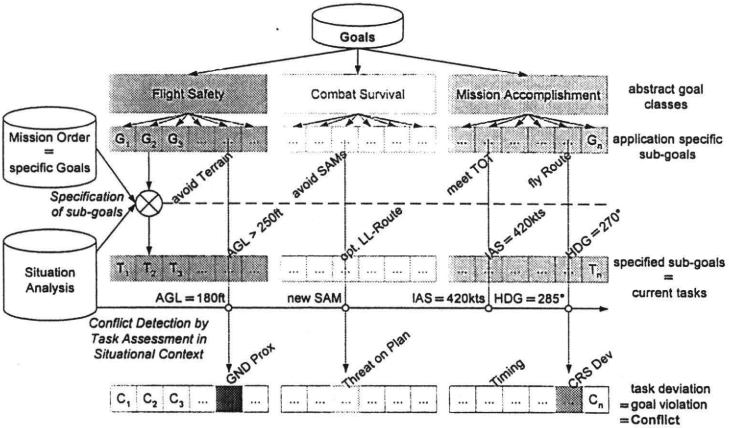


Figure 3 Goal – task model for conflict detection

EXPERIMENTAL DESIGN

In order to prove the approach, the TMM has been implemented as a functional prototype and integrated in the flight and scenario simulation environment. It has undergone an experimental evaluation with operational personnel in spring 2001. The following sections give details on the apparatus, scenario and tasks as well as the subjects.

Apparatus, scenario and tasks

For the evaluation of the Tactical Mission Management System a comparative study was chosen. Two different simulator set-ups were configured, on the one hand representing the basic functions of a reference combat aircraft cockpit (e.g. Tornado) and on the other hand demonstrating the TMM functions and displays (Figure 5). Following the experimental procedures the pilots had to perform a dedicated test mission with each of the cockpit configurations. Figure 4 shows the phases of the test mission located in the south-west regions of Germany i.e. (1) tactical transit, (2)

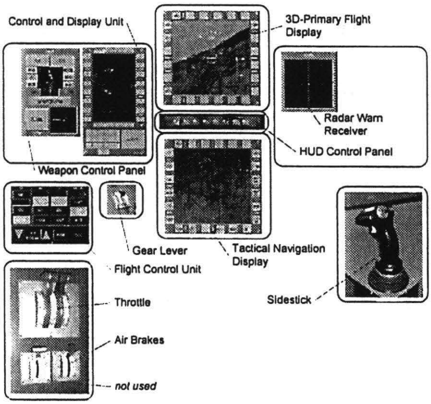


Figure 5: Cockpit in the TMM configuration

low-level ingress, (3) attack, (4) low-level egress and (5) tactical transit. During the low-level phase the mission was supported by computer-generated units such as SEAD-forces for suppression of enemy air defence and AWACS. Using the TMM-configuration the aircraft was participant of a tactical data-link network providing data on other participants and surveillance information. During the mission the tactical situation (i.e. hostile SAM sites) was supposed to change several times forcing the pilots to react accordingly (e.g. route adaptation, re-planning, threat avoidance), thereby workload being imposed on the operator.

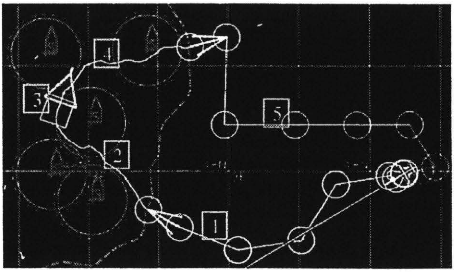


Figure 4: Scenario of the test mission for system evaluation

Subjects

The subjects were four German Air Force pilots (partly flight instructors) from the Fighter Bomber Wing 34, Memmingen at an age of 30 to 38 years. Their flight experience ranged from a total of 900 to 3000 flying hours on Tornado and other NATO combat aircraft. During a one day familiarisation phase the pilots had the opportunity to train the handling of the simulator and the interaction with the TMM before the test mission had to be performed.

EVALUATION RESULTS

The main scope of the evaluation of the Tactical Mission Management System was to account for improvements in comparison to the reference system in terms of the following categories of measurements:

- Measurement of pilot performance
- Evaluation of system performance of the TMM
- Evaluation of pilot's workload
- Measurement of pilot's situation awareness
- Subjective ratings concerning TMM performance

The following sub-sections report on the specific results.

Pilot's performance

The investigation of the pilot's performance has been performed under the consideration of the three abstract goals as defined earlier in this paper: flight safety, combat survival and mission accomplishment.

With regard to *flight safety*, the area of low-level flight guidance has been investigated. During the experiments it could be observed that pilots frequently took the risk of dangerous ground and obstacle proximity in order to avoid military threats. Therefore, it was investigated how often certain given above-ground-level minima were violated while performing low-level flight. Table 1 gives the results of the comparison between the reference system (REF) and the Tactical Mission Management System (TMM). The assessment of ground collisions and the frequency of AGL minima violations make clear that the TMM caused a significant risk reduction by a better ground separation. In a deeper investigation of the terrain following performance [Schubert & Schulte, 2001] it is evident that the flown vertical profile becomes significantly smoother (i.e. less vertical acceleration, less variation in altitude) in the TMM configuration. Thereby, the pilot's comfort level could be increased.

Table1: Violations of Minima during low-level flight

REF	*< 0 ft AGL*	*< 50 ft AGL*	*< 150 ft AGL*
Pilot 1	1	5	39
Pilot 2	0	0	17
Pilot 3	1	3	19
Pilot 4	1	9	36
TMM	*< 0 ft AGL*	*< 50 ft AGL*	*< 150 ft AGL*
Pilot 1	0	0	2
Pilot 2	0	0	5
Pilot 3	0	0	5
Pilot 4	0	1	23

An important feature of the TMM is the ability of situation-dependent in-flight re-planning of the mission plan for threat reduction. In order to quantify the effect of this assistance function in terms of *combat survival* the mean threat exposure has been evaluated along the flown trajectories. Obviously, a massive reduction effect on the threat exposure could be noted by use of the TMM. It should be emphasised that the improvement of threat avoidance with the TMM could be achieved in combination with a much better ground separation (see Table 1).

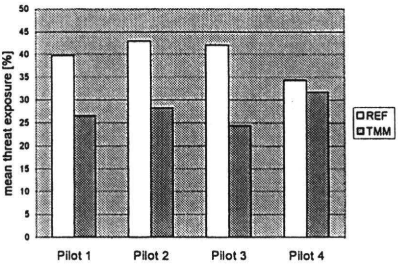


Figure 6: Mean threat exposure in comparison

Table 2: Global criteria of flight safety and mission accomplishment

	Pilot 1		***Pilot 2***		***Pilot 3***		***Pilot 4***	
	REF	*TMM*	*REF*	*TMM*	*REF*	*TMM*	*REF*	*TMM*
SAM shots [#]	-	-	-	-	1	-	1	-
ΔTOT [s]	L 8.0	E 0.4	E 0.3	E 2.1	E 2.3	E 0.2	L 3.8	L 2.1
Target hit	OK	OK	OK	OK	OK	OK	OK	OK
Destination reached	OK	OK	OK	OK	OK	OK	no	(OK)
Fuel on Board @ Touch down [%]	4.4	14.3	9.1	13.8	4.3	12.5	0.0	8.8
ACO violations	5	-	5	-	5	-	8	-
Mission accomplished	no	OK	OK	OK	no	OK	no	(OK)

Table 2 provides a collection of criteria concerning *mission accomplishment*. The results make evident that the pilots performed notedly better with the TMM than without. Due to a better threat avoidance with the TMM the SAM shots could be reduced. Performance criteria such as meeting the Time-over-Target (TOT), hitting the target or reaching the destination could be fulfilled by all pilots quite well. Only pilot 4 did not reach the destination with the reference system due to a flame-out condition. In general it was found that fuel consumption could be decreased significantly with the TMM during the mission. Another observation was made concerning the number of violations of the Airspace Co-ordination Order (ACO) routing, which could be totally eliminated by the use of the TMM. So, the risk of being hit by friendly fire was minimised.

System performance

One of the most important features of the TMM is the pilot assistance in optimising a threat minimal route under a dynamically changing hostile threat theatre. Figure 7 shows the total and mean threat exposure computed with an underlying worst-case scenario. Comparing the threat exposure of a direct routing (1st column) with the result of the low-level route planner of the TMM (3rd column) makes the advantage obvious. The total threat accumulation could be decreased from about 8500 to 5700 *%km*. Due to the longer flight trajectory, the effect on the relative threat exposure is even more noticeable (55 down to 30%). The columns 2 and 4 in Figure 7 show the threat values of the actual flown trajectories, again under consideration of a worst-case scenario. It is obvious that the co-operation between the system and the pilots yields another improvement in terms of threat avoidance due to synergetic effects.

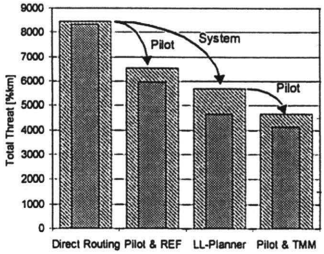


Figure 7: Threat exposure reduction

Workload and situation awareness

The TMM was designed to reduce the operator's workload by providing functions to support a better situation awareness and particular automation functions. During the experiments measurements of situation awareness and workload were conducted. Therefore, the experiment was stopped at dedicated points of time in order to perform the NASA Task Load Index (TLX) and the Situation Awareness Global As-Assessment Technique (SAGAT) [Endsley, 1988]. The evaluations were conducted four times each experimental run (reference sys-system and TMM). The measuring points were the task situations 1 (Transit Ingress), 2 (Low-level Ingress), 4 (Low-level Egress) and 5 (Transit Egress) according to Figure 4. Figure 8 and Figure 9 show the results of the assess-

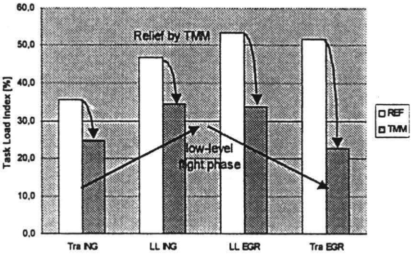


Figure 8: NASA TLX results over mission phases

assessments averaged over the four subjects. Concerning the NASA TLX rating (see Figure 8) it was found that the overall workload could be reduced massively by use of the TMM with an expected slight increase of workload during the low-level phases of the mission. The situation awareness assessment was based upon the evaluation of a total of 26 multiple-choice questions concerning situational features. Figure 9 shows the weighted results. An increase in situation awareness of about 10 to 15%, in particular during the early mission phases, can be observed.

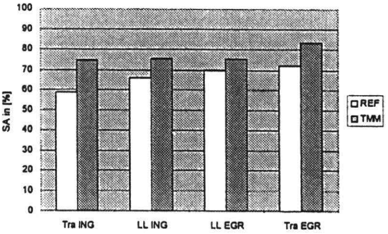


Figure 9: SAGAT results over mission phases

Subjective ratings

For a further evaluation of the system performance the experimental subjects had to give subjective ratings by the use of questionnaires in a de-briefing session at the end of the two day evaluation period. The main aspects of the survey were the evaluation of

- system performance of the TMM,
- acceptance of the TMM by the user, and
- overall assessment.

The rating scales covered a range of values from 1 to 7 and were each labelled by a pair of antithetic terms (e.g. good – bad; agree – disagree). The following paragraphs report on some selected results.

Figure 10 shows the pilots' evaluations of the quality and performance of the assistance functions offered by the TMM. The overall assessment can be regarded as very positive. Although, there can be identified some minor objections caused by unfamiliar display of timing and system information, insufficient training with the system and some shortcomings in pilot's behaviour modelling. Despite these (easy to remove) deficits the assessment of the implemented prototype was almost optimal.

	good						poor
situation gathering: ACO	● ●	● ●					
situation gathering: Route	●	● ● ●					
situation gathering: Timing	●	●		●	●		
situation gathering: System	●		●	● ●			
Assessment SAM-Threat	●	● ●		●			
Planning according to ATO	●	● ● ●					
ACO Routing	● ●	● ●					
Low-level Routing	● ●	● ●					
Low-level Guidance	● ● ●	●					
Mission Re-planning	● ● ●	●					
Time Management		● ●	● ●				
Monitoring Threats	● ●	● ●					

Figure 10: Evaluation of assistance quality

Figure 11 shows some selected but representative results of the overall assessment given by the pilots. The subjects fully agreed with the hypothesis that the TMM provides a better *big picture* in terms of global situation awareness. TMM is qualified to increase mission efficiency according to the pilots. The operators regarded the presented technology of Tactical Mission Management and crew assistance to be absolutely necessary, suited and adequate.

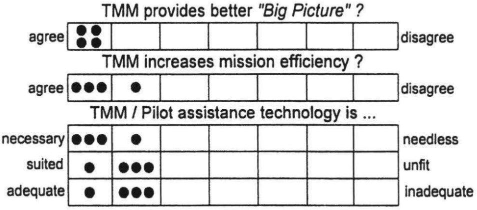


Figure 11: Overall assessment of the TMM

CONCLUSIONS

After having gathered years of experience in cognitive flight-deck automation and crew assistance, this paper presents the results of an exceedingly successful effort in the field of military combat aircraft. The functional breakdown of the *Tactical Mission Management System* has been derived from a model of human information processing, in order to approach a co-operative automation principle. A laboratory prototype system has been evaluated with fighter-bomber pilots in simulator trials. Besides the fact that the system was very well appreciated by the pilots, objective measures evidence a significant increase in performance in terms of threat avoidance and mission efficiency. These results could be achieved in conjunction with a noticeable reduction of the terrain collision risk and the operator's workload. Therefore, the system approach is highly recommended for application in the advancement and automation of combat aircraft.

REFERENCES

Billings, C.E. (1997), *Aviation Automation. The Search for a Human-Centered Approach.* Lawrence Erlbaum Associates: Mahwah, NJ.

Endsley, M.R. (1988), Situation Awareness Global Assessment Technique (SAGAT). National Aerospace and Electronics Conference.

Onken, R. (2000), Kognitive Systeme als Assistenten zur Leistungsverbesserung bei der Fahrzeugführung. *Automatisierungs- und Assistenzsysteme für Transportmittel.* Braunschweig.

Rasmussen, J. (1986), *Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering.* North-Holland.

Schubert, A. & Schulte, A. (2001), Automatische Generierung eines pilotengerechten Tiefflugprofils mit empirischer Modellparameteradaption für die Flugführung. *Deutscher Luft- und Raumfahrtkongress 2001.* Hamburg.

Stütz, P. & Schulte, A. (2000), Evaluation of the Cockpit Assistant Military Aircraft CAMA in flight trials. *3rd International Conference on Engineering Psychology and Cognitive Ergonomics.* Edinburgh.

COGNITIVE COCKPIT SYSTEMS ENGINEERING: PILOT AUTHORISATION AND CONTROL OF TASKS

R.M. Taylor

Defence Science and Technology Laboratory, Human Sciences,
A2 Bldg, Rm G007, Ively Gate, Ively Road, Farnborough, Hants GU14 OLX, United Kingdom.
Tel: 44 (0)1252-455784. Fax: 44(0)1252-455048. Email:rmtaylor@dstl.gov.uk

SUMMARY

A technology demonstration project on intelligent pilot aiding investigated systems engineering requirements for cognitive automation and interfaces in the complex, highly dynamic work environment of the single seat, fast jet military cockpit. The project integrated technologies for functional state assessment, task knowledge management and knowledge based decision support to provide context sensitive aiding. A framework was developed for pilot authorisation and control of tasks (PACT) through a tasking interface with variable levels of autonomy. The PACT system enabled the pilot to delegate responsibility for tasks whilst maintaining executive control through a set of pilot-computer contracts with a practical set of autonomy levels. This paper describes the philosophy and functioning of the PACT system and considers the implications for cognitive load and mitigation of risk using a decision ladder approach to control task analysis.

INTRODUCTION

The UK Ministry of Defence through DERA have recently completed a three year programme of applied research aimed at providing proof-of-concept demonstration of cognitive cockpit technologies for use in future envisioned air systems (Taylor, Howells and Watson, 2000). The aim was to allow the pilot, in control of the aircraft or of an uninhabited air vehicle (UAV) *"to concentrate his skills towards the relevant critical mission event, at the appropriate time, to the appropriate level"*. The Cognitive Cockpit (COGPIT) project sought to achieve this by providing cognitive system architecture for control of action with the following characteristics:

- Enabling the pilot to delegate responsibility for tasks whilst maintaining executive authority.
- Enhancing adaptiveness through context-sensitive aiding with the appropriate balance of feedback and feed-forward control strategies.
- Providing automation for reactive, opportunistic, skill-based and rule-based feedback control
- Supporting proactive, planful, strategic, tactical knowledge-based feedforward control.

The project involved integration of technologies for pilot state monitoring and automated decision support. The cognitive engineering methods used for requirements anlaysis are described elsewhere (Taylor et al, 201). The work required the development of a system with variable levels of autonomy that could be easily understood, operated and controlled in flight. The aim of this paper is to describe the system developed to provide flexible levels of autonomy, and to provide understanding of it's method of operation with particular reference to cognitive automation and interaction issues.

COGNITIVE COCKPIT FUNCTIONAL ARCHITECTURE

The COGPIT architecture couples on-line monitoring of the pilot's functional state and on-line task knowledge management and decision support for context-sensitive aiding, deriving information to mediate the timing, saliency and autonomy of the aiding. Three principle agents with different tasks can be distinguished as comprising the COGPIT system.

- A *Cognition Monitor* (COGMON) is responsible for monitoring the pilot's physiology and behaviour to provide an estimation of the pilot's functional state.
- A *Situation Assessor Support System* (SASS) is responsible for monitoring the aircraft situation and outside environment and recommends actions.
- A *Task Interface Manager* (TIM) is responsible for monitoring the mission plan, deciding automation and managing the cockpit interface.

The TIM module is concerned with on-line analysis of higher-order outputs from COGMON and SASS, and other aircraft systems. A central function for this system is maximisation of the goodness of fit between aircraft status, 'pilot-state' and tactical assessments provided by the SASS. These integrative functions enable this system to influence the prioritisation of tasks and, at a logical level, to determine the means by which pilot information is communicated through the TIM and the associated cockpit interfaces. Overall, this system allows pilots to manage their interaction with the cockpit automation, by context-sensitive control over the allocation of tasks to the automated systems (Bonner, Taylor and Miller, 2000).

A simplified representation of the processes performed by these agents, in support of updating the mission plan, is shown in Figure 1. The aim of the COGPIT architecture was to increase system adaptiveness by enabling changes to be made to the mission plan in response to changes in the situation. Information from monitoring the environment, the mission plan and the pilot provided inputs into the processes of re-planning the mission, automating tasks, deciding automation and configuring the cockpit. These processes then provided inputs into updating the mission plan. The architecture of the adaptive cockpit involved the functions and flow of information and control illustrated in Figure 2.

Figure 1. COGPIT Agents, Processes and Tasks

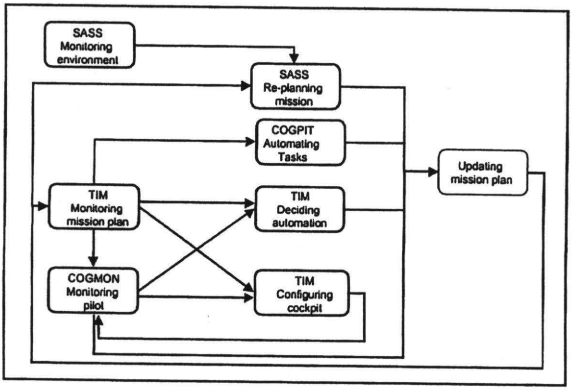
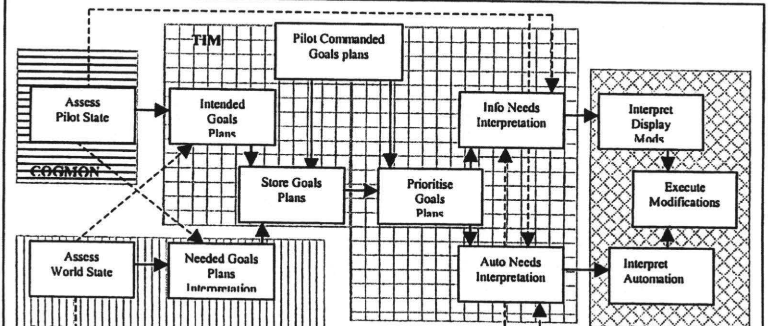


Figure 2. Flow of information across functions



TASKING INTERFACES

The idea of a tasking interface exploited the lessons learnt from the US Army's RPA program (Miller, Guerlain and Hannen, 1999). It arose from the need to be able to predict pilot expectations and intentions with reference to embedded knowledge of mission plans and goals. The aim was to provide an adaptive or "tasking" interface that allowed the operators/pilots to pose a task for automation in the same way that they would task another skilled crewmember. It afforded pilots the ability to retain executive control of tasks whilst delegating their execution to the automation. A tasking interface necessitated the development of a cockpit interface that allowed the pilot to change the level of automation in accordance with mission situation, pilot requirements and/or pilot capabilities. It was necessary that both the pilot and the system operated from a shared task model, affording communication of tasking instructions in the form of desired goals, tasks, partial plans or constraints that were in accord with the task structures defined in the shared task model.

Allowing pilots to choose various levels of interaction for the tasks they are required to conduct can mitigate the problem of unpredictability of automation. TIM utilised the monitoring and analysis of the mission tasks provided by the SASS combined with the pilot state monitoring of the COGMON to afford adaptive automation, adaptive information presentation and task and timeline management. The TIM functional architecture comprised modules for goal-plan tracking and for interface, timeline, automation and task management utilising a blackboard for goal-plan tracking information. Figure 3 shows the overall architecture for the TIM modules.

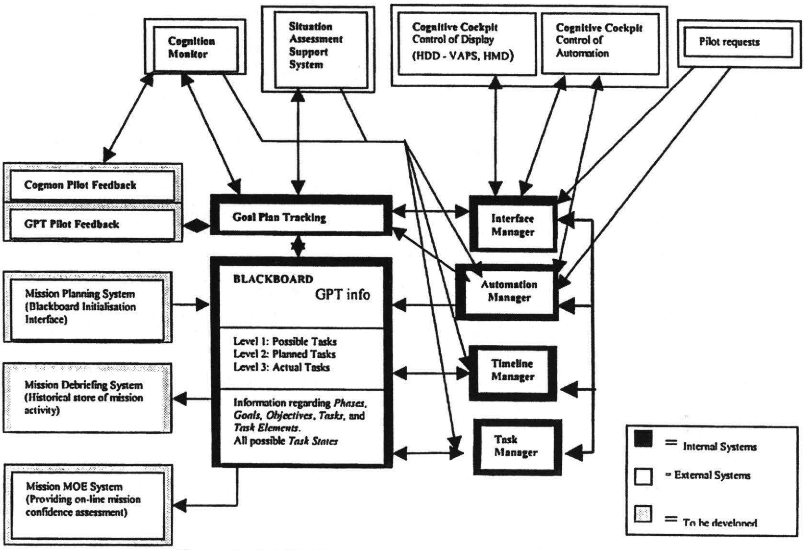


Figure 3. COGPIT system architecture for TIM functions.

LEVELS OF AUTONOMY

The building of trust between the pilot and the computer automation system has been identified as a key issue in increasing the capability for cognitive automation. Trust is built when consistency and correctness is observed in the computer system's decisions and actions. Two important guidelines for building trust have arisen (Reising, 1995):

Define the Prime Directives. These are overall governing rules which bound the behaviour of the aiding system, and yet provide a logical structure for aiding system to act in a rational and reliable manner, avoiding arbitrary behaviour, so that the pilot does not experience any surprises e.g. Asimov's Laws of Robotics.

Specify the Levels of Autonomy. These also bound the behaviour of the aiding system by limiting its decision authority for the performance of specific sub-functions to a set of system configurations specified and set by the pilot.

Providing flexible levels of autonomy for the performance of tasks and functions is a key requirement for implementation of the tasking interface concept. The human factors literature on automation and autonomy has identified the main constraints and design drivers. Sheridan and Verplanck (1978) first proposed 10 possible levels of allocation of decision making tasks, or levels of autonomy, between humans and computers. More recently, Parasuraman, Sheridan and Wickens (2000) have considered the application of automation to a four-stage model of independent information processing functions (information acquisistion, analysis, decision selection and action implementation). In doing so, they have sought to apply a revised set of levels of autonomy. Both the original and the revised redefined levels of automation are listed for comparison in Table 1.

Levels of Automation of Decision and Action	
1978 Original Set *Sheridan and Verplanck (1978)*	**2000 Revised Set** *Parasuraman, Sheridan & Wickens (2000)*
10. Computer does the whole job if it decides it should be done, and if so tells human, if it decides human should be told.	10. The computer decides everything and acts autonomously, ignoring the human.
9. Computer does the whole job and tells human what it did. The computer decides whether or not human should be told.	9. The computer informs the human only if it, the computer, decides to.
8. Computer does the whole job and tells human what it did only if human explicitly asks.	8. The compute informs the human only if asked.
7. Computer does the whole job and tells human what it did.	7. The computer executes automatically, then necessarily informs the human.
6. Computer selects action, informs human in plenty of time to stop it.	6. The computer allows the human a restricted time before automatic execution.
5. Computer selects action and implements it, if human approves.	5. The computer executes the suggestion if the human approves.
4. Computer selects action and human may or may not do it.	4. The computer suggests an alternative.
3. Computer helps determine the options and suggests one, which human may or may not follow.	3. The computer narrows the selection down to a few.
2. Computer helps by determining the options.	2. The computer offers a complete set of decision alternatives.
1. Human does the whole job up to the point of turning it over to the computer to implement.	1. The computer offers no assistance. The human must make all the decisions and actions.

Table 1 Comparison of Sheridan's original and revised levels of automation.

The DARPA/USAF Pilot's Associate (PA) programme, 1985-1992, is the foundation project in intelligent pilot aiding. The PA program provided a practical implementation of intelligent pilot aiding based on prime directives and levels of autonomy (see Taylor and Reising, 1998, for review). A summary of the PA design approach underpinning the levels of autonomy is shown in Table 2.

Pilot's Associate Design			
Operational Philosophy	**PA Capabilities**	**Operational Relationships**	**Modes for Levels of Autonomy**
The pilot is in charge - i.e. the pilot shall always have the capability to act according to his desires.	PA could not act on its own.		
PA's plans may be: Approved or rejected explicitly with little effort Approved or rejected pre-mission Approved or rejected implicitly by pilot action, or Ignored with predictable results	PA could make recommendations. PA could take actions based on pilot discretion.	OR2. The activity is performed automatically by the PA OR7. PA may perform an action only if various conditions are met.	**Associate.** In Associate mode, under full dynamic function allocation (DFA), the proposed system maintains advisory functions and accepts pilot allocated tasks, but also takes over tasks as the context demands.
The PA must operate in a predictable manner. The PA is required to monitor the pilot, not the other way around.	PA could fly the aircraft tactically on autopilot. PA could take action based on interpreting pilot intent.	OR6. PA has been given authority to perform, but with pilot consent.	**Assistant.** In Assistant mode, the PA would maintain advisory functions and also assume responsibility for tasks explicitly allocated to it by the pilot.
The PA must notify the pilot of key mission events (as defined and set by the pilot).	PA could diagnose malfunctions, identify mis-communications, & determine correct response.	OR5. PA may prompt the pilot. OR4. PA may remind the pilot.	**Advisor**
The effort required of the pilot to control the PA must be less than the effort saved by the PA. PA shall save more effort for the pilot than it creates - it shall be responsive to the pilot and not demanding of his resources.	PA could deal with ambiguities in human speech in the context of the mission.	OR3. PA may remind the pilot, if the pilot asks, or has authorised such.	**Standby**
		OR1. The pilot must perform the activity	**Inactive**

Table 2. Pilot's Associate Design Approach for Levels of Autonomy

PA design was guided by a top-level operational philosophy based on the pilot being in charge (Lizza and Banks, 1991: Lizza et al 1992). The goal of the PA was to provide consistently correct information, and to aid the pilot's decision making by helping to manage workload, reduce confusion, and simplify tasks. This led to the philosophy of the PA as an intelligent subordinate to the pilot, with specific capabilities for decisions and actions (Lynch et al 1995). These top level requirements led to specific operational relationships (ORs) for discrete PA sub-functions interactions, with increasing degrees of automation and autonomy (Krobusek et al 1989). From these ORs, pilot selectable levels of autonomy (LOA) were obtained for groups of functions governed by the required pilot operational relationship and interaction. Four discrete LOA modes were proposed, namely: Inactive, Standby, Advisor, Assistant, Associate. Each LOA mode was associated with tailorable functional clusterings for flexible responding to avoid too rigid automation imposed by design (Yadrick et al 1992). These modes were aimed to provide bounded, communicable structure for delegated levels of authority, minimising mode confusion, and building trust and confidence. Human factors research indicates that the required control structure should be cognitively simple, and not complex. Pilots tend to view computer autonomy simply as either automatic, with or without status feedback; semi-automatic, telling what will happen and asking permission to proceed; or advisory, providing information only (Lynch et al 1995).

PILOT AUTHORISATION AND CONTROL OF TASKS

Following from the above, a key development under the COGPIT project was the framework devised for providing only the necessary and sufficient levels of autonomy for the TIM Task Manager. The resultant COGPIT framework is known as the system for pilot authorisation and control of tasks, or PACT (Table 3). PACT is based on the idea of *contractual autonomy*. Borrowing an aircrew term from co-operative air defence, the idea is that the pilot forms a *contract,* or set of contracts, with the automation using the PACT system by allocating tasks to PACT modes and levels of automation aiding. The contract defines the nature of the operational relationship between the pilot and the computer aiding during co-operative performance of functions and tasks. Autonomy is limited by a set of contracts, or binding agreements, made between the pilot and the computer automation system governing and bounding the performance of tasks. Through PACT contracts, the pilot retains authority and executive control, while delegating responsibility for the performance of tasks to the computer.

Table 3. Bonner-Taylor PACT System for Pilot Authorisation of Control of Tasks

Primary Modes	**Levels**	**Operational Relationship**	**Computer Autonomy**	**Pilot Authority**	**Adaptation**	**Information on performance**
AUTOMATIC		**Automatic**	Full	Interrupt	Computer monitored by pilot	On/off Failure warnings Performance only if required.
ASSISTED	**4**	**Direct Support**	Advised action unless revoked	Revoking action	Computer backed up by pilot	Feedback on action. Alerts and warnings on failure of action.
	3	**In Support**	Advice, and if authorised, action	Acceptance of advice and authorising action	Pilot backed up by the computer	Feed-forward advice and feedback on action. Alerts and warnings on failure of authorised action.
	2	**Advisory**	Advice	Acceptance of advice	Pilot assisted by computer	Feed-forward advice
	1	**At Call**	Advice only if requested.	Full	Pilot, assisted by computer only when requested.	Feedforward advice, only on request
COMMANDED		**Under Command**	None	Full	Pilot	None performance is transparent.

PACT succeeds in reducing the number of automation or autonomy modes required to three - namely, fully automatic, assisted or pilot commanded - with a further four secondary levels nested within the semi-automatic, assisted mode, which can be changed adaptively or by pilot command. The PACT system uses military terminology for categories of support for Army land forces military operations (At Call, Advisory, In Support, Direct Support) to afford usability and compatibility with military user cognitive schemata and models. It provides realistic operational relationships for a logical, practical set of levels of automation, reduced from ten to six levels of autonomy, with progressive operator/pilot authority and computer autonomy supporting situation assessment, decision making and action (Figure

4). Mission functions and tasks, at different levels of abstraction allocated individually or grouped in related scripts or plays, can be set to these levels in a number of ways:

- Pre-set operator preferred defaults,
- Operator selection during pre-flight planning,
- Changed by the operator during in-flight re-planning, probably using Direct Voice Input commands (Dru-Drury et al, 2001).
- Automatically changed according to operator agreed, context-sensitive adaptive rules.

Figure 5 illustrates a set of mission functions and tasks with PACT contract levels arranged along a timeline in a hypothetical task network.

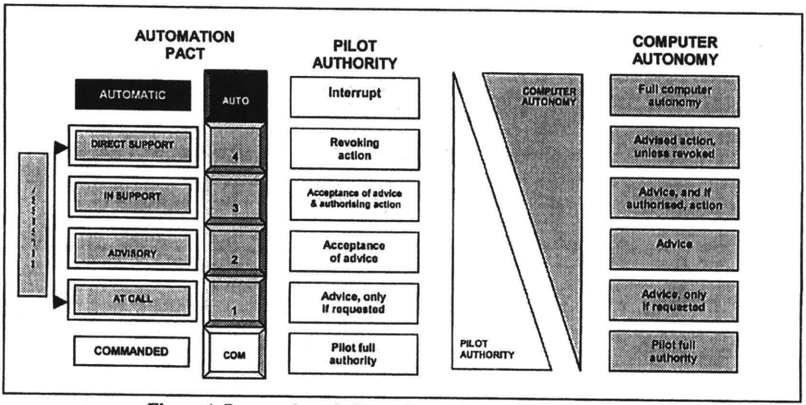


Figure 4. Progression of pilot authority & computer autonomy

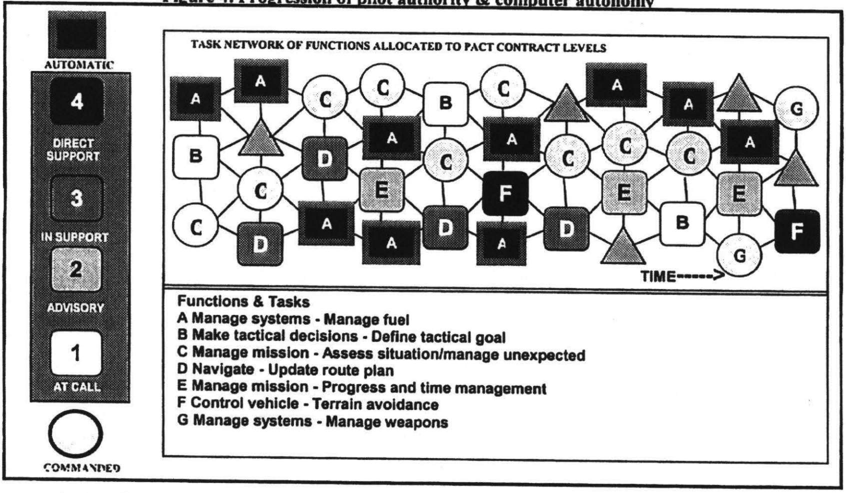


Figure 5. Task network of functions and tasks set to PACT contract levels

IMPLEMENTATION

In the COGPIT implementation, PACT levels are triggered adaptively, in accordance with PACT contracts, in response to contextual input from COGMON, SASS and TIM mission goal-plan tracking (GPT). The COGPIT intervention strategy is illustrated in Figure 6. The intention is to monitor and manage the variability in performance through a barrier system approach (monitor, detect, correct, reflect performance), and through appropriate cognitive streaming interventions (join, break, divert cognition). TIM feedback and feed-forward control messages are used with appropriate multi-modal intervention saliency (background, hinting, influencing, directing, compelling) developed to reduce cognitive bias with decision support systems (Taylor and Dru-Drury, 2001).

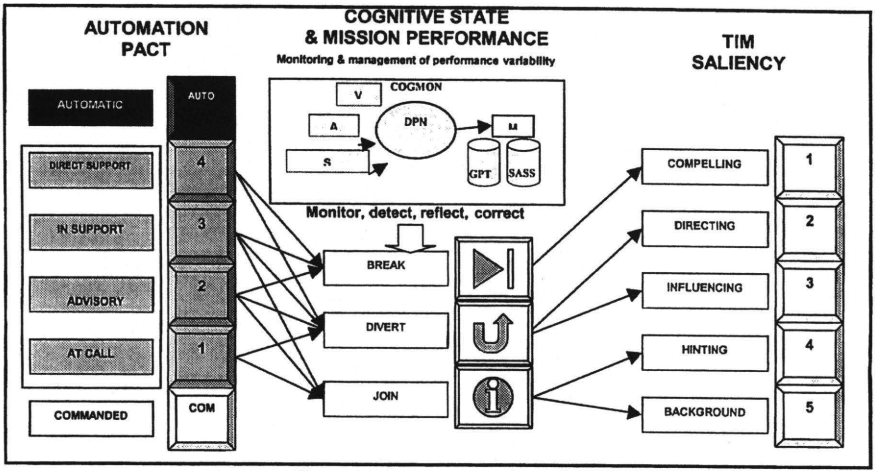


Figure 6. TIM intervention strategy

For the COGPIT demonstration, all the tasks in the mission scenario (operation of aircraft Defensive Aids System & re-route following pop-up threat) were pre-allocated to *possible* PACT level contracts by the pilot SME. The individual task PACT levels (defaults and contingencies) were set to mitigate the risks to achievement of the individual task goals. The TIM Task Manager distinguished between pending, active and completed tasks for the current scenario/vignette. Individual tasks progressed from pending, to active and to completed as the scenario progressed. An example of the COGPIT TIM Task Manager display is shown in the lower right window of the TIM control station in Figure 7.

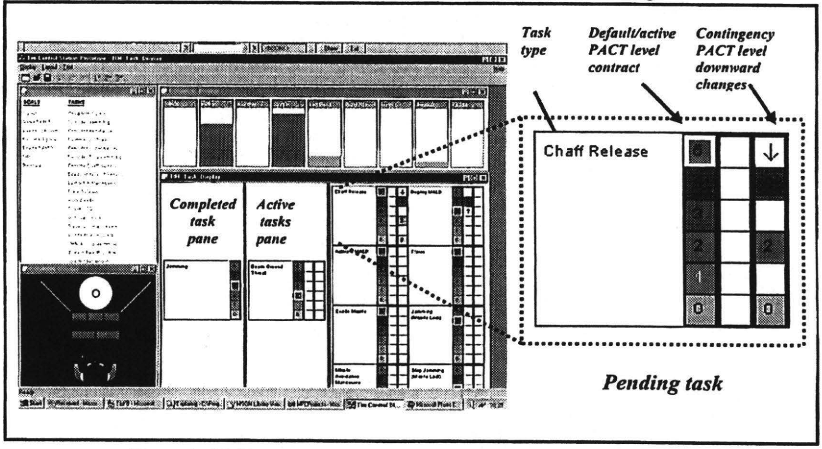


Figure 7. COGPIT Implementation of the TIM Task Manager display.

The left-most pane of the Task Manager display in Figure 7 shows the most recently completed task ("Jamming"), whereas the active ("Beam ground threat") and pending panes ("Chaff release"; "Activate MALD" etc) show multiple tasks as necessary. Individual tasks are represented as boxes containing their name together with their associated *possible* default and contingency PACT level contracts, which range from 0 (Commanded) to 5 (Automatic). The "Chaff Release" task is highlighted with its default PACT level display set for 5 (Automatic). Two additional PACT level displays are available to the right of this default display to show contingencies for how the PACT level can change by increasing (↑) or reducing (↓) PACT levels according to pre-set contracts. In this example, when the default PACT level is set at 5 it is possible under certain circumstances (e.g. Chaff remaining < 30%) for the PACT level to change down to 4, 2 or 0 (but not 3 or 1). Explanatory information on the circumstances for triggering changes (contract details) is stored and available for inspection. PACT levels that are unavailable are also indicated (reverse contrast caption) as shown for PACT level 1.

ANALYSIS

The PACT system is designed to support the pilot's cognitive work. The support ranges from providing advice to providing action. The cognitive work required can be represented in terms a perception-assessment-decision-action (PADA) decision ladder in state flow transition diagrams. Control task analysis (Vicente 1999) has been used to identify the structure of the cognitive work performed by the pilot and by automation at each PACT level. Figure 8 provides an example of the control task analysis for PACT Level 3 Assisted-In Support, expressed in PADA decision ladder terms. On the basis of control task analyses for each PACT level, estimates of the resultant or residual pilot cognitive load were identified in consultation with the pilot SME for different degrees of pilot critical involvement (immediate acceptance, critical acceptance, independent analysis). A simplified characterisation of the levels of estimated PADA cognitive work for Automatic, Assisted and Commanded PACT levels is illustrated in Figure 9, with SME identified risks for Automatic and Commanded PACT levels mitigated by the Assisted PACT levels. This analysis discovered that unlike the potential for immediate acceptance of advice on situation assessment, status, options and plans, authorisation of action was unlikely to occur, without critical appraisal or independent analysis. This may limit the reduction in cognitive load arising from automation of advised action (In Support, Direct Support). Increased concern about the validity of automated action seems reasonable during early familiarisation with the system, and probably will continue until trustworthiness of the system can be established.

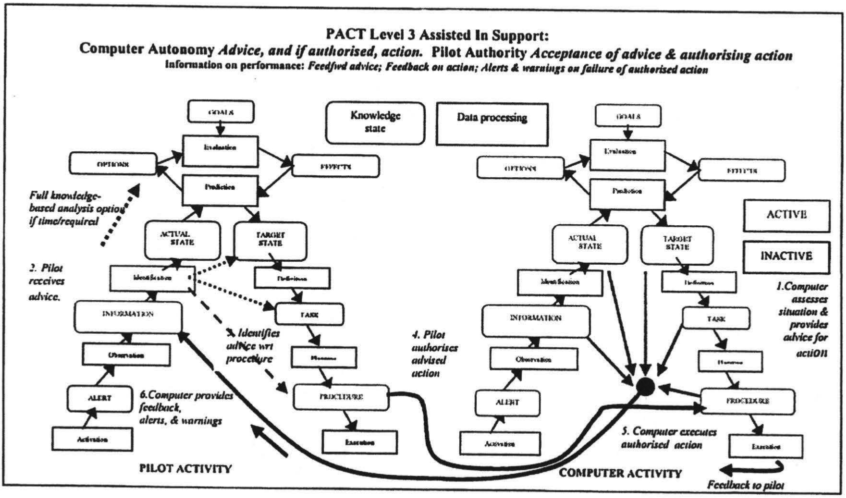


Figure 8. Control task analysis for PACT Level 3 Assisted-In Support

CONCLUSIONS

The final customer milestone demonstration of the integrated COGPIT system took place at DERA CHS in March 2001. It was deemed highly successful by the military customer. For future air systems with a mix of manned and unmanned platforms, it seems likely that advanced adaptive interface technologies, such as the TIM and PACT systems, will be a requirement for operator/pilot control of multiple UAV and Unmanned Combat Air Vehicles (UCAVs).

The work shows how useful assistance in the management of cockpit interfaces, tasks and automation can be provided by a tasking interface system based on a shared task model. The development of an effective TIM, with which pilots can interact easily, has been critical for the successful integration and acceptance of the outputs of the COGMON and SASS sub-systems. The technical specification of a tasking interface for this type of system is a major task, particularly as the functional components require iterative development, precluding early definition of inputs and outputs. Although it is relatively easy to track tasks instantiated in a mission plan, it becomes very difficult to track and support tasks that deviate from the intended plan requiring the system to infer likely pilot intent.

The PACT framework developed for pilot authorisation of control of tasks provides a simplified, practical set of flexible levels of contractual autonomy. This enables the pilot to delegate responsibility for tasks to the computer through a set of contracts that limit autonomy and bound the behaviour of the aiding system, while maintaining the pilot's authority through executive control. This is consistent with the principles for intelligent pilot aiding developed under the DARPA/USAF Pilot's Associate (PA) programme. Control task analysis, cognitive loading and risk analysis provides useful tools for understanding and modelling the functioning of the PACT system. The PACT framework seems sufficiently robust and useful to be applicable to other systems and environments requiring cognitive control with variable levels of autonomy, such as the control of multiple uninhabited vehicles.

ACKNOWLEDGEMENTS

The author acknowledges the significant contributions made to the development of these ideas by Chris Miller and Mike Bonner, and to the implementation by Simon Matthews, Tom Diethe and Paul Baker.

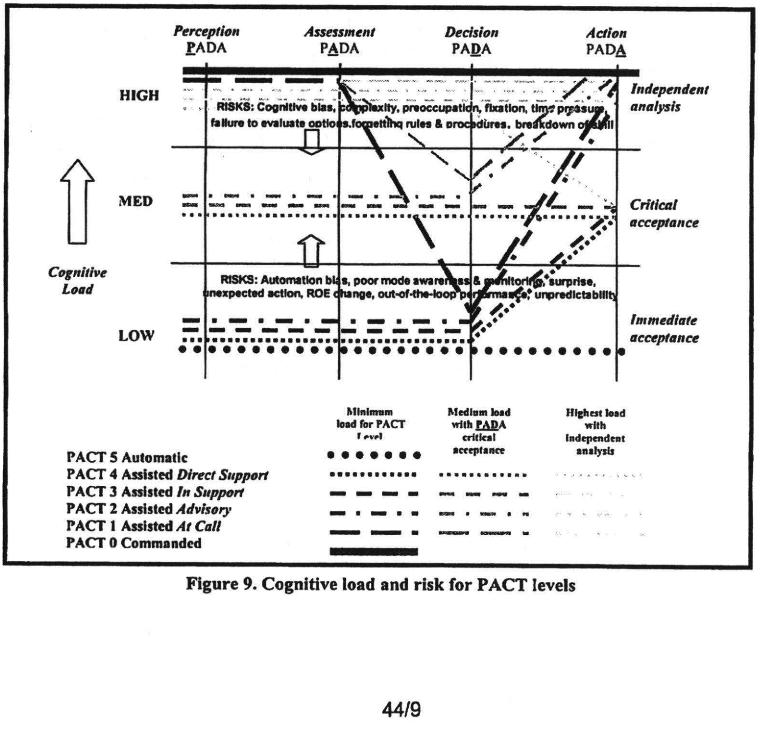


Figure 9. Cognitive load and risk for PACT levels

REFERENCES

1.Bonner, M., Taylor, R.M. and Miller, C. 2000. Tasking interface manager: Affording pilot control of adaptive automation and aiding. In P.T. McCabe, M.A. Hanson, and S.A. Robertson (Eds.), Contemporary Ergonomics 2000, pp 70-74. London: Taylor and Francis.

2. Dru-Drury R., Farrell, P.S.E., and Taylor, R.M., 2001. Cognitive cockpit systems: Voice for cognitve control of tasking autoamtion. In D. Harris (Ed), Engineering Psychology and Cognitive Ergonomics, Aerospace and Transport Systems. Vol 5, Chapter 12, pp 99-107. Aldershot: Ashgate.

3. Krobusek R.D., Boys R.M., and Palko K.D. Levels of autonomy in a tactical electronic crewmember. WRDC-TR-89-7008, Wright Patterson AFB, OH. 1989.

4. Miller, C.A., Guerlain, S. and Hannen, M. 1999. The Rotorcraft Pilot's Associate Cockpit Information Manager: Acceptable behaviour from a new crew member. Proceedings of the American Helicopter Society, 55th Annual Forum, Montreal, Quebec, May 25-27, 1999.

5. Lizza C.S., and Banks S.B. The Pilot's Associate: a co-operative knowledge based application, IEEE Expert, June 1991.

6. Lizza C.S., Rouse D.M., Small R. and Zenyuth J.P. Pilot's Associate: An Evolving Philosophy. WRDC-TR-92-3078, p149-154, Wright Patterson AFB, OH. July 1992.

7. Lynch R., Arbak C, Barnett B., and Olson J. Displays and Controls for the Pilot/Electronic Crewmember Team. WL-TR-96-3039, p 130-138, Wright Patterson AFB, OH., Dec 1995.

8. Parasuraman R., Sheridan T.B, and Wickens C.D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Sytsems, Man, and Cybernetics. Part A: Systems and Humans, Vol 30, No 3, pp 286-297. May 2000.

9. Reising J.M. 1995. Must the Human-Electronic Crew Pass the Turing Test? WL-TR-96-3039, p 103-108. Wright Patterson AFB, OH., December 1995.

10. Sheridan, T.B. and Verplank W.L., 1978. Human and computer control of undersea teleoperators. Technical Report. MIT Man-machine Systems Laboratory, Cambridge, MA.

11. Taylor,R.M., Abdi, S., Dru-Drury R., and Bonner M.C. 2001. Cognitive cockpit systems: Information requirements analysis for pilot control of cockpit automation. In D. Harris (Ed), Engineering Psychology and Cognitive Ergonomics, Aerospace and Transport Systems. Vol 5, Chapter 10, pp 81-88. Aldershot: Ashgate.

12. Taylor,R.M. and Dru-Drury R. 2001. Cognitive cockpit systems: The effects of reliability and saliency of aiding information on map situation assessment. In D. Harris (Ed), Engineering Psychology and Cognitive Ergonomics, Aerospace and Transport Systems. Vol 5, Chapter 11, pp 89-97. Aldershot: Ashgate.

13. Taylor, R.M., Howells, H. and Watson, D. 2000. The Cognitive Cockpit: Operational requirement and technical challenge. In P.T. McCabe, M.A. Hanson and S.A. Robertson (Eds.), Contemporary Ergonomics 2000, pp 55-59. London:Taylor and Francis.

14. Taylor R.M. and Reising J. 1998. The Human Electronic Team: Human Computer Collaborative Team-working. In Collaborative Crew Performance in Complex Military Operations, RTO-MP-4, pp 22-1 to17. NATO Research and Technology Organisation, Neuilly-sur-Seine, CEDEX, France, December 1998.

15. Vicente, K.J. 1999. Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-based Work. New Jersey: Lawrence Erlbaum.

16. Yadrick R., Judge C., and Riley V. Decision Support and Adaptive Aiding for Battlefield Interdiction Mission. WRDC-TR-92-3078, p122-126, Wright Patterson AFB, OH. July 1992.

A Formal Method for Integrated System Design:

How to Incorporate Cognitive Engineering Principles in a Systems Engineering Method

Michael Feary
MS262-4
NASA Ames
Research Center,
Moffett Field,
CA, USA
94035-1000

Lance Sherry
Honeywell
Inc./Rand Corp.
1200 S.Hayes St
Arlington, VA,
USA
22202-5050

Everett Palmer
MS262-4
NASA Ames
Research
Center,Moffett
Field,CA, USA
94035-1000

Peter Polson
Department of
Psychology
Campus Box
345
University of
Colorado

Summary

This paper describes a formal methodology (referred to as the Situation – Goal – Behavior (SGB) Methodology) that has been used in the design of software logic based avionics systems. It is expanded to incorporate some cognitive engineering guidelines and principles for use in design of complex systems. The SGB method is referred to as *formal* because the tool associated with the method has algorithms that are used to assist requirements specification. The purpose of this paper is to describe the application of the method as a framework for integrating the requirements for interface design, training, and procedure design within the system engineering method. This paper first describes the method, how it can be used, and then moves on to case studies which have shown the method to be useful for incorporating information requirements in the design process.

The Situation – Goal – Behavior (SGB) Method

The Situation – Goal – Behavior (SGB) method is a means of organizing and providing a common framework for all of the necessary groups involved in the design process of software based systems. The SGB method is similar to many decision-making behavior modeling environments, but with a different representation and approach to the problem of organizing decision-making software. The primary focus of this paper is to integrate the use of cognitive engineering guidelines and principles within a systems engineering method early in the design cycle.

The main benefit of the SGB table format from a cognitive engineering perspective is the ability to use the table as a type of description language for discussing design implementation across the engineering, training, operational (flight standards, etc.), interface, and procedure design groups. This early integration is necessary for the development of systems which are easier to train, learn and operate. (Billings, 1996). The ultimate goal of the use of the methodology would be to have a design change implemented based on a training requirement, for the purposes of reducing training, while early in the design process.

Proper task specification and the communication of the goals of the system in automated systems is crucial in complex, automated system development (Rasmussen, 1994). It is, therefore, the intent of the SGB method to involve the users in specifying the mission, or the tasks that the system will need to accomplish. If the task specification is successful and complete, the design engineering task of making the system function properly should be straightforward. Additionally, the training and interface development should be task-oriented as well, if it were possible to put the task specification in a framework that is accessible to the operational community (users, training developers, interface developers, etc.) in the design cycle the result should be a more "user-centered" system.

How the SGB Method Works

The SGB method uses a tabular representation of a systems engineering model which represents the decision logic of an automated system (Table 1). The table is composed of Situation Inputs and States, Situations, Goals, Behaviors, and Behavior Outputs.

In a typical design cycle for a large complex system, the customer requirements for a new system would ideally be presented by the customer or a representative of the customer from the operational community (users, training, and procedure development representatives), to provide the desired Goals and Behaviors for the system. As the Goals and Behaviors for the system are defined, the engineering design team can provide the inputs required by the system to determine the situation and perform the correct behavior. The *formal* systems engineering portion of the method includes some algorithms to ensure that all possible combinations of inputs are covered, and that no two combinations are the same, which would lead to a non-deterministic system. Typically, this process will iterate several times, with engineers asking the customer what behavior they might prefer for situations that were not defined originally. To initially demonstrate the method and the use of the table, an example SGB table for a very simple digital wristwatch which has one button and an 8 character wide display is provided (Table 2).

Table 1. SGB Table. The X-axis of the table shows the Situation, Goal and Behavior descriptors, and the Y-axis shows the Situation input and input-state names.

		Goal 1		Goal 2
		Situation A	Situation B	Situation C
Input I	state a			
	state b			
Input II	state a			
	state b			
	state c			
		Behavior Description 1		
Output	state a			
	State b			

Table 2. Example of SGB table usage

		Display Date		Display Time	
		Change to date mode	Stay in date mode	Change to time mode	Stay in time mode
Button	Pushed	true		true	
	Not pushed		true		true
Current goal	Time mode	true			true
	Date mode		true	true	
		Display Date		Display Time	
Display	Date	true			
	Time			True	
Calculate	p:Date	date			
	m:Time			time	

Situation Inputs and States

Inputs and Input states are developed by the design engineers to try to meet the Goals specified by the customer. The Inputs may represent a detected action by the user, an input that is automatically detected by the system via a sensor, the result of calculation elsewhere in the system or other means. Input states are used to quantify the input to enable the automation to interpret the input value. In Table 1, Inputs I and II are represented by two and three states respectively, however in actual tables there is no limit to the number of states per input. An example of this will be shown in the following section.

How Inputs are Defined. As a cognitive engineering requirement, these inputs should be meaningful to the operational community. This is easily met in a simple system, such as watch, by using the button label. However, it can be a challenge with a larger, complex system with multiple sensor inputs, and results of calculations elsewhere in the system which have no physical attributes to provide inherently operationally meaningful names.

If the inputs can be given operationally meaningful names, then the interface design groups can design feedback to the user on the accurate goal of the system. The example in Table 2 is a simple illustrative example and because it does not have much software decision-making logic, the input names are intuitive and operationally meaningful. However in larger, more complex systems, inputs may not have names that are easy to understand and operational, especially if the input refers to a portion of computer code, for which names may be arbitrary. It is important to make the inputs operationally meaningful so that the situations resulting from the combinations of inputs are operationally meaningful. If the situations are operationally meaningful, users should be able to answer the questions "Why is it doing that?", and "What is it going to do next?". Unfortunately, in many of today's systems, these questions are asked all too frequently.

Another requirement is to identify the inputs which are visible to the operator. Ideally, all inputs would be visible to the operator, but the inputs which are not must be justified and agreed upon by the different groups. Additionally, the situations which are defined by the inputs which are not visible to the operator will have to be accommodated. These situations will naturally lead to confusion, because there is a difference between the information the operator is using and the information the system is using.

Formal analyses (e.g. model checkers) can perform an evaluation of the situations defined in the system that the user can see (via controls and displays) versus those situations which the user cannot see because they are based on system information which is not displayed or requested by the user (operator).

Situations

Situations are a description of the combination of inputs intended to make the table readable and easily understood by the customer. In Table 1, Situation A can be seen as a description of the combination of Inputs I and II. The Situation descriptions can be very valuable to the training, interface and procedure development communities later as they define the tasks of the system, and as the system is reorganized. During the design cycle the situation description becomes the foundation for organizing the hierarchy of goals.

How Situations are Defined. The Situations in the table provide a description of each of the events that can occur in the system. In the watch example, the system needs to be able to cope with at least two situations to display the date. First, if the system is currently displaying the time, then pushing the button will display the date, but the system also needs to know the current state so that pushing the button changes the situation, otherwise pushing the button would cause the watch to always switch to date mode or always switch to time mode.

Goals

Goals refer to the mission objectives that the customer will be responsible for. In complex systems, goal descriptions are a hierarchical grouping of the situations by the operational community and are representative of the task to be accomplished. In this way, Goals are used to organize what the design team thinks is most important for the user to know.

How Goals are Defined. Goals or Behaviors are generally the first requirement given by a customer. In the example in Table 2, the customer for this simple wristwatch has specified two goals and behaviors, Display Date and Display Time. In this simple example, the Goal and the Behavior are the same, however the usefulness of the Goal representation becomes evident in the design of larger and more complex systems, when decisions need to be made about what priority should be given to different system behaviors. The interaction can be seen between the Goal and the Behavior in the form of an "m:" preceding the "Time" behavior output. This "m:" refers to a sub-mission, or lower level table, which describes functions that the design team has decided do not need to be immediately visible to the operator. In this case the "m:" refers to a table which describes how to set the time on the watch.

Behaviors

The Behaviors refer to how the system will perform to meet to associated Goal. The customer defines the required behavior for the system, and the engineering portion of the design team uses the outputs and output states in a similar fashion to the inputs and input states to meet the performance objectives. The Behaviors are natural language descriptions that are written by the operational group of the design team and should closely match the goal.

How Behaviors are Defined. The behaviors are used in conjunction with the Goals to organize the system around cognitive engineering principles. When the desired behaviors have been specified by the customer, the design team can set about organizing the behaviors based on importance and level of interaction required by the operator. Those behaviors which the operator can have little effect on in the system are candidates for being made less visible and those behaviors which the operator needs to know about to operate the system correctly need to be made more visible. This is accomplished through the use of hierarchy with multiple tables, as shown in Table 2 with the "m:" preceding a behavior output. In the method the highest level tables are designed to be the most visible. This process needs to be carried out with the training and interface design representatives.

The table does not automatically provide the best information requirements for a particular system. These decisions require operational experts within the design team. What the tool does is allow the different members of the design team to communicate with each other in a format that forces them to be precise about what they intend. Later design decisions and environmental requirements may force modifications to the information requirements, but having the decisions documented in this format allows the changes to be made easily. Once all of the information requirements for the display have been identified, the interface design team can apply other human factors guidelines to aid in the design of the display. An example of this is the concept of *label following* which has shown that users will tend to respond to a task by using an interface object that most closely resembles the task, even if it is not the correct action. (Franzke, 1995).

Another cognitive engineering practice is to match the input devices to the behaviors to avoid functional overload. A functionally overloaded input device is one for which the behavior of the input is dependent upon the state of the system at the time the input device is used. This functional overload may easily become cognitive overload and error prone for the operator, if that operator has poor feedback, is improperly trained, has memory lapses from fatigue, etc. (Sherry et al., 2001)

Behavior Outputs

Combinations of behavior outputs are used by the engineers to meet the requested behavior by the customer, and give feedback to the user.

How Behavior Outputs are Defined. From a systems engineering perspective, the output states can be given different classifications within the SGB tool, examples of these include assignment to another table (shown by the "m:" preceding a Behavior Output) or, assignment of primitive functions (e.g. an equation) where the table stops and leads to a calculation that can be made by the automation. The primitive function is represented by the "p:" in the Date behavior output in Table 2. In this case, it refers to the watch calculating the date.

Summary of Cognitive Engineering Guidelines and principles using the method

- Develop, identify and organize Goals for the level of interaction
- Develop operationally meaningful inputs
- Identify/ensure inputs are visible to the operator
- Match the Goals/behaviors with display design
- Match the input devices to the behaviors to avoid functional overload

Application to aviation automation

Aircraft automation, particularly the automation surrounding vertical flight guidance, has been cited as an area of training difficulty and a source of confusion during operation. A number of incidents have been attributed to a lack of crew understanding of what the automation was doing.

As part of a NASA program, the pieces of the SGB method were evaluated using aircraft automation on a modern, commercial transport category aircraft very similar to the Boeing MD-11 aircraft. To determine where the method should be focused, a survey was distributed to Boeing MD-11 line pilots to assess where pilots thought they were having difficulty, and where they would like the most help with the automation. (Feary et al., 1998) More than 75% of the pilots surveyed felt that aspects of the Vertical Flight Guidance system were trained inadequately, including the FMS Speed Logic, PROF (Vertical Navigation Mode), and the interpretation of the Flight Mode Annunciator (FMA). The SGB Methodology was used to design a new interface, procedures and training material for the Vertical Flight Guidance system. These changes were evaluated experimentally (Feary et al., 1998), and examples from these experiments should demonstrate the utility of the method for some applications.

Table 3 shows a small fraction of the Vertical Guidance for a modern commercial air transport. (Honeywell Cockpit Pilot's Guide, 1994). The Goals, Situations, inputs, input states, and behaviors are representative of those used in the larger table, although the inputs have been reduced to correspond with the different behaviors depicted. Specifically, Table 3 shows 3 behaviors used in the descent phase of flight.

Before discussing the meaning of Table 3, some background domain information needs to be defined. The input state Reference altitude refers to the altitude specified by the pilot using a knob on the Autopilot control panel, generally referred to as the Mode Control Panel (MCP).

Starting with the input column, the input *VG Type* refers to the level of automation being used.

Vertical Navigation (VNAV)/ Profile(Prof) mode refers to the names of the most highly automated modes in current aircraft. In *VNAV/Prof* (the names differ for different aircraft manufacturers), speed and altitude targets are specified in the Flight management Computer (FMC) and flown automatically. Airmass – VNAV/Prof refers to slightly less automated mode for which the speed targets are automatically generated by the FMC, but the altitude targets are not. Finally the AFS mode refers to one of the least automated modes, for which the pilot selects both speed and altitude targets.

Table 3. Example of SGB table of a vertical guidance system. In the table, "1" refers to a true state.

Goals		Airmass	Descent	Late	Descent	Descent Path	Overspeed
Inputs	Situations/ Input States	Aircraft is Descending (without both Prof and FMS Speeds)	Aircraft is descending early of D/A Path and Prof/FMS speed engaged	A/C is level late of the D/A Path level at the ref. Alt and the ref. alt	Aircraft is descending late of D/A Path and Prof/FMS speed engaged	Aircraft exceeds speed tolerance while descending on D/A path	Aircraft is level with a speed that exceeds the speed tolerance when ref. Alt is lowered and a/c captures D/A path
VG Type Altitude	VNAV /Prof			1	1	1	1
	Airmass – VNAV/Prof	1	1				
	Airmass - AFS						
Aircraft Altitude	Above distance Referenced D/A path			1	1		
	below distance Referenced D/A path						
Aircraft Speed	Overspeed for D/A path					1	1
	Within speed tolerance for D/A path	1	1	1	1		
Aircraft Altitude	Within D/A Path capture region						
	Not Within D/A Path capture region	1	1	1	1		
Reference Altitude	Has not changed						
	Has changed		1	1			1
Behaviors		Airmass Descent to the D/A path D/A path speed	Referenced recapture using the descent profile	Airmass Descent the D/A the late profile	Referenced to recapture path using descent speed	Airmass Descent D/A path path descent	Referenced around the at the D/A speed profile
Altitude Target	M:Climb/Cruise						
	M:Descent/Approach	Descent/ Altitude	Approach Target	Descent/ Altitude	ApproachTarget	Descent/ Altitude	ApproachTarget
Speed Target	M:Late descent			Late Speed	Descent Target		
	M: Descent/Approach					Descent/ Speed	Approach Target
	M: Airmass Descent	Airmass Speed	Descent Target				
	P: engine-out						
Speed/ Altitude Control Mode	P: THRUST |HOLD						
	P: PITCH ||IDLE	PITCH |	| IDLE	PITCH|	| IDLE	PITCH |	| IDLE
	M: D/A DESCENT SPEED/Altitude						

The Speed/Altitude Control Mode refers to how the automation will control the airplane to meet the guidance requirements. For example an aircraft may pitch for a certain rate of climb or descent, or the aircraft may pitch for a particular airspeed. In another, unique case the automation may pitch the aircraft to remain on fixed earth-referenced path referred to as the Descent/Approach(D/A). The D/A path takes into account Air Traffic Control (ATC) speed and altitude restrictions and computes a descent path such that the engines should remain at idle thrust, resulting in the most fuel efficient descent possible should the aircraft stay on the path.

The behaviors depicted in the table exist to deal with exception cases which would cause an aircraft to deviate from the D/A path. The Airmass Descent case accounts for scenarios during which the FMC has not computed a D/A path, this may happen if the pilot does not enter a destination in the FMC, or for the short time while the FMC is computing the path. The Late descent behavior accounts for a scenario for which the aircraft is not allowed to descend until passing the optimum descent path (due to ATC constraints), or a case in which the aircraft has too much energy to stay on the optimum descent path (an unforecast tailwind would cause this). The Path Descent Overspeed case accounts for a scenario similar to the Late Descent case, however in this case the aircraft the aircraft starts on the path, but once reaching a certain speed threshold faster than the assigned speed target, the aircraft flies away from the D/A path at a predetermined speed.

The behaviors shown in Table 3 are all different, however the display in many modern commercial aircraft are very similar for the different behaviors, and in at least one they are exactly the same presentation to the pilot with little to no displays available to differentiate between them.

Case Study 1: The Behavior based Flight Mode Annunciator (FMA)

The number of distinguishable display events should be directly related to the number of different Goals/Behaviors of the system. This is required for users to be able to answer the question "What is it doing now?", which is a question frequently asked by users of complex, time-critical and highly dynamic systems (e.g. aircraft automation) today.

That simple guideline was used to evaluate an automation display of a modern transport category airplane, and several cases of non-compliance were found. To evaluate the impact of this non-compliance on one automation display, referred to as the FMA, Feary et al., (1998) conducted a study with a modified display in a certified simulator with actual airline pilots who were current and very experienced in the type of aircraft being evaluated.

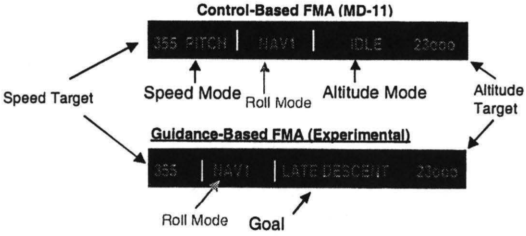


Figure 1. Control vs. Guidance-Based FMA

The experiment involved a change in the organization and wording of the Flight Mode Annunciator display, as it is seen in the cockpit of a modern, transport category aircraft. To provide more conceptual system information on the display, the new display annunciated the Goals of the system. These Goals were based on the Behaviors in the SGB representation. The current display was compared to the new display (Fig.2) and, combined with a training package, distributed to participants in both conditions (to equalize the group). The displays were tested by 27 current line pilots, with at least 1 year of experience on the MD-11 (an airplane with the same display and vertical guidance to the experimental display), flying a short flight in a Fixed Base Simulator. The pilots in the new display condition showed significantly less errors ($p>.03$) for the experimental scenario.

Case Study 2: Modification of a functionally overloaded Control Device

An example of functional overloading was observed on the MCP of a modern, transport aircraft. The current MCP for the aircraft under investigation uses one button *PROF/VNAV* to engage a higher level of automation, and depending upon the situation of the automation at the time, pushing the button may result in one of at least 13 different behaviors. It is interesting to note that in at least one aircraft, the FMA displays the letters PROF to refer to descending on the D/A path, while at the same time using a button labeled PROF to refer to changing between more and less automated modes of flight. By clearly defining the behaviors in the system, the complete SGB tables of the vertical guidance system (not presented here) show this contradiction clearly. Additionally, using the table as a starting point for further analysis, it can be seen that descending on the D/A path is fundamentally different than any of the other modes in that it has a fixed earth-reference point. A proposed solution is shown in Figure 2, resulting in the simple addition of a DES PATH button to refer to the unique mode, and some modification of the logic to restrict automatic switching to other descent modes. (Sherry et al., 2001)

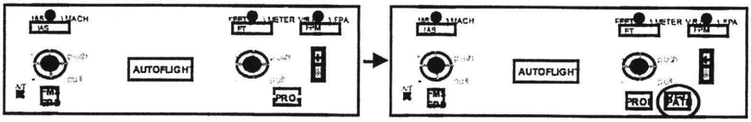


Figure 2. Comparison of modern Autopilot Control Panel with MCP designed according to the cognitive engineering principles. A separate input device (DES PATH button) provides the option to arm the capture and tracking of the FMS optimum path. This button has only one behavior – to capture and track the path.

Current and Future Work

At the time of writing an initiative has begun to link the Situation Goal – Behavior - Method to task analytic techniques to evaluate usage and procedure techniques. This work will consist of building a model of a pilot's conception of a portion of the Vertical Guidance System and making a formal comparison with Situation – Goal – Behavior model of the same system. It is hoped that the task analytic model will reveal areas of possible pilot error, and be able to predict areas of the system for which operators lack conceptual understanding.

New versions of the tool are being developed and evaluated, and parts of the tool are publicly available. Requests can be sent to mfeary@mail.arc.nasa.gov or Lance.Sherry@cas.honeywell.com.

References

Billings, C., 1997, *Human-Centered Aviation Automation*, Mahwah, NJ, Lawrence Erlbaum Associates.

Feary, M., Sherry, L., Palmer, E., and Polson, P (1998) Aiding Vertical Guidance Understanding. NASA Technical Memorandum 1998-112217. NASA Ames Research Center, Moffett Field, CA, USA.

Franzke, M. (1995), Turning research into practice: characteristics of display-based interaction, in Proc. CHI'95 Human Factors in Computer Systems, ACM, pp. 421-428

Honeywell (1994). *MD-11 Cockpit Pilot's Guide* (Revision 1). Honeywell Document C29-1101-01-01. Phoenix, AZ., USA

Irving, S., Polson, P., and Irving, J. (1994) A GOMS Analysis of the Advanced Automated Cockpit. In Proceedings of CHI 94 conference. Boston, MA.

NTSB (1999) Safety Recommendation. A-99-39 through 44.

Rasmussen, J. (1994) *Cognitive Systems Engineering.* New York, USA: John Wiley and Sons, Inc.

Sherry, L., Feary, M., Polson, P., Mumaw, R., and Palmer, E. (2001) A Cognitive Engineering Analysis of the Vertical Navigation (VNAV) Function. NASA Technical Memorandum 2001-210915. NASA Ames Research Center, Moffett Field, CA, USA.

Vicente, K. (1999) *Cognitive Work Analysis.* Lawrence Erlbaum Associates.Mahwah, NJ.

INFERENCE MAKING DURING INCIDENT ANALYSIS IN AIR TRAFFIC MANAGEMENT

Cyril BARRIQUAULT & Janine ROGALSKI
Laboratoire Cognition et Activités Finalisées
ESA 7021
2, rue de la liberté- 93526 St Denis
E-mail: cbarriquault@univ-paris8.fr
rogalskij@univ-paris8.fr

ABSTRACT

The main issue of the paper is to contribute to the design of experience feedback (EF) systems in the field of air traffic control (ATC). A method is proposed for analysing incident reporting by investigators belonging to the controllers community 1) for identifying elements possibly shaping their activity, and 2) for proposing tools for supporting them in their inquiry task. Such a support system should help not only to find sharp-end causes but also to prevent failures by identifying possible defenses at various levels in the air-traffic system. A preliminary study led us to consider two models for coding verbal reports of an expert investigator. The results confirm that 1) significant attention was paied to potential elements; 2) filters effects are due to the reporting system, which does only consider factual elements. Consequences are derived for a methodology for incidental situations analysis.

Key words: incident analysis, safety model, protocol analysis.

INCIDENT ANALYSIS AND THE EXPERIENCE FEEDBACK PROCESS

Air Traffic Control (ATC) is one of the safest systems in terms of crash probability. The rather well controlled technical aspect contrasts with the uncertainty of context—aviation is a very fast evolving domain (organisation, economy)—and particularly with the variability of human activity. In the same time, air-traffic increase renders danger more unacceptable for the public.

Safety management in EF—as in almost all risky domains—is largely based on experience feedback (EF). In fact, the role of episodic knowledge issued from problematic situations was demonstrated as a component in individual expertise development in process control (Baerentsen, 1996). This includes the cases where episodic knowledge comes from others' experience; as a consequence, EF is also useful for developing collective and organisational competence. Three categories of episodes can be defined in EF: cases of normal situations, near misses, severe incidents and accidents. The notion of *near miss reporting* as a safety tool was particularly developed in van der Schaaf's study dealing with the domain of chemical process control (van der Schaaf, 1992).

Woods and al. (1994) stressed the necessity of not focusing only on accidents or severe incidents: incidental situations provide useful information for safety management. From another point of view, they claimed also that, behind individual human error, organisational determinants are of first importance for identifying possible failures in the safety system as a whole; in the same line, Rasmussen (1997) proposed models for analysing actual operators' behaviour in terms of behaviour shaping features, beyond the individual human factors. Two main levels are concerned with EF—whatever its shape—: the level of safety management and the level of front-line operators training.

EF consists in recording critical events and constructing an explanation, in order to create an organisational track. Such a track is later used to define actions in order to prevent risks at the various organisational levels. The initial phase of incident analysis is an essential one, as it is the first step in safety process. EF first aimed at a better understanding of "human errors", their suppression, or, protection against them. Nowadays, this tool is used to communicate to the whole system the main lessons and possible defences inferred from an incident inquiry. However, as it is based on observed incidents, the weakness of EF is its present difficulty to help protect the system against extremely rare and catastrophic scenarios. The actual problem is how to choose a relevant scenario from the wide database.

The purpose of this paper is to study the process by which ATC investigators develop representations about incidental situations from initial information gathering to final reporting. In ATC, incidents are analysed by investigators at the local operational level. They have an expertise as air traffic controllers, and are in charge of safety briefings in case of severe near misses (as reported by pilots or as identified in computer-based records). They also receive direct information from controllers reporting near misses.

The studied situation is intermediate between the near-miss reporting system studied by van der Schaaf (1992), or the near miss reporting in aviation (Aviation Safety Report System), and the incident reporting in the UK Railways as implemented in the CIRAS system (Davies et al., 2000). In the former cases, operators themselves are reporting incidents encountered in their activity; in the later, operators are interviewed by a psychologist (not belonging to the organisation).

THE FIELD STUDY: THEORETICAL FRAMEWORK AND METHOD

To analyse incidental situations, ATC investigators view audio- and video-recordings (radar images, communications) and they interview controllers. From this information phase, they build an explanation of the observed facts. Such an analysis takes about one month. An analysis report is then written and distributed to two types of actors: local operators and managers. Finally, data on the incident are stored in a computer-based information system, for statistical analyses. Two safety models help us coding the verbalisations and writen reports expressing such a construction: a *reactive* model and a *proactive* one. These models are derived from van der Schaaf's model.

- *Reactive and proactive models for near-miss analysis*

On the one hand, some risk models explain an event with a short number of real facts. That is a "cause oriented" or reactive paradigm (Rasmussen, 1990; Paries & Amalberti, 1997). Some models frame a search of large causal factors encompassing individual, organisational and contextual factors (Leplat, 1985; van der Schaaf, 1992; Reason, 1997). On the other hand, some authors (Wioland, 1997; Doireau & al., 1997) consider that it is important to focus on successful incident measures. The idea is to collect error identifications or recoveries in order to reinforce these natural safety barriers. These models help modelling a causal and temporally marked description of the past.

From the point of safety management, these models are reactive in that they are considering causal analysis after the identification of an incidental situation. They are quite directly developed within the scope of experience feedback in the plain meaning of the words. They correspond to reactive regulation, when considered in the general scope of systems regulation. Figure 1 sketches the model we design in such a reactive orientation.

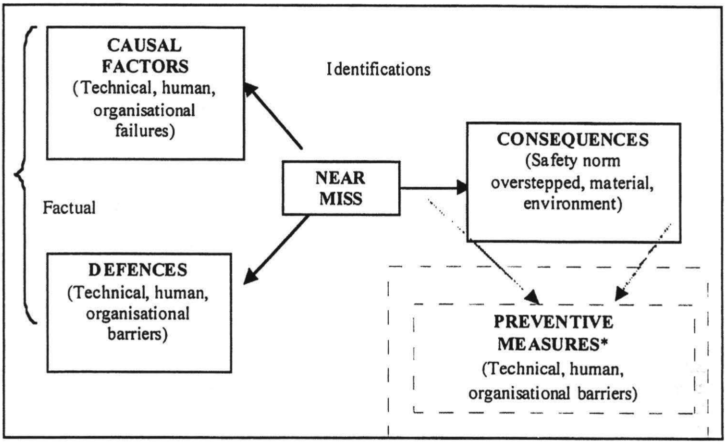


Fig. 1: Schema of actual risk inferred by inquirers from a *factual* event: a near miss.
On the basis of an incident, inquirers analyze facts in order to build a causal tree. They express defences that have protected the system against the catastrophic issue. Finally, they talk about means of prevention against the kind of encountered case: the preventive measures (* non analyzed in this paper).

However, when analysing the investigators' analysis process (Barriquault, 1999), we noticed that they do not limit themselves to this factual description: they also formulate "what if" questions. "What would have happened if defences were not functioning?" "Which defences can we imagine if this scenario were realised?" In order to take into account this kind of analysis, we have been interested by "potential-risk" approaches and proactive models.

These theoretical and practical frameworks are not only based on real factor investigation but on "possible world" construction. There are three focus of study. Firstly, proactive models built in the human-factors community are derived from Flanagan's critic incident method (1954). The method is a clinical anticipation of accidents based on in-deepth observations and interviews. Otherwise, human reliability models are based on probabilistic analysis (Vuillemeur, 1988; Fadier, 1990; Kirwan, 1994; Johnson & Bottier, 1999). For instance, the failure tree consists in assessing the probability of occurrence of a feared event. Finally, safety managers recently suggest the interesting concept of *possible scenarios* construction for prevention or precaution (Llory, 1996; Godard, 1997; Pariès & al., 1999).

The goal is to look for particular scenarios and try to avoid them even if only rare cases support them and with a weak probability of realisation. We call this process an inference of *Potential Near Miss (PNM)*. After that phase, the technical system can be protected against non-established risk by improving safety barriers: what are possible organisational decisions? Which kind of human safety management is possible? We call this construction of possible facts the *counterfactual thinking*. Counterfactual thinking is a cognitive simulation of alternative outcomes. It consists in imagining causal processes that conduct to better or worst outcomes: possible facts that contrast with actual ones (Hesslow, 1988; Hilton, 1990; Roese, 1997).

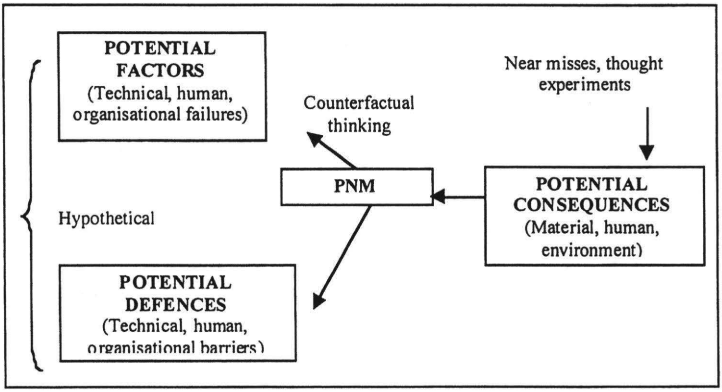


Fig. 2: Schema of potential risk inferred by *counterfactual reasoning*.
From accident precursors (conjunction of incidental factors, serious incident) or from thought experiments, investigators isolate a potential severe risk (high injuries and damaging issues). Afterward, they infer 1) non-actual but possible scenarios (Potential Near Misses) which could realize this risk and 2) means to avoid it: the potential defences.

- *The field situation*

The two models are used for analysing two incidental situation analysis by an expert ATC investigator: a "classical" near miss (two planes becoming too near), and an atypical incident (closing the control centre due to a tempest). We monitor complete sequences of his inquiry phases: from the event report and data gathering to the report writing. In the report, investigators have to fulfil: a) a "rough fact presentation" (transcripts of communications, aircraft trajectories, etc.); b) their analysis, and c) the preventive actions they proposed to be taken. The last point in the report is a categorisation of the incident to be included in the database.

We asked the investigator to explicit his first explanation after the initial information gathering phase. Afterwards, interviews consisted in asking "what are the new collected information, the causes which appear as relevant, and the actions to be taken". Data analysis is a comparison between the successive states of incident representation as expressed by the investigator (interviews, written report).

- *The coding scheme*

Hoc & Amalberti (1999) devise a coding scheme—based on a cognitive activity model—in order to track diagnosis in a dynamic environment. The causal analysis activity of an ATC investigator can be seen as a kind of diagnosis. But, its particularity is that this diagnosis is not made "on line" but *post factum*: the obtained verbalisations are series of explanations during a long period of time (1-3 months). We adapt their coding schema in the following way: we keep the "predicat-(<argument>) format; we categorise verbalisations as *identifications*, *inference processes* and *information expressions* with regards to the factual reactive model, while *counterfactual thinking* is linked with the proactive model. *Metacognition* is also expressed (as in the Hoc & Amalerti coding scheme). Finally measures to be taken can be expressed in near miss analysis: this process is not analysed here as it only concerns the final reports and the data base. Figure 3 resumes the key points in the coding scheme.

Both models define a <LEVEL> argument:

Examining the elements of a causal chain ("causal field", Mackie, 1974), five levels of facts may be distinguished by the inquirers:

- *Technical*: it concerns the technical part of the system that explains the analysed event.
- *Individual*: a sentence like "an initial mistake: the pilot doesn't say that he has a technical problem" (radio problem) is coded IDENT (causal factor, individual).
- *Collective*: it concerns aspects of co-ordination, co-operation, and collaboration inside a team or between several teams.
- *Organisational*: these arguments deal with the main decisions (economical, human and technical resources, activity monitoring, and procedures).
- *External factors*: these factors are meteorology, traffic load, etc.

The near-miss models also defines a <TYPE> argument.

Types of identifications:

In the causal field characterising an event (fig.1), one can separate elements contributing to the event realisation from those impeding a negative evolution. Some elements can have reinforced the "bad issue" and, on the contrary some others have reinforced the robustness of the system against the latent catastrophe. The former are coded as *defences*, and the later as *causal factors*. Causal factors are real facts that explain the story of a near miss. A defence is a fact that has contributed to preserve the technical macro-system against a catastrophic issue. Moreover, a *consequence* is an event caused by the near miss. Lastly, inquirers stress general characteristics of the situation (conflicts for instance) that we call *problems*.

- *Causal factor* (fact that contributes to the incident realisation):

The most general factors analysts identify are failures, gaps from a norm or from prescribed procedures, or a lack of defence. Hart & Honore (1959) define it as an "abnormal event".

- *Defence:*

we code here every king of recovery activity achieved by operators, technical safeguards or safety procedures cited by inquirers.

- *Consequence:*

It is an observed modification of the state of the world after the near miss. It concerns facts considered as pertinent by an enquirer: technical, environmental or human breakdown (cf. figure 1).

- *Problem:*

A problem is the identification of a critical situation (conflict, hazardous context, etc.).

Types of counterfactuals:

In the counterfactual model (fig.2) the types are transposed in order to express the fact that there are potential elements (and not factual ones).

- *Potential factor*:

They are facts that could lead to a catastrophe.

Example: *"Nevertheless, if this kind of problem have been met during a load period of traffic, it would have lead to a catastrophic situation"* is coded COUNTERFACT (potential factor, external factor). The external factor is here the traffic load.

- *Potential defence*:

They are potential barriers that could prevent the PNM realisation. Every level (individual, organisational, technical) can be involved.

- *Potential consequence*

Example: *"if the plane have been immediately descended from 350 to 310, with the same reaction time as the other, in fact, they would have crashed"* is coded IDENT (potential consequence, individual). (The level is individual, as in the ATC operative language, pilots are spoken about as "planes" in all cases where the plane dynamics is concerned.

- *PNM* is a potential event.

Example: *"it makes me think about the fire risk"* is coded COUNTERFACT (PNM, external factor). PNM is the counterpart of Problem in the case of factual causal reasoning (identifications).

1. *Identifications* is coded IDENT (<TYPE>, <LEVEL>)
2. *Definitions of Counterfactuals* coded COUNTERFACT (<TYPE>, <LEVEL>)
3. *Hypothesis inferences* are coded INFER (<NATURE>) Example: "*We could find a frequency problem*" is coded INFER (frequency problem).
4. *Enunciation of information* is coded EN_INFO (<OBJECT>, <LEVEL>) Information that is not explicative in the causal process is coded EN_INFO.
5. Enunciation of general *elements about the enquiry* itself or the inquirers task coded META (<CONTENT>)
6. *Decisions of Action* are coded DECI_MES (<NATURE>, <LEVEL>) It concerns expressions about *preventive measures* to be taken.

Figure 3. Global content of the coding scheme for actual and potential near-miss analysis.

In the present paper we will only consider the first argument for identifications and counterfactuals, and the predicat (that is the cognitive action expressed) in the case of the other categories used for verbalisations and written report analysis. More detailed coding would be used later for analysing the whole analysis process, taking into account the dynamics of this process. This will require to go beyond this coding scheme, in order to manage redondancies and changes of abstraction levels in the "granularity" of the analysis.

RESULTS

Results concern the two contrasted cases analysed by the same expert ATC investigator. Tables 1a and 1b respectiveley present the frequences of the main analysis processes in the successive phases of the investigator's activity for the atypical event and for the classical near miss. (Metacognition that are very scarce in each phase and decisions of action which only appear in the final report and the data base are not reported here, and remain few).

Table 1a. Frequences of the main reasoning processes for the verbal reports (1-3), the final report and the data base, for the atypical event analysis (tempest).

Tempest		IDENT	COUNTER FACT	EN_INFO	INFER
	1	5	0	1	1
	2	31	4	24	14
	3	28	7	12	0
	Final report	24	7	5	2
	Data base	17	0	0	1
	Total	105	18	42	18

Remarks:

- There appear to be two main analysis phases before the final report. Counterfactual thinking does appear in these phases, and also in the final report; it constitute a significant cognitive activity (more than 20% of the main reasoning processes expressed by the investigator).

- Identification of evenemential elements is the central process, followed by the enunciation of information during the phases where the investigator construct his mental representation of what happened and could have happened.

There is no track of this activity in the data base, which is clearly centered on identification of evenemential elements.

Table 1b. Frequences of the main reasoning processes for the verbal reports (1-3), the final report and the data base, for the classical near-miss analysis (loss of a plane pressurisation, leading two planes being too near)

loss of pressurisation		IDENT	COUNTER FACT	EN_INFO	INFER
	1	27	1	26	2
	2	0	0	6	0
	3	7	2	7	0
	4	11	12	0	1
	5	8	4	10	2
	Final report	32	4	11	2
	Data Base	2	0	0	0
	Total	87	23	60	7

Remarks:

- Three main phases are present for the classical near-miss analysis. As previously, counterfactuals appear mainly before the final report; it also constitutes a significant cognitive activity (27% of the main reasoning processes expressed by the investigator).

- As for the atypical event identifications and information enonciation were dominantly expressed. (Distribution along the analysis phases varies widely.)

- There are very few elements in the data base, possibly because the controllers' activity was not directly involved.

Tables 2a and 2b present the frequences of causal / potential types of identifications / counterfactuals: factors, defences and consequences, respectively for the atypical incident and for the classical near miss.

Table 2a. Frequences of types for the atypical incident (tempest) depending on the analysis phase.

Tempest		Factors		Defences		Consequences	
		Causal	Potential	Actual	Potential	Actual	Potential
	1	3	0	2	0	0	0
	2	15	1	15	1	1	2
	3	10	3	16	2	2	2
	Final report	10	5	14	0	0	2
	Data base	10	0	7	0	0	0

Remarks:

- Factors and defences were arguments of similar weight for causal / actual reasoning.
- Consequences were more frequently considered in considering potential near miss, however potential factors were dominant.

Table 2b. Frequences of types for the classical near-miss depending on the analysis phase.

Loss of pressurisation		Factors		Defences		Consequences	
		Causal	Potential	Actual	Potential	Actual	Potential
	1	23	0	2	1	2	0
	2	0	0	0	0	0	0
	3	7	1	0	1	0	0
	4	11	2	0	6	0	4
	5	8	1	0	3	0	0
	Final report	30	0	2	2	0	2
	Data base	2	0	0	0	0	0

Remarks:

- Causal factors are the dominant type of identifications (actual / causal reasoning); a possible reason might be the fact that the decisions of ATC were deeply involved in this near miss, while defences are dominant as regards counterfactual thinking. Nevertheless, as for the atypical event, consequences remain non neglictable as regards potential elements.

CONCLUSION

This study highlights the fact that managers may be ignoring precious knowledge for experience feedback, particularly, the expertise of investigators to identify some potentially severe scenarios: the thought experiments carried out by an expert incident investigator are a significant part of his activity, while—at the present— they are not stored in the database.

Most of EF systems only exploit a small part of the inquirers' expertise: their ability to isolate a limited number of relevant causes in order to record a logical causal chronic of an event. The hidden skills (i.e. actually unexploited) of these analysts contain essential elements for prevention. We demonstrate in this paper that an expert investigator is able to:

- Identify, in the analysed event, *efficient barriers* (organisational, collective, individual) which stopped the potentially accidentogenic situation.
- Isolate *potential barriers* (not observed but inferred) for an increased robustness of the system.
- Isolate *potential consequences* and particularly *potential accident scenarios* against which the system must absolutely defend itself.

It has to be stressed that in the specific case of Air Traffic Control, classical near-misses are planes becoming too near. The shift from near-miss to the accident is a discontinuous process: it has a real catastrophe as a certain consequence: aerial collision is one of the worst situations, involving hundred of human deaths.

Security improvements will be obtained by strengthening current methods (systematic analyses and their wide diffusion, training, etc.). But, over-focusing on micro-conditions of incidents does not seem sufficient to increase safety. Indeed, the difficulty—if not the impossibility—to select relevant catastrophic scenarios from large databases is the main weakness of current EF (Amalberti, 1995, 1999, Amalberti & Barriquault, 1999, Barriquault, 1999). In this study, studying an expert ATC investigator's activity in near-miss analysis reveals clearly his ability to identify accident precursors and to choose very relevant scenarios for precautionary measures. Therefore, these skills should be promoted for safety management. It would help managers to fill the gap between near misses and a potential accident.

A first step in supporting near-miss analysis could be to use the twofold model: the reactive model (factual incident analysis) and the proactive model (potential near-miss analysis) for defining a method for near-miss analysis. Such a method might orient the investigators' activity with a structured guide-line. A second step would then be to organise training to use such a method in a flexible way, in order to adapt the content and level analysis to the specific incident at hand.

The results indicate that the structure of the data base may produce filtering effects on incident analysis, which is the the first and crucial phase in the process of experience feedback. Introducing—within the data base—the possibility for sharing the counterfactual thinking processes of investigators in charge of near-miss analysis could enlarge the possible preventive measures to face future problematic situations.

Acknowledgements

We thank controllers and inquirers of the French air traffic control for their cooperation.
This study has been done with the help of the Centre d'Etude de la Navigation Aérienne and in collaboration with the Institut de Medecine Aerospatiale du Service de Santé des Armées.

References

Amalberti, R. (1995). Paradoxe des grands systèmes à risques. *Performances Humaines et Techniques, 78*, 45-55.

Amalberti, R. (1999). Les effets pervers de l'ultra-sécurité. *La Recherche, 319, 66-77.*

Amalberti, R. & Barriquault, C. (1999). Fondements et limites du retour d'expérience (Bases and limits of experience feedback). *Annales des Ponts, 91*, 67-75.

Baerentsen, K. B. (1996). Episodic knowledge in system control. In B. Homqvist, P. B. Andersen, H. Klein & R. Posner (Eds.), *Signs of work. Semiosis and information processing in organisations*, (pp. 283-323). Berlin: Walter de Gruyter.

Barriquault, C. (1999). *La compréhension naturelle des incidents (Natural incidents understanding)*. Mémoire de DEA, Saint-Denis : Université Paris8.

Davies (2000). Confidential incident reporting on the UK Railways: the CIRAS system. *Cognition, Technology & work, 2,* 17-125.

Doireau, P., Wioland, L. & Amalberti, R. (1997). La détection des erreurs par des opérateurs extérieurs a l'action : le cas du pilotage d'avion. *Le travail Humain, 60,* (2), 131-153

Fadier E. (1990). Fiabilité humaine : méthode d'analyse et domaines d'application. *in* J. Leplat, G. de Terssac (Eds), *Les facteurs humains de la fiabilité dans les systèmes complexes.* Marseille : Octarès.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51,* 327-358.

Godard, O. (1997). *Le principe de précaution dans la conduite des affaires humaines.* Association NSS-Dialogues, édition de la maison des sciences de l'homme.

Hart, H.L.A. & Honoré, A.M. (1959). *Causation in the law.* Oxford: Clarendon Press.

Hesslow, G. (1988). The problem of causal selection. In D. J. Hilton (Ed.), *Contemporary science and natural explanation.* The Harvester Press.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin, 1,* 65-81.

Hoc, J.M., & Amalberti, R. (1999). Analyse des activités cognitives en situation dynamique : d'un cadre théorique à une méthode. *Le Travail Humain, 62,* 97-129.

Johnson, C.W. & Botting, R.M. (1999). Using Reason's model of organisational accident in formalising accident reports. *Cognition, Technology & Work, 1,* 107-118.

Kirwan, B (1994). *A guide to practical human reliability assessment.* London: Taylor and Francis.

Kouabenan, D. R. (1999). *Explication naïve de l'accident et prévention (Naive accident explanation and prevention).* Paris: PUF.

Leplat, J. (1985). *Erreur humaine. Fiabilité humaine dans le travail.* Paris : Colin.

Llory M. (1996). *Accidents industriels : le coût du silence; opérateurs privés de parole et cadres introuvables.* Paris, L'Harmattan.

Mackie, J.L. (1974). *The cement of the Universe: A study of causation.* Oxford: Clarendon Press.

Mandel, D.R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology, 71,* 450-463.

Mill, J.S. (1872/1988). *Un système de logique.* Bruxelles : Mardaga.

Pariès, J. & Amalberti, R. (1997), Aviation Safety Paradigms and Training Implications, in In Sarter, N. & Amalberti, R. (Eds). *Aviation Cognitive Engineering,* New York: LEA.

Pariès, J., Merritt, A. & Schmidlin, M. (1999). *Developpment of a Methodology for Operational Incident Reporting and Analysis Systems.* Final Report – Convention DGAC 96/01, Paris : DEDALE.

Rasmussen, J. (1990). Learning from experience? How? Some research issues in industrial risk management *in* J. Leplat, G. de Terssac (Eds), *Les facteurs humains de la fiabilité dans les systèmes complexes.* Marseille : Octarès.

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science,* 27(2/3), 183-213.

Reason, J. (1997). *Managing the risks of organizational accidents.* Aldershot: Ashgate.

Roese, N. (1997). Counterfactual Thinking. *Psychological Bulletin,* 121, (1), 133-148.

Schaaf, T. W. van der (1992). A framework for designing near miss management systems. In T. W. van der Schaaf, D. A. Lucas, & A. R. Hale (Eds.), *Near miss reporting as a safety tool* (pp. 27-34). Oxford: Butterworth-Heinemann.

Vuillemeur, A. (1988). *Sureté de fonctionnement des systémes industriels, fiabilité, facteurs humains, informatisation.* Paris: Eyrolles.

Wagenaar W.A. & Groeneweg, J. (1987), Accident at sea: multiple causes and impossible consequences. *International Journal of man-machine studies, 27,* 587-598.

Wioland, L. (1997). *Etude des mécanismes de protection et de détection des erreurs, contribution à un modèle de sécurité écologique.* Thèse de doctorat, Université Paris V.

Woods, D., Johannesen, D., Cook, R. & Sarter, N. (1994). *Behind human error: Cognitive systems, computers and hindsight.* Technical report. Dayton Ohio: CSERIAC, WPABFB.

Aviation II

Future Air Traffic Management:
A Perspective on Distributed Automation

Thomas Prevot*, Everett Palmer, Nancy Smith and Todd Callantine*
*San Jose State University
NASA Ames Research Center
Moffett Field, CA 94035-1000, USA
tprevot@mail.arc.nasa.gov (phone: 1 650 604 2441)

Abstract

In this paper we present our perspective on some key issues regarding current trends in Air Traffic Management (ATM). We review results and observations from integrated air ground simulations that we conducted over the last few years at NASA Ames Research Center. In those high fidelity simulations we introduced new flight deck and/or ground automation and procedures for human-machine and human-human interaction to experienced pilots and controllers and evaluated/observed the impact. We present some specific and general findings as well as potential benefits and problems. We describe our current approach to investigating distributed air ground traffic management in the framework of recent, current and upcoming projects.

Introduction

Several concepts for a safer and more efficient air traffic system are currently being developed. One characteristic of all concepts is the continued introduction of advanced decision support tools for air traffic controllers and managers and the increased use of advanced automation. Flight deck automation is already at a fairly advanced state and existing systems as well as novel approaches are researched in depth. Most of the flight deck research is conducted in isolated aircraft simulators with scripted stimuli from the external environment. This may result in some aircraft automation such as vertical navigation (VNAV) working so poorly in busy air traffic that it cannot be used – thus eliminating the potential efficiency benefits.

Air Traffic Control (ATC) facilities are only at the beginning of the automation age. Integrating new display and control systems, decision support automation and data link into ATC facilities can be compared to introducing "glass cockpits" instead of "steam gauges". There is a multitude of possible effects that new automation may have on how operators interact with their computer tools and each other. These effects need to be identified and carefully regarded when modifying a system as complex and safety critical as the air traffic system. While looking at the local impact of new automation in well-defined specific experiments is a required step, it is just as important to understand the impact of this automation on the interaction of pilots and controllers in an operational environment.

Future Air Traffic Management in the Arrival Environment

If current airspace operations remain unchanged, increasing traffic demands are expected to compromise both on-time performance and safety. Coping with these increasing airspace capacity requirements will require substantial modifications and improvements to current-day operations. One approach to addressing this problem is to give airlines more freedom in scheduling and selecting preferred traffic routes while continuing to assign responsibility for separation and arrival planning to the Air Traffic Service Providers (ATSP). ATC-oriented approaches focus on airspace restructuring and/or development of new tools for air traffic managers and controllers that enable them to manage air traffic more safely and efficiently. Tools like COMPAS [1], URET [2] and CTAS TMA and FAST [3] are being developed and fielded in several ATC facilities.

We investigated this ATC-oriented approach within NASA's Terminal Area Productivity (TAP) program from several angles over the last four years. The first section of this paper summarizes this research. The second section introduces our current work in the more radical concept of Distributed Air Ground Traffic Management (DAG-TM).

Terminal Area Productivity (TAP) Research

In the ATM portion of the TAP program, we investigated the integration of future ground-based ATM decision support systems and Flight Management System (FMS) equipped aircraft within the terminal area. The experiments and demonstrations focused on increasing airport capacity for arriving traffic by using the Center TRACON Automation System (CTAS) for generating efficient trajectories, data link for communicating those trajectories to the aircraft and FMS equipped aircraft for flying them precisely.

We looked at the problem of aircraft arrival rushes into major airports. The goal was to provide a safe, highly efficient flow of traffic from enroute into TRACON airspace that reliably delivers aircraft to the runway threshold, while maintaining as much flight crew flexibility and authority as reasonable. Successful planning and execution of an efficient arrival flow requires a thorough understanding of all aircraft, operator, traffic management and spacing constraints, and involves coordination between controllers, flight crews, dispatchers and traffic management. We envision a human-centered system in which controllers and pilots use procedures, flight management automation and decision support tools to actively manage arrival traffic. We targeted a future air traffic system controlled and managed by the Air Traffic Service Providers (ATSP) and expected to be operational in the 2010 time frame [4].

The operational concept for achieving efficiency enhancements over today's operations is to plan an efficient arrival stream ahead of time and then execute the "arrival plan" as precisely as possible. We introduced a "multi-sector arrival planner" ATC position to bridge the gap between traffic managers, dispatchers and sector controllers. The planner's tasks involve creating the most efficient schedule and sequence for all arriving aircraft and conflict-free flight paths that meet this schedule. The planner coordinates the generated flight paths with the sector controllers using a graphical coordination tool. After reviewing the proposed flight path, the sector controllers issue appropriate clearances to the flight crews. Flight crews follow the cleared flight path precisely using their flight management automation. Sector controllers are responsible for maintaining separation and adjusting the arrival plan to new circumstances. Automation and procedures are designed to help with all these tasks.

This TAP concept is more strategic than today's systems but the controllers are actively involved in every step of the process of developing and executing a traffic flow plan for the arrival rush. Even though it significantly changes the roles of the stakeholders, it does not change their responsibilities. Observations and pilot and controller feedback from simulations demonstrating this concept can be found in [5,6,7,8] and are summarized below.

TAP-ATM Simulations

In addition to several interface reviews and engineering tests the following experiments were conducted in order to investigate the different aspects of this concept:

1. Full mission flight simulator study of the human factors of flying CTAS descents in the Terminal Area conducted at NASA Ames Research Center
    a. Use of data link with different pilot interfaces in the Terminal Area
    b. Use of Flight Management Automation (LNAV/VNAV) in the Terminal Area
    c. The impact of a Vertical Situation Display to help with these tasks
2. Part task flight simulator study of arrival time errors when flying CTAS descent clearances conducted at NASA Langley Research Center
    a. Trajectory prediction accuracy between FMS and CTAS
    b. Arrival time errors at the Final Approach Fix for FMS-managed descents vs. vectored arrivals
3. Initial demonstration of CTAS/FMS operations with controllers conducted at NASA Ames Research Center
    a. Acceptance and usability of operational concept
    b. Controller interaction with advanced automation tools
4. Main demonstration of CTAS/FMS operations with pilots and controllers conducted at NASA's Ames and Langley Research Centers
    a. Acceptance and usability of operational concept
    b. Controller interaction with improved automation tools
    c. Pilot controller interactions in a strategic ATM environment
    d. Flight crew factors in the CTAS/FMS environment

Experiments Focused on Ground or Air Side

1. The first flight deck oriented full mission simulation demonstrated that data link usage in the terminal area was acceptable and desirable for flight crews. A streamlined FANS-type CDU datalink interface was acceptable to the flight crews. Most crews preferred a Boeing 777 like data link implementation that reduced heads-down time in the cockpit. Flight crews could successfully use the lateral flight management function LNAV to the final approach fix. Using the vertical flight management function VNAV close to the ground was a concern to pilots [6]. A Vertical Situation Display (VSD) prototype was introduced to help using FMS automation closer to the ground and received high ratings by the flight crews. Significant workload or performance differences could not be found between conditions with and without the VSD [7].

2. A flight simulation at NASA Langley Research Center found that arrival time errors at the final approach fix can be significantly reduced when flying TRACON trajectories with FMS guidance rather than heading vectors. Again the streamlined FANS data link interface on the CDU was found to be acceptable for TRACON operations.

3. The initial demonstration of CTAS/FMS operations with controllers demonstrated the potential for increasing the efficiency of arrival streams by using the CTAS tools for planning and monitoring. The designed controller interface with the automation and the data link was acceptable, but could use further improvements. Too much information in the standard data block, a clumsy and complicated route trial planning interface and the three-button mouse were mentioned as some of the main shortcomings. The operational concept received very positive feedback and the controllers were enthusiastic about its potential.

The main demonstration of CTAS/FMS operations with pilots and controllers in the loop

We conducted a set of simulations combining all the different elements. We staffed two full mission flight simulators at Langley and Ames, 3 to 5 center controller positions, 3 TRACON controller positions, and 9 pseudo pilot positions, each of which handled multiple aircraft. The flight simulator at Ames was additionally connected to the Crew Activity Tracking System (CATS) for model-based on- and offline evaluation of task performance [13].

Most of the prior findings and observations held true during these tests and all subjects were very impressed with the potential of a futuristic ATM system like the one they participated in. However, several issues were raised that did not come up in any of the previous experiments.

Flight crew perspective

To study flight crew factors under integrated CTAS/FMS operations with controllers-in-the-loop we included the NASA Advanced Concepts Flight Simulator (ACFS) in the distributed simulation. The ACFS is FMS-, VSD- and datalink-equipped. Eight qualified flight crews received a briefing on ACFS cockpit systems, FMS Arrivals and Transitions, and data link operations. Each crew then flew six descents, alternating between two scenarios. The 'Center scenario' started at cruise altitude outside the Center airspace, and the 'TRACON scenario' started at the TRACON boundary.

In both scenarios, crews first established data link communications. In the Center scenario, crews received forecast winds via data link. They could also downlink a preferred descent speed to CTAS. The controller could issue a data link message or voice clearance to modify the cruise and descent according to a CTAS advisory, or modify the lateral route. The controller then issued an FMS descent clearance to begin the descent on the FMS trajectory. Speed and/or route adjustment could occur in the low altitude sector by voice or data link. In both scenarios, the TRACON feeder controller issued a clearance to fly an FMS Approach Transition to a given runway. In the TRACON scenario, the final controller sometimes issued a route modification clearance via data link. In both scenarios, the final controller then cleared the aircraft for the approach and handed it off to the tower controller for landing.

We evaluated crew performance on each descent using measures that address the operational concept, specifically, the ability of crews to precisely follow a flight plan using the aircraft's FMS, and to coordinate air and ground operations via data link. We used CATS to analyze digital data from the ACFS; videotape was used to confirm and analyze key observations in greater detail. Crew acceptability of the proposed procedures was evaluated with a questionnaire.

We obtained data for 22 Center scenarios and 23 TRACON scenarios. In 60% of 45 flights, the lateral portion of the route was flown entirely in LNAV mode; this measure reflects positively on the success with which controllers were able to issue FMS clearances without resorting to vectors. During times when the flights were cleared on FMS routing (which, at the very least, included the descent on the FMS Arrival in the Center scenario), crews complied with 82% of speed restrictions and 93% of altitude restrictions. Lastly, of the 80 data link messages crews received, 96% were handled in a correct and timely manner. As in the previous ACFS study, however, crews often switched to a tactical control mode instead of VNAV whenever other tasks assumed a higher priority than monitoring the automation.

More detailed analysis identified several issues that deserve slight modifications or further training. First, the data link message text designed to cue the entry of a preferred descent speed needs clarification. Second, some pilots were confused about how far they were cleared on charted routing, indicating a need for separate charts for FMS routes to a particular runway. Third, if controllers issued a voice clearance while the crew was responding to a data link clearance (or vice versa), crews had to request clarification as to which portions of each to comply with. Pilots sometimes also found it confusing to have a check-in call simply acknowledged with no mention of a recently issued datalink clearance. Controllers sometimes also issued ambiguous clearances on check-in, such as a clearance 'direct to' the same waypoint that is already active in compliance with a previous FMS arrival clearance; some pilots were uncertain whether such a clearance should be interpreted as a cancellation of the FMS Arrival. Fourth, one pilot thought that data link clearances were guaranteed to be flyable, negating the procedural requirement to review the clearances carefully before accepting and executing them. Fifth, we designed FMS Arrivals and Transitions to be flown in VNAV with the last charted altitude restriction set as the limiting target altitude. Pilots from airlines whose policy is to 'step down' the altitude target to the most constraining altitude at times were unwilling to set the last charted altitude as the limit altitude. Finally, some crews over-committed to an "expect" clearance by re-programming the FMS route. This resulted in increased workload when a clearance different from the "expect" clearance was issued.

The questionnaire covered FMS procedures, charts, FMS clearance phraseology, automation usage, data link clearances, and data link response procedures. Again pilots found workload under the CTAS/FMS integration concept to be slightly lower than in current-day operations; however, more monitoring is required. The FMS procedures as a whole were acceptable, but the experiment FMS arrival charts required some improvements. Using LNAV mode to fly precise lateral routing was acceptable, even at low altitudes in the TRACON airspace. On the other hand, pilots gave VNAV generally lower acceptability and comfort ratings. Pilots generally viewed data link usage positively. However, some pilots did not know whether the data link speed clearance phraseology meant flying a Mach value in the descent until the Mach/CAS transition, or whether the CAS should be flown immediately. Performing FMS edits in the TRACON airspace also elicited a range of opinions. Pilots who over-committed to "expect" clearances found FMS edits less agreeable than those who left route discontinuities in the route until they were actually cleared on the routing.

Overall, from the flight crew perspective, procedures developed for FMS and data link operations can work in concert with CTAS tools. In general pilots found the concept favorable, and with some modifications and additional pilot familiarity, the concept appears especially promising. A more detailed description can be found in [8].

Controller perspective

Our simulation scenarios were based on the northwest arrival stream into Dallas Ft. Worth, which currently experiences at least two major arrival rushes every day. The main scenario was derived from recorded traffic and weather data from a day with IFR weather conditions in spring 1999. Traffic loads in different scenarios ranged from moderate to more than current day peak rush demand.

From a controller's perspective the arrival scenario develops as follows: Aircraft arrive at the center's airspace on direct routes or in-trail. The ground automation (CTAS) estimates feeder fix arrival times for these aircraft. The CTAS Traffic Management Advisor (TMA) software automatically creates an initial sequence for these aircraft, taking all airport flow control constraints into consideration. The planning controller evaluates this sequence and interacts with the TMA and conflict probe to adjust the flow for spacing and scheduling. This task is supported by the CTAS Descent Advisor (DA) software, which assists the controller in creating flight paths (route and/or speed modifications) that meet the scheduled time at the feeder fix. If no significant delay has to be absorbed (~5 minutes or less), an early modification to the aircraft's cruise speed and perhaps its descent speed is usually sufficient. This flight path modification is communicated to the flight crew (by voice or data link), who set up their FMS accordingly. After an arrival clearance is given to fly the FMS computed path, aircraft automation is used to follow the plan precisely. Pilots and controllers thus know when the aircraft will start to descend and where it will be at any given time. If aircraft are data link equipped, the FMS flight path is transmitted to the ground system, and the controller can inspect it for any significant differences from the ground-predicted trajectory.

At least six controller positions managed the arrival flow in our simulation: the arrival planner, high and low altitude sector controllers in the Center; and one TRACON controller to pick up the flow managed by the center controllers; as well as two more TRACON controllers managing a second arrival flow that was initialized at the meter fix. All center positions were equipped with a TMA timeline, a conflict prediction list, access to the DA advisories and a trajectory preview tool that allowed controllers to quickly preview the predicted traffic situation to any given time in the future.

Each session took three days for the controllers. Center controllers were trained for one and a half days on the CTAS tools and FMS arrival procedures. Three or four data collection scenarios were run during the last two days of each session.

All subjects stated that the overall concept is very promising and bears a great potential for improving traffic flow into, out of, and across congested areas.

When It Works, It Works Well...

After three days of training and simulation runs, participant controllers were capable of handling complex arrival rushes. In these runs, almost the maximum throughput was achieved for the one test runway, with efficient FMS descents for about 35 consecutive aircraft.

In several runs, the three Center controller participants (Planning, High, and Low sectors) successfully handled the arrival traffic flow. During these runs, the majority of aircraft received FMS descent clearances and benefited from almost undisturbed descents into the TRACON. Most aircraft arrived at the metering fix within 15 seconds of their scheduled time and an efficient TRACON feed was provided without imposing extensive workload on the controllers.

At the same time radio frequency congestion was reduced by replacing many tactical clearances with a few strategic ones.

...But the Strategic Plan May Fall Apart

In some runs controllers reverted to tactical control of the traffic. The strategic FMS arrival plan was disturbed or even fell apart. Successful implementation of the arrival plan is sensitive to good planning and aircraft compliance with the planned flight path.

The role of the arrival planner became increasingly important with the complexity of the arrival rush. The planning job required very good skills in traffic management and control, and proficiency with the tools. Arrival plans that set up aircraft well within their performance limits and used similar descent speeds among aircraft were generally easier to handle for downstream controllers. If the plan did not provide sufficient buffers against separation loss for the sector controllers, they were likely to change it or not execute it. Aircraft that did not comply with their clearance or did not receive the descent clearance on time often caused significant problems for the controllers in implementing the arrival plan. Because of the use of high-energy FMS descents, non-compliance or late descents typically required controllers to vector the problem aircraft to meet the TRACON restrictions.

One problem of FMS arrivals is the increased compression effect created by high-energy FMS descents. In today's environment controllers adjust speeds and altitudes step by step to maintain consistent states between aircraft. Aircraft performance on idle FMS descent profiles varies significantly by aircraft type, weight and descend speed. This adds complexities to the task that do not exist in today's environment.

Data Link

Data link in this concept needs to be viewed from several angles. Even though the concept does not require the availability of data link per se, passive data exchange seems to be very helpful. Controllers had different opinions and showed different behavior for issuing clearances via data link. Some liked it because it cut down on verbal communication and was easy to use. It was in fact so easy to use that controllers sent more speed updates to the aircraft than they would have issued by voice, causing some confusion in the aircraft. Other controllers did not like to have to wait for the data link response, which is delayed compared to the immediate readback they receive in the voice environment. They stated that having to continuously monitor the data link status indication in the data block was an additional task, whereas by using voice they did not have to closely monitor the aircraft for a while after giving the instruction.

Dealing with the Automation

The shift between manual flight control and automated flight management in modern aircraft has been discussed and researched in depth. Our 2010 scenario requires controllers to use and trust the automation in the aircraft and on the ground to manage a more complex arrival problem than could be controlled without the automation's support. Similar automation issues arise for controllers as for flight crews, including the potential for mode confusion, clumsy entry procedures, problems with shifting between tactical and strategic control, and difficulty maintaining the "big picture" as situation complexity increases.

Distributed Air Ground Traffic Management

We try to apply some lessons learned for ongoing and upcoming work in NASA's Distributed Air Ground Traffic Management (DAG-TM) research project [9]. DAG-TM is targeting a free-flight environment in which flight crews play a more active role in the decision making process. Instead of simply executing controller instructions, crews will have some freedom in requesting and selecting flight paths. Advanced on-board automation for conflict detection and resolution will impact pilots' behavior, thus affecting controller behavior and putting more requirements on ground automation and information sharing

The DAG project's Concept Elements (CE) 5 *En Route Free Maneuvering* [10] and 11 *Terminal Arrival: Self Spacing for Merging and In-Trail Separation* [11] give flight crews in fully equipped aircraft some or all of the responsibility for separation, thus changing the role of air traffic controllers and flight crews. Concept Element 6 *En Route Trajectory Negotiation* [12] addresses the issue of negotiation of strategic trajectories.

Previous and ongoing research in free flight and Cockpit Display of Traffic Information (CDTI) will be combined with our ongoing research work. Advanced flight deck prototypes will be integrated into the simulation environment.

Two Extremes in DAG Arrival Management

The DAG concepts encompass a variety of possible ways to manage arrivals ranging from uninterrupted free-flight to fully ground-controlled. Two extremes are described below.

Free-flight to the threshold

One extreme has the flight deck responsible for path planning and separation from the aircraft throughout the arrival. The aircraft arrives at the Center in free flight and is responsible for separating itself from other traffic. Traffic flow management constraints for entering the terminal area are made available to the flight crew, who adjusts their terminal arrival plan (i.e. FMS descent trajectory) accordingly. When approaching TRACON airspace, the flight crews select the aircraft that they want to trail to the threshold and select the proper merging and spacing parameters. They then follow the lead aircraft to the runway.

Ground (ATSP) controlled arrival

The other extreme in arrival management is very close to the concept demonstrated in our previous TAP research. When entering the terminal airspace free flight is cancelled for the arriving traffic. Ground based traffic managers create the schedule and arrival trajectories and communicate those to the aircraft. The aircraft can at any time downlink flight path requests that the ATSP may or may not accept. The controller determines candidate aircraft for self-spacing approaches and appropriate spacing intervals and issues clearances to self-space.

Responsibility for separation and trajectory planning remains on the ground throughout the arrival phase. The flight crew receives more strategic FMS and spacing clearances than in today's tactical environment.

Designing for DAG Arrival Management

Free flight to the threshold will require additional aircraft equipage, which may include Required Time of Arrival (RTA) capabilities, Cockpit Display of Traffic Information (CDTI), conflict detection and resolution algorithms, self-spacing and merging algorithms, etc. Ground controlled arrivals do not use the aircraft capabilities in the most efficient manner and put the entire flow management burden on the controller. The future air traffic system will manage arrivals in a way that lies somewhere between the two extremes, possibly gradually moving from ground-controlled to more free-flight.

Research and operational practice will show which concept appears to be most appropriate. The amount of free flight vs. ATC control can depend on the traffic situation, facility practice, aircraft equipage, and airline preferences. It may be different between facilities and even time of day. We believe that the air traffic system should be designed to accommodate all possible modes of operation between the extremes. Therefore all enabling technologies have to be developed, integrated and evaluated, including

- CDTI with airborne conflict detection and resolution
- FMS with RTA capability
- On-board merging and spacing tools
- ADS-B and CPDLC data link communication
- Traffic Management advisory tools
- Ground-based conflict detection and resolution
- Ground based tools for trajectory generation with meet time constraints

Most of these technologies are already available in more or less isolated research prototypes. We are currently in the process of integrating them at NASA Ames Research Center to create a simulation environment that allows researching these issues.

Initial Arrival Concept for DAG

We are developing an arrival concept that provides the flexibility to adapt the amount of self separation to traffic flow management constraints and other requirements. We initially intend to keep the free-flight airspace separate from the ground-controlled airspace. The boundary can be specified as an arc around the meter fix or the nearby arrival gate or a simple altitude floor. This can be adjusted for traffic complexity. In very low traffic situations, the free flight area may be as close to the airport as the meter fix itself.

The arrival scenario begins with aircraft arriving at the Center in "free maneuvering mode". The flight crews are responsible for separation. Traffic management constraints at the metering fix are communicated from the planner utilizing the CTAS TMA to the flight deck as arrival information. The flight crew is expected to plan their flight path to arrive at the metering fix close to the expected time, if scheduling is required. The flight crew will also be told where the free flight boundary currently ends and when to check in with the controller. The arrival planner keeps evaluating the situation using Descent Advisor tools and tries to create an arrival plan for the ground-controlled airspace that he or she relays to the sector controllers. When the sector controller receives the check in from the free maneuvering aircraft, he or she cancels free flight and issues the arrival clearance to the aircraft based on aircraft preference and arrival plan. Aircraft are expected to fly the arrival clearance to the meter fix precisely. The CTAS TRACON tools (Final Approach Spacing Tool FAST) aid the TRACON controllers in determining proper aircraft pairs for receiving in-trail spacing clearances. Separation responsibility remains with the controller throughout the TRACON.

This scenario allows us to investigate most aspects of the relevant DAG-TM concept elements and builds on our previous arrival research. Recent discussions with controllers and pilots gained positive feedback. Initial demonstrations are planned for fall 2001.

Concluding Remarks

The concept of strategic arrival management demonstrated in the TAP research appears to be very promising. The DAG research moves from a ground-controlled environment to a more distributed environment with possibly shifting separation responsibilities. NASA Ames is currently preparing a research environment to investigate DAG-TM with all major technologies integrated. Initial concepts and scenarios have been defined and discussed with pilot/controller focus groups.

References

1. Voelckers, U. (1991) Application of Planning Aids for Air Traffic Control: Design, Principles, Solutions, Results. in *Automation and System Issues in Air Traffic Control.* J.A. Wise, V. D. Hopkin & M.L. Smith (Eds.). Springer Verlag, Berlin, Heidelberg
2. *URET User Request Evaluation Tool*, CAASD, URL: http://info.caasd.org/PDF/URET.html
3. *Erzberger, H., Design Principles and Algorithms for Automated Air Traffic Management, AGARD Lecture Series 200 Presentation, Madrid, Spain, Paris, France, and Moffett Field, California, USA, November 1995. See also* URL: http://www.ctas.arc.nasa.gov
4. Palmer, E.A. Williams, D.H. Prevot, T. Romahn,S. Goka, T. Smith, N. and Crane,B (1999) An Operational Concept for Flying FMS Trajectories in Center and TRACON Airspace in *Proceedings of the 10th Int. Symposium on Aviation Psychology* May 3-6,1999, Columbus, OH.
5. Prevot, T., Crane, B., Palmer, E.A and Smith, N.(2000) *Efficient Arrival Management Utilizing ATC and Aircraft Automation*. HCI-Aero 2000, Toulouse, FR
6. Crane, B, Prevot, T. and Palmer, E.A. (1999) Flight Crew Factors for CTAS/FMS Integration in the Terminal Airspace in *Human Computer Interaction Vol 2* (L. Erlbaum Ass.), Mahwah, NJ, 1999
7. Prevot, T. and Palmer, E.A. (2000) Staying Ahead of the Automation: A Vertical Situation Display Can Help *SAE-2000-01-5614.* 2000 World Aviation Conference, October 10-12, 2000, San Diega, CA
8. Callantine, T., T. Prevot & E.A. Palmer (2001). Flight crew factors under integrated CTAS/FMS operations with controllers-in-the-loop. *The Fourth International Air Traffic Management R&D Seminar ATM-2001*, Santa Fe, NM
9. DAG Concept Description (2001) NASA Ames Research Center, 2001
10. Philips, C. T. (2001*) Detailed Description for CE-5 En Route Free Maneuvering*, NASA-98005 RTO-41
11. Sorensen, J. A.. (2001) *Detailed Description for CE-11 Terminal Arrival: Self Spacing for Merging and In-trail Spacing* NAS2-98005 RTO-41
12. Couluris, G.J. (2001) *Detailed Description for CE-6 En Route Trajectory Negotiation* NAS2-98005 RTO-41
13. Callantine, T. (2000). *A glass cockpit crew activity analysis tool.* SAE Technical Paper 200-01-5522. Warrendale, PA: SAE International.

Learning, automation and interaction: the case of Australian air traffic control

Dr. Christine Owen, University of Tasmania, GPO Box 252-66, Hobart TAS, 7015, Australia. Email: Christine.Owen@utas.edu.au

This paper presents a detailed analysis of individual and collective cognitive work processes in one type of work activity — air traffic control (ATC). The presentation is based on completed PhD research findings investigating the influence of organisational contexts on workplace learning. The research used data collected from 100 semi-structured interviews conducted in three ATC workplaces, observations of ATC work practices and the author's active participation as a training provider and process facilitator within a range of organisational change events. This paper draws on these data to demonstrate how changing organisational structures and cultures, enabled and constrained learning and action involving individual, social and distributed cognition. During the study a semi-automated air traffic control system was introduced (The Australian Advanced Air Traffic Control System or TAAATS) giving the author the opportunity to compare and contrast the impact of technological change on learningful work activity.

Introduction

This paper uses learning and organisational theories to examine automation and interaction in the air traffic control (ATC) workplace. Its success rests on whether the reader is convinced by two messages: First, that inherent in ATC work are practices associated with continuous learning and that examining everyday work practice using learning theory can provide valuable insights for workplace designers and human factors professionals. Second, that in examining human factors issues in workplaces, it is important to widen the focus beyond the interaction between the human and the machine to include broader contextual factors that nevertheless influence work activity. Learning can be formally organised (e.g., in a training program) and it can also occur as part of everyday work activity. A broad definition of learning is therefore used in this paper to capture both formal and informal learning processes. Learning is understood as the processes involving the transformation of experience through reflection, conceptualisation and experimentation, which leads to an increased capacity, in individuals, groups and organisational systems to act in the environment [after Kolb, 1984].

The importance of learning in the aviation industry is underscored by the observation that between 60-80 per cent of all accidents and near-miss incidents are as a result of human error [Hartel & Hartel, 1995]. In many of these cases [Hartel & Hartel, 1995], the information needed at the time was available somewhere within the organisational system, but had either not been passed on, or had been passed on incorrectly. In these contexts, what is needed is the development of "requisite variety" [Weick, 2001] through a culture of conscious inquiry [Westrum & Adamski, 1999]. According to Weick [2001], "requisite variety" is needed in complex systems because the complexity (variety) that occurs within those systems exceeds the individual capability of the people required to manage them. "When people have less variety than is requisite to cope with the system, they miss important information, their diagnoses are incomplete, and their remedies are short-sighted and can magnify rather than reduce a problem" [Weick, 2001 p. 331]. Requisite variety can be enhanced through the skills of conscious inquiry enabling people to pool the capability they have.

The skills of conscious inquiry, includes behaviours such as: being able to ask relevant and thought-provoking questions, sharing observations, seeking alternative perspectives, assertively challenging a particular opinion, seeking clarification, and sharing information through processes of consulting and collaborating. Requisite variety is then enabled because such learning-related behaviours make possible the expansion of the number and variety of experiences, and this increases the number of possible options that might be available. The link between work practice and informal learning immediately raises the problem of work contexts and how these influence behaviour. This is because all of these behaviours occur in a workplace context and such contexts influence the degree to which learning-related behaviours will be present.

In organisational theory terms, ways of operationalising contexts have most commonly been conceptualised in terms of organisational structures and cultures. An organisation's structure is based on a formal system of (i) interlocking roles and relationships between roles (i.e., differentiation: the degree to which job roles are specialised or generalised), (ii) levels of accountability (hierarchy), (iii) policies and rules (the degree to which activities are formalised), and (iv) coordination and control or 'integrative' mechanisms [Tosi, Rizzo & Carroll, 1995]. An organisation's structure is also evident in the (v) external relationships between the organisation and other organisations and (vi) the way the work is physically organised. In addition to the formal systems, every organisation has a set of interrelated informal systems that influence behaviour. These are described as organisational culture or sometimes as "communities of practice". Organisational culture is defined as a set of understandings shared by a group of people that are largely tacit among members [Louis, 1986]. Cultures tie the actions of individuals to a particular group. Values, beliefs, attitudes and norms are used by members of certain groups to justify certain decisions and behaviour. Elements of culture are particularly important as they provide a means of identifying processes of informal learning.

This paper draws on a qualitative research study undertaken into a particular high-technology, high-intensity and high-reliability (High-3) work environment in the aviation industry: air traffic control. High-technology organisations are those that involve work mediated by computers and other forms of technology. High-intensity work occurs in contexts that have time pressures and sometimes a strong sense of urgency. The dynamic "real-time" nature of the work contains elements that once initiated cannot be stopped and must be managed as events evolve. Within such organisations the focus is also on extremely reliable operations, because an error can potentially lead to unacceptable consequences. The study utilised a qualitative design and involved interviewing a stratified sample of 100 air traffic controllers, in some cases on multiple occasions, across three ATC Centres in Australia. The research question addressed incuded: "In what ways do organisational structures and cultures enable and constrain learning in the workplace?" and "In what ways do organisational changes, such as the introduction of complex technology, influence workplace learning?" The basic premise of the research was that the lived experiences of people at work are significantly influenced by their contexts (in organisations most commonly conceptualised as structures and cultures), and that these changing contexts have implications for learning, and are in turn reproduced or transformed by people. The methods used and the process of data analysis has been described elsewhere [Owen, 1999].

Although using a qualitative research design and examining the role of context is unusual in the aviation industry, there have been increasing calls to do so. For example, the European Organisation for the Safety of Air Navigation [1996, p. 12) recommended greater use of qualitative research processes in aviation psychology, for exploring job tasks in ATC. Moreover, one of the conclusions of the FAA's research into human factors and automation in ATC [Wickens, Mavor & McGee, 1997], was that "there is a lack of research data that would permit identification of the specific mechanisms by which formal and informal organizational contexts within the FAA interact and how they affect organizational climate and controller performance" (p. 7). This paper addresses these calls.

A conceptual framework for evaluating the influence of contexts on work practice

The definition of learning, described earlier, draws attention to the importance of experience, reflection, conceptualisation and experimentation as necessary components of the process of learning. It also emphasises the continuous and iterative nature of learning. In changing organisational environments, such a model is particularly useful because it addresses 'how' learning occurs, rather than 'what' is learned. This is valuable because needed job content in workplaces will change in the future and will change across contexts.

The main elements of the processes involved in learning, and the ways in which these are in evidence in the structures and cultures of the ATC workplace, are summarised in Table 1. It should be noted that although 'experience' is likely to have embedded within it the processes of reflection, conceptualisation and experimentation. For the purposes of this paper, these phases are introduced separately, something akin to a snapshot.

Table 1: The linkage between elements of context and processes involved in learning

LEARNING PROCESSES	ATTRIBUTES	STRUCTURAL ANALOGUES	CULTURAL ANALOGUES
Experiencing	Being • corporeally (Body) • cognitively (Mind) • affectively (Self) • socially (Environment)	Physical work organisation. Dimensions of work experience: temporally, complexity, affectively, socially.	⇐ Marked, remembered, accounted for and generated through ⇓
Reflecting	Observing narrating/ noticing; pattern seeking; labeling	Work organisation/ Integration structures [e.g., Job roles (+) Teams (+ -)	Collective remembering (war stories) (+).
Conceptualising	Thinking about, sense making; schemas used for interpretation	Integrative structures (e.g., boundary spanning functions); Job Roles	Collective remembering/ schemas
Experimenting	envisaging/enacting; developing choices validating/testing.	Integrative structures (e.g., teams) Centralisation./Formalisation .	Norms of practice (+ -); War stories (+).Collective schemas

(+ : enabling; -:constraining)

The first component in the learning cycle is experience. In the workplace context, experiences are structured by the artefacts used in work organisation, such as the way the work is physicaly organised, the resources available, as well as the policies and rules governing activity. However, experience also involves perception, implies consciousness and always comes with meaning. Given that interpretation of meaning is the foundation for culture, culture is thus always embedded in the interpretation of experience. The structuring and interpreting of experience will influence opportunities for learning in certain ways because structures will make certain opportunities available and not others and cultures will focus attention on particular interpretations of the experience and not others. In terms of experience, these cultural elements are marked, remembered and accounted for through reflection and conceptualisation on experience, often leading to experimentation.

Work is experienced principally in four ways: corporeally, cognitively, psychologically and socially. That is, work may be intense on the body; easy or demanding on the mind; an expression of, or alienation of, the self and it always occurs in a social environment (although some forms of work organisation are more socially organised). In ATC, the ways of experiencing work give rise to a number of dimensions that can be used to analyse work practice . These are temporal, complex, affective and social dimensions. For example, the temporal dimension of experience is emphasised in the "real time" dynamic of the job. Complexity is evident in the nature of job tasks that require higher order thinking and problem solving. An affective dimension is also evident in, for example, the importance given to individual and collective decision making as an expression of self and group. Finally, a social dimension is evident in the interdependence of job tasks. In this paper, the term "interaction" is used to mean interpersonal communication between actors in the work system.

Given the temporally demanding nature of the work controllers used a number of aids which also include using their bodies as a resource. Controllers, for example, talked about developing their own "body clocks" to aid the controller to establish accuracy with timing. In ATC, the temporal experience of work intersects with the social experience as air traffic controllers use shared spaces and shared objects to monitor the flow of traffic. Shared objects and displays facilitate the process of what Resnick [1993, p. 10] called "referential anchoring". In ATC work, the physical organisation of the workspace means that any controller's workload is visible, allowing other controllers to referentially anchor their work in relation to what can be anticipated. For example, the build up of paper strips (each representing an aircraft) on a neighbouring console provided a common referent that can be used by controllers to monitor what work is building up that is likely to be coming their way. Because the information about the activity of work is publicly available, team leaders and colleagues can also anticipate when a controller working at the console is likely to need help. Controllers also referentially anchor their work by developing what they call the "third ear". That is, they develop a capacity to undertake their work and listen out for what is happening around them. The third ear becomes an important resource for synchronising temporally and complexly demanding work when working interdependently. What is important to note here is that the way work is organised shapes opportunities for experience in certain ways and that these organisational structures influence the transformation of experience into learning by enabling or constraining the transition to the next moment in the cycle — reflection.

The second component in the learning process is reflection. In learning theory terms, reflection is an essential ingredient for learning. Reflection involves attending to the salient features of an experience, marking and noting those features through labeling and pattern seeking. Reflection can occur individually and collectively. When the experience of work is temporally demanding and complex, there is little time for reflection at the workface. In ATC work, it was found that work that is temporally demanding requires, for example, delaying reflection and conceptualisation processes of learning, thus inhibiting opportunities for immediately learning at the console. Work organised such that it involves a complexity of decision making and tasks often makes the explication of rules and their application difficult and opaque, thus inhibiting both the formal learning of such tasks and the identification of strategies to enhance learning these kinds of activities. Part of the difficulty with workplace environments where individuals have to respond immediately to changes in those environments under periods of pressure and intensity is that often they do not encourage reflection in any systematic way. When work is intense and reflection is delayed strategies to enhance reflection need to be structured into either work practice in job roles (e.g., taking on related jobs) and/or enhanced through the utilisation of other resources, such as audio and visual aids that can be used later.

There were, however, organisational structures that did enhance reflection in informal learning though this was not their primary purpose. It was found that work that was organised so that performance required different roles surrounding the same work activity (such as taking on the role of instructor) allowed the controller to view the work from a different vantagepoint. In all interviews conducted with instructors, informal learning for the controller performing the role of instructor was enhanced since the instructional role required observing and reflecting on the controlling job (in order to facilitate the learning of the trainee). Teamwork was also found to be an important structure enabling reflection, provided certain conditions were met. These included a team-culture based on open communication patterns and norms of practice that involved sharing information and inquiry. This finding is supported in the team-work literature where structures of team-work were found to increase decision making effectiveness [Brannick Salas & Prince, 1997] and that teams with better communication practices also achieved higher levels of performance [Bowers, Jentsch, Salas, & Braun, 1998].

In the field study reflection was also observed to occur culturally, as part of shared remembering evident in the telling of war stories. In these contexts, cultural processes of reflection, found in shared remembering through narration, become particularly important as a compensatory mechanism for the ways in which structures limit this process of reflection. Structures and cultures also enable and constrain the next process in the learning cycle, conceptualisation.

The third component in the learning cycle, conceptualising, is the process of making sense of what has occurred, to interpret reflections on experience and to generalise these interpretations to new settings. For Resnick [1993], interpretation of experience is based on schemas that both enable and constrain individuals' processes of sense-making. A schema provides an interpretive framework that allows reasoning to proceed [Resnick, 1993]. As an interpretive framework, a schema is often based on past history, sets up expectations about what will be important, and therefore, will help guide what is attended to, what is perceived, what will be remembered and what will be inferred. Schemas are not purely individual constructions but are heavily influenced by the kinds of beliefs and reasoning schemas available in the individuals' surrounding culture [Resnick, 1993]. Thus, schemas enable and constrain both individual and collective opportunities for learning. Indeed, within the human factors literature, "shared mental models" has been a recent focus of research investigating performance in High-3 organisations [e.g., Helmreich & Merritt, 1998; Orasanu, 1995; Reason, 1998; Westrum & Adamski, 1999]. Such research has pointed to the ways in which unstated interpretations (mental models or schemas) become assumptions upon which action is based, sometimes with devastating results. The conclusion of such research is that practices associated with continuous inquiry are crucial in High-3 environments and procedures have been developed (the basis for crew resource management) to ensure that mental models are shared, making work more reliable.

In the data, norms of practice and communication were mediated by the kinds of conceptualisations controllers individually and collectively held about work performance. Where controllers had experiences that had resulted in "Lone Ranger" conceptualisations and worked in groups where norms of practice were not based on teamwork, informal learning as part of everyday work practice was inhibited. This was supported also by individual or social identities that were embedded in behaviours that did not include practices of inquiry. However, other contextual features were found to enhance learning. The data examined showed how changes in cultures, for example, from working as "Lone Rangers" to "Team-players", provided positive results for both accredited and informal learning in the ATC workplace.

In addition to culture, conceptualisation is influenced also by organisational structures when the activities of explaining, accounting for, and pattern generating are built into job tasks and roles. In High-3 workplaces traditional means of learning, such as the trial and error (generally associated with experimentation), are not available. Nevertheless, opportunities for experimentation are evident in a range of ways.

The fourth phase in the learning cycle is experimentation. According to Kolb [1984], learning is limited if an individual formulates concepts to generalise to other settings, but fails to test their validity. Testing the validity of conceptualisations, based on reflections on experience, can be done through evaluating past experience and envisaging new alternatives to be put into action immediately or some time in the future. Envisaging new alternatives may occur also in thinking about past actions (reframing). In this case an expansion of the range of choices available might be made though they may or may not be acted upon in the future. Therefore, "experimenting" is emphasised because it can involve reframing actions that occurred in the past, action to be taken in the present and also it can mean developing choices to put into action in the future. Experimentation then, refers to developing choices and envisaging new ways of acting. These choices and alternatives are tested out mentally and/or practically through developing alternative plans of action for the future and acting on those alternatives when appropriate.

Organisational contexts enable and constrain opportunities for experimentation for both individuals and groups. Structures such as team-work, for example, increase possibilities for experimentation because they enable a shared continuity of experience to occur across team-members and thereby enable the experience to be used as a resource for inquiry to generate increased possibilities for action. Similarly, the degree of formalisation within a workplace may limit the capacity for individual experimentation, though policies and rules formalising work activity may also embed behaviours aimed at generating alternatives into job tasks and roles. Organisational culture influences individual and group opportunities for experimentation to the degree that such practices are enabled and constrained by collective norms of practice, shared conceptual schemas that account for how the world works as well as shared capacities for collective remembering.

Organisational change and its impact on learning and interaction

The rest of this paper will discuss the second research question: "In what ways do organisational changes, such as the introduction of complex technology, influence workplace learning?" This question provides insights into the linkage between automation and interaction. The term 'interaction' is used here to mean social interaction between actors in the work system.

Change has the capacity to produce both positive and negative outcomes. It is argued that the history of previous change provides important insights into likely future directions for this particular workplace well as others with similar features. This is because automation is not introduced in a vacuum but instead is frequently a part of a longer trajectory of technological development. Examples of the changes discussed are summarised in Table 2.

Table 2: Examples of changes occurring within the ATC workplace

EXAMPLES OF CHANGE (S= STRUCTURE; C= CULTURE)	IMPLICATIONS	DIMENSIONS OF EXPERIENCE
Greater traffic demands (S)	Increased intensity	Temporal
Increased diversity in traffic system (S) Technological change (shift to more abstract & symbolic interpretation) (S)	Increased variety Shift from sentient to intellective knowing	Complex
Changes in recruitment patterns (S)	Shifts in occupational identities	Affective

Changes in values/histories (C)	(through) changes in valuing work and work's meaning	
Technological change leading to reduced sociality of work Changes in access to narratives (C)	Decline in sentient interdependence / Increased need for systems understanding	Social

Greater traffic demands have increased the intensity of work, thus altering the temporal experience of that work. This is leading to attempts to reduce complexity through increased regularisation of procedures. In turn, for some controllers, these changes are perceived as deskilling, because there is a requirement to follow specific rules and to rely less on their judgement (change in the affective experience of work).

One of the most significant changes to the cognitive complexity of ATC work involved the introduction of radar (although in Australia radar coverage still only exists over a small proportion of controlled airspace). Historically, radar controllers used what were called "bright" displays, which consisted of a radar screen with small plastic tags, called "shrimp-boats". The controller would write the details associated with the aircraft (such as its call sign and flight level) on the shrimp-boat and then physically move the shrimp-boat along the display as the aircraft completed its flight trajectory. In the 1980s, technological improvement resulted in the information the controller had needed to hand-write being automatically made available on the radar screen. In terms of work practice, these technological innovations resulted in two significant changes: the controller no longer needed to physically move the aircraft, pointing out where the aircraft was going to be; and the reduction in the amount of communication needed between controller(s) and pilots, since the information was more easily available and did not have to be requested. These technological improvements have since had some unintended consequences for those involved in formal learning programs. For example, in on-the-job learning, the technological enhancements result in data being automatically presented to the trainee in such a way that some of the cognitive processes of the trainee become opaque to the instructor. This is because the trainee no longer has to actively seek such information from other actors in the aviation system. Instructors have to develop news ways to understand the trainee's thinking and problem solving.

The move from shrimp-boats to a computer-mediated display changed the relationship between thinking and action for controllers because these cognitive processes were now mediated by symbols and data available on a screen. The symbols being interpreted, however, have remained until recently a "real-time" representation of the air traffic pattern the controller was working with at the time. Over the past five years, Australia has introduced a new ATC system called The Australian Advanced Air Traffic System (TAAATS). TAAATS involves a new computer-based system which integrates advanced flight processing technologies with radar and Automatic Dependent Surveillance (ADS) processing to provide controllers with graphical displays of all known aircraft information.

The implementation of TAAATS involves a greater degree of abstraction and interpretation than was necessary in the old system because there is an increased variety of symbols for the controller to interpret, thus increasing cognitive complexity. Under the current technological development the interpretation of symbolic data generated by the computer takes on even greater importance, as real-time work is distinguished from that which has been planned and projected. Controllers working radar sectors, for example, have on the screen a real-time representation of the traffic pattern in the air. This is in contrast to the controller working on a non-radar sector who have the same screen, though the representation will be a computer generated image of where the computer has positioned the aircraft (based on the information provided to it through the pilot's flight plan and updated situation reports). This increased variety in the interpretation of symbols is also exacerbated by a greater variability in the standards and procedures to be applied to aircraft with differing technologies (and demonstrated on the computer screen by different symbols). Some aircraft, for example, have their own forms of surveillance and radar detection equipment on board (such as collision avoidance systems), but not all aircraft flying through controlled airspace have these technologies. Aircraft (and airlines) investing in such technologies wish to gain full advantage of such systems which means a desire for smaller degrees of separation and greater flexibility in terms of, for example, flight path trajectories.

Automation has also changed the ways in which controllers interact with each other and with other stakeholders in the aviation system. Automated handoffs of aircraft between controllers, or automated messages between controllers and aircrew, for example, remove the requirement for active coordination. These changes have positive consequences, for example, by reducing the time taken to complete these work processes. However, there are also negative consequences, for example, in terms of the potential loss of individual and collective situation awareness.

These technological changes discussed require modifications in thinking and interpretation of data. They increase the variety of information available (and "noise" in the system, since there is more information to process and filter out what is important and what is not) and they will require different skills in interrogating data.

Table 3: Changes to the sociality of ATC work

SOCIALITY OF ATC WORK	HISTORICAL PURPOSE	CHANGE
Third ear	To monitor information transfer	Decline - limiting of direct communication
Body clock	To monitor ebb and flow of work	Decline - use of tech. aids, alerts
Referential anchoring. paper strips	To plan and anticipate what work may be building	Decline - automatic computer data transfer
Physical organisation of work	Capacity to observe and reflect on others work.	Decline in social cognition. Less easily visible physically.
Cultural artefacts - narrative	To share the experiences of others collective memory	Reducing.

Access to the social dimension of experience, important for the performance of work, are changing at the new data interface. As outlined in Table 3, these ways of experiencing ATC work had a purpose in enabling successful work practice. Historically controllers built up their understanding through sentience - i.e., the sentient know-how gained through working physically in concert with others. "Interdependent sentience" was built up through the use of the body and other artefacts. Interdependent sentience was also available through social cognition [Resnick 1993] where the work involved physical peripheral cues that were publicly available and access gained through working in close proximity. The primary means of developing interdependent sentience, as practiced through the "third ear", the "body clock" and referentially anchoring work based on the activity of those nearby, are no longer available. Work within the TAAATS system has less opportunity for social cognition [Resnick, 1993]. Controllers are still able to observe the work of others, though not using the typical strategies outlined in Table 3. It is possible, for example, with a computer-windows menu for the controller to scan nearby sectors in order to view what nearby traffic may be anticipated. However, the loss of physical peripheral information, needs to be replaced by other cues, because the complexity and automation within the socio-technical system still demands awareness about the interconnectedness of roles, relationships and responsibilities.

Other contextual changes and their linkage with interaction

The means by which people acquire knowledge as part of their everyday work practice, through structures and cultures, have also undergone change and these have implications for both learning and interaction. When individuals undertake their work tasks they learn not only about those tasks, but to some degree, about the ways in which those tasks impact on the work of others. In organisations where people move through various job roles, for example, there is an increased likelihood that those individuals will gain a greater understanding of the "requisite variety" [Weick, 2001] within the system. The ways in which people acquire knowledge informally through work activity in air traffic control for example has undergone some dramatic changes. In ATC, work re-organisation has lead to the closure or reduction of some services, the introduction of new accountability and decision making structures, (such as the introduction of teams) as well as changes to recruitment practices and career paths (such as "streaming" where controllers are recruited and trained to operate on one specialised stream of work such as approach or enroute control). Clearly these practices have been introduced for the betterment of the aviation system. And in most cases, this intention has been achieved. However, what these changes have also done has been to remove from the organisational system a variety of opportunities that employees have had in the past to learn about the aviation system and their impact on other actors operating within that system.

Conclusion: Where to from here?

A key question for future research is how will people working in tightly coupled, interdependent, computer-mediated environments gain the information necessary do successful work? When working is no longer reliant on the body and easily visible on physical artefacts for its representation, new cues need to be found. Unlike interdependent sentience, where physical cues could be noted implicitly, and interpretations based on tacit understandings, when work is mediated by a collectively shared information system, communication needs to be made deliberately explicit [Orasanu, 1995]. In understanding their intellective work, operators need to recognise the interdependence of that work. The interpretative processes required depend, therefore, on creating and sharing meaning through inquiry and dialogue. If the information is not socially available, more direct means need to be found to explicitly share and test out understandings and possible plans of action. This requires, therefore, a shift in emphasis from social to shared cognition. That is, these emerging work environments require communicative practices associated with enhancing and expanding the amount of relevant information available so that individuals can successfully undertake their work in concert with others. At the heart of intellective interdependence are the practices of inquiry that support informal learning.

It is concluded that contexts enable and constrain both formal and informal learning in important ways. The implication of this finding is that we need to evaluate the impacts of the contexts we create on possibilities for learning in order to create educative work environments where learning is a continuous process. Workplaces and learning endeavours developed with this aim would:

- provide opportunities for rich individual and collective involvement in work activity that involves the processes of reflection and conceptualisation on experience and intentional experimentation;

- maximise the ways in which transitions between the processes of experience, reflection, conceptualisation and experimentation occur for individuals and groups;
- embed the transfer of insights from engagement in learning within and between individuals and groups in organisational structures and cultures;
- prioritise the role of group communication and interaction aimed at developing shared understanding and expanding possibilities for future action;
- constantly evaluate the impact of work structures and cultures from the perspective of learning.

The material presented in this paper suggests that when considering appropriate human factors interventions in complex systems, it is necessary to widen the focus beyond human-machine interaction and to consider the ways in which other organisational contexts are implicated. Doing so may yield some important insights as to how work activity supports practices associated with informal learning.

References:

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C.C.. (1998). "Analyzing communication sequences for team training needs assessment." *Human Factors* 40 (4): 672(8)

European Organisation for the Safety of Air Navigation. (1996). *Model for task and job descriptions of air traffic controllers* (Released Issue). Brussels: Eurocontrol.

Brannick, M. T., E. Salas, & Prince, C. Eds. (1997). *Team performance assessment and measurement, theory, methods, and appplications.* New Jersey, Lawrence Erlbaum Associates.

Hartel, C. E. & Hartel, G. F. (1995). *Controller resource management – What can we learn from aircrews*? U.S. Department of Transportation, Federal Aviation Administration.

Helmreich, R. & Merritt, A. C. (1998). *Culture at work in aviation and medicine: National, organizational and professional influences.* Aldershot UK: Ashgate Aviation.

Hutchins, E. & Klausen, T. (1996). Distributed cognition in an airline cockpit. In Y. Engestrom & D. Middleton (Eds.), *Cognition and communication at work.* New York: Cambridge University Press.

Kolb, D. (1984). *Experiential learning.* Englewood Cliffs: Prentice Hall.

Louis, M. R. (1986). An investigator's guide to workplace culture. *Perspectives, 3/5*, 73-93.

Merritt, 1995

Orasanu, J. (1995). *Situation awareness: Its role in flight crew decision making.* Paper presented at the Proceedings of the Eighth International Symposium on Aviation Psychology, Ohio.

Owen, C. (1999*). Learning in the workplace: The case of air traffic control.* Unpublished PhD Dissertation, School of Secondary and Postcompulsory Education, University of Tasmania, Australia.

Reason, J. (1998). *Managing the risks of organizational accidents.* Aldershot UK: Ashgate.

Resnick, L. (1993). Shared cognition: Thinking as social practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition.* Washington, DC: American Psychological Association.

Tosi, H. L. Rizzo, J. R. & Carroll, S. J. (1995). *Managing organizational behaviour.* Cambridge, MA: Blackwell.

Weick, K. E. (2001). *Making sense of the organization.* Oxford, Massachusetts, Blackwell Publishers Inc..

Westrum, R. & Adamski, A.J., (1999) Organizational factors associated with safety and mission success in aviation environments. In D.J., Garland, J.A. Wise and V.D. Hopkin, *Handbook of aviation human factors*, London: Lawrence Erlbaum Assoc.

Wickens, C. D. Mavor, A. S. & McGee, J. P. (1997). *Flight into the future: Human factors in air traffic control.* Washington: National Academy Press.

Land-based Transportation

A Concept for a Learn-Adaptive Advanced Driver Assistance System

Matthias Kopf, Julien Simon
BMW Research
D - 80788 München

There is no doubt that cars and their systems will become more complex and that overall demands on the driver will increase. Therefore, solutions to resolve this conflict must be found. This paper presents work on the conceptualisation of a learn-adaptive, self-explaining Advanced Driver Assistance System. The results of a long-term study of Active Cruise Control uncovered many positive aspects of the system but also certain difficulties. These were largely associated with the system's limits, operational usage, and it's use in particular environmental conditions. In this paper, two multimodal help systems are suggested to help drivers develop appropriate conceptualisations of the system's behaviour and reduce the learning time before a steady state of usage is achieved.

Introduction

The Adaptive Cruise Control System (ACC) was the first of a series of Advanced Driver Assistance Systems (ADAS), recently introduced in the market. The ACC system is the extension of a conventional cruise control system and does not only keep a fixed speed but adapts also, by means of a radar sensor, the distance to a preceding car. A schematic of an ACC-system is shown in Fig. 1. For a detailed description of the BMW ACC system see [Prestl et al. 2000].

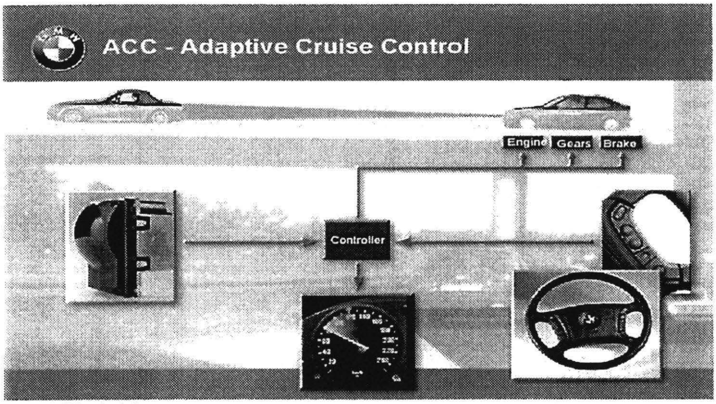


Fig.1: ACC principle

In the near and farther future much more of ADAS will show up. Possibly the next system to be marketed will be a Stop & Go Assistant, capable of handling the speed range between zero and about 40 km/h - currently not covered by the ACC system. Another system pending is the Heading Control system, which assists the driver in the lane-keeping task. In the more distant future, several combinations of assistance for longitudinal and lateral control with varying degrees of automation are to be expected. Experimental studies have demonstrated that all these systems significantly reduce the driver's workload and thus contribute to an increase in traffic safety [Nilsson, 1995].

However, the more functionality and sensors are integrated in an assistance system - with their specific limits - the more complicated the system is likely to become for the driver. The studies conducted hint to the fact that even rather simple assistance systems have to be learnt and that comprehension problems could produce critical traffic situations [Stanton and Marsdon, 1996]. In particular, take-over situations are known to be critical and sometimes very demanding for the driver. Take-over situations are typically situations in which the ADAS reaches its limits forcing the driver to take over control, possibly within a very short space of time. The only means, as yet, to overcome these comprehension problems is by providing an increasingly complex and comprehensive user manual. However, firstly, such a complex system can not be optimally represented in a manual and secondly, it is well known that the user manuals are not read by all users.

Therefore, it is necessary to find additional ways of making these ADAS understandable. One of these possible ways will be presented in this paper.

The basic vision is a "self-explaining ADAS" which needs no additional explanation by means of a written manual. To this end, the ADAS is equipped with an additional tutorial module, which gives the individual driver hints and explanations according to his learning state and the current traffic situation. In order to give personalised information, the system incorporates a personalisation component, which can be, for example, a reading device for a personal card. In addition, the system has to be connected to the environmental sensors as well as to the ADAS system itself. Depending on the complexity of the supported system, the tutorial module comprises also an additional man-machine interface like a display or speech input/output.

Besides the challenge of designing such a system to be useful for the driver and speeding up the learning process, there is another challenge with respect to driver acceptance. It is known that people adopt various strategies by approaching new technical systems and that these strategies are influenced by several personal characteristics e.g. the locus of control, [Wandke, 1999] or the motivation for learning. Thus, a "self explaining ADAS", and in particular the tutor module, has to be adaptable to these individual characteristics.

The proposed strategy to cope with both these challenges consists of eight steps:

1. Preparing the theoretical background in the area of learning technical systems.
2. Evaluating data from an long-term ACC study with respect to learning behaviour.
3. Defining the learning goals and extracting the most important learning hurdles.
4. Specifying a tutor module for an ACC system.
5. Implementing the module in a driving simulator, equipped with an ACC system.
6. Evaluating the system experimentally in the simulator.
7. Redesigning the system and implementing in an experimental car.
8. Evaluating the system in the experimental car.

This paper reports on the first 5 steps of this procedure.

Theoretical background

The basis for an analysis of users information needs in the aim of improving interactions with an ADAS such as the ACC system, is to get an understanding of the users conceptualisations of the task. These internal 'mental models' of the system provide predictive insights and help to explain users interaction and understanding of the system. Mental models are evolving models, they are renowned to be incomplete and not technically accurate but always functional [Norman, 1983]. Through the interaction with the system, the user continues to modify his or her mental model until a workable model is achieved. A major purpose of a mental model is to enable the person to understand and anticipate the behaviour of a system [Williams et al., 1983]. This means, in the case of the ACC system, that the model must have predictive power for the drivers about the state of the system, the feedback and the reproducibility of particular ACC situations. A significant determining factor of user's mental models is the difference in the user's technical background, previous experiences with similar systems and the structure of the human information processing system [Endsley, 2000].

A conceptually sound design will convey an image to the user that is consistent, cohesive and intelligible [Norman, 1983]. Norman has qualified this product characteristic package, the "system image" and asserts that it should be distinguished from the conceptual model upon which it was based and the mental model one hopes the user will form of the system. A self-explaining system should then be consistent with this system image, aiming to explain the underlying conceptual model to the user thereby keeping the system image consistent with that model and consequently, the user's mental model also. For a consistent mental model, the conceptual model that is taught to the user must fulfil three criteria: learnability, functionality and usability [Norman, 1983]. All three parts are addressed and explained in more depth below, in the specifications of actual situation-dependent speech and visual outputs.

The transformation of users' mental models or the way in which the user acts upon incoming information about the system state can be analysed within the framework of Rasmussen's Skill, Rule and Knowledge (SRK) classification [Rasmussen, 1979]. In this model, Rasmussen identifies three general task situations and shows the type of response required of the operator. At the "skill based level", the user's task operations correspond to normal routine, consisting mainly of automated routines and unconscious responses with no relation to the physical understanding or knowledge of the process [Rasmussen, 1979]. The "rule-based level" corresponds to the perception of familiar but abnormal situations which may be easily classified according to past experience. The "knowledge-based" level is entered if the situation is unfamiliar. It represents an initial problem-solving stage in which "... the user must evaluate the situation according to a mental model..." Rasmussen asserts that: "The control of human activity shifts from level 2 through to level 1 when an instructor is active" [Rasmussen, 1979].

The SRK model thus provides a basis on which to clarify differences in behaviour upon which the design and specifications of the help outputs can be based within the various learning stages.

However, Rasmussen sees behaviour as being driven not solely by goals but also by experience. He postulates that the form of information or the *product* of the change in mental context is a signal, a sign or a symbol respective of the operators skill, rule or knowledge based behaviour [Rasmussen, 1983]. The SRK classification therefore allows the history of learning and the current hierarchical structure of types of behaviour to be viewed simultaneously. Figure 2. attempts to demonstrate the interaction between the learning sequence (horizontal axis) and the hierarchical functioning of the levels (vertical axis). Items in brackets are those which, at the given learning level, have not yet developed [Penelope et al., 1988].

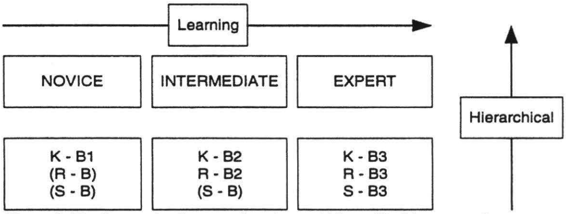


Figure 2. The interaction between learning and hierarchical interpretations of the SRK classification [Penelope et al., 1988].

This description attempts to show that at the novice level, users will tend to have a (small) repertoire of the knowledge based behaviours. At the intermediate level, they will acquire more rule-based knowledge and by the expert stage, the SRK levels will be qualitatively different from when they were being learnt, which is indicated by the numerals. For example, expert knowledge-based behaviour at K - B3 will operate differently from the way it does for the novice at K - B1. The knowledge base itself might be considerably refined and enriched but more importantly, the knowledge will be used in a very different way, influenced by such things as timing, trust and confidence in the system. The tutor will help the drivers move through the stages quicker, taking account of the qualitative change in the users knowledge.

Description of the long-term study

In order to gather appropriate long-term interactions with the system, extensive data and video analyses were conducted on an explorative study involving five participants driving for 2,5 weeks an average of 3,500 kilometres. The key questions of the study were:

- When (in what situations) and how (with what settings) is the ACC-system used after extended usage of the system?
- How do people go about learning the system capabilities and limits?
- How does the discovery of limits of the system affect future operational use and driver behaviour?

The participants were aged between 28-55, had very different occupations and social backgrounds, owned different vehicle types and drove yearly between 10 and 90.000 kilometres. The road ahead and the participants were recorded during the entire driving time along with all related car data. Subjective analyses consisted of semi-structured interviews and questionnaires.

The ACC System was equipped with a radar sensor capable of detecting objects as far as 150m away. The ACC was functional between 30km/h and 180km/h and was limited to a maximum braking capacity of $-1{,}6 m/s^2$. All the ACC related functions are operational through the knobs integrated in the multifunction steering wheel (see figure 3.)

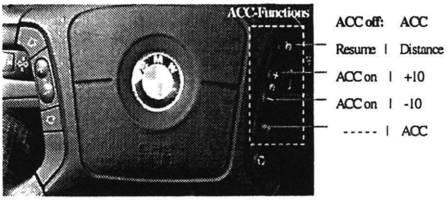


Figure 3. Functions in the Multi-Functions Steering wheel.

Feedback is given in the display of the onboard computer situated underneath the tachometers and via LED's around the tachometer. The overall measurement structure can be seen in Figure 4.

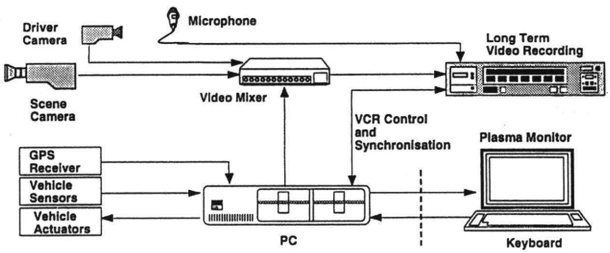


Figure 4. Overall hardware set-up in car.

Identification of the learning problem

In-depth analysis of the usage of the ACC system over the complete test period revealed stages of use hindered by three main problematic areas:

- The operational usage of the system.
- System limits.
- Use of the system in various potentially troublesome environments i.e. type of road or adverse driving weather conditions: poor visibility, rain, snow and fog.

These three areas form the core of the self-explaining system.

In the graph below (see figure 5.) the average amount of times the ACC was switched on per kilometre (or frequency per km driven) is calculated and the changes over time separated into quarters of total driven km) demonstrated.

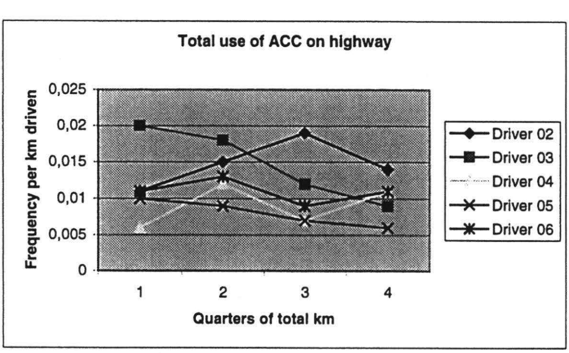


Figure 5. Total frequency of times ACC is switched on over the total driven kilometres.

Two levels of interpretation are possible from the graph. Firstly, two different types of usage strategies can be observed: a cautious approach with a steady increase in usage over time (driver 02) and a very early sustained usage of the system, with a gradual fall over time (drivers 03, 06, 04). Interestingly, both 'types' reach a balanced use in the fourth quarter that might be characterised as representing the drivers preferred 'optimal usage' - although extended analysis would need to be conducted to establish this. Secondly, the graph demonstrates the existence of different stages in system usage over time. The stages, represented through the quarterly time periods can be viewed as representing a preliminary stage of getting to know the system - learning to operate it, a testing phase - learning the system limits, and a familiarisation stage - learning to use the system most appropriately in particular environments. In terms of their representation in the SRK classification, all but driver 04 seemed to have reached the last, skill-rule behaviour, in the last quarter of the experiment, however, the resulting changes in the users mental models in terms of users interactions with the system must be analysed further.

At the interaction level, therefore, 'taking-over situation' – situations in which the driver must take over the control from the ADAS – were evaluated. These situations broadly cover the system limits over the regulation task. They are broadly dividable into two categories of situations. Category A includes the type of situations which the system is not designed to cope with in which driver intervention is always necessary i.e. stopping at traffic lights, the approach to standing vehicles, decelerating behind another vehicle to a speed below 30km/h, in steep curves and in the event of man objects (bridges, rail-tracks...). The situations in category B includes the type of situations in which driver intervention is sometimes necessary i.e. close cut-in of another vehicle or of the ACC vehicle, the deceleration of the leading vehicle, the approach to another moving vehicle and in the case of uncertain object detection (cyclists, motorbikes...). The latter category, therefore, clearly represents the most learn-intensive of the two requiring more 'learning with experience'.

The following graph (figure 6) shows an analysis of 'hard braking' situations. The first criterion for these situations was a braking force higher than -2,8m/s^2: a deceleration rate in excess of the ACC system's programmed deceleration capability. The second was that only the situations where the driver intervened directly after the beginning of the traffic situation which caused the need for deceleration were included. Analysis of the graph shows that in situations requiring drivers to intervene only sometimes (category B), there was a trend towards 'testing the limits of the system' at the beginning, followed by a certain apprehension of the system capabilities in the second (drivers 02, 03, 04) or third quarters (drivers 05, 06) and ending, in the last quarter, in a more balanced, personalised 'steady usage' of the system. In terms of Rasmussen's SRK classification, the stages reflect the operational behaviour of the drivers: from a knowledge based to a skill based behaviour – in which the knowledge of the system is quantitatively different from when he/she was still learning. One of the main design goals of the tutor module is to reduce the absolute number and the high increase of these situations at the beginning of the learning process.

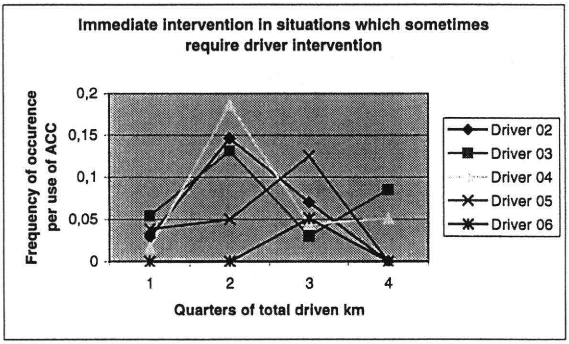


Figure 6. A hard braking analysis of drivers intervention rates in situations which sometimes require 'taking over'.

Taking over situations can further be analysed in terms of turning the system off via a moderate use of the brakes. It is hypothesised that this occurrence will decrease over time as drivers learn the situations which the system can handle and the situations in which they have to actively take-over control by braking (see figure 7.).

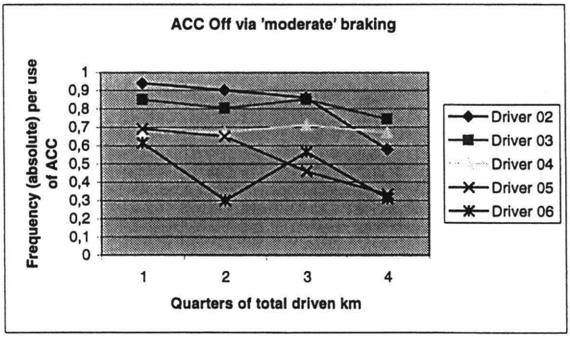


Figure 7. Switching off the ACC by moderate braking over time.

Switching the system off deliberately via the on/off button should increase accordingly, as the drivers' ability to predict situations increases (see figure 8.). This hypothesis is confirmed by the following graphs with respect to the drivers 2, 5 and 6. Drivers 3 and 4 show a different behaviour: Obviously, their strategy, which didn't change over the experiment, was to switch off the system mainly by moderate braking and only to a small extent by the on/off button. In this respect, it seems that they didn't "learn" anything or didn't want to learn.

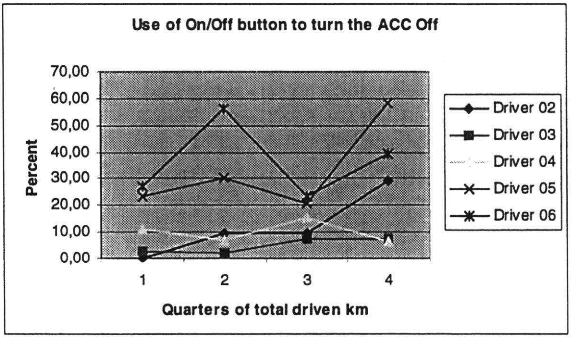


Figure 8. Switching the ACC off using the on/off button over time

Specification of help measures

After analysis of the system's usage, it is evident that intricate use of the system is not without it's problems. In the view of improving the effectiveness of the Man-Machine-Interaction (MMI), tackling issues related to comfort and satisfaction as well as optimising the safety of driver assistance system such as ACC, two 'tutor methodological concepts' to be tested and evaluated are proposed: a passive, 'pull mode' and an active 'push mode' tutor.

The in-vehicle on-line 'tutors' combine the information from the CAN-BUS (supplying vehicle data like velocity, activation of controls etc), Global Positioning System (GPS), light and rain sensors, ACC MMI operations and the system's current state to disclose timely and personally adapted explanatory help and advice to the driver. Additionally, a situation monitoring module will be integrated able to detect the situations in which the limits of the system are reached.

Based on two different methodological backgrounds, the passive tutor will provide explanations on demand only – 'pull mode' - while the active tutor will provide active explanatory feedback – 'push mode'. These two modes remain very much in line with BMW's philosophy of keeping the driver in control. In the passive mode, the driver must actively request information using the 'help button', while in the active mode, the driver must pre-select the guided tour mode to receive information about the system.

Both tutors feature similar functionalities i.e. both tutor modes have speech output as main feedback source, will issue a short introduction upon first usage of the most important information regarding general use and aims of the ACC... However, the passive tutor aims to provide additional explanations in lieu of incomprehensible or confusing situations, while the active tutor represents a self-contained 'guided-tour' of the system in an attempt to replace the

manual altogether. This dichotomy of modes is supposed to support the two most important learner types: the courageous, self-confident "trial and error" learner and the cautious, help-seeking learner, respectively.
As outlined above the main difficulties associated with the usage of the system can be categorized into three core areas. On this basis, three general user-centred goals were outlined:

- Achieving an optimal level of interaction with the system.
- Reducing the number of encountered high demand 'taking-over' situations.
- Obtaining a clear situation-specific understanding of the system.

The first goal, to reduce the number of encountered high demand situations implied, in the first instance, defining these situations practically. Most of these situations, corresponding to the inherent limits of the system, were derived from the main situational difficulties expressed by the five participants of the long-term study. The limits of the system can be characterised by the technical limits of currently available radar sensors for series production and the implemented restrictions for braking and accelerating capabilities of the system. The limited far-range sensing capabilities, for example, will make the approach to a vehicle at very high difference velocity a high demand situation requiring the driver to take-over the control. Further, the cone shaped nature of the radars results in late recognition of close range cut-ins from other cars and create potential difficulties when a detected target is 'lost' in a curve or a target is only detected very late in a curve. On the basis of these situation characteristics, algorithms can be deduced to identify and filter, in terms of registered car data values, the 'experienced' situation and determine the a-posteriori explanation needed by the driver to build a successful model of the system's behaviour.

The second goal, to achieve optimal operational usage of the system, entails the achievement of an efficient, effective, 'error free' usage of the system in minimal time. The difficulty of defining the criteria for optimal usage due to its subjective nature advocated the need for an analysis of the participant's operational 'learning curves'. Dependent on the drivers learning progress, the system will correspondingly adjust the explanations to be issued. Criteria used for identifying the actual learning state of the driver are, for example, errors committed, specific situations encountered or the help messages already put out. In the active mode, explanations of the specific function of each button will be automatically triggered upon detection of experienced operational difficulties. The order in which outputs will be issued through the multimodal tutor interface are prioritised, according to the learning state, usage pattern of the driver and, of course, the 'status' of the ACC system. During initial use of the system, the tutor will automatically issue the basic explanations twice - with a 'repeat' possibility. In the passive tutor mode, the 'help button' will light up when a button in the ACC interface is pressed. On request, the help button will issue the information of the last selected button's functionality.

The third goal of the tutor help-system aims to optimise the use of the ACC system with respect to safety in particular environments and environmental conditions. Combined data from the GPS, from the light sensors, rain sensors and the 'operational use' of the system is transferred to the tutor module which, in turn, issues the appropriate advice, explanations and support to the driver when inappropriate usage of the ACC system is detected dependent on road type, traffic conditions, at night and in particular weather conditions such as fog or heavy rain.

Experiment in Driving Simulator

The learning effects and the acceptance of the passive and active tutor modes are to be tested in the BMW driving simulator. The structure of the embedded tutor module is shown in the figure 10. The study will aim to assess the ability of participants in both tutor groups to operate the system effectively but also optimally in different weather conditions and their ability to deal with difficult situations will be tested against a control group. The measurement of participants ACC knowledge and skill acquisition will be based on their knowledge of different system components and devices; system principles (mental model without procedural skills); their ability to apply solutions to different problems and on their ability to generalise knowledge and skills beyond the information provided explicitly by the tutor (mental model with procedural skills).

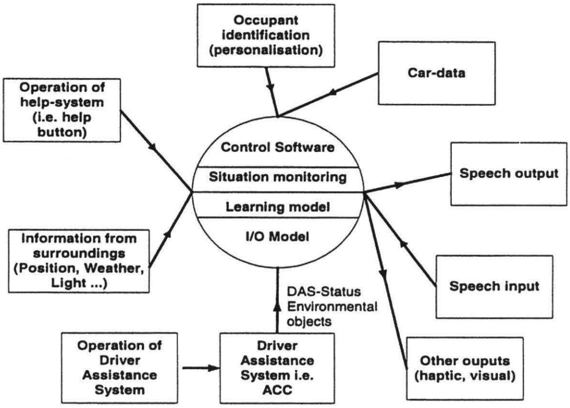


Fig. 10: Structure of the embedded tutor module

Conclusion

The work presented in this paper gives some insight into users learning process of ADAS and particularly, the most important learning obstacles. It has outlined a concept how the drivers learning curves could be shortened in the view of improving acceptance and safe use of ADAS in future. This research will be continued in the aim of implementing the tutors into a real vehicle and test the learning effects on both the system and the drivers in real traffic.

References

1. Endsley M.R. 2000, Situation Models: An Avenue to the Modelling of Mental Models, Proceedings of the IEA 2000/HFES 2000 Congress, San Diego, California, USA.

2. Nilsson L. 1995, Safety effects of adaptice cruise control in critical traffic situations. In Proceedings of the Second World Congress on Intelligent Transport Systems: "Steps Forward", Volume III, pp. 1254-1259. VERTIS.

3. Norman D.A. 1983, Some observations about mental models. In D. Gentner and A.L. Stevens (Eds.), Mental Models, (pp. 7-14) Lawrence Erlbaum, Hillsdale, NJ.

4. Penelope M.Sanderson and Kelly Harwood. 1988, The skills, rules and knowledge classification: a discussion of its emergence and nature, In L.P. Goodstein, H.B. Anderson, and S.E. Olsen (Eds.), Tasks, Errors and Mental Models (pp. 21-34), Taylor and Francis.

5. Prestl W., Sauer T. Steinle, J. and Tschernoster O. 2000, The BMW Active Cruise Control ACC, SAE Technical Paper Series, SAE 2000 World Congress, Detroit, Michigan, 2000.

6. Rasmussen J. 1979, On the Structure of Knowledge – A Morphology of Mental Models in a Man-Machine System Context, Riso Report M-2192 (Roskilde, Denmark: Riso National Laboratory).

7. Rasmussen J. 1983, Skills, rules and knowledge: Signals, signs and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 257-266.

8. Stanton, N.A. and Marsden P. 1996, From fly-by-wire to drive-by-wire: safety implications of vehicle automation. Safety Science **24** (1), pp35-49.

9. Wandke, 1999: Personenvariable der Situationskontrolle - Ihre Bedeutung für die Bewertung der Interfacegestaltung in Mensch-Maschine-Systemen. 3. Berliner Mensch-Maschine-Werkstatt, Fortschritt-Berichte VDI, Reihe 22, Nr. 1. VDI-Verlag Düsseldorf, 2000

10. Williams, M.D. , Hollan, J.D., and Stevens, A.L., (1983). Human reasoning about a simple physical system. In D. Gentner and A.L. Stevens (Eds.), Mental Models (pp. 131.153) Lawrence Erlbaum, Hillsdale, NJ.

Contribution of a 'comprehensive analysis' of human cognitive activity to the Advanced Driving Assistance devices design

Thérèse Villame
RENAULT Technocentre, Research Division, Ergonomics and Human-Machine Interaction Department , Guyancourt, France (therese.villame@renault.com)
Jacques Theureau
Centre National de la Recherche Scientifique / Université de Technologie de Compiègne, France (jacques.theureau@hds.utc.fr)

For some years now the cars manufacturers, in relation with the development of in-board electronics field, have been involved in the design of *advanced driving support systems*. The first point to underline is the necessity to describe in some details the *drivers' cognitive activities* as experienced by them with and without these systems, in order to find the network of their *bodily, situational* (including the considered support system) *and cultural determinants*. The second point is the necessity to develop these empirical studies of drivers' cognitive activities all along the design process, and to integrate their results at every step of it. It is what we call a '*comprehensive analysis' of human cognitive activity*, of which we will outline the practical results for the design process.

Key-Words : Advanced Driving Assistance, Ergonomics, Human-Machine Interaction, Cognitive activity, Situated Cognition, Embodied Cognition, Situation Awareness

Introduction

In Europe a key-step in the design of advanced driving support devices had been the program named Prometheus (PROgraM for European Traffic with Highest Efficiency and Unprecedented Safety) which gathered together, between 1986 and 1994, cars manufacturers, car industry suppliers and transport research laboratories for working on the concept of 'intelligent vehicle'. Advanced driving support systems are now a reality and some of them, such as navigation systems which help the driver to find and follow an itinerary, are available on vehicles recently launched on the market. The development of such systems obviously requires technical abilities. It also requires takling the question of the relevance of the technical choices made to the users' needs. Indeed, the quality of this relevance determines, first how the drivers will accept and appreciate the new driving system, second the quality of the human-machine interaction and finally the conditions of the integration of the new driving assistance into the driving activity. Involved in the design process of several new advanced driving support systems as ergonomists, we take here the opportunity to make a synthesis of the lessons learnt from this experience (see [4], [5] and [9]).

1. Main ergonomic questions of driving support systems design

Research and development is being conducted on more and more aspects of car driving in order to design driving support systems of varying degrees of sophistication. In this context, it is useful to formalise the distinctions between these systems, as their design raises different and specific ergonomic questions. This section is devoted to this. Based on our experience and knowledge, driving support systems can be classified into two principal categories, depending on their purpose and the relationship between the system and driver that this purpose creates, each category raising particular ergonomic questions:

- *Information systems*, the main purpose of which is to provide information to the driver in more or less sophisticated ways and using different media, for example, traffic information systems or driver warning systems. More advanced systems can also be placed in this category, for example, systems which aim to provide step by step guidance to the driver for a specific type of action, such as navigation systems which assist the driver in planning and following a route, or systems giving precise instructions to help the driver perform a given manoeuvre. This category of system essentially raises ergonomic questions in relation to: defining the most appropriate information for transmission to the driver in order to help him; defining how this information should be presented (in particular warning information) and, as it becomes more widespread, how the display media should be shared between providers or systems; sharing the attention resources of drivers, which comes back in particular to the question of defining the most appropriate moments - or the ones to avoid - for transmission of this information.

- *Interactive systems*, the purpose of which is to assist the driver with one or more particular aspects of driving. These are mainly vehicle control support systems (regulation of speed, distance or the lateral positioning of the vehicle etc.) or active safety systems such as braking or trajectory control systems, etc. These systems, some of which intervene directly in the activity of the driver or reinforce his action, raise very specific questions in terms of their relationship to the driver, questions which it is not always easy to immediately answer. They sometimes result in a significant modification of the interactions between the driver, his vehicle and the environment, resulting in the following questions being asked of ergonomists, in particular: the conditions required for drivers to accept a new system or device; how the functions between the driver and a given support system should be distributed and in particular, how control of the interaction should be shared; how the system should intervene in the driver's activity so that it is integrated as effectively as possible, and constitutes real assistance and not a constraint, even if only occasional.

Of course, the boundary between these two categories of system is not clear cut. In fact, it is better to think of there being a continuum between these two categories and that a given type of system will sometimes be an information system and sometimes an interactive one, depending on the purpose of the assistance devised by its designers. Thus, a manoeuvre support system can be purely informative, for example giving the driver information on the distance separating him from another vehicle during a reversing manoeuvre or on how the wheels are positioned. It can also be designed more as an interactive system if the system directly controls the dynamic of the vehicle or movement of the steering wheel for example. Similarly, a system whose purpose is to assist the driver in staying in his lane of traffic may be "only" informative if its action is "limited" to warning the driver when he veers from the lane; it will be more of the interactive type if, once the warning is given, the system also acts directly on the steering wheel. There is also no clear cut boundary between the two main groups of ergonomic questions mentioned above. They are in fact relevant for all systems, whatever their category, but carry different "weight". For example, the question of how information should be presented to the driver is obviously also of relevance for interactive systems, but to a lesser degree. Similarly, the question of acceptance of a new system and that of its integration into the entire driving process is applicable to information systems as well as interactive ones, but is more crucial for the latter as the resulting modification of driver / system interaction is more significant and of another type, particularly with respect to safety concerns.

For our part, we have mainly worked on driving support systems that we have described here as interactive, that is, which directly intervene in the driving process. The examples supporting our ideas in what follows are therefore particularly relevant to these systems and the ergonomic and design problems associated with them. Be that as it may, our purpose here is to discuss the fact that, whatever category of driving support system the devices we are looking to design come from, it is clear from examining the literature in this area (or the different sometimes visible technological developments which their creation has given rise to) that their design is associated with only one particular view of human activity, which may be coined as classical. However, recent scientific progress in the cognitive sciences, neuroscience, psychophysiology and psychology has contributed to the emergence of new paradigms regarding humans in activity. These paradigms are particularly interesting in light of the reappraisal of technological problems related to the design of innovative and appropriate driving support systems. If these new paradigms are taken into account, we in fact obtain a different view of human activity from the approach which still currently traditionally dominates the area of human cognition in general and, as a result, that of car driving. This different view of human cognitive activity also results in a different view of the role of technical devices in this activity. We will therefore now examine current scientific knowledge of human cognitive activity and its relevance for driving support systems design.

2. Current state of scientific knowledge of human activity and driving support systems design

In order to analyse a human activity such as car driving, and to evaluate a support system for this activity, we need to have an understanding of human activity in general and of the role that tools and other artefacts play in it. For some years now, the fields related to the study of human activity (cognitive sciences, psychophysiology, psychology, cognitive anthropology, robotics and ergonomics) have seen the development of a new emphasis on *action* and the historical, material and physical *situation* in which it takes place. We now speak of situated action, of situated and distributed cognition, of autonomous robotics, enaction and constructivism. A brief summary of this new perspective will be very useful in providing the background to our questions and analyses.

2.1. Towards an alternative vision of human activity

Firstly in the area physiology, as [1] has shown, we have moved from a physiology of reaction - or reflexology - to a physiology of action. That is, we have moved from a physiology in which stimuli are given in order to study responses, to a physiology in which we study the action produced endogenously by an animal or person on the basis of his involvement in the world at a given moment. This has become possible by the progress made in methodology. For fifty years researchers worked on anaesthetised animals by giving them stimuli and recording their reflexes. Since then we have moved from this reflexology to a physiology of anticipation, planning and action, in which action and perception cannot be separated, in which there is no perception without action. In other words, contrary to what was previously thought, the brain does not transform passive sensory information into reconstructions of objects in the world. The brain pre-specifies the objects that it wants to analyse and constructs the world on the basis of hypotheses. These ideas put forward in modern neurophysiology and experimental psychophysiology actually reflect a biological reality. In cognitive sciences, physiology, cognitive psychology and ergonomic cognitive anthropology, another parallel revolution is also occurring, namely the "reincarnation" of cognition. This is the fact that [1], [8], [3] and [6] (see also, more specifically, [2] and [7]) all respectively insist on, to quote just one work in each of these four disciplines. We are leaving an era characterised by, on the one hand, muscular energy, or humans as "human motors", and on the other, a formal study of cognition, or humans as "human computers". After the disappointments of traditional artificial intelligence, a new approach to human cognition is tending to replace the ideas of symbolic representation or mental image with those of active constitution of the perceptive experience and of coupling with a concrete situation.

In the classical approach to human cognition, cognitive activity is divided into a sequential chain of information processing operations ended by the execution of an action. This is easily understood if one sees the cognitive activity of humans as information processing. The modelling of perception, reasoning and decision-making are thus directly based on how computers operate. In the new approach that we are describing here, the actions of the person are present from the outset, in the very constitution of the perceptions themselves. Similarly, reasoning and decision-making are understood in terms of a gradual transformation of the concrete situation (material environment, bodily posture, internal memory and external inscriptions, motivations, etc.). This new approach has important consequences for how we consider the role of artefacts in human activity and in particular for understanding the use of information processing technologies. These technologies are central to the support systems that we have examined.

Many support systems are *information* processing systems which deliver *information* to the driver. However, if we no longer compare human cognition to a simple information processing operation, these two uses of the concept of information mask a confusion between sensation, perception and interpretation which should be uncovered and clarified. Information for technical computerised calculation or communication systems is a clear and well-defined concept. This is information in the sense used by Shannon, obtained by discretisation (and coding). Perception by a human cognitive system is very different. It is not the simple passive analysis of sensory data. It is created over time as a synthesis of a heterogeneous set of sensations obtained via various sensory processes on the basis of the actions undertaken. When engaged in a perceptive activity, humans are not aware of the sensations received, but their attention is placed on the object perceived. For example if we use a stick to explore our environment, we will perceive the relief of the ground at its end, where our exploratory action of moving the stick determines the changes in sensation at the level of our hand. But as soon as our attention is placed on this perception, we will not be aware of the movements and vibrations of the stick in our hand. In the same way, engaged in a visual perception, we are not aware of the extremely variable sensations received on our retina. It is at this point that one leaves the levels of analysis covered by experimental physiology and psychology for those which are covered by phenomenological psychology and ergonomic cognitive anthropology. What is perceived is not always the concrete situation in which the human activity is directly taking place. These can also be signs which then have to be interpreted, for example, the screen of the speedometer, or the terminal of a geographical guidance system. This interpretation depends on the experience of the person involved, but also on the circumstances in which he finds himself and his state at that particular time. Either the interpretation is direct: the signal has an unambiguous meaning for the person because its variations have a direct causal link with the state of the situation (for example the rev-counter, or an audible or kinaesthetic signal triggered by leaving the road) which they have learnt to evaluate. Or the interpretation requires the person to use a symbolic system: the link between the signs and their meaning is then arbitrary, relative to other signs, and conventional (for example the milometer or an audible speech message). It should be noted that even in the first case, the perception of the signal must sequentially precede the understanding of its meaning: for example, the attention must first be placed on the rev-counter, before the current state of the motor can be understood. However as we know, drivers who are even just a little bit experienced use this indicator very little. Its use is better replaced by the direct perception of the relationship between action on the accelerator and sensory feedback (noise of the motor, acceleration, resistance of the steering wheel, various types of stress feedback, etc.). In fact, in this sensory-motor loop, the attention is placed directly on the motor functioning without requiring intermediate focusing on the dashboard.

2.2. Towards an alternative view of driving support systems design

In the classical approach to human activity, support systems are often developed using prior modelling of this activity which is understood as a series of information processing operations. Support systems are then developed which artificially perform some of these operations, and which are thus intended to replace the human operator. The tools are therefore essentially designed as artefacts replacing human activity. In contrast, in an approach which places the priority on action and physical involvement in the situation, the tools are firstly designed to allow modifications in the perceptive and operating possibilities of the human operator. Indeed, if perception is created through the activity of the operator, it depends directly on his ability to act and feel (what is known as the "proper body"). The perceptive organs are understood to be coupling devices since they allow a relationship between action and sensation to be established through coupling with the environment. On this basis, any tool, if it is correctly understood, can be seen as a coupling device which is integrated into the "proper body" of the perceiving person. For example, the white stick used by blind people is a coupling device giving them access to tactile perception a little way ahead of them. In this perspective, support systems, like any other tool, are artefacts designed to increase the possibilities of human activity, and not to replace it.

When it operates as a coupling device, the tool becomes transparent for the human operator since it participates in his perceptive activity. At the moment when it is used to perceive and act, it is not actually perceived itself. Appropriation of a tool corresponds to how successfully this process is "integrated into the proper body", in other words whether the driver "embodies his car". He will thus perceive the road under "his" wheels, the gravel he is driving over or the edge of the pavement which he has just touched (and no longer consciously perceives the relationships between the feeling of vibration of his seat or steering wheel and his motor commands). Attention is directed outside the vehicle (and no longer towards the vehicle). Appropriation of a new tool is not immediate. It is achieved through learning repeated relationships between action and sensation through this device. It is successful when the device becomes transparent: the tool is used directly for the activity and is no longer the object of a learning activity.

In general, internal "information", that is, a signal emitted by the car (symbols on the dashboard, forms projected on screens), creates a dissociation between the driver and his vehicle. The perception of this form, its recognition and its interpretation imply a spatialisation-exteriorisation of the dashboard before it is interpreted. Interpretation of the situation becomes confused and fragmented when the user has to quickly switch from perception with a coupling device to perception of this device itself (like trying to perceive with a stick, and at the same time perceiving the stick in your hand). A rule governing the acceptability of a tool states that not only must it not hinder direct perception (the user does not need to look away from the road significantly), but also that it should be possible for it to become "invisible" in the course of its use. In other words, using the formula of the 19th century psychologist, A. Bain, who said “thinking is holding yourself back from acting”, such a tool must not force you to think when you should be acting.

3. "Comprehensive analysis" of human cognitive activity: theoretical and methodological principles

We developed our studies in driving support systems design over the last few years along this new approach to human activity. In this section and in the next one, we will describe how we did in practice when looking more specifically at car driving.

3.1. Understanding activity in order to design appropriate support systems

In the driving support systems design projects, we systematically start from the basic idea that, in order to be effective and accepted, and constitute real assistance and not an additional constraint for the user, as is unfortunately sometimes the case, a technical system must correspond to the essential characteristics of the activity which it is dedicated to. In order to do this, detailed knowledge of these characteristics and what determines them is required, thus making it possible to do more than simply describe the phenomena observed. This point of view determines the particular conditions governing knowledge of this activity, in other words the conditions for collecting and analysing data on that activity. It is in this context that we have developed the concept of “comprehensive analysis” of the activity, with which specific theoretical and methodological principles are associated, as we will now describe and illustrate using examples from studies conducted during our involvement in different design projects.

3.2. Collecting qualitative and quantitative data on situated driving activity

Taking into account the construction of the action in the situation, and considering action and perception as inseparable in this construction, our approach gives priority to the study of drivers' activity in a natural driving situation. For us this is a basic condition for understanding the complex and dynamic character of the activity of driving, and its eminently contextual dimension. We believe in fact that driving is largely created as a function of circumstances, never possible to fully anticipate and constantly changing. In addition, driving is multi-sensory and the driver is also almost permanently interacting with other drivers. In order to take account of all these characteristics and of the construction of driving in relation to a given situation, we feel it is essential to put drivers in real driving situations and to consider their point of view on how they carried out the activity, in order to collect “explanatory” data on it. Our studies were mostly based on field studies on the open road during which a combination of quantitative and qualitative data were collected, firstly in relation to these general characteristics of the activity of driving. For example, we were very systematic in collecting data on the dynamic of the vehicle and of certain other vehicles with which the driver was interacting (speed, acceleration, use of the brake, deceleration methods, combinations of speeds used, etc.), on the behaviour of the driver (manoeuvres, positioning in traffic lanes, action carried out on the vehicle and/or particular equipment, etc.), and on the context encountered by the driver (traffic, infrastructure, manoeuvres of other drivers, etc.). Secondly, we also collected data in relation to the characteristics specific to the particular dimension of the activity that we wanted to provide assistance with. It was thus possible to collect data on lateral veering or the immediate repositioning of the driver in his lane of traffic in the context of a study conducted for the design of a "Lane Keeping" type system. Relative speed and relative distance data were collected more particularly in the context of studies on management of speeds and distances. Similarly, data on distance in relation to an obstacle or another vehicle were collected more specifically for studies looking at how manoeuvres are carried out. In all cases, important emphasis is given in the studies that we conduct to the point of view of the driver himself on his activity, as an access to his involvement in the driving situation. This emphasis takes the form of collecting verbal data while the activity is actually being carried out and/or in an autoconfrontation situation (the driver watches a film of his journey, the latter being systematically recorded, and comments on it to clarify his actions after the event). It should be noted in this regard that in the perspective we adopt, the point of view of the driver and of the observer cannot be identical and it is therefore necessary to develop special methodological principles which allow these to be articulated, a question which we examine elsewhere.

3.3. Defining the questions about the drivers' activity to be answered for support system design

Developing such a study is not, however, self-evident and preparatory stages are necessary in order to determine the different methods involved. To do this, our first step in this procedure is to identify as precisely as possible the requirements of the designers themselves. We identify the questions they ask themselves, the elements which they need to progress in the project, in particular regarding the interaction between the driver and the future or close-to-completion system. The objective is then to agree on the questions which the ergonomic study must attempt to answer and in what form, but also to clearly ask those which are not directly within its remit. Thus, in the context of our involvement in the development of a manœuvre support system, a prior ergonomic study was conducted (that is, before any introduction of a support system), the principal objectives of which were to contribute to the definition of the functions to be developed, to determine the methods of interaction between the system and the driver, and how the information would be given back to drivers, as well as defining the most relevant criteria for evaluating already existing intermediate technological solutions and the prototypes which would then be developed. This prior work, devised to closely reflect the concerns of the designers, determines the kind of study protocol produced, and defines in particular the study situation selected.

3.4. Defining the situations under study

Conducting field studies on the open road is not sufficient to guarantee the natural character of the drivers' activity. The following stage is to therefore select the experimentation situation which will allow the activity we want to examine to take place as naturally as possible, while also allowing us to obtain the response elements relevant to the technological concerns of the designers. In this perspective, we feel it is also particularly important to put drivers in a realistic driving situation and one which will reflect, as far as possible, the probable future situation in which the system under design will be used. For example, in the context of our involvement in a project to design a speed and distance control system for urban and suburban situations, in consultation with the designers, we decided to ask a panel of drivers to make their usual home-work and work-home journeys and at their usual time. Regarding the drivers (who were company employees, largely for reasons of confidentiality), they were mainly selected on the basis of not being involved in the project which resulted in the study, or even in any similar technological project conducted by the company. For example also, in the context of evaluating a prototype distance control system, a first study consisted of asking pairs of drivers to make a relatively long motorway trip, representative of the type of journey and use for which this system is designed (for example a holiday journey). During a later study, drivers completed a journey of several hundred kilometres several times that they actually had to make in the context of their professional activity. In addition, in order to reinforce the natural character of the driving during our studies, we do not give particular instructions to drivers in terms of specific tasks to be completed or performances we want them to attain. The drivers are in fact encouraged to drive as they would normally. For example, in the context of a study on overtaking manoeuvres, it is the drivers themselves who decided whether or not to overtake a vehicle, and at what moment and how to perform the overtaking manoeuvre. We do, however, give them specific explanations of the methods used to collect the verbal data, by explaining to them what it is essential for us to collect, in order to then have enough elements to allow us to understand the phenomena observed. This type of methodological choice, closely reflecting our theoretical hypotheses on human cognitive activity, obviously does not allow us to make precise comparisons between drivers. However, the objective is in fact to identify, for example relative to the cognitive mechanisms used in the management of speed and distance in urban and suburban contexts, or relative to the use of a particular support system, firstly the determining regular features, in particular contextual ones, and secondly the typical specific features, from among a very diverse panel of drivers.

4. Main practical results provided by a comprehensive analysis of the activity

The practical results of such "comprehensive" studies of drivers' activity in natural situations, in particular when they take place sufficiently upstream of the design process and are not simply of an exploratory nature, lie in the fact that they constitute a solid and essential base of knowledge which provides elements on several of the dimensions involved in design issues. These elements are even more valuable since they will be useful throughout the design process, at each of its stages, and also can be transferred, at least in part, to other design processes. In this concluding section, we will present the main ones.

4.1. Identifying the characteristics of situations in which support systems are used

Being positioned a long way upstream of the design process, one of the first practical results of a comprehensive analysis of drivers' activity is to assist in the identification and specification of the most relevant typical situations for the probable future use of the system being developed. At a later stage when a prototype system is evaluated, the analysis is used to validate and complete identification of these situations and above all to characterise them in as much detail as possible in relation to realistic use of the system. In particular, it allows us to define the sometimes critical nature of some of these situations. In the two cases, both the quantitative and qualitative results produced by comprehensive analysis of drivers' activity allow us in particular to understand the diversity of these situations, as well as the constraints created by the road context. For example, in the case of a study for a project to design a speed and distance control system for urban and suburban contexts, conducted before any introduction of a support system, two major elements were revealed in terms of the characteristics of the situations of use which the future system would have to be able to deal with. On the one hand, this study showed that there are no standard urban or suburban journeys but in fact a huge variety of them. In addition to this simple observation, our study helped to define the sources of this diversity in great detail, and, for a given panel of drivers, the combinations of diversity within a particular journey. From a methodological point of view, it can be seen that we were able to obtain this result because we chose to ask our panel of drivers to make their respective usual home-work journeys, thus placing them in diverse and therefore real situations, rather than imposing on all of them a journey that we had identified beforehand (and perhaps wrongly) as a typical urban or suburban journey. On the other hand, given that the purpose of this speed and distance control system for urban and suburban contexts is to manage speeds and distances in dense to very dense traffic situations, it was possible in the context of the same design project to identify and characterise different types of traffic jam. This was important because, as we were able to demonstrate on the basis of various objective parameters and using the drivers' points of view of their activities, the elements of the situation to be taken into account, i.e. the constraints for drivers, particularly in relation to risk-taking and time factors, were actually different depending on the type of traffic jam in question. For the design, these results were of direct relevance to the question of how to manage the approach of a traffic jam as well as manage a succession of traffic jams, the latter possibly being of different types. For example also, in the context of a study devised to evaluate the advanced prototype of a speed and distance control system, we were able to determine those situations where use of the system was particularly costly for the driver, for example notably because of some of the laws of control implemented. In addition to this observation, it was the data produced from the detailed analysis of the coupling between drivers, the system and the environment which allowed us to fully understand what was happening in these specific situations and to thus determine precise recommendations for the development of the system.

4.2. Defining functional and interaction requirements

Upstream of the design process, this characterisation of the situations to be taken into account regarding future use of the system thus represents a first contribution to the functional definition of this system, by providing information on the functions necessary for it to meet the requirements of drivers in terms of the situations they will be confronted with. Data on the activity produced from a comprehensive analysis of that activity also help to specify the modalities of driver / system interaction throughout the design process. The objective is in fact, on the basis of the previously identified defining characteristics of the activity, to help determine what the optimum modes of dialogue are in order to implement the functions of the system and what information feedback is most relevant in relation to its operation. We also aim to provide guidance for designers regarding the actual relationship between the driver and the system. For example, with respect to speed and distance control systems for motorway journeys, which act directly on the longitudinal control of the vehicle (that is, which accelerate, decelerate and even brake themselves depending on the speed and distance of the vehicle in front in the same lane), the crucial aspect is not so much the level of information to be supplied to the driver, but the laws of behaviour which will be implemented in the system. The important thing here is to fully understand the dimension of the activity which will be computer-supported (in our example, regulation of acceleration and deceleration), but also the global context in which it takes place, so that, to continue with our example, deceleration is managed in different but always appropriate ways for the smooth running of the driver's activity, depending on whether he gets closer then stays behind the vehicle, or whether he gets closer and then overtakes it.

4.3. Defining situations and evaluation criteria

When the prototypes start to be developed and can be evaluated, this knowledge of the factors determining the aspect of driving to be computer-assisted provides specific criteria which can be used to judge how accurately the system developed meets the needs of drivers, and also to identify the directions which should be focused on in order to optimise driver / system / environment interaction. As in a classical ergonomic evaluation, these criteria will be in terms of both the usefulness of the system in question and its "usability", that is, its ease of use. A particular aspect of it is how quickly drivers get used to and learn the system developed. In the case of innovative driving support systems, we should add two fundamental elements for judging the appropriateness of the system: firstly, the degree to which drivers will successfully appropriate to themselves the new system, which depends in particular on the degree of confidence it gives them when used in a driving context;

secondly, how compatible use of the system is with the determining characteristics of driving as a whole. In particular, the new system must be in line with and respect this activity in all aspects of its components, otherwise it will become a hindrance or constraint (or even partial hindrance) for the driver instead of assisting him. This is why it is important to know these determining characteristics in as much detail as possible and, preferably, as far upstream in the design process as possible, so that the optimum conditions for successful adoption of the new system can be determined and fully integrated into the driving process. In addition, for all these aspects, the knowledge produced from a comprehensive analysis of the activity can be used, during the evaluation phases, to select the most relevant situations for analysis, to identify the criteria to focus on, and to guide the interpretation and understanding of the results of this evaluation. For example, in the context of a project to design an intelligent cruise control system for urban and suburban contexts, the prior study meant we were in a position to guide the choice of journeys to be made by the drivers used in order to evaluate a prototype of the system, during which its ability to manage speeds and distances on different types of journeys with different types of traffic jam could be tested. To conclude this point, we would also point out that the knowledge produced from a comprehensive analysis of the activity can be used to not only identify the most relevant criteria for evaluating the system covered by the design project, but is also useful for evaluating families of systems, existing intermediate solutions and solutions under development.

References

[1] BERTHOZ A. (1997) - *Le sens du mouvement*, Odile Jacob, Paris.

[2] PINSKY L., THEUREAU J. (1987) - Description of visual "action" in natural situations, in O' Regan, J. K., Levy-Schoen A. eds., Eye mouvements: From physiology to cognition, *Selected/edited proceedings of the 3rd European conference on eye mouvements*, Dourdan (France), Sept, Elsevier, Amsterdam.

[3] PORT R.F., VAN GELDER T. (1995) – *Mind as motion*, MIT Press, Cambridge.

[4] SAAD, Th. VILLAME (1996) - Assessing new driving support systems : contribution of an analysis of drivers' activity in real situations. *Third annual world congress on Intelligent Transport Systems*, Orlando.

[5] SAAD, Th. VILLAME (1997) - Behavioural adaptations to a new driving support system. *Proceedings of the 13th triennial Congress of the International Ergonomics Association*, Tampere, Finland.

[6] THEUREAU J. (1992) – *Cours d'action : analyse sémio-logique*, Peter Lang, Berne.

[7] THEUREAU J. & FILIPPI G. (2000) - Course of action centred design, chapter 4, in Heath, Luff & Hindmarsh eds., Workplace studies, Cambridge Univ. Press.

[8] VARELA F. (1989) – *Connaître: Les Sciences cognitives*, Seuil, Paris.

[9] VILLAME T. (1997) - Assessment of the influence of an ACC system on driving activity : methodological issues for data collection and analysis. *Intelligent Vehicle Information System Safety Test and Evaluation Workshop*, IHRA-ITS, Berlin.

A SIMULATION-BASED STUDY ON NIGHT TRAIN OPERATOR'S TRACK INSPECTION PERFORMANCE BY USE OF COGNITIVE-PERCEPTUAL MODEL

Kenji Itoh
Tokyo Institute of Technology
e-mail: ken@ie.me.titech.ac.jp

Masaki Seki
Central Japan Railway Company

Takehiro Hoshino
Central Japan Railway Company

ABSTRACT

This study presents a simulation approach employing a cognitive-perceptual model to discuss a train operator's performance of obstacle detection on the rails of a Japanese high-speed railway. As its application, we specifically focused on a "Track Inspection Task" which is operated after completion of daily track maintenance operations with a special purpose tack maintenance train, i.e., so-called "Track Inspection Train". This model described a train operator's perceptual and cognitive process during track monitoring and inspection, and it allows us to simulate his behaviour as well as to track states during train operations. As input data to reproduce his visual function, the model required a train operator's attention allocation patterns and eye-gaze distribution for each location, both of which were obtained by task observation in reality, as well as experimental data of the obstacle detection rates with eccentricity of visual field and its size. In the present paper, a great number of simulation runs were carried out to estimate the operator's obstacle detection rates under various operating conditions, changing the size of an obstacle, its location on the track, driving speed, and so forth. Based on these simulation results, we discuss missing detection of obstacles that may cause a critical accident of the bullet train.

Keywords: track inspection, track maintenance train, eye tracking analysis, cognitive-perceptual model

INTRODUCTION

Human factors play a crucial role for railway safety as in other high-tech man-machine operations such as aviation, ship navigation and nuclear power plant control. It is well-known that human errors are the predominant cause of accidents and incidents in those areas (Margetts, 1976; Miller and Swain, 1987), and their rates are largely affected by an individual operator's factors such as work experience, skills, knowledge, and workload and fatigue as well as working conditions. Train operations share many of these characteristics.

There are many types of trains operated in the railway: high-speed trains, commuter trains, subway trains, freight trains, track maintenance trains, and so forth. In this study, we focus on the *track maintenance trains* operated to maintain tracks and rails for the high-speed train (Shinkansen) in Japan. Track maintenance trains are operated by a team consisting of a supervisor and a driver both working in a driving cockpit under more changeable and more stressful conditions in comparison with those of normal passenger trains. For example, these trains have no traffic signals available when they are operated in operation interval break during the night time after the last evening train and before the first bullet train on the following morning. Thus, this requires the train crew to make go/stop decisions based only on their own perception and judgement. In this study, we have a special concern with "*track inspection task*" since the track monitoring – one of the most important activities comprised of this task – exhibits train operators' activities common to all the types of track maintenance trains. The major purpose of this task is to ensure there are no obstacles on the track and the surrounding environment after completion of daily track maintenance constructions so that high-speed trains can be running safely on the rails. When the supervisor finds an obstacle during his monitoring process, he orders a driver to stop the train and picks it up from the track though there is actually very few occasions with something on the track. The choice of focus on this task was motivated not only by the above-mentioned difficulty of operating condition but also by the necessity of operation/task. This task is of critical importance since it serves as the final check for safety driving of bullet trains after daily track maintenance operations. In addition, the accident and incident rate of track maintenance trains is higher than that of passenger trains, though it is actually very low in terms of the absolute number and rate of accidents and incidents.

As can be easily understood from the above task description, the track inspection task is performed in skill-based manner (Rasmussen, 1986), and its quality and efficiency depend highly on the operator's visual perception and attention allocation on the track and its surroundings as well as on operational conditions. For tasks of man-machine operations having the above-mentioned characteristics, the eye-tracking technique has been employed in investigating operators' cognitive and perceptual performances with various ergonomics purposes, e.g., interface design (Itoh, 1998; Itoh et al., 2001b). Furthermore, operators' cognitive processes elicited by the eye-tracking analysis were modelled so that computer simulations could be performed for the purpose of safety assessment (e.g., Itoh et al, 1998; Itoh et al., 2001a). A cognitive simulation approach like in these studies may have suggested its potential abilities for large-scale risk analysis of man-machine operations due to its advantage of cost and time savings.

In the present paper, we apply the cognitive simulation approach to risk analysis for track inspection task, i.e., estimating risks of missing obstacles left on the tracks and its surroundings. A cognitive-perceptual model for the track inspection task was constructed based on the task analysis employing the eye-tracking analysis (Itoh et al., 2000). As for the input data to the model, we obtained a database of their visual perception functions through an experiment with ten train operators as subjects. To uncover derailing risks of bullet trains by an obstacle on the track, its detection rates were estimated by a number of simulation runs of the cognitive-perceptual model under various operating conditions, changing driving speed of track maintenance train, size of obstacle, its position on the track and so forth. Based on these results, we discuss risk factors for missing obstacles in the track inspection task.

OPERATOR'S PROCESS FOR TRACK INSPECTION

Observation of Track Inspection Task

We observed four track inspection sessions, each of which was carried out in the same driving area on a different day using a different scenario. In each session, a supervisor performed the track inspection task while the train was running approximately 40 km distance on the outbound track and the same distance on the return track at 3:00-5:20 am. Each scenario included a supervisor with or without geographic knowledge on the inspection area, eye-tracking recording time either on outbound or on return trip, and so forth. The supervisor's eye-tracking data was recorded for approximately one hour on either one way using an ASL 4000 eye-tracking system. In Session IV, in particular, four boxes (15x25x8.5cm) were put at different locations dispersed throughout the inspection area to observe the supervisor's actual monitoring process in detecting an obstacle.

Attention Allocation of Track Monitoring

A general pattern of a supervisor's track monitoring process is generated, aggregating eye-movement data from all three sessions involving the supervisor having the geographic knowledge (Itoh et al., 2000). Figure 1 indicates the transition network of the supervisor's attention allocation during the track inspection task. In this figure, the size of each circle represents the percentage of total gaze time at each location on the track and surrounding environment. The bigger a circle is, the more attention was paid to that location during the task. The thickness of the arrowed arc between two locations indicates the frequency of transition in terms of the relative percentage over the total number of eye-movements. The thickest lines in this figure, i.e., lines between the distant position on the running (own) track and the side of the running track, and between the distant position and the parallel track, represent 5-10% of transition over all the attention shift. The thinnest lines mean 0.5-1.0% of transition. No line is provided for a smaller percentage of transition than 0.5% in the network.

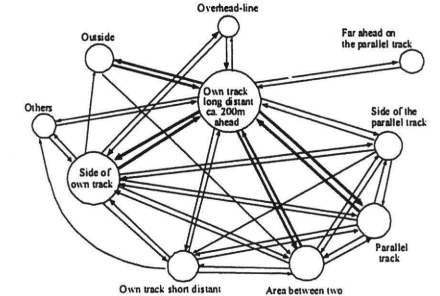


Figure 1 Transition pattern of train operator's attention during track monitoring

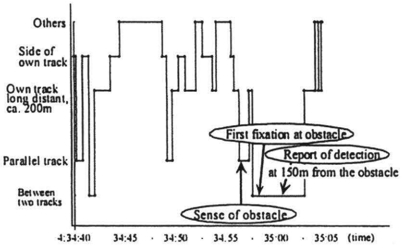


Figure 2 Operator's eye-movement in detecting an obstacle

The most frequently attended location was a distant region on the running track (approximately 200 metres ahead), on which a head light of the train focused. The supervisor gazed at this location for about 30% of time during the task. From this fact, this location seems to be the *home-base* of his attention for track monitoring. Besides this location, he also frequently looked at the side of the running track for about 20% of total duration. In this location, markers indicating distance from Tokyo and a telephone box were placed at regular intervals and shinning spots sometimes appeared due to the reflection of the train's light. Therefore, it is natural to deduce that his fixation shifted automatically to this location, rather than that he paid critical attention to this region. His attention was also shifted frequently to the area between the two tracks (9.3%), the parallel track (8.9%) and a near-point (approximately 50 metres ahead) on his own running track (7.1%). Attention in the direction of the other parallel track was relatively infrequent, totalling less than 20%. Regarding gaze times at specific locations on the track - no figures are provided due to space limit -, their distributions were highly skewed and biased towards shorter duration. The mean gaze time was approximately one second for every location, and its standard deviation was about one second.

From these results, the supervisor's track inspection process can be conjectured to be a home-base monitoring strategy for inspecting the states on the tracks and surroundings. In this strategy, his home-base of attention is the distant region of the present running track approximately 200 metres ahead. He monitors the states on his tracks around his home-base by a short look-at each gazed location (taking less than a few seconds). That is, his attention shifts from the home-base to another location such as the side of his own track, the parallel track, etc. and then returns to the home-base of monitoring.

Obstacle Detection Process

Using eye-tracking data recorded in Session IV, the supervisor's actual detection process of an obstacle was analysed (Itoh et al., 2000). Transition of fixation points is shown in Figure 2 for one of the four boxes left on the track. This was the case where a box was put at the most difficult place to detect, i.e., on the centre-pathway between two tracks (one metre below the track level).

Up until about 10 seconds before his report of the obstacle detection, the supervisor was inspecting the states in his track and its surroundings in a usual home-base monitoring as shown in Figure 1. About 5 seconds before the detection, he obtained a sense of something left on the centre-pathway when he attended to the parallel track. He later described this situation as follows: "When gazing the parallel track he got a visual image of its surrounding in his peripheral vision. He performed a pattern matching of his acquired visual image with the visual template stored in his memory as geographic knowledge. As a result of the pattern matching, he supposed there be a discrepancy between his visual acquisition and his geographic knowledge". Thus, at this moment, he might catch a hypothetical visual image of obstacle in his parafoveal or peripheral vision, i.e., a few to ten degrees in eccentricity. Then, he shifted his attention to the location where he sensed the discrepancy, i.e., on the centre-pathway, to test his hypothesis concerning the obstacle. After fixating properly at the box, he reported its detection approximately 150 metres before the obstacle. The other three boxes examined in Session IV were also detected by the same process.

COGNITIVE-PERCEPTUAL MODEL

Modelled Process of Obstacle Detection

In this section, we build a model of track maintenance train operator's cognitive-perceptual process in detecting an obstacle mentioned in the last section. The obstacle detection is modelled straightforward as a relation between clearness of its visual image and his visual acuity. The image clearness depends on the location and size of an obstacle, and viewing distance from the operator. Thus, it can be defined by a combination of visual angle of the obstacle and the place of its visual image on retina, i.e., the eccentricity from his fixation point. Regarding the human visual acuity, it is well-known to be reduced rapidly with the eccentricity from fovea, and therefore it can be represented as a function of these two variables. When the visual acuity clear enough to percept a particular obstacle, the operator can detect it. According to this activity modelling of obstacle detection, the image clearness of a particular obstacle is changing time to time because of changes in viewing distance and eccentricity of the obstacle's image from the fixation point by his eye-movement while the train is running.

The process of obstacle detection is modelled stochastically using random digits to reproduce the operator's attention allocation during track monitoring and to determine whether he can detect an obstacle and where he does. The train operator's fixation point during track monitoring is generated from time to time by a random digit following the attention allocation pattern. Fixation duration at any particular location is also determined based on its distribution. Process of his "hypothesis generation" on an obstacle is modelled to be initiated according to the clearness of its visual image, depending on its size, viewing distance from the operator and his fixation point, i.e., its visual angle and eccentricity from fovea. A clearness index can be calculated based on these variables as well as his visual acuity at the eccentricity from the fovea. A index value exceeding its lower limit activates to generate a hypothesis on an obstacle. The hypothetical obstacle is tested by comparing the index value with a random number uniquely distributed ranging between 0 and 1. If the random number is less than the clearness index, the hypothesis is supported and thus the operator's attention shifts to the obstacle. Then, the visual clearness is tested again while he gazes the obstacle in his foveal vision. If its index value exceeds a random number selected at the new fixation point, the model decides to detect the obstacle. In case of selecting a larger random number than the clearness index in the hypothesis testing of the obstacle, the next eye fixation is generated following the attention allocation pattern, and the operator continues to perform usual track monitoring process.

Databases of Visual Detection

As mentioned above, the cognitive-perceptual model requires the train operator's attention allocation pattern during track monitoring and distribution of gaze duration for each location in order to simulate the operator's behaviour during track inspection task. The both data input to the model were obtained in the above-mentioned experimental sessions by use of eye-tracking recordings (Itoh et al., 2000). Particularly, the operator's attention allocation pattern was shown in Figure 1.

In addition to these data, a database on the operator's visual performance, i.e., detection rates for different-sized obstacles with several levels of eccentricity of visual field are required to input to the model. In order to obtain these data, an experiment was carried out with ten track maintenance train operators as subjects. In the experiment, each subject identified objects from driving cockpit of a track maintenance train in the dark at an identical luminance level to the actual driving condition on the track. All the visual objects were rectangle shaped and covered with fluorescent paint on their surfaces similar to that attached on tools used for usual track maintenance operations. Four different sized objects were presented at five different locations 100 metres ahead from the subject, i.e., 0, 2, 4, 6 and 8 degrees in eccentricity from the fixation point, randomly in order with five repetitions. Each subject was asked to report whether he could identify an object or not, while he fixated at a mark attached on another track maintenance train parked also about 100 metres straight ahead.

As a result of the experiment, detection rates are shown in Figure 3 for each combination of object size and its location. As can be seen in this figure, the detection rate was reduced with increase of eccentricity and with decrease of its size, but not linearly for the both factors, although there were also individual differences in the detection rate. To generate input data to the model, the detection rate was approximated as a function of the eccentricity from the fixation point and the visual angle of an object (in degrees), as shown in Figure 4 for several example conditions. A detection rate for any condition of size and location was obtained by this function in a simulation run.

SIMULATED OBSTACLE DETECTION

Simulation studies were conducted, applying the cognitive-perceptual model presented in this paper to various operational scenarios, changing driving speed, size of obstacle and its location on the track, and so forth. One

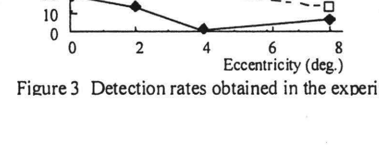


Figure 3 Detection rates obtained in the experiment

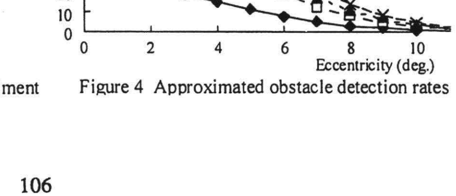


Figure 4 Approximated obstacle detection rates

thousand simulation runs were performed using a different seed of random digit for each scenario. An obstacle detection rate for each operational condition was derived from simulated performance obtained by these 1,000 trials. As an example result of simulation studies, Figure 5 indicates a distribution of train's position, i.e., distance between the train and the obstacle, at the moment of obstacle detection as well as a missing rate of detection for different sized obstacles. In the case shown in this figure, the obstacle was left in the centre of the running (own) rails while the train is running at the maximum speed limit, i.e., 60 km/h. As can be seen in this figure, the detection rate for a not-very-small obstacle put at the operator's home-base of attention, i.e., in the centre of the own rails is estimated 100% in driving at the speed of 60 km/h. For the smaller obstacles examined in this study, there were not many cases of missing detection, i.e., 21% and 3.5% for 5x5x5 and 10x10x10 cm boxes, respectively. However, the size of obstacle is found to critically affect the detection timing when the operator can detect it even if it is left at his home-base of attention. For example, the operator was expected to detect a 40x40x40 cm obstacle more than 200 metres ahead in about 90% of cases. This means that the train can stop enough before the obstacle anytime. In contrast, for the smallest, 5x5x5 cm box, the simulated operator detected an obstacle at the distance shorter than 50 metres in all the detecting cases, and less than 100 metres in about 90% cases for a 10x10x10 cm box. In these cases, it is impossible for the operator to stop the train before the obstacle.

Changes in the detection rate with running speed are depicted in Figure 6 for several sized obstacles which are left in the centre of the own rails, assuming the operator's attention allocation pattern is not changed with the speed. The smaller an obstacle left on the track, the larger the effect of train speed on the detection rate becomes. This figure also indicates that if an obstacle is big enough, i.e., no smaller than 20x20x20 cm, the detection rate is not decreased independent of speed increase up until 120 km/h. This means that there is no or little critical risk of derailing of bullet trains due to an obstacle on the track. That is, there exists no big obstacle close to the running rails after the track inspection task is completed, even if the speed limit of the "track inspection train" is increased to 100 or 120 km/h.

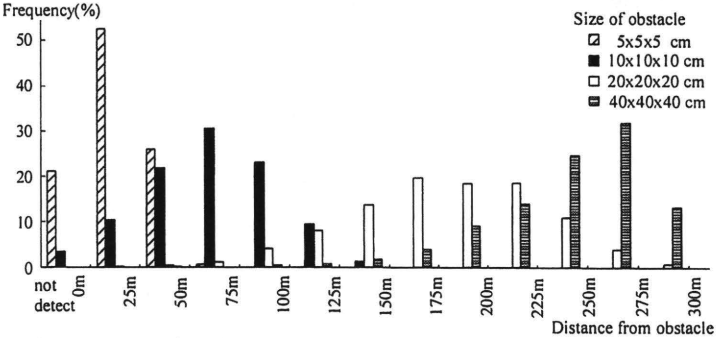


Figure 5 Distance from the train when detecting an obstacle (running speed: 60km/h)

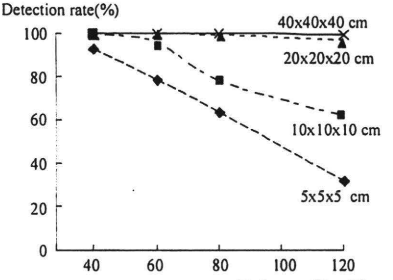


Figure 6 Detection rates at different running speed

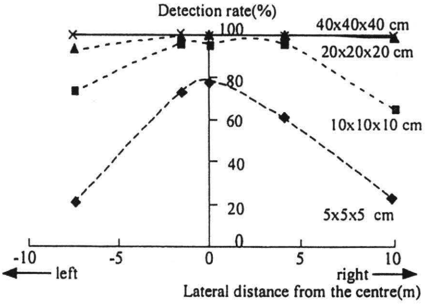


Figure 7 Detection rates of obstacles in various locations

To examine the effect of obstacle's location on the operator's inspection performance, Figure 7 illustrates detection rates for different sized obstacles which are left at different positions ranging from the side of the own track, i.e., 7.5 metres left from the centre of the own rails, to the parallel track, i.e., 10 metres right, when running at the speed of 60 km/h. For larger obstacles, i.e., no smaller than 20x20x20 cm box, the detection rate is not reduced greatly even for ones far from the own running rails. In contrast, for smaller obstacles, the detection rate is affected much more by its location on the track. For example, the simulated operator's performance indicated to detect the smallest obstacle (5x5x5 cm) left in the centre of the own rails in about 80% cases while the detection rate was estimated to be decreased to about 20 % for the same sized box left on the parallel track.

CONCLUSION

In this paper, we modelled the train operator's track monitoring and obstacle detection processes based on the task analysis applying the eye-tracking data (Itoh et al., 2000) as well as on experimental data collection of their visual functions. This model allows us to simulate the operator's perceptual behaviour during track inspection task while the track inspection train is running. Applying this model, obstacle detection rates were estimated under specific operational conditions to discuss derailing risks of the bullet train due to missing detection of an obstacle on the track. Major findings in this simulation study are as follows: There is no or very little derailing risk of the bullet train by an obstacle under the present operational condition of track inspection task, e.g., driving speed, etc., unless the vigilance level of track maintenance operator goes down greatly. The operator can detect a big obstacle, which may cause derailing of a train, at the position distant enough to stop the train before it. For an object small enough not to interfere with driving of the bullet train, 20% or higher probability of missing detection was suggested to exist, depending on its location on the track.

As we did not mention in this paper, we conducted additional simulation studies applying different databases of operator's visual functions, which were also obtained in the former study (Itoh et al., 2000). For example, the attention allocation pattern of an operator having *no* geographic knowledge of the inspection area was found to be different from that of the knowledge-holding operator who was focused on in this paper. Thus, we performed a series of simulation runs by applying the same cognitive-perceptual model, only switching the database of attention allocation pattern to the operator's having no geographic knowledge - although his cognitive process for obstacle detection might also be somewhat different from that of the knowledge-holding operator's. As a result of these simulations, the detection rate of the knowledge-less operator was lower for obstacles left far from the own running rails than that of the knowledge-holding operator. With increase of the running speed, the knowledge-less operator's detection rate was decreased more rapidly than that of the usual operator. Regarding an effect of running speed, it was found to have little difference in attention allocation between 40km/h and 60km/h in the former study (Itoh, et al., 2000), and therefore we used the same database which was obtained in sessions running at the speed from 40 to 60 km/h throughout all the simulation runs independent of the driving speed ranging from 40 to 120 km/h. However, it may be reasonable to exist a difference in attention allocation pattern between 60km/h and 120km/h or faster.

In a future study, further data collection is required to obtain more precise estimations of obstacle detection rates corresponding to specific operational conditions. In addition, the model should be improved in that appropriate cognitive process is built in according to individual human factors and operational conditions, as mentioned above. We believe it is possible to apply a simulation approach with a cognitive-perceptual model presented in this study to risk analysis of any type of man-machine operations which are routinely performed in normal situations. This is made possible by developing an appropriate model of a task under study based on the descriptive approach of cognitive task analysis (Vicente, 1999).

ACKNOWLEDGMENTS

We would like to acknowledge Kunihiro Kondo, Masahiro Kawagoe, and Minoru Hiranaga of Central Japan Railway Company, as well as Toshitatsu Ishii and Jushiro Takahashi of Nihon Kikai Hosen, Co. Ltd, and Akikazu Umoto and Norimasa Nogawa of Futaba Tetsudo Kogyo, Co. Ltd. for their cooperation in this project. We also thank Hirotaka Aoki, Shinji Akiyama and Takashi Naono, Tokyo Institute of Technology, for their assistance with eye-movement recording and its data analysis. John Paulin Hansen, the IT University of Copenhagen, and Gunnar Hauland, Risoe National Laboratory provided useful ideas to the eye-tracking data recording and analysis.

REFERENCES

Itoh, K. (1998). A Cognitive Modelling Scheme for Task Analysis and Its Application to Computer Editing Task, *International Journal of Cognitive Ergonomics*, 2(3), 269-295.

Itoh, K., Arimoto, M. and Akachi, Y. (2001b). Eye-tracking applications to design of new train interface for the Japanese high-speed railway, *Proceedings of the 4th International Conference on Engineering Psychology and Cognitive Ergonomics*, New Orleans, August.

Itoh, K., Hansen, J.P. and Nielsen, F.R. (1998). Cognitive modelling of a ship navigator based on protocols and eye-movement analysis, *Le Travail Humain*, 61(2), 99-127.

Itoh, K., Tanaka, H. and Seki, M. (2000). Eye-movement analysis of track monitoring patterns of night train operators: Effects of geographic knowledge and fatigue, *Proceedings of the 14th Triennial Congress of the International Ergonomics Association, IEA 2000*,Vol.4, pp.360-363, San Diego, July-August.

Itoh, K., Yamaguchi, T., Hansen, J.P. and Nielsen, F.R. (2001a). Risk analysis of ship navigation by use of cognitive simulation, Cognition, Technology & Work, 3, 4-21.

Margetts, B.D. (1976). Human error in merchant marine safety, The National Research Council, Washington DC.

Miller, D.P. and Swain, A.D. (1987). Human error and human reliability, In G. Salvendy (Ed.), *Handbook of human factors* (pp.219—250). New York: John Wiley & Sons.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*, North-Holland, . New York.

Vicente, K.J. (1999). *Cognitive work analysis: Toward safe, productive and healthy computer-based work.* Lawrence Erlbaum Associates, Mahwah, NJ.

Postersession I

Detailed Sub micro-Traffic Simulation as an Aid for Testing Fully and Partly Automated Mobile Systems

Marita Irmscher *, Thomas Jürgensohn **
* Ford Werke AG, Spessartstr. D-MC/Pz-PEZ 2, Germany, mirmsche@ford.com
** MIT-E40-291, Cambridge MA 02139, USA, thomasj@mit.edu

Abstract

In order to augment safety and comfort for the driver more and more driver information and support systems are being developed and applied in modern vehicles. This means in many cases that parts of the driving task are taken over by some kind of automatic system. However, this new quality of human-machine interaction makes it necessary to consider more than before cognitive and individual characteristics in driver behavior. One reason lies in the fact that every driver is different and that for optimum assistance a system should be able to adapt to the needs and wishes of any particular driver at any time.

Another reason is the necessity to predict the effects of the employment of driver assistance on overall traffic safety and efficiency. For this purpose, and of course to study design variants of assistant systems, traffic simulations can be conducted. To render results more life-like and investigate different situations as easily as possible it is desirable to include different driver types.

In this paper a driver model with individual characteristics is presented that can be used for simulations of different vehicles and vehicle equipment in a typical traffic situation. The modelled driver behaviour depends both on situational knowledge and internal dispositions described by a number of motives that account for individual differences.

1 Individual Driver Behaviour

Drivers are a very heterogeneous group compared to other machine operators. This is due to the various purposes of driving which range from professional occupation to leisure and hobby. Therefore a very strong emotional component in driving besides the primary transport purpose can be observed that influences the traffic behaviour of drivers. Determining factors of driving behaviour are thus not only traffic rules and situational requirements (external factors), but also the momentary mood and the purpose of driving (internal factors). Other influences like sex, age, driving experience and skill are personal characteristics of the driver. To describe these influences two major questions arise:

- Which are the main factors that determine the driving behaviour of a person in a given traffic situation?
- In which behavioural traits do individual differences show up?

As the revelation of these questions is an important issue in the investigation of accident causes, both experimental studies and theoretic modelling have been put forward in traffic psychology. In this context a close correlation to motivation psychology can be seen. Motivation psychology focuses on causes of human action, relations between motives and action as well as interpersonal behaviour differences in general. Because findings of both disciplines about influences on driver behaviour will be applied to setting up a simulation model, some fundamentals are presented in the following.

2 Influencing Factors on Driver Behaviour

Within the frame of studies on traffic safety conducted in the seventies, several factors that describe individual differences in driver behaviour were identified (Berger et al., 1974; Utzelmann, 1977). Näätanen & Summala (1976) introduced the terminus *extra motives* for behavioural influences besides the primary motives of transport and survival. The extra motives can be integrated in a list of fundamental causes of human action as set up in motivation psychology – so called elementary motives – (Murray, 1938; McClelland, 1954), as shown in the following table 1.

Table 1 influencing factors on the driver

Extra motives (Näätanen & Summala, 1976)	Influencing factors on driving behavior (Utzelmann, 1977)	Elementary motives (Murray, 1938)
Risk taking	Challenge of driving close to personal and physical limits	Achievement
	Safety und relaxing by calm driving (glide)	
Show off	Rivalry	Power
Driving fun	Joy of expert driving (piloting)	Play, independence
Exhibition	Demonstration of power	Exhibition

This table can be supplemented by further factors like hurry, fear, frustration, aggression and comfort, which are also found within the elementary motives. Because of this coincidence between influencing factors on driving behaviour and motives it makes sense to analyse and describe effects of individual dispositions on driver actions in a situation on the basis of motivation psychology.

3 Manifestation of Motivation Influences on Driver Behaviour

Driving behaviour means goal oriented acting in a traffic situation and requires performing at least the following subtasks:

- Prescribe and control the goal (e.g. desired speed),
- Observe and perceive environment and traffic
- Recognition and interpretation of a situation,
- Decision between several action alternatives with respect to the goal,
- Motoric execution of the action (actuate steering wheel, gas or brake pedal),
- Communicate with other traffic participants.

These subtasks can be classified into *perception, mental information processing and motor action*. According to Huguenin (1988) *attitudes* (which can be set equal to motives in this context) act on information processing, on decision criteria in a situation as well as the sensomotory driving, all of which can be correlated with the above subtasks. On the other hand, taking the definition of motivation as a "force that *generates* goals and *initiates and moderates* behaviour" (Madsen, 1974) a meaningful structuring of the action would be that into *goal generation*, action *initialisation* (e.g. decision processes that result in an action) and action *moderation* (way and intensity of performance). Motivation influences can then be described using knowledge about each action element found in driving behaviour.

In order to model motivated driver behaviour for a computer simulation, quantitative relations between motivation and behaviour have to be formulated. Dörner, who follows an approach that is both pragmatic and based on fundamental psychological theories, is tackling this very difficult issue.

3.1 Motivated Action – the PSI-Model of Action Regulation

Schaub (1993) and Dörner (1998) developed a theory describing an autonomous system that exchanges material and information with its environment and has to adapt to changing external conditions. Using perceived external and internal states, goals and intentions matching the situations are generated. For the selection of a realized intention criteria like the following are used.

- Motive strength = function of deviation and time of deviation
- Selection pressure = function of competence, motive strength, urgency and time pressure

However, the modeller is confronted with the problem to find meaningful relations (that is, mathematical functions) between motivation and action. One possible solution for this difficulty is the application of learning methods like artificial neural networks. Fuzzy Methods are another possibility with the description of system behaviour using natural language to formulate rules. Compared to artificial neural networks Fuzzy Methods have the advantage that a priori knowledge about the system or the process can be used when setting up the model. This approach has been chosen to model driver behaviour in a traffic situation and will be explicated in the following.

4 Fuzzy–Modelling

Fuzzy Methods were employed in artificial intelligence mainly because of their inherent capability to describe vagueness, imprecision, and uncertainty of human knowledge and action. Both the decision making of a driver in a traffic situation and the motoric execution of the driving task can be described by means of fuzzy rule bases (Fuzzy Decision Making and Fuzzy Control).

Fuzzy Control and Fuzzy Decision Making can be depicted by a common scheme, as shown in figure 1.

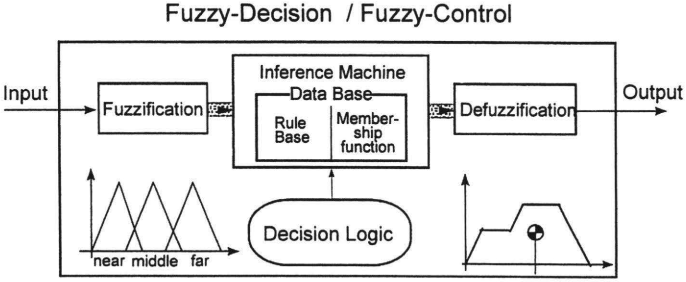

Fuzzy-Decision / Fuzzy-Control

Figure 1: Flow scheme of Fuzzy Control and Fuzzy Decision Making

The first step is the fuzzification of crisp input values by mapping them to terms of linguistic variables using membership functions. For example terms like "far", "near", "very near" can be introduced to describe the linguistic variable *vehicle distance*. The membership functions that are prescribed by the modeller to describe to which extent an input value correspond to a linguistic variable.

Linguistic variables for input and output are then correlated by means of a rule base, which is made up of rules like the following

IF *vehicle distance* very near **AND** *relative velocity* positive large **THEN** brake very large

Each rule is evaluated in the inference machine (decision logic), then the rules are aggregated to get an overall output value. In the case of fuzzy decision making each action alternative is described by means of attributes, which are expressed as linguistic variables with corresponding benefit functions assigned to them. The evaluation of the rules gives an overall benefit for each action alternative. The alternatives are then ordered according to benefit to determine a choice.

Now the findings of motivation theory and traffic psychology are integrated into a fuzzy model of driving behaviour in the following manner.

5 A model of individual driving behaviour

The driver model is established as part of a microscopic traffic simulation, where the model moves and behaves autonomously together with other traffic participants within a virtual simulation environment Traffic situations arise from the behaviour of the observed driver-vehicle–units and are defined mainly by mutual distances, relative velocities and the interpretation of the observed constellation by the drivers. The investigated situation is a typical overtaking constellation on a highway, because individual influences become very obvious here and can be studied in detail.

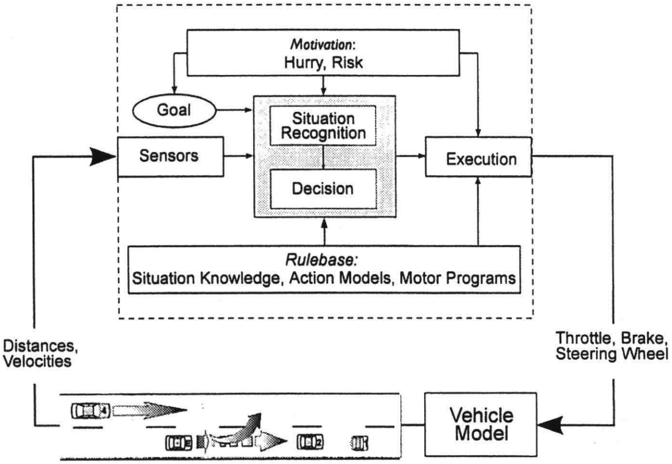


Figure 2: Model structure of a microscopic traffic simulation with a model of individual behaviour

Driver behaviour is structured into three action components *sensoric interface*, *cognition* (situation recognition and decision making) and *motor action*. This is a scheme also seen in cognitive modelling and traffic psychology. Motivation effects the goal setting, cognitive processes and motor action. Two influencing factors, hurry and risk were considered. Of course this does not cover the variety of individual differences in driving, but very clearly differences in driving styles can be shown this way.

Since driving is a highly trained activity, knowledge can be described in form of rules and quasi-automated action sequences, so called motor programs.

Sensory of drivers is modelled through the fuzzification of the calculated vehicle data, thus accounting for imprecise human perception.

The overtake situation shown is structured into two hierarchical decision problems, which are described by fuzzy decision-making. The subsequent motoric execution of the selected action alternative is modelled through generalized motor programs (Schmidt, 1986) and fuzzy control. The decision process steered by the current situation and driver motivation is shown in the following figure 3.

The first decision is whether or not to overtake. Next the start time for the overtake manoeuvre, which is set equal to the lane change, has to be chosen. As long as there is no overtaking, lane keeping is modelled by means of fuzzy control. As long as the driver decides not to overtake but to follow the vehicle in front of the own vehicle, velocity and distance are controlled using gas and brake pedals, modelled as a decision between gas and brake and a control process.

The main decision to overtake can be interpreted as *direction* of the action, correspondingly the lane change decision as action *initiation* and the execution of the lane change or car following as action *moderation*. According to the definition of motivation these are exactly the components of action, which are subject to motivation. In the proposed model individuality is therefore considered by means of motives or motivation factors for hurry and risk in direction and initiation of the action *overtaking*, as described in the following section.

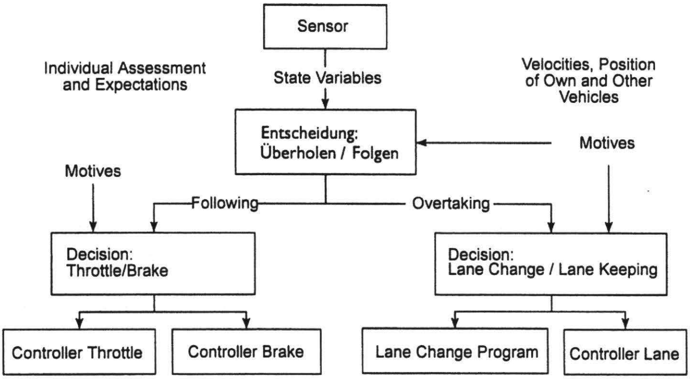


Figure 3: Overtaking as a multilevel decision process

5.1 The decision process

Decisions are characterized by the existence of various action alternatives in a given context. Each alternative is distinguished from the others by a set of attributes. The decision is made by evaluation of the attributes with respect to the benefit for the actor. The decisions in the context of overtaking as shown in figure 3 are assigned the following attributes:

Table 2: action alternatives and attributes of the fuzzy driver model

Overtake/ Follow	Lane change / Lane keeping	Braking /Accelerating
TF: maximally tolerated following time	**AV/AH**: vehicle ahead/behind swerves out	**AG**: danger of collision to vehicle ahead
FVL: approaching a slower vehicle ahead	**SV/SH**: safety gap ahead /behind	**SV**: safety gap ahead
FVR: approaching a slower vehicle on right lane	**FHS**: approach of a vehicle on target lane	**G**: own velocity
GÜ: accepted over speed while overtaking	**FVL**: vehicle exists on right lane in front	**A**: own acceleration
SPR: position of own car (right/left lane)	**SPR**: position of own vehicle (right lane)	
FzgV: vehicle ahead exists	**QS**: lateral stability	
	LH: flashing of vehicle F4	
	BL: front vehicle has set blinker	

For each attribute of action alternatives, a benefit function is defined and described by fuzzy sets for the process of fuzzy decision making as shown in figure 4.

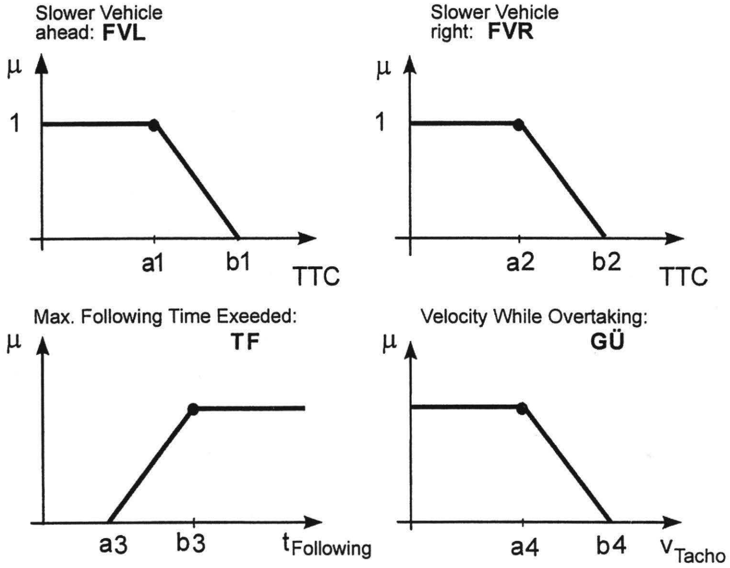


Figure 4: benefit functions of lane change attributes

The benefit functions used can be interpreted in the following manner:

The condition *approaching a slower vehicle ahead* (FVL) is truer with decreasing TTC (Time to Collision) to the vehicle ahead. The benefit to overtake such a vehicle is large. Beginning with a certain value of TTC this benefit begins to decline and becomes zero if there is no more perceivable velocity difference (this is expressed by a large value b1 for the TTC).

There is only a benefit to overtake after a tolerated following time a3. The benefit increases with longer following times until a maximum value is reached at the maximum tolerated following time b3. Then the benefit remains on this level.

During overtaking a higher speed as the desired speed (or the legal speed limits) is accepted for a short time to shorten the overtaking (GÜ). In order to avoid overtaking with very high speed or very low speed difference, the benefit of very high overtaking speed is set to zero.

To model individual differences in the decision making for overtake and lane change, the parameters a_i and b_i of the benefit functions are prescribed dependent on motivation factors. Accepted distances in space and time to a vehicle ahead were found in Rekersbrink (1994). For the tolerated following time ***TF*** and the accepted overtaking speed ***GÜ*** parameters were derived from highway driving experiments.

The benefits of the attributes are combined to the benefit of the action alternative in a given situation by means of linguistic rules like the following:

IF my car is on the right lane
AND speed of car ahead is less than my desired speed
AND tolerated following time has passed
THEN overtake.

This rule can be formalized by correlating the attributes of the action alternative by means of an operator γ

$$SPR \, \gamma \, FVL \, \gamma \, TF \rightarrow \ddot{U}$$

The γ-operator is parametric and can be adjusted to a value between the optimistic (risky) OR-correlation and the pessimistic (defensive) AND-correlation. The parameter is chosen according the prescribed risk behaviour (Truskawa, 1999).

The defuzzification of the obtained benefit function of an action alternative makes use of its maximum. If there is a distinct value above a threshold set to 0.5, the corresponding action is chosen. In case that no alternative achieves the required benefit value, a default action is executed. This is following for a vehicle on the right lane, and lane change to the right for a vehicle in the left lane.

After the decision for a possible action alternative has been made the action is executed. This includes both lateral and longitudinal control of the vehicle, hence operating gas or brake pedal and the steering wheel.

5.2 Simulation results

To demonstrate individual differences in the traffic behaviour of drivers the development of the described traffic situation in time was simulated with varied prescriptions of hurry and risk behaviour of a driver. The results are shown as a number of "snapshots" marking some interesting simulation times (figure 5)

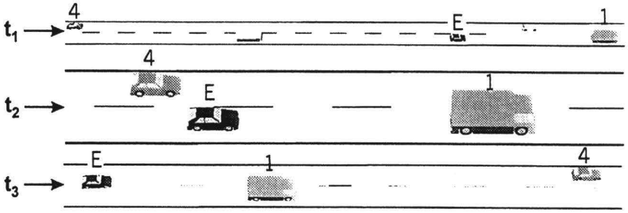


Figure 5: Simulated traffic constellation for a typical overtaking situation with a driver (E) who is not hurried and defensive.

At time t1 driver E is driving behind a slower vehicle 1. After a certain following time the tolerance limit is reached, but at this time the fast driving vehicle 4 has approached on the left lane (t2). E slows down, lets vehicle 4 pass and starts overtaking afterwards (t3).

With the same start constellation but a different setting for E's risk behaviour the situation develops differently (fig. 6). Because of the higher desired speed due to more risk taking E decides to overtake at time t2 and continues driving with his desired speed. As a consequence vehicle 4 (also a risk taking driver!) approaches to a very short distance (t3 and t4). E does not accelerate and changes back to the right lane after the overtaking manoeuvre (t5).

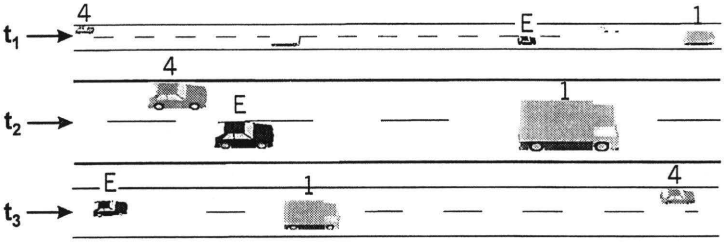


Figure 6: Simulated traffic constellation for a typical overtaking situation with a driver (E) who is not hurried (hurry=0) but risk taking (risk=1).

In the last example a hurried and risk taking driver was considered (fig. 7). Until time t4 there is no difference to the preceding example. But at time t5 the hurried driver decides to accelerate on the left lane and continue the overtaking.

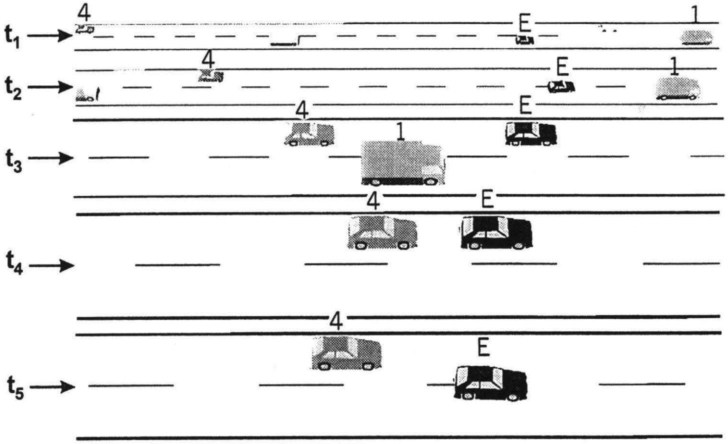


Figure 7: Simulated traffic constellation for a typical overtaking situation with a driver (E) who is hurried (hurry=1) and risk taking (risk=1).

These results show that for only two varied influencing factors a traffic situation can develop quite differently due to changed driver behaviour. The results are plausible. The chosen approach is therefore a very economical possibility to model individuality in traffic simulations by means of parameters that describe driver characteristics.

7 Summary

Individuality in driver behaviour is formalized by means of two influencing factors (motives): hurry and risk taking. An individual acting driver model was implemented in a microscopic traffic simulation. Possible applications are simulations in the field of accident investigations or the studies of driver assistant systems in traffic situations, where individuality of drivers is of particular importance. Another application in this context would be the definition of "intelligent" traffic for design studies. Simulated traffic constellation for a typical overtaking situation with a driver (E) who is not hurried but risk taking.

8 Literature

Berger, H.-J.; Bliersbach, G.; Dellen, R. G.; Kroj, G.; Pfafferott, I. (1974). Psychological research of the safety belt (Psychologische Forschung zum Sicherheitsgurt und Umsetzung ihrer Ergebnisse). Schriftenreihe Unfall- und Sicherheitsforschung im Straßenverkehr, Berichte der Bundesanstalt für Straßenwesen, Heft 2

Dörner, D. (1998). A blueprint of the soul (Ein Bauplan für eine Seele). Reinbek: Rowohlt

Huguenin, R.D. (1988). Driver behaviour in traffic (Fahrerverhalten im Straßenverkehr). Braunschweig: Rot-Gelb-Grün

Irmscher, M. (2001). Modelling and simulation of motivational influences on driver behaviour (Modellierung und Simulation von Motivationseinflüssen auf das Fahrerverhalten). VDI Fortschrittberichte, ZMMS Spektrum Bd. 12. Düsseldorf: VDI-Verlag

Jürgensohn, T. (1997). Hybrid Driver Models (Hybride Fahrermodelle). ZMMS-Spektrum Bd. 5, Sinzheim: pro universitate

Jürgensohn, T.; Jung, R.; Willumeit, H.-P. (1997). The handwriting of steering. ATZ Automobiltechnische Zeitschrift 99, Nr. 4, p. 216-219

Madsen, K.B. (1974). Modern Theories of Motivation. Kopenhagen: Munksgaard

McClelland, D.C. (1985). Human motivation. London: Scott, Foresman & Co.

Murray, H.A. (1938). Explorations in personality. New York: Oxford University Press

Näätanen, R.; Summala, H. (1976). Road user behavior and traffic accidents. Amsterdam: North Holland Publishing Company

Rekersbrink, A. (1994). Traffic flow simulation with the aid of fuzzy-logic and a concept of possible collision times. Dissertation, Universität Karlsruhe

Schaub, H. (1993). Modelling of action organisation. Bern: Huber

Schmidt, R.A. (1985). The search for invariance in skilled movement behavior. Research Quarterly for Exercise and Sport, 56, p. 188-200.

Truskawa, S. (1999). Entwicklung eines Modells zur Nachbildung des Fahrerverhaltens mit Hilfe von Fuzzy Decision Making. TU Berlin, Diplomarbeit am Institut für Straßen- und Schienenverkehr

Utzelmann, H.D. (1977). Features of driver behaviour and their connections to motivational conditions (Merkmale des Fahrerverhaltens und ihre Zuordnung zu motivationalen Bedingungen). Symposium über Sicherheit im Straßenverkehr. p.127-136.

Factors influencing recovery from failures

Lisette Kanse* & Tjerk van der Schaaf
Eindhoven University of Technology, Department of Technology Management,
Pav. U42a, PO Box 513, 5600 MB Eindhoven, The Netherlands.
tel. +31 40 247 2441; e-mail: L.Kanse@tm.tue.nl

Summary

The prevention of errors and other failures has always been a central theme in safety and reliability management. However, additional benefits could be gained from focussing on what can be done after a failure has occurred, but before this leads to negative consequences. This contribution examines the factors that influence the processes people follow to recover from failures. Six human performance and reliability experts rated a list of recovery influencing factors with regard to main category, recovery process phase influenced, and (less successful) strength of influence. Based on agreement found between the raters we have attempted to structure the list. We found that a simple taxonomy does not do justice to the complex nature of the influencing factors. Therefore, our future efforts will concentrate on the development of a generic influence diagram. A field study is planned to provide insight in the differences in importance of each of the recovery influencing factors.

Introduction

Traditionally, the focus of safety and reliability management has mainly been on the prevention of errors (i.e. human failures) and other failures. However, it is worthwhile to also consider the possibilities for recovery after a failure has occurred; in other words to focus on how the negative consequences of the failure can be avoided. After all, what one really wants to prevent are these negative failure consequences, not necessarily always the failure itself. Some failures, especially unforeseen ones, are difficult or even impossible to prevent; or prevention may not be the most cost-effective way of dealing with them. And in many cases, after the occurrence of a failure, there is still a chance to recover from it, through the timely and effective application of countermeasures aimed at avoiding the negative failure consequences. Depending on their effectiveness and timeliness, the countermeasures can be completely successful, so a near miss results, or only partially. In the last case the end result still is an accident, but the severity of the outcomes has been reduced by the countermeasures.

In this paper, the detection that a failure has occurred, combined with problem diagnosis and the identification and application of countermeasures, will be referred to as the recovery process. General agreement exists among researchers in the domain of human reliability (e.g. Zapf & Reason, 1994; Rizzo et al., 1995; Kontogiannis, 1999; Van der Schaaf, 1988) about the different types of recovery actions that can be distinguished. These action types correspond with the phases recovery processes go through:

- actions aimed at failure or problem *detection*,
- actions aimed at *explanation* of the causes,
- and the planning and execution of both ad-hoc and structural *corrective actions*.

Research (e.g. Kanse & van der Schaaf, 2000; Rizzo et al., 1995) has shown that to correct a problem, an explanation of the causes is not always necessary. There are cases where, after the detection of a failure, the corresponding corrective actions are immediately known and implemented. An attempt to locate the causes of the problem may or may not follow at a later stage. In many recovery processes (Kanse & Van der Schaaf, in press), after detection and some initial counteractions, new actions aimed at explanation and additional (long-term) corrective actions are performed. Several repetitions of the explanation phase and the correction phase are possible.

Most of the existing research (e.g. Sellen, 1994; Reason, 1990) regarding recovery processes has focussed on the detection phase; which is important since, after all, no corrective actions will be initiated for failures that remain undetected. In an initial attempt to gain more insight in the *entire* recovery process, including the explanation and correction phases, we have performed both a review of the relevant literature and a field study at a chemical process plant. In the field study, we collected data about near misses and incidents via the plant's standard near miss reporting forms, and held follow-up interviews with those involved in the reported events (Kanse & Van der Schaaf, in press). In our analysis of the collected events, we attempted to establish which mechanisms were involved in the recovery processes. Our study taught us that for an in-depth understanding of recovery processes, analysing only the process mechanisms was not enough, we also needed to analyse which factors actually influenced the processes. In other words, next to knowing *how* recovery processes work, we also need to know *why*: what causes recovery processes to turn out the way they do? After all, recovery processes can take many different forms. Variations are possible in the amount and types of actions needed until successful recovery has been achieved, the amount of time it takes to reach that state, the number of people involved, and so on. Some recovery processes are not as successful as others, either.

By reviewing the work of other researchers who have studied recovery processes in the past, we identified a large number of factors that are found or believed to influence recovery. The field study provided evidence for the influence of many of those factors. We also identified a number of additional factors that had influenced some of the recovery processes we analysed, but that were not mentioned in the literature. All of these factors, along with an indication of their origin, are presented in the next section. In the third section we describe the study we performed to structure this list of recovery influencing factors, so that it would become useful for the analysis and, even more important, the support of recovery processes. The first results of this study are presented and discussed in the fourth section. The fifth and final section contains our more general conclusions.

Factors that influence recovery

Table 1 contains an overview of the factors which are found or (in some cases where there is no scientific evidence yet to prove the influence) expected to influence recovery processes. For each of the factors, we have indicated where we found evidence or arguments for the influence of that factor on recovery: in the literature, or in our earlier field study, or both. To save space, only the most important sources (preferably those with empirical evidence) are listed for those influencing factors mentioned by many different researchers. For most of the listed factors, both negative and positive extremes of the factor are possible (e.g. one can possess little, or a lot of knowledge about the plant and process). One of these extremes corresponds with a negative influence on recovery and the other with a positive influence. But again, to economize on space, in table 1 we have only listed a more or less neutral version of the factors.

Table 1. Recovery influencing factors

influencing factor	**source**	**found in field study?**
seriousness of potential negative consequences if problem is not recovered	Cacciabue et al (1993)	yes
time available for recovery, until the earliest possible occurrence of negative consequences		yes
time available for recovery task(s), considering all other tasks requiring attention		yes
number of failures involved in preceding failure process	Cacciabue et al (1993)	
type of failures involved in preceding failure proces	Cacciabue et al (1993)	yes
simplicity and recognizability of preceding (combination of) failures	Kontogiannis (1999)	yes
expected amount of effort required for recovery	based on expectancy theory by Vroom (1964)	
performance phase in which the immediate result of the failure process is detected (during planning phase / while carrying out the action / when outcome of the action is observable)	Sellen (1994)	yes
presence in area where problem occurs because of other tasks		yes
observability of failures/problems	e.g. Rasmussen & Vincente (1989); Kontogiannis (1999)	yes
quality of feedback 1) about result of normal production-related actions 2) about occurring problems 3) about the result of implemented recovery actions a) from control systems, installations and equipment b) from co-workers c) from the rest of the company	e.g. Hale & Glendon (1987); Bagnara et al. (1988); Rizzo et al. (1995); Sellen (1994); Reason (1990); Wioland & Amalberti (1998); Wioland (1997)	yes
speed of feedback 1) about the result of normal production-related actions 2) about occurring problems 3) about the result of implemented recovery actions a) from control systems, installations and equipment b) from co-workers c) from the rest of the company	Embrey & Lucas (1988); Mason & Redmon (1992); Sellen (1994); Rizzo et al. (1995) (for (3))	yes
control over setting of feedback mechanisms such as pre-alarms	Rizzo et al. (1995)	yes
traceability of problem causes 1) in control systems, installations and equipment 2) among co-workers 3) in rest of company	Kontogiannis (1999)	yes
reversibility (possibility to return from problem state/situation/setting to earlier states/situation/setting)	e.g. Rasmussen & Vincente (1989); Kontogiannis (1999)	yes
availability of materials, tools and equipment needed for recovery tasks		yes
availability of support systems such as expert systems, decision support systems, search aids and memory aids	e.g. Amalberti (1992); Embrey & Lucas (1988); Frese (1991)	yes

influencing factor	**source**	**found in field study?**
possibility to use back-ups from important data and settings	e.g. Frese (1991)	
possibility to test recovery strategies	Kontogiannis (1999)	
availability of channels for improvement suggestions		yes
quality and user-friendliness of interface with process	e.g. Reason (1990); Bagnara et al. (1988); Rizzo et al. (1995); Frese (1991)	yes
availability, quality and usability of procedures and work-instructions 1) for normal production situations 2) for disturbances, problems, or emergencies	Woods (1984) (provides guidelines for (2))	yes
functioning of planned, automatically operating recovery mechanisms (such as barriers, defences, spares, trips)	Cacciabue et al (1993); Hollnagel (1999); Svenson (1991); Hale & Glendon (1987); Perrow (1984) (especially for tightly coupled systems)	yes
quality of problem explanation	Embrey & Lucas (1988)	
time used to establish an initial problem explanation	based on Waller (1999) who found that early engagement in adaptive responses to problem events is crucial	
time used to arrive at the right problem explanation		yes
supervision	Embrey & Lucas (1988) (supervisors performing checks); Edmondson (1996) (coaching behaviour)	
selection-, training-, and competency assessment plan	Hale & Glendon (1987); Frese (1991); Kontogiannis (1999); Sellen (1994)	yes
knowledge and skills 1) regarding plant and process 2) regarding failures in general and corresponding problem solving strategies 3) regarding specific problem and corresponding recovery	Hale & Glendon (1987); Hutchins (1996); Duncan (1987) (support for (2))	yes
experience 1) at plant and with process 2) regarding failures in general and corresponding problem solving strategies 3) regarding specific problem and corresponding recovery	Hale & Glendon (1987); Hutchins (1996); Curry & Gai (1976)	yes
self efficacy	logical complement to factor immediately below	difficult to assess
team efficacy	inspired by Edmondson (1999)	
awareness of limits of own knowledge and possibilities	in reaction to efficacy-factors	yes
awareness of limits of knowledge and possibilities present in team		yes
attitude towards failures and recovery 1. from individual 2. from team, collective 3. from management	Van Dyck (2000); Edmondson (1996)	yes
management attitude & organisational policy towards safety		yes

influencing factor	source	found in field study?
feeling of responsibility for maintaining safety, quality and production in general 1. personal 2. with team		yes
feeling of responsibility for this specific problem 1. personal 2. with team		yes
critical individual attitude towards assumptions and chosen strategies	Embrey & Lucas (1988); Kontogiannis (1999)	
critical collective attitude towards assumptions and chosen strategies	Embrey & Lucas (1988); Kontogiannis (1999)	
concentration and vigilance (vs. fatigue)	Curry & Gai (1976)	
motivation to perform well	Curry & Gai (1976)	
stressfulness of the work environment	Cacciabue et al (1993)	
stress coping ability / immunity to stress	related to factor immediately above	
performing checks for the occurrence of failures	Embrey & Lucas (1988); Kontogiannis (1999)	yes
preparedness in dealing with and responding to expected problems	Frese (1991); Van Dyck (2000)	yes
availability of good strategies for dealing with failures	Kontogiannis (1999); Frese (1991); Van Dyck (2000)	yes
alertness on possible and probable failures and problems	Embrey & Lucas (1988); Van Dyck (2000)	yes
intuition, suspicion, distrust	Van Dyck (2000); Kontogiannis (1999)	yes
division of tasks and responsibilities in organisation	Orasanu et al. (1998)	yes
shift schedule	relevant if factor below plays a role	
shift work coping ability	influences fatigue	
control over work pace	Mason & Redmon (1992)	
individual opportunity to learn by making errors	Van Dyck (2000); Frese (1991)	yes
collective opportunity to learn as a group from errors	Van Dyck (2000); Frese (1991)	yes
availability of co-workers in same team	needed to make cooperation with them possible	yes
availability of colleagues in other departments		yes
cooperation between co-workers from the same team (including communication)	Hutchins (1996); Kontogiannis (1999); Frese (1991); Van Dyck (2000); Sasou & Reason (1999); Wioland & Amalberti (1998); Wioland (1997)	yes
cooperation between people from different parts of the company (including communication)	same as factor above	yes

Method

The list of recovery influencing factors presented in the previous section is not only very long, but also rather unstructured. A quick glance through the list shows that some of the factors seem to be closely related to each other. Some factors even influence each other, causing an indirect effect on recovery processes. The influencing factors have not been grouped yet based on similarities. For many of them, no indications are available as to which recovery process phase(s) the factor influences. In addition, there is no insight in how important each of the recovery influencing factors are for recovery processes: some of the factors may have a much stronger influence than others.

Transformation of this list of recovery influencing factors into a more structured format, in which the problems mentioned above are solved, will contribute to its use in the future for the analysis and support of recovery processes. To obtain a systematically structured list, we have chosen a study design that relies on shared opinions on how to categorise each influencing factor. The rationale behind this approach is that it is better to have a categorisation supported by several individuals than one exclusively based on the researchers' own ideas.

The six individuals who participated in the study aimed at categorising and structuring the recovery influencing factors were subject matter experts in incident and accident analysis. A proper understanding of the mechanisms involved in accidents and near misses was necessary for the categorising task which they were asked to perform. Three of the participants were researchers in the domain of human performance and reliability. The three other participants were members of the Safety, Health, Environment & Quality department of the chemical process plant where the field study was performed earlier. Their plant's near miss database, with which all three are well acquainted, served as a shared frame of reference for them during their participation in the study. The three researchers based their categorisations on a smaller subset of this database, i.e. the cases analysed during the earlier field study, combined with their own experiences.

For each recovery influencing factor, we asked the participants three questions:
1. to which main category does the factor belong;
2. which recovery process phase is or phases are influenced by the factor;
3. and how strong is the factor's influence on recovery.

Regarding the first question, the categories from which the participants could choose were: factor related to failure process, technical factor, human factor, informal organisation factor, formal organisation factor, and other type of factor. Where they felt that a factor belonged to more categories, they were asked to distribute 100% over the applicable categories. The rationale for using these six main categories is as follows:
Over the last decades, a variety of classification systems have been developed for errors and other failures, or (more general) root causes of incidents and accidents. The main categories technical, human, and organisational factors have often been used for the classification of failures (e.g. Van der Schaaf, 1992; Hoyos & Zimolong, 1988; Kjellen & Larsson, 1981; Reason, 1991), usually with a variety of subcategories. As a first step towards the development of a similar classification system for recovery (instead of failure) influencing factors, we wanted to examine if a distinction between the same three main categories was applicable. Instead of having just one category of organisation factors, we decided to distinguish two different types: the formal organisation factors, i.e. factors embedded in the official organisational structure and procedures; and the informal organisation factors, i.e. factors resulting from social, interpersonal and group processes not formally established in the organisation. As can be seen from the literature on organisations and organisational behaviour, for a company to be successful, it needs both its formal and its informal processes to work well (e.g. Mintzberg, 1983; Hall, 1987). The informal processes often compensate for difficulties arising in the formal organisation, which makes them important especially for unplanned, ad-hoc recovery actions. On the other hand, the formal organisation provides procedures and built-

in recovery mechanisms for failures that are foreseen. These provisions correspond with the barriers or defences as described by Svenson (1991) and Hollnagel (1999). We added another category: factors specifically related to the failure process from which recovery is needed, i.e. characteristics of the failures themselves, their causes and their consequences, to separate these from all other influences. We expected that this category would be useful to gain additional insight in the relationship between failures and recovery; a relationship to which Embrey & Lucas (1988); Bagnara et al. (1988), and Sellen (1994) had drawn our attention. Finally, we included a rest category 'other factors' for those factors that do not fit in any of the other categories.

With regard to the second question, the participants could choose one, two or all of the recovery process phases described in the introduction: detection, explanation, or correction. Two additional possible answers were added, not to be used in combination with any of the answers mentioned above: 'none of the process phases', or 'I don't know'. Regarding the third question, about the strength of the influence on recovery processes, we asked the participants to indicate their choice on a 7 point scale, with categories ranging from no influence at all (1), to extremely strong influence (7). The purpose of including this question in the study was to identify which influencing factors are the most important and to compare the strength of the influence among the factors.

The influencing factors were presented to the participants in a random order. Only factors with just a slight difference between them were kept together, so that these differences would be noticeable. The participants were asked to rate two lists of influencing factors. In the first list all the factors were formulated as positive influences on recovery (e.g. good feedback, a lot of experience in dealing with specific problems/disturbances). The second list contained the same factors but formulated as negative influences on recovery (e.g. no or poor quality feedback, lack of experience in dealing with specific problems/disturbances). The advantage of having these two related lists is twofold: a consistency check is possible between ratings on the first and second list; and with regard to the strength of the influence, we can determine if a strong influence of the positive version of an influencing factor always corresponds with a strong influence of the negative version, and vice versa, or not. The same set of written instructions, complete with an example, was given to each participant before they started the categorising task. A researcher was present during the entire time they were working on the task. One reason for this was that the participants could ask questions in case of possible difficulties. Another reason was that not only these difficulties could be recorded, but also the decision process followed by the participants, who were encouraged to think aloud during the task. The list of positive influences, once categorised by the participant, was taken away before he or she could start with the list of negative influences. An individual session was organised for each participant to perform the categorising task. To avoid participants from influencing each other, they were asked not to discuss the categorising task with each other until all of them had completed the task.

To measure the agreement between the participants with regard to their categorisations of the influencing factors, some of the Q-methodology principles and techniques were used. Q-methodology (Brown, 1986; McKeown & Thomas, 1988) focuses on the relationship between a relatively small number of respondents regarding their judgement about a large number of subjects (as opposed to the more common R-methodology where the relationship between variables across persons is studied). In most of the Q-methodology techniques correlations between persons across subjects are used as association measure. However, in our study this is not the most suitable measure; not only because of the nominal nature of two of our rating scales, but also because high correlations do not necessarily correspond with high levels of agreement between participants (Krippendorff, 1987). Instead, we used the kappa-statistic (Cohen, 1960) as a measure of nominal scale agreement, the multi-rater variant (Fleiss, 1971; Siegel & Castellan, 1988), and for the ratings of the strength of the influence also weighted kappa (Cohen, 1968, Landis & Koch, 1977; Schouten, 1982), so that ratings differing only slightly between two raters would also count towards agreement. The kappa statistic provides a measure of agreement that is corrected for agreement exclusively based on chance.

Results & discussion

In this section we will only present *some* of our most interesting initial research findings. Within the scope of this contribution it is not possible to include all our findings, and, furthermore, we have not yet completed all the analyses of the data gathered in this study. First, we will present and discuss a few important statistics we calculated based on the ratings from the participants in our study. After that, we will show how we used the results of our study for our attempt to structure the list of recovery influencing factors.

We calculated the kappa statistic for agreement between multiple raters (Fleiss, 1971; Siegel & Castellan, 1988) for the participants' ratings for each of the three questions separately. With regard to the main category to which each of the influencing factors belongs, a kappa of 0.55, significant at the $\alpha=0.01$ level, was obtained as a measure of overall agreement between the six participants beyond chance agreement. According to guidelines provided by Landis & Koch (1977), this corresponds with a *moderate agreement* beyond chance. With regard to the phases influenced by each of the factors, the value of kappa was 0.66, significant at the $\alpha=0.01$ level, corresponding with a *substantial agreement* beyond chance. To accommodate the possibility of multiple answer categories chosen for one single influencing factor, per factor we calculated the agreement proportions for each of the answer categories separately, and averaged these to obtain overall agreement on the factor.

The ratings for the strength of the influence of each factor on recovery were made by only four participants; two of the researchers felt that they were not familiar enough with the plant's near miss database to be able to make the required judgements. A very low value was obtained for the normal multi-rater kappa statistic, 0.03, indicating no agreement beyond chance. But in this value, two ratings which are only slightly different from each other, a very common situation when using ordinal categories, do not count towards agreement. We tried to overcome this problem by using a weighted kappa statistic for each of the six possible pairs of raters. Exact agreement obtained a weight of 1, the largest possible difference a weight of 0, and all other differences weights in between these values, with smaller differences weights closer to 1 and larger differences weights closer to 0, following a linear formula described by Schouten (1982). We obtained weighted kappa values between 0.04 and 0.22 for the six possible pairs of raters, which are still very low. We had to conclude that *no overall agreement* beyond chance existed among the participants with regard to the strength of the influence of the factors on recovery. Even though they had the same near miss database as a shared frame of reference, the participants probably kept different examples of recovery processes in mind when they rated the strength of the influence of the factors, which explains the lack of agreement. For some of the influencing factors, however, there was sufficient agreement among the raters about the strength of the influence to allow specific conclusions regarding those factors. With regard to 'feedback' as influencing factor, feedback from control systems, installations and equipment was consistently considered more important than feedback from co-workers, which on its turn was considered more important than feedback from the rest of the company. For feedback regarding general production-related actions and regarding occurring problems, quality was considered slightly more important than speed (i.e. the time it takes before feedback becomes available). One participant's ratings, however, indicated the opposite; but this person rated the speed of feedback-factors assuming an already satisfactory quality. Our conclusion is that both quality and speed of feedback are vital, in most cases, one without the other is not much use for successful recovery. For feedback about the result of recovery actions, speed was considered slightly more important than quality. The often restricted amount of time available for recovery is a possible explanation for this finding; once a recovery process is started one needs to know quickly how effective the countermeasures are, so that the recovery process can be completed as soon as possible and other tasks can continue. No notable differences were found between ratings for influence strength of the positive influences (i.e. good quality and fast feedback) and the corresponding negative influences (i.e. poor quality and slow feedback).

To see if there were any substantial differences between each participant's ratings for the positive influencing factors and the negative influencing factors, we calculated a kappa-like statistic for 'intra-rater reliability'. This was done similar to how one would calculate inter-rater reliability for a single pair of raters, i.e. by comparing a participant's ratings for the positive influencing factors to his or her ratings for the corresponding negative influencing factors. Two values were calculated for each participant, one for the recovery process phases influenced by the factors, and one for the main category to which the factors are assigned. The resulting values ranged from 0.70 to 1, with three values qualifying as substantial agreement, six values as almost perfect agreement, and three as perfect agreement (again according to the guidelines from Landis and Koch (1977)). Given the lack of overall agreement on ratings for the strength of the influence of the factors on recovery, and the difficulty the participants had to perform this part of the task, we decided not to calculate intra-rater reliabilities for these ratings. Unfortunately, this was the only rating where we actually hoped to find some more substantial differences between the positive and negative influences.

Since the highest overall agreement on ratings was found for the recovery process phases influenced by the factors, we used these ratings as the basis for structuring the list. For each recovery process phase, we listed and counted the factors for which four or more raters had indicated that they influenced that particular phase. Out of the total of 172 influencing factors (86 positive influences and 86 negative influences), raters agreed that 81 factors influenced the detection phase, 112 factors influenced the explanation phase, and 115 factors influenced the correction phase.

The factors influencing the *detection* of failures or the resulting problem situation were (formulated as neutral versions of the influences, similar to table 1): simplicity and recognizability of preceding failures, performance phase in which failure is detected, presence in problem area for other tasks, observability of failures, quality & speed of feedback, control over feedback settings, interface design, quality & availability of procedures for problem situations, available defences, selection & training, knowledge & experience, individual- team- and management attitude towards failures and recovery, feeling of responsibility for safety, alertness, fatigue, motivation to perform well, stress, stress coping ability, performing checks, preparedness for problems, intuition, shift schedule, shift work coping ability, and the opportunities to learn from errors.

Factors influencing the *explanation* phase were: potential failure consequences, time until negative consequences can occur, time available considering other tasks, number & types of preceding failures, recognizability of failure, expected amount of effort required for recovery, performance phase in which failure is detected, feedback, traceability of causes, availability of support systems, possibility to test countermeasures, interface design, time used to establish an initial problem explanation, supervision, training & selection, knowledge & experience, self and team efficacy, awareness of own & team limitations, individual- team- and management attitude towards failures and recovery, feeling of responsibility for safety and for specific problem, critical attitude towards chosen strategies, motivation to perform well, stress, stress coping ability, preparedness for problems, strategies for dealing with failures, shift schedule, shift work coping ability, control over work pace, opportunities to learn from errors, and the availability of and cooperation with co-workers and colleagues.
The factors that influence *corrective actions* were: potential failure consequences, time until negative consequences can occur, time available considering other tasks, number & types of preceding failures, expected amount of effort required for recovery, performance phase in which failure is detected, feedback regarding countermeasures taken, reversibility, availability of materials and equipment needed for recovery, availability of support systems, availability and use of back-ups, possibility to test countermeasures, availability of channels for improvement suggestions, quality & availability of procedures for problem situations, available defences, quality of and time used for problem explanation, supervision, training & selection, knowledge & experience, self and team efficacy, awareness of own & team limitations, individual- team- and management attitude towards failures and recovery, feeling of responsibility for safety and for specific problem, critical attitude towards chosen strategies, motivation to perform well, stress, stress coping ability, preparedness for problems, strategies for dealing with failures, division of tasks and responsibilities, shift work coping ability, control over work pace, opportunities to learn from errors, and the availability of and cooperation with co-workers and colleagues.

Obviously, these lists overlap. In each of the three lists, a subdivision was made based on the main category to which every factor was assigned (at least three out of the maximum possible six points needed to be given to that category). Of the factors influencing detection, 5% were assigned to the main category of failure process-related factor, 17% were considered technical factors, 37% human factors, 21% informal organisation factors, and 14 % formal organisation factors. The corresponding percentages for the factors influencing explanation are 7%, 12%, 30%, 28%, and 19%, respectively; and 6%, 10%, 33%, 21%, and 24% for the factors influencing correction. The remaining percentages are either factors that were assigned to the 'other type of factor' rest category, or factors about which the raters did not show a sufficient level of agreement with regard to the main category to which the factor belongs.

A closer inspection of the ratings for main category to which a factor belongs, showed that many factors which were mainly assigned to the categories technical, human, or informal organisation factors, were also partly assigned the category of formal organisation factors. This is an indication of the indirect nature of the influence of the formal organisation on recovery. Other examples of these indirect influences are obvious just by going through the list: for example, the levels of knowledge and experience of the organisation's personnel can be influenced via training and selection. Given the existence of these indirect influences, combined with the overlap in assignments of factors to recovery process phases, in hindsight we feel that a taxonomy with main and sub categories to classify recovery influencing factors is probably not the most useful tool for the analysis and support of recovery processes. Therefore, our next steps in structuring the recovery influencing factors will include the development of a model in which this overlap is avoided and indirect influences are visible. Insights gained in the initial structuring attempts as described above can serve as a basis for this development process.

Conclusions

Based on ratings made by six experts in the field of human reliability and accident and incident analysis, we have attempted to evaluate and structure a list of factors which have been found to (or are expected to) influence recovery processes. There was sufficient overall agreement beyond chance between the six participants with regard to the recovery process phases influenced by each of the factors and the main category to which each factor belongs. No overall agreement beyond chance was found between the four participants who also rated the strength of the influence.

One problem we encountered in our study design, was that participants found it difficult to indicate the strength of the influences on recovery, performing the categorising task based on memory alone. A better (and more objective) way to find out how important each of the influencing factors is, would be to analyse and count which factors actually played a role in the recovery processes, for a large set of near misses and incidents, representative for the organisation where the study is performed. This way, we can measure two aspects of the importance of the influence of a factor: *how strong* the influence of the factor is when it plays a role in a recovery process; and *how often* a particular factor actually plays a role in recovery processes in the organisation. An additional benefit of the proposed field study is the possibility to assess whether there are differences in the strength of the influence and frequency of occurrence between the positive and the negative extremes of the recovery influencing factors. The data from our current study did not allow us to draw any conclusions about this.

We have structured the list according to the recovery process phase(s) each factor influences, and organised the influencing factors per phase according to the main category to which they belong. By doing so, we found that a simple taxonomy of recovery influencing factors, with main- and subcategories comparable to taxonomies for failure factors, does not do justice to the complex nature of those influencing factors. After all, one factor can influence several phases of the recovery process, and indirect influences are also quite common. For our future attempts to clarify and model these influencing factors we are planning to develop a generic influence diagram. This diagram will be centred around the three different types of recovery actions: detection, explanation, and correction; using arrows to indicate direct and indirect influences. The insights gained in our current study will continue to be useful in our future modelling efforts.

References

Amalberti, A. (1992). Safety in process control: an operator-centred point of view. *Reliability Engineering and System Safety.*

Bagnara, S., Ferrante, D., Rizzo, A., & Stablum, F. (1988). *Causal analysis in error detection and recovery: when does it occur?* Paper presented at the international conference on joint design of technology, organization and people growth. Venice, Italy, October 12-14, 1988.Brown, S.R. (1986). Q Technique and Method: Principles and Procedures. In: W. Berry & M. Lewis-Beck (Eds.): New Tools for Social Scientists. Newbury Park, CA: SAGE.

Cacciabue, P.C., Carpignano, A., & Vivalda, C. (1993). A dynamic reliability technique for error assessment in Man-Machine Systems. *International Journal of Man-Machine Studies, 38*(3), 403-428

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement, 20*(1), 37-46.
Cohen, J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213-220.
Curry, R.E., & Gai, E.G. (1976). Detection of random failures by human monitors. In: T.B. Sheridan and G. Johannsen (Eds.): *Monitoring behavior and supervisory control*, pp. 205-220. New York: Plenum press.
Duncan, K.D. (1987). Fault Diagnosis Training for Advanced Continuous Process Installations. In: J. Rasmussen, K. Duncan, & J. Leplat (Eds.): *New technologies and human error*, pp. 209-221. New York: John Wiley & Sons.
Edmondson, A. (1999). Psychological safety and learning behavior in workteams. *Administrative Science Quarterly, 44*, 350-383.
Edmondson, A.C. (1996). Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *Journal of Applied Behavioral Science, 32*(1), 5-28.
Embrey, D.E. & Lucas, D.A. (1988). The nature of recovery from error. In: L.H.J. Goossens (Ed.): *Human recovery: Proceedings of the COST A1 Seminar on Risk Analysis and Human Error.* Delft: Delft University of Technology.
Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378-382.
Frese, M. (1991). Error management or error prevention: Two strategies to deal with errors in software design. In: H.J. Bullinger (Ed.): *Human aspects in computing: Design and use of interactive systems and work with terminals*, pp. 776-782. Amsterdam: Elsevier Science Publishers.
Hale, A., & Glendon, A.I. (1987). Individual behavior in the control of danger. *Industrial Safety Series, Vol. 2*. Amsterdam: Elsevier.
Hall, R.H. (1987). *Organizations: structures, processes, & outcomes* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Hollnagel, E. (1999). Accidents and barriers. In: *Proceedings from CSAPC '99*, Seventh European Conference on Cognitive Science Approaches to Process Control, September 21-24, 1999, Villeneuve d'Asq, France, pp.175-180.
Hoyos, C.G., & Zimolong, B. (1988). *Occupational Safety and Accident Prevention: Behavioral Strategies and Methods.* Advances in Human Factors / Ergonomics, 11. Amsterdam: Elsevier.
Hutchins, E. (1996). *Cognition in the wild.* London: The MIT Press.
Kanse, L., & Van der Schaaf, T.W. (2000). Recovery from failures – Understanding the positive role of human operators during incidents. In: D. De Waard, C. Weikert, J. Hoonhout, & J. Ramaekers (Eds.): *Proceedings Human Factors and Ergonomics Society Europe Chapter Annual Meeting 2000*, Maastricht, Netherlands, November 1-3, 2000, pp. 367-379.
Kanse, L., & Van der Schaaf, T.W. (in press). Recovery from failures in the chemical process industry. *International Journal of Cognitive Ergonomics.*
Kjellen, U., & Larsson, T.J. (1981). Investigating accidents and reducing risks – A dynamic approach. *Journal of Occupational Accidents, 1981*(3), 129-140.
Kontogiannis, T. (1999). User strategies in recovering from errors in man-machine systems. *Safety Science, 32*, 49-68.
Krippendorff, K. (1987). Association, agreement, and equity. *Quality and Quantity, 21*, 109-123.
Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
Mason, M.A., & Redmon, W.K. (1992). Effects of immediate versus delayed feedback on error detection accuracy in a quality control simulation. *Journal of Organizational Behavior Management, 13*(1), 49-83.
McKeown, B., & Thomas, D. (1988). *Q Methodology.* Quantitative Applications in the Social Sciences Series, 7(66). Newbury Park, CA: SAGE.
Mintzberg, H. (1983). *Structures in fives: designing effective organizations.* Englewood Cliffs, NJ: Prentice-Hall.

Orasanu, J., Fischer, U., McDonnell, L.K., Davison, J., & Haars, K.E. (1998). How do flight crews detect and correct errors? Findings from a flight simulation study. In: *Proceedings of the human factors and Ergonomics society 42nd annual meeting*, 1998, pp. 191-195.

Perrow, C. (1984). *Normal accidents: living with high-risk technologies.* New York: Basic Books.

Rasmussen, J., & Vincente, K.J. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man-Machine Studies, 31*(5), 517-534.

Reason, J.T. (1990). *Human Error.* Cambridge, UK: Cambridge University Press.

Reason, J.T. (1991). Too little and too late: a commentary on accident and incident reporting systems. In: T.W. van der Schaaf, D.A. Lucas, & A.R. Hale (Eds.): *Near miss reporting as a safety tool*, pp. 9-26. Oxford: Butterworth Heinemann.

Rizzo, A., Ferrante, D., & Bagnara, S. (1995). Handling human errors. In: J.M.Hoc, P.C. Cacciabue, & E. Hollnagel (Eds.): *Expertise and Technology: Cognition and Human Computer Interaction*, pp.195-212. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Sasou, K., & Reason, J. (1999). Team errors: definition and taxonomy. *Reliability Engineering & System Safety, 65*, 1-9.

Schouten, H.J.A (1982). Measuring pairwise interobserver agreement when all subjects are judged by the same observers. *Statistica Neerlandica, 36*(2), 45-61.

Sellen, A.J. (1994). Detection of Everyday Errors. *Applied Psychology: An International Review, 43*(4), 475-498.

Siegel, S., & Castellan Jr., N.J. (1988). *Nonparametric statistics for the behavioral sciences (2nd ed.).* New York: McGraw-Hill.

Svenson, O. (1991). The Accident Evolution and Barrier function (AEB) model applied to incident analysis in the processing industries. *Risk Analysis, 11*, 499-507.

Van der Schaaf, T.W. (1988). Critical incidents and human recovery: Some examples of research techniques. In: L.H.J. Goossens (Ed.): *Human recovery: Proceedings of the COST A1 Seminar on Risk Analysis and Human Error.* Delft: Delft University of Technology.

Van der Schaaf, T.W. (1992). *Near miss reporting in the chemical process industry.* PhD thesis, Eindhoven University of Technology.

Van Dyck, C. (2000). Putting errors to good use: Error management culture in organizations. PhD thesis, Amsterdam: KLI.

Vroom, V.H. (1964). *Work and Motivation.* New York: John Wiley.

Waller, M.J. (1999). The timing of adaptive group responses to nonroutine events. *Academy of Management Journal, 42*(2), 127-137.

Wioland, L. (1997). *Étude des mécanismes de protection et de détection des erreurs: Contribution à un modèle de sécurité écologique.* PhD thesis, Université Paris V.

Wioland, L., & Amalberti, R. (1998). *Human error management: towards an ecological safety model a case study in an air traffic control microworld.* Paper presented at the 9th European conference on cognitive ergonomics. Limerick University, Ireland, August 24-26, 1998.

Woods, D.D. (1984). Some results on operator performance in emergency events. *Institute of Chemical Engineers Symposium Series, 90*, 21-31.

Zapf, D. & Reason, J.T. (1994). Introduction: Human errors and error handling. *Applied Psychology: An International Review, 43*(4), 427-432.

Evaluation of a Temporal Display for Presentation of Threats in Air Combat

Lena Linde, Dan Strömberg, Göran Pettersson
Swedish Defence Research Agency
**{lena.linde,dan.stromberg,
goran.pettersson}@foi.se,**

Rikard Andersson, Jens Alfredson
Saab AB
**{rikard.andersson,
jens.alfredson}@saab.se**

Abstract

The purpose was to evaluate the feasibility of a temporal display (TD), presenting information about time-constrained events (weapon emission or evasive action). BVR air combat was simulated. Two fighter pilots were used as subjects. Two conditions for presentation of tactical information were compared. In one a tactical situation display (TSD) and a multisensor display were used. In the other condition the above temporal display replaced the flight data display (FDD). Pilot performance (no. of weapon emissions, no. of hits) was recorded. Subjects also rated workload (WL) and situation awareness (SA), and answered queries in written and oral form. Precision (no. of hits/ no. of emissions) was found significantly higher when the temporal display was available. Neither WL nor SA was affected by the presentation of the temporal display. Subjects complained about, e.g.:

- Inappropriate color coding
- Inadequate symbol resolution, insufficient display size to enable focus on a specific area.

The TD was tested in a generic cockpit environment. Performance figures in the simulation cannot be compared to performance figures in any specific aircraft.

Introduction

It is crucial for operators of complex human-machine systems to maintain situation awareness (SA). According to Endsley (1995) situation awareness has three levels.

- Perception of relevant elements and details of a situation
- Comprehension of their meaning
- Prediction of the outcome of the situation in the near future

Human information processing (e.g. Wickens, 1984) could roughly be divided into four phases as shown in the upper part of Figure 1. According to Parasuraman, Sheridan & Wickens

(2000) information processing in a human-machine system can be decomposed into four corresponding phases as shown in the lower part of Figure 1.

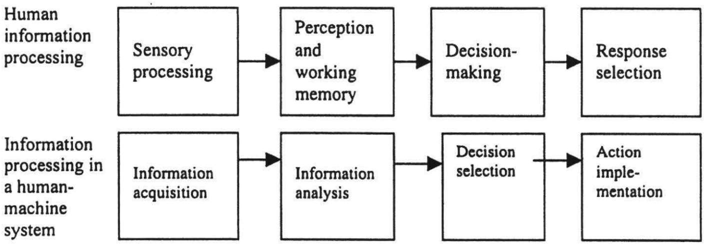


Figure 1: Human information processing and information transmission in a human-machine system.

Coury & Boulette (1992) investigated the relationship between time stress, display format and performance. The formats were either a digital display with three system values, or an integrated, polygon representation of the same values. The integrated display gave higher performance under time stress. However, the benefit of the integrated display under time stress decreased with increasing uncertainty in the data (i.e. data values near the bounder of a category). The time pressure is extremely high in air combat and "temporal awareness" is of utmost importance for a fighter pilot. His/her ability to establish the point of time, at which the enemy may release their weapon is decisive.

The information, required for decision selection, in a BVR (beyond visual range) air combat task is complex. According to Petterson, Axelson, Jensen, Karlsson & Malmberg (1999) the information may roughly be divided into three levels. Each level is based on a set of category and variable values as shown below.

Level 1: Classification of objects

- Friend or hostile
- Class (fighter, attack etc)
- Nationality
- Aircraft type

Level 2: Temporal Situation Assessment

- Spatial information (azimuth, altitude)
- Speed, direction, (visualized by vectors with varying length and direction)

- Weapon-range, sensor-coverage (e.g. spatial launch boundaries represented as cones)

Level 3: Action zones

- Time boundaries for launching weapon

The information system in a modern fighter aircraft for acquisition and analysis (i.e., for classification and for assessing spatial-temporal data values) is complex. There may, for example, be a command and control center (CCC), a Data Link (DL) for providing information from other aircraft, and radar and IR sensor in the own aircraft.

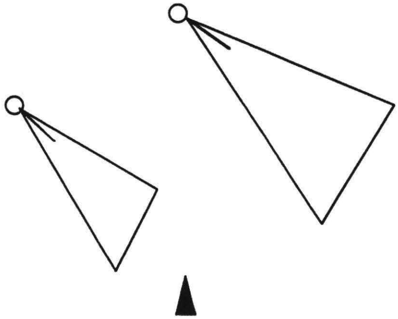


Figure 2: Simplified form of the cone representation of threat possibilities of hostile aircraft (empty circles) against own aircraft (black triangle). A simplified interpretation of the cone symbol is that the length of the cone corresponds to the range of the weapon of the firing aircraft.

Tactical information in a combat aircraft system is mostly presented on a tactical situation display (TSD) comprising a map, with superimposed symbols of hostile and friendly aircraft. Direction and speed data are often indicated by vectors on the aircraft symbols. Aircraft altitude is represented numerically. One way of visualizing a launch boundary, is to use a cone-like symbol, which can be superimposed on the TSD. Examples of "cones" are shown in Figure 2. An interpretation of such a "cone" is that any object, which is within the "cone" when the robot has reached its goal, can be hit. The scenario in a BVR combat situation mostly involve several potentially threatening hostile aircraft. The threat possibilities of the own forces can also be visualized by "cones". But if the cone representation is used in all cases, it is confusing and requires time-consuming, serial scanning. The aim of the so called *temporal display*(see Figure 3), which was investigated in the present experiment, is twofold:

- To give the pilots a more *comprehensive visualization of all threats*, by showing them in one integrated representation, than attained by the cone representations. [Supporting level 2 SA in terms of the model of Endsley (1995)].
- To *reduce the number of variables* required of mental estimation of time margins [Supporting level 3 SA in terms of the model of Endsley (1995)].

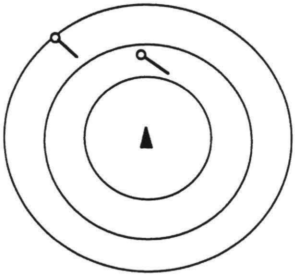


Figure 3: The temporal display (TD). The circles represent varying degrees of threat, which hostile aircraft constitute. Thus, indirectly TD also shows the urgency of counter-measures against them. The triangle in the center represents own aircraft. The symbols in the upper and lefts parts represent hostile aircraft (a threat). The direction of threats on the temporal display corresponds to their real direction.

A technical descriptions of the TD is given by Petterson, Strömberg, and Roldan-Prado (1998). TD consists of four concentric circles, representing varying degrees of threat as shown in Figure 3 (See also, Strömberg and Alfredson, 1997). The circles display time-constrained threatening events, which may be exposed on the own aircraft. In simple terms; the urgency of counter-measures decreases from the inner to the outer circles. The TD only shows tactical information from a defensive point of view (i.e. "What can the enemy do to me?"), but, in contrast to the "cones" no offensive information ("What can I do to a particular threat?").

The purpose of this study was to investigate the feasibility of TD for BVR air combat, using dynamic, simulated BVR combat scenarios. Presentation of tactical information in the conventional way, which comprised a TSD, was compared to a presentation, which in addition comprised TD. The comparison included objective performance and ratings of workload and SA. Six scenarios, comprising BVR combat, were constructed. They were run on a UNIX workstation.

Method

Subjects: Two fighter pilots took part. Both had more than 1000 flight hours, of which at least 150 hour on a multi-role aircraft made in Sweden.

Prestudy and pretraining: A prestudy with two subjects (non-pilots, but well-trained on the simulator) was run to calibrate the final experimental setup. The two subjects in the final study were given four hours training, including instruction and demonstration of the TD.

Scenarios: Six different scenarios (initial configurations of a simulated air combat) were constructed using simulation software at Saab AB. Each began with a *target selection phase* and ended with a *combat phase*. The following criteria were used for the scenarios.

- Their difficulty should be high enough to make form of presentation possibly decisive.
- They should represent clearly different scenarios with a similar degree of difficulty.

Two experienced fighter pilots (not the subjects) examined each scenario and found the criteria fulfilled. There were around ten hostile aircraft, divided into two or three groups, in each scenario. "Own" forces always comprised two aircraft controlled by the two subjects.All hostile aircraft flew towards "own" aircraft. Each scenario contained different unexpected events (e.g. sudden shifts in altitude or speed).

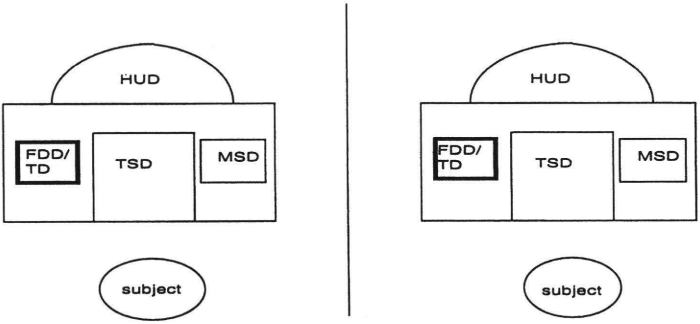


Figure 4: Alternative display configurations. HUD=Head-Up Display, MSD=MultiSensor Display, TSD= Tactical Situation Display, FDD=Flight Data Display, TD=Temporal Display. The bold square indicates the instrument panel part, that was different in the two experimental conditions.

Procedure: There were two experimental conditions. In one of them tactical information was shown on the TSD (condition TSD). In the other the TD was used in addition to TSD (and MSD) (condition TSD+TD). The alternative configurations are shown in Figure 4.The TD replaced the flight data display (FDD) under condition TSD+TD. The two subjects were tested simultaneously, forming a fighting wing. Each scenario was run for a maximum of 4 minutes. Before each run the subjects were allowed to study each scenario, come to agreement about how to execute the task, and settle their roles as leader or wingman. There were six different scenarios. Each scenario was run once under condition TSD, and once under condition TSD+TD. The order of scenarios within each condition was randomized. To minimize the possibility of learning effects, confounding effects of display condition, the following order between conditions was used:Run 1,2,3: TSD+TD, Run 4,5,6,7,8,9: TSD, Run 10,11,12: TSD+TD.

Collection and analysis of data: To assess objective performance number of discharged shots, hits, and hits on own aircraft were counted for each run. A modified "Bedford rating scale" (Roscoe & Ellis, 1990) was used to assess subjective workload and situation awareness (also VINTHEC, 1997). Besides, the subjects rated their own performance after each run. When they had completed all runs, they were also given a questionnaire, containing open and multiple choice questions. In addition, they were interviewed by the authors in the end of the experiment. Successful performance in BVR air combat is to a large extent based on efficient collaboration of the fighting wing. Therefore, number of hits and discharged shots was computed on a wing basis (i.e. the total number of both pilots). Precision of the discharges was assessed as number of hits/ number of discharged shots.

Results

Objective performance: Table 1 shows mean no. of hits, mean no. of discharged shots and mean precision for six runs (that is, for six scenarios) for each experimental condition. First, hits and discharges for each subject were added to obtain the wing performance. Then, a precision value for each run was computed by dividing the hits by the total number of launched shots. Mean precision in Table 1 refers to mean of six runs of these quotients. Wilcoxon's signed, ranked test for two dependent samples (N=6) was used to test the statistical significance of the differences in objective performance between experimental conditions. The dependent samples were the six scenarios presented under two experimental conditions. The two-tailed P-values of the difference between number of hits, and difference between number of launched shots were non-significant. There was a significant difference in precision between the two experimental conditions (two-tailed $P<.05$). As seen in Table 1, whereas about 1 of 5 discharges resulted in hits *without* TD, approximately 1 of 2 discharges resulted in hits when the presentation comprised TD. There were totally 3 hits on own aircraft (i.e. losses). Two of them occurred in the first part of condition TSD+TD. Both pilots were "shot down" within an interval of one minute.

Table 1: Mean no. of hits/run, discharged shots/run and mean precision/run (see above) for each experimental condition. Figures within brackets are standard deviations. All figures refer to performance of the fighting wing.

Condition	$M_{\text{no. of hits/run}}$	$M_{\text{no. discharged shots/run}}$	$M_{\text{precision/run}}$
TSD	1.17(1.17)	5.5(1.38)	.19(.20)
TSD+TD	2.0(1.26)	4.17(.75)	.52(.33)

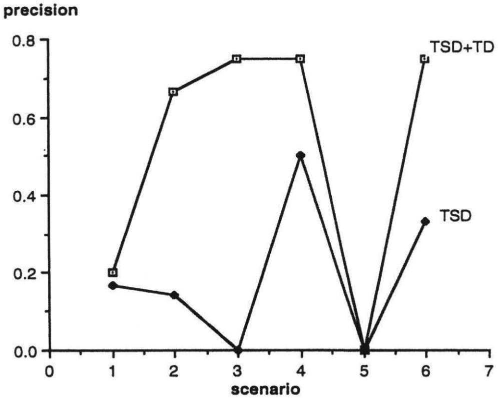


Figure 5: Precision (wing performance) with and without the TD for each scenario.

In Figure 5 precision is plotted against scenario number for presentation conditions with or without the TD. Except for scenario no. 5 precision is higher for condition TSD+TD. Table 2 presents discharges/run and precision for each subject. It shows, first, that both subjects discharged fewer shots/run when TD was available and, second, that the higher precision under the TSD+TD condition is attributable solely to the performance of subject B. One must observe that, since number of hits in specific runs was 0 in several cases (subject B under the TSD condition, in particular), the rounding error for precision of individual subjects is high. One must also observe that the "fairness" of these individual performance scores are dependent on the role that the pilot had on most runs.

Table 2: Mean precision and discharges/run for each subject under the two experimental conditions.

	$M_{\text{Discharges/run}}$		$M_{\text{precision}}$	
Condition	Subject A	Subject B	Subject A	Subject B
TSD	2.7	2.8	0.35	0.06
TSD+TD	1.8	2.3	0.30	0.58

Runwise subjective ratings: Wilcoxon's signed, ranked test for two dependent samples (N=12; 6 scenarios and 2 subjects) was used to test the statistical significance of the differences in objective performance between experimental conditions. No significant differences in subjective workload or SA were obtained between experimental conditions.

Questionnaire and informal interview: Two comments by both subjects, especially stressed during the informal interview, were:

- The legibility of TD was poor. The symbols were small, and gave a "messy" overall picture when they were concentrated in one area. Association between a certain symbol on TSD and the corresponding one on TD was hard to establish. The subjects also suggested improvement of the color coding to increase legibility.
- The informational content was not sufficiently relevant. The usability was limited by the the fact that computed threat levels only was based on information from the CCC, and not on information from the air-borne sensorsystem.

Discussion

The present study comprises only two subjects. The fact that air combat has a collaborative nature with a specific division of subtasks among the two pilots in a wing, and that the superordinate goal is to put as many hostile aircraft out of action as possible, objective performance was assessed on a wing basis. The statistical testing of precision comprised analysis of variability related to presentation condition in relation to variability among scenarios, and not as usually variability pertaining to the experimental manipulation in relation to within or between subject variability.

Ratings of workload and SA were apparently unaffected of whether TD was presented, or not. This may be been due to one or more, of three different reasons:

- One is insufficient legibility, e.g. due to too small symbols, inappropriate color coding etc. Both subjects complained about the "messiness" of the configuration of symbols on TD.
- One purpose of TD is to decrease pilot workload by reducing the demand for "mental computations" (by directly grading the threats into a four classes). The fact that the pilots did not rate workload lower, when TD was presented, may either have been due to the that TD comprised *new* information in addition to the one on TSD, or that the form of the information on TD was not good enough to reduce workload.
- A final reason for lack of increase of subjective SA under the TSD+TD condition may have been the fact that the informational content on TD was found irrelevant and/or the quality of information too low. As noted the subjects commented on the fact that only data from the CCC system, and not the air-borne sensorsystem, was used to estimate threat degree.

Even though the subjects did not rate workload lower and SA higher, objective performance, in terms of precision, was found to be higher when the TD was presented. This difference in performance must be interpreted very cautiously, due to the small number of observations. Despite these facts, two reasons may tentatively be offered:

- The representation of threat degrees in one integrative, closed figure (the circles) facilitated the pilot's evaluation of the overall threat situation from a defensive point of view and thereby the initial prioritization of targets to be attacked. Expressed differently, it was easier to quickly *compare* different threats on the TD than to compare two or more cones. That is, the TD (but not the cones) allowed a direct representation of the threat possiblity of one object *in relation* to the threat possibility of the other objects.
- The TD supported collaborative work, by providing a *mutual perception and understanding* of the threats, ie. a common "blackboard" of the tactical situation (Orasanu, 1994; Artman, 1999; Hoc, Loiselet, & Amalberti, 2000). Therefore the TD may have required *less verbal negotiation* about how to prioritize threats than the cone representation during the first phase of the scenarios. (It might be mentioned that the

verbal communication was recorded during all runs, but that these data not yet have been analyzed.)

References

Artman, H. (1999) Situation awareness and co-operation within and between hierarchical inits in dynamic decsion-making, *Ergonomics*, 42(11), 1404-1417.

Coury, B.G .& Boulette, M.D. (1992) Time Stress and the Processing of Visual Displays, *Human Factors*, 34(6), 707-725.

Endsley, M. (1995) *Towards a Theory of Situation Awareness in Dynamic Systems*, Human Factors, 37(1), 32-64.

Hoc, J.-M., Loiselet, A., & Amalberti, R. (2000). *Cognitive Cooperation in Dynamic Situations: Lessons Drawn from Fighter Aircraft Piloting and Air-Traffic Control.* Paper presented at the HFES 2000 Congress, San Diego.

Orasanu, J.M. (1994) Shared problem models and flight crew performance, in N. Johnston, N. McDonald and R. Fuller (eds.) *Aviation Psychology in Practice*, Avebury Technical.

Parasuraman, R., Sheridan,T., & Wickens, C. (2000) A Model for Types and Levels of Human Interaction with Automation, *IEEE Transactions on Systems, Man and Cybernetics- Part A: Systems and Humans*, 30 (3), 286-297.

Pettersson, G., Axelson, L., Jensen, T., Karlsson, M., & Malmberg, A. (1999) *Multi-source Integration and Temporal Situation Assessment in Air Combat*, In Proceedings from the Conference on Information, Decision & Control, Adelaide, Australien, febr 8-10.

Petterson, G., Strömberg, D., & Roldan-Prado, R. (1998) *Temporal Decision Support and Datafusion in BVR-combat*, In Proceedings of The International Conference on Data Fusion (Eurofusion 98), Great Malvern, UK.

Roscoe, A.H. & Ellis, G.A. (1990) *A subjective rating scale for assessing pilot workload in flight: A decade of practical use*, Royal Aerospace Establishment, Technical Report 900 19.

Strömberg, D. & Alfredson, J. (1997) *Temporal Displays for Fighter Pilots*, In Proceedings of the 6th International Conference on Cognitive Science and Process Control (CSAPC), Baveno, Italy, September 1997.

Wickens, C. (1984) Processing Resources in Attention. In R. Parasuraman & D.R. Davies (Eds.) *Varities of Attention, Academic* Press, NY.

VINTHEC (1997) *Dynamic Measures of Pilot Mental Workload (PMWL), Pilot Performance (PP), and Situational Awareness (SA).* Technical Report: VINTHEC-WP3-TR01.

Developing Lecturer-Student Interfaces in a Distant Training Approach by Cognitive Ergonomics

A. Perassi, P.C. Cacciabue***

* Polytechnic of Turin, Dept. of Aerospace Engineering
C.so Duca degli Abruzzi 24,
10129 Turin, Italy

** European Commission, Joint Research Centre
Institute for Systems, Informatics and Safety
21020 Ispra (VA), Italy

ABSTRACT

This paper considers an IT structure by which knowledge can be stored and transferred between academia professors and students for distant training. An important and detailed work is carried out by experts in IT to develop the necessary programming environments and software by which the data, information and knowledge can be stored and retrieved. The core of the problem, however, lies on the cognitive approaches utilised by the "Knowledge Administrator" of the IT platform to develop the interfaces between tutors and students and the actual formats of the training material. The approaches utilised for such cognitive processes and the experiments carried out in order to develop a prototype distant training course are described in some detail.

INTRODUCTION

The need to capture knowledge and information of experts, that is not easy to formalise, and its use for expanding efficiency and quality of work has been recognised in the late 70ies and was one of the major goals of the various techniques and approaches related to Artificial Intelligence (AI) [Cohen and Feigenbaum, 1986; Engelmore and Morgan, 1988]. Nowadays, the limits of AI have been identified and formal methods exist to elicit knowledge [Davenport and Prusak, 1998].

While there are still severe controversy and debate on the terms *information*, usually based on numerical data, and *knowledge*, based on information, culture and cognitive psychology, a possible way to exploit the know-how derived from AI studies is then more focused on:

1. *Technical approaches* used by different solution-providers;
2. *Management* of companies' numeric information and business performance; and
3. *Transfer* knowledge and information intra- and inter-organisations.

In particular, the transfer of knowledge is not only a transfer of information but also a two-way exchange of messages addressed to a specific operative goal, with communication barriers, noise, feedback and well-defined procedures [Baldwin et al., 1999; Perry et al., 1999; Rist et al., 2000].

The development of Information Technology (IT) has lead to new possibilities of training and transferring knowledge at a distance. In the industrial environment, this implies reducing time losses, efficiency and cost in activities such as maintenance, customer services and business management. The key issue then becomes the optimal combination IT techniques with knowledge management approaches for satisfying organisation needs and objectives.
This is the goal of a European Commission funded project, called ACCESS-Maints[1] (Advanced Cross-Communication Environment providing Support Services to dispersed **main**tenance and technical support Engineers), which is described and discussed in this paper.

The standpoint of the project is the definition of Corporate Knowledge (CK) as the set of values, beliefs, assumptions, symbols and behaviour that identify an organisation in relation with other organisations and govern the ways in which relationships, business and practices are carried out. The project started in early 2000, with the identification of the real needs of users to manage, represent, store, transmit and share parts of the Corporate Knowledge and on the practices adopted to access this knowledge [Thoburn, et al., 1999].

Two different European industrial environments have been chosen as reference-users and test bases of the experimental part of project, namely the Aerospace Division of a large European industry and a Machinery manufacturer. Another particular user has been selected from the academic environment. The objective of the latter is the development of a distant training approach that suits both lecturer's aim and student needs, for the delivery of "standard courses" and "stages" (training on the job periods with academic goals) dedicated to students and/or young professionals c/o local industries.

In this paper, we will describe the technical aspects of the Access-Server-Platform, i.e., the IT structure on which the knowledge is stored and transferred, with particular attention to the academic environment. The major purpose of the academic demo scenario is mainly addressed to investigate relations between professors and students and how training is carried out for the purpose of coaching students towards industrial environments, in order to improve current academic training with knowledge management tools and methodologies.

THE ACCESS-SERVER-PLATFORM

The Access-Server-Platform (ASP) is substantially made of two databases, i.e., the Multimedia Data Base (MMDB) and Knowledge Data Base (KDB). The MMDB is built with basic chunks of information and can be structured and developed quite straightforwardly by standard algorithms. On the other hand, the development of the KDB is much more complicated and requires the direct contribution of a Knowledge Administrator (KA), or of the group of persons to whom the role of KA is assigned (Figure 1).
Central concept or assumption to the ACCESS-Maints project is the definition of Knowledge. The following definition of Knowledge was selected: "*..the intellectual capital resident within an organization. The facts, experiences or competencies known by a person or group of people,*

---

[1] ACCESS-maints, EC-FWP V – IST Programme - Project IST - 1999 - 11763

or held within an organisation (in human- as well as machine systems), gained by individual or shared experiences, training or education.(e.g., understanding what a set of information, facts or situations mean and how to predict and make judgements on them...ability to organise a system of information beliefs and values in order to produce interpretations and decisions)".

The ASP architecture identifies a number of "roles" or "actors" that should operate on the organisation and on the platform for developing the content of CK and for managing the architecture. These actors may be compared with more general figures/roles defined in the Knowledge Management (KM) literature [Delphi Group, 1997; Frappolo and Tucker, 1999; Baldwin et al., 1999].

In general, the literature defines three basic figures or actors in the realm of Knowledge Management:

1. *Knowledge Holders*, i.e., those who have knowledge and know how to use it.
2. *Knowledge Brokers*, i.e., those who know where knowledge is in an organisation, who needs it, and those who have it. They have the big view of Knowledge distribution and management.
3. *Knowledge Seekers*, i.e., those who search for and need Knowledge to improve their performance, efficiency, safety etc.

These three figures of Knowledge Holders-Brokers-Seekers are very general and represent mainly the organisational needs and potential operators at KM level. However, when building a formalised tool, such as the case of the ASP, it is necessary that all activities, i.e., conceptual ones and more IT and implementation oriented functions are fully considered in assigning roles and duties to all contributors. Consequently in the ASP, more profiles of roles are identified, in order to cover all aspects of the development and maintenance of a system, such as the one proposed in the ASP (Figure 1).

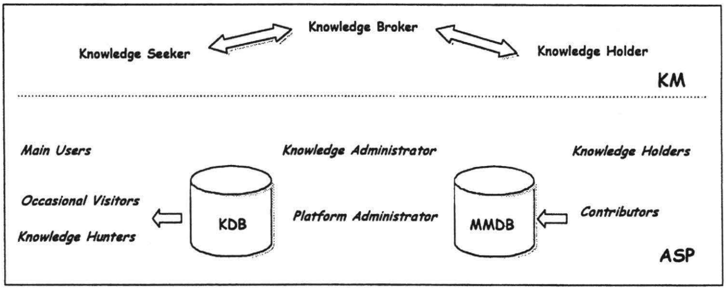


Figure 1. Roles of actors in general KM architectures and in the ASP

The main roles assigned to people in the ASP architecture may be described at theoretical level as follows:

- *Knowledge Holders*, match perfectly the same role of the general definition already given, i.e., those who have knowledge and know how to use it.

- *Contributors* are those who actually implement data and information in the Multi-Media Data-Base.
- *Knowledge Administrators* are those who play the role of knowledge broker within the organisation, but also contribute to the implementation of Knowledge in the appropriate form in the Knowledge Data-Base. Consequently, a knowledge administrator differs from a knowledge broker as the former is expected to actually perform the implementation in the ASP of chunks of information and associated attributes. The role of Knowledge Administrator (KA) may be assigned to a single person, or, most likely, is covered by a group of people with the necessary expertise in knowledge management and, possibly, with experience within the organisation.
- *Platform Administrators* have the role of managing and keeping the ASP, by implementing upgrades, new structures and software material and packages that may be necessary for the maintenance of the platform.
- *Main Users* are the persons to whom the platform is addressed. They are the ultimate users of the KM chain. Their contribution is instrumental in support of the development, content and layout of interfaces of ASP.
- *Occasional Visitors* are members of the organisation or external observers, such as customers, who may get infrequent access to the ASP and/or may need information that could be stored in the ASP and may be utilised for a specific problem.
- *Knowledge Hunters* are persons with valuable experience and expertise in the domain who may wish to contribute to the improvement and expansion of the KDB with their own knowledge. However, they may not act directly on the KDB, as this would compromise the quality control of the KDB, and they should always filter their intended contributions to the KDB through the Knowledge Administrator.

THE CASE STUDY OF ACADEMIC ENVIRONMENT

Transfer of academic knowledge and courses

Nowadays, there are many differences between *degree certification* and *master degree*: these differences regard mainly topics widening, number of hours available to carry out lessons and each qualifying course (e.g. aerospace, mechanics, managerial, nuclear etc...). It is important to go through each single issue. The *degree courses* are characterised by a similar numbers of master degree checks but the topics are more restricted, as tutors have only 40 hours at their disposal. Therefore, they have both to limit number of subjects and investigation level. The current number of hours implies that it is very difficult to carry out an interactive lesson and, quite often, students have no possibilities to address questions during the lessons, in order to clarify their doubts. The degree courses permit students to gain a general overview on different technical topics, but do not foster a specific and complete knowledge as master degrees.

Both for master degree and degree certification courses, the investigation level changes between, for example, aerospace and managerial engineering. Thus, while aerospace courses deepen all mechanics subjects aspects, managerial ones focus mainly on economics.

Some topics characterize one specialization field as, for example, "Human Factors and Aeronautic Legislation" for aeronautical and space engineering.

Tutors, therefore, have to collect data, organize and structure them in relation to the course requirements. The most important point is to understand which are the methods and the way in which professors use to carry out their lessons, which is their logical format and cognitive model, and why they adopt it. Each professor has a different mental model that he/she uses to structure lessons and it is not so easy to capture it in order to create a useful tool for students.
The knowledge provided to students is, nowadays, transferred by voice and comes both from teachers' heads, thanks to experiences acquired, and a lot of reading like articles, books and publications. Moreover, such knowledge base would be useful to teachers also outside classroom didactics and could be used to store and retrieve knowledge for other tasks such as remote courses preparation, research activity and, in particular, exchanges with other academic institutes. This process could also help in verifying the effectiveness of courses, and suggest possible advisable changes.

Within this realm, the Polytechnic of Turin (Polito), as representative of Academy of the ACCESS-Maints project, presents two scenarios to be analysed:

- Students degree course;
- Students companies stage.

The first scenario will be analysed and tested as it is more practical due to all ASP platform capabilities and it could allow capturing and enhancing the knowledge and how it is possible to transfer and optimise it. In order to reduce the ambitiousness and to become more practical, it was decided to analyse only one course, "Human Factors and Aeronautical Legislation". The second scenario will be considered in order to activate innovative links between Universities and enterprises for student's stage.

The definition of these scenarios has been performed in order to satisfy the requirements of the final users (students), integrating them within the ACCESS-Maints supporting methods, tools and procedures. The final goal is to have a demo platform showing the real possibilities of the Knowledge implementation in the academic environments.
The basic aim for this scenario is to be referred on design, development and delivery activities for a multimedia course implementation and innovative work practices enabling to capture, transfer as well as sharing professors' knowledge across students. This latter function will be defined in order to train students on practical aspects and issues when entering actual industrial work domains. The purpose, here, is to transfer them with all the tacit knowledge the professors may have gained during their contact and personal experiences with industries and operational environments. Nowadays, this is a crucial issue because of professors have no enough resources/time and dedicated training programmes to get students acknowledged and aware of work environment requirements and demands outside the academy or university.

The academic scenario will imply:

- Elicitation and selection of professors' knowledge and translation into electronic format, through a methodological suggestion;
- Capture methodologies that professors use to create a course;
- Organisation and structuring of new course programs in accordance with professors' demands (limited number of hours, etc...);
- Identification of valuable knowledge to be provided to the students;
- Effectiveness of training produced/accessed with ASP support tool;
- Test of ASP functions.

In this perspective, the major objective is to give students a good level of technical and scientific knowledge to complete profitably the specific course, to answer to possible students questions in order to clarify their doubts and to support auto learning.

The tool has to be able to store all material necessary to build a very complete course, structure it, collect data into a MMDB, organise information in a logical format (elicitation of knowledge), and identify a knowledge path to construct knowledge database structure.
In this context one of the greatest problems is transferring, cataloguing and recording all implicit knowledge gained through experience which is usually shared (in implicit mode) during the classroom lessons. In this sense a questionnaire or interview for knowledge elicitation is useful.

The main objectives is the development and then the availability of a new kind of course to be submitted to the students: this course will permit a time savings and better learning by the use of ASP system either as knowledge repository or learning system. Moreover, it will support students in their job, increasing the knowledge transfer trough the elicitation of tacit knowledge and structuring explicit knowledge.

One of the most innovative aspects of this scenario is the introduction of a new concept of course, from classroom to multi media: it allows students to share knowledge on practical aspects and issues of real working environments beyond academy realm. An interactive course could be a future for this kind of scenario. Moreover, an exchange of Knowledge Objects among Universities could be a very innovative part to be tested only in the exploitation plan.

The contribution of the ASP to academic training

The implementation of the concepts and formalisms of the ACCESS Server Platform (ASP) with respect to the issue of training academic courses leads to the consideration of the following actors (Figure 2):

- *Lecturer*: who has the knowledge and knows how to use it, he/she could be the professor;
- *Contributor*: inserts raw data into the MMDB, could be a information technology expert or technicians;
- *Knowledge Administrator*: He/she could be the interviewer or professor himself. The Knowledge Administrator has to know the type of knowledge to be stored/transferred (he must OWN some of this knowledge), the knowledge identification/storage methods (he must be a knowledge expert) and the knowledge management and structuring methodologies;
- *Platform Administrator*: is a computer technician in charge of database and network administration.
- *Students*: are the persons who will access the knowledge database in read only mode; in the first Polito demo Scenario the students, as readers, will be the real final users.

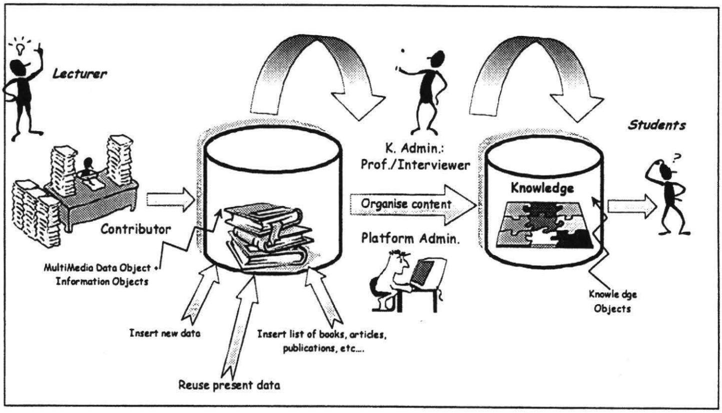


Figure 2. Generic overview of the ACCESS Server Platform

Lecturers, therefore, have to collect data, organise and structure them in relation to the course requirements. The most important point is to understand which are the methods and the way that lecturers use to carry out their lessons, which is their logical format and (cognitive) model of student, and why they adopt it. Each lecturer has different mental model that use to structure the lessons and it is not so easy to capture it in order to create a useful student's tool.

The knowledge provided to students is, nowadays, transferred by voices and comes both from teachers heads, thanks to experiences acquired, and a lot of reading like articles, books and publications.

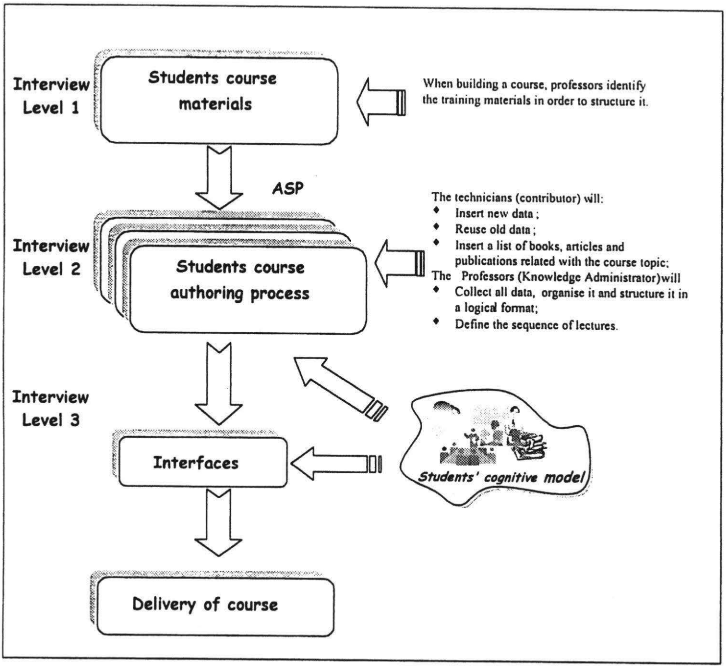
Moreover, such a knowledge base would be useful to teachers also outside classroom didactics and could be used to store and get knowledge for other tasks such as remote courses preparation, research activity and, in particular, exchanges with other academic institutes in order to verify the courses effectiveness, and the possible advisable changes on it.

Data Allocation and Implementation

The processes by which such distant training courses are then developed differ according to the type of course. In particular, "standard courses" and "stages" require slightly diverse initial approaches that account for the contribution of two different "experts", either the university lecturer or/and the industrial tutor. Figure 3 describes the process in the case of "standard courses":

1. The aims of the lecturer to approach training must be discovered through appropriate interviews carried out by the *knowledge administrator* (level-1 interviews).
2. Then, by means of level-2 interviews, modes and contents of course delivery are combined for the actual building of the course.

3. The implementation of the courses is then carried out through an interface that is devised according to level-3 interviews to students, by which user needs and expectations of trainees are considered.



When building a course, professors identify the training materials in order to structure it.

The technicians (contributor) will:
- Insert new data ;
- Reuse old data ;
- Insert a list of books, articles and publications related with the course topic;

The Professors (Knowledge Administrator)will
- Collect all data, organise it and structure it in a logical format;
- Define the sequence of lectures.

Figure 3. Logical process for building a "standard course" utilising the ACCESS Server Platform

For the development of the courses and interfaces the mental model of students behaviour and cognitive process is considered.

The Polito demo scenario is most about a new kind of multimedia course in which all the technical knowledge, presented in the scenario, will be available to students for consulting. The ASP platform has to allow the transfer of knowledge in two basic forms:

1. Tacit knowledge that is the experiential, subjective and personal knowledge hold by professors and substantially more difficult to convey. It is important to find a new way to share tacit knowledge;

2. Explicit knowledge that is already codified and contained in available documents such as articles, publications, topics related books.

It is important to underline that tacit knowledge may vary to al large degree in terms of contents and associated values, being constantly revaluated reviewed and endowed of new meanings, concepts and vision. Moreover, knowledge is not a static source-unlike an information base of documents that can be captured with relative ease, but true knowledge is, in large part, found in the sophistication of the methods and attitudes by which knowledge can be consistently renewed.

The following actions will be done:
- Identification of relevant tacit knowledge to be explicit;
- Organisation and structure of knowledge database;
- Links identification.

Moreover, to realise this process it will be necessary to:
- Properly identify knowledge holder to interview;
- Capture implicit knowledge from knowledge holder through semi-structured interview;
- Elicit the knowledge.

Multimedia Database

ASP will provide a Multimedia Database in which MMDOs and Information Objects (IOs) will be stored. It is important to remember that raw data are encoded in different digital format and added with Primary attributes to obtain MultiMedia Data Object. Then MMDOs, with the addition of Secondary attributes, give rise to Information Objects.

The steps to be made in order to create a MMDB are the following:

- Identification and collection of relevant Multi Media Data Objects to be stored: the raw data for "Human factors and Aeronautical Legislation" course is widely documented, but the documentation is spread in many books, articles and publications;
- Digitalisation of paper format: some further knowledge is dispersed among human factors experts in working notes format;
- Identification of attributes value to characterise each MMDO: each file has to be stored with the related attributes, such as pre-defined attributes (file name, creation date, etc...), and secondary attributes. The free text attributes will be created in accordance with results of first interview.
- Creation of MMDOs and IOs and their insertion in MMDB.

Knowledge database

The creation of Knowledge database structure is very difficult part since it is mostly related to tacit knowledge; for example it could be the "professors experiences" to define a sequence of courses to his/her students according to the understanding of the discipline they have reached.

Knowledge in teaching is not the courseware material but the way in which professors advise their students to learn it, according to the feedback received from students.
As students may miss parts of the course, it results difficult for them to assimilate some parts. In this case a tool that gives them a multimedia course, with a sort of navigation path in course materials, represents teaching knowledge. It is important to underline that if some parts of

course material is missing or couldn't be implemented, the tool could help students to complete and enrich their knowledge.

Moreover, another CK could be Frequently Asked Questions (FAQ) by the students and associated FAQ answer, could be important to permit to students to do question in order to clarify their doubts and then collect them.

CONCLUSIONS

This paper has discussed the problem of knowledge transfer within (intra) and amongst (inter) organisations in order to improve working processes in different environments. The "ACCESS-Maints" project that supports this research covers manufacturing and aerospace industry as well as academia, and utilises an Information Technology approach that is based on the most recent techniques for transfer and management of "data", "information" and "knowledge", in many different forms. We focused on the application to the academic environment. The state of the art of implementation of the concepts and IT structure proposed (ASP - Access-Server-Platform) in the project has been discussed.

Further improvement of the theoretical assumptions and proposed models is necessary before the ASP can be validated and considered ready for practical application to other environments. However, the results obtained so far show that the ASP presents sufficient flexibility to cope with diverse environments with rather different needs and exploitation requirements.

REFERENCES

Baldwin, L. P., Paul, R. J., and Williams, D. S. (1999). Managing soft knowledge in a Hard Workplace. *International Journal of Cognition Technology and Work*, CTW, **1**, 3, pp. 133-141. SpringerVerlag, London.

Davenport T. H., and Prusak L., (1998). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Publishing.

Delphi Group, (1997). *Delphi on Knowledge Management*. The Delphi Group, Boston, MA.

Engelmore E., and Morgan T., (1988). *Blackboard Systems*. Addison-Wesley, Workingham, England, UK.

Cohen, P. R., and Feigenbaum E. A., (1986). *The Handbook of Artificial Intelligence* - Volume III. Addison-Wesley, Reading, MA.

Frappolo, C., and Tucker, M., (1999). Search Engines in the Age of Knowledge. *Intelligent Enterprise Magazine*, **2**, 18.

Perry M. J., Fruchter, R., and Rosenberg, D., (1999). Co-ordinating distributed knowledge: Study into the use of an organisational memory. *International Journal of Cognition Technology and Work, CTW*, **1**, 3, pp. 142-152. SpringerVerlag, London.

Rist, T., Martin J.-C., Néel F. D., and Vapillon, J., (2000). On the design of intelligent memory functions for virtual meeting places: Examining potential benefits and requirements. *Le Travail Humain*, **63**, 3, pp. 203-225.

Thoburn, J. G., Arunachalam, S., and Gunasekaran, A., (1999). Difficulties arising from dysfunctional information systems in manufacturing SMEs - case studies. *International Journal of Agile Management Systems*, Vol. **1** (2), pp. 116-126.

DESIGNING ADVANCED SUPERVISORY AND CONTROL SYSTEMS: AN INTEGRATED AND HUMAN-CENTRED APPROACH

Piccini M.°, Cacciabue P.C.*

°Politecnico di Torino, Dipartimento di Energetica
C.so Duca degli Abruzzi, 24
10129 Torino, Italy
Phone: +39 011 564 4475 - Fax: +39 011 564 4499
Email: mipiccin@polito.it

*EC, Joint Research Centre
Institute for Systems, Informatics and Safety, Human Factors Sector
21020 Ispra (VA), Italy
Phone: +39 0332 789869 - Fax: +39 0332 785813
Email: pietro.cacciabue@jrc.it

Abstract: the design of control systems and human-machine interfaces in the field of complex and safety critical environments is today an open issue, in spite of the high technological evolution of the last years. The increasing use of automation has improved efficiency, safety and easiness of operations but, at the same time, has complicated operator's situation awareness and has changed the nature of their possible operative errors. The research project described in this paper is an attempt to develop a methodological framework to support designers of control systems and human-machine interfaces, focusing in a particular way on the need of a deeply recursive approach related to the implementation of the systemic and human aspects in the design process of a human-machine system intended as a *Joint Cognitive System*. A validating case study is presently going on.

INTRODUCTION

The extended use of automation in modern technology has improved efficiency, safety and easiness of operations but, at the same time, has complicated operator's situation awareness, because of the frequent opaqueness of the autonomous automatic systems [Cacciabue and Hollnagel, 1998]. As a consequence, possible "human errors" have changed their nature, acquiring a more cognitive character.

Human factors engineers and researchers agree in pointing out that a great number of problems related to control of complex systems are design related and are not inherent to the deficiencies of advanced technologies [Mitchell, 1999; Sheridan, 1999; Woods, 1993]. The role of Human Factors (especially concerning human behaviour modelling and cognitive ergonomics) has then become fundamental in the context of design of control systems and interfaces, as well as in performing safety and reliability assessments.

The research project described in this paper represents an attempt to develop a methodological framework to support designers of control systems and human-machine interfaces, by conveying modern theories of supervisory/cognitive control and human-centred design principles [Hollnagel and Cacciabue, 1999; Piccini, 2001]. A particular attention has been devoted to the need of a recursive approach related to the implementation of the systemic and human aspects in the design process of a human-machine system intended as a "*Joint Cognitive System*" [Hollnagel, 1998].

The idea of supplying designers of control systems and interfaces with guidelines made up of "theoretical" aspects and, above all, more "practical" issues, is not new. Indeed several guidelines are available in this field. Some authors argue that designers are often confused and even the best guidelines are sometimes ignored. It has however to be emphasised that these guidelines primarily address low-level physical or generic Human-Computer or Human-System Interaction (HCI and HSI), rather than higher level cognitive issues. When guidance is given beyond low level, its nature is often conceptual and there is no consensus about the related implementation in an actual design.

All these factors often create problems to designers, that require more or less detailed guidelines during their activity, helping them to cope with the new human-related problems arising from technological developments of advanced automation. The heart of the problem has been indicated by Mitchell [1999], by observing that the fast development of advanced technologies in complex systems and their implementation in the field of control systems and human-machine interfaces, have not been followed by contemporary deep knowledge of their implications and consequences. In summary, while generally accepted and widely applicable design knowledge is available for anthropometric and ergonomic problems, design knowledge of supervisory control systems and automation, is less detailed and more conceptual and design experience is often not applicable across a wide range of applications.

REFERENCE PARADIGMS AND BASIC ASSUMPTIONS

The proposed research activity aims at proposing a methodological framework, intended to be a means offered to designers to exploit and apply available practical elements and most validated theoretical support.

The approach is based on two fundamental assumptions, coming from cognitive systems engineering and cognitive ergonomics. Firstly, it is assumed that interactions between humans and automated control systems must be considered in terms of a *Joint Cognitive System*. Secondly, it is recognised that human behaviour (and consequently all the possible erroneous actions) is deeply influenced by the socio-technical context, in which it develops [Hollnagel, 1998].

The notion of *Joint Cognitive System* (JCS) comes from the theories of supervisory control [Sheridan, 1999]. According to these theories the operator remains at the top of the system and can substitute the automated control system when a particular situation requires it: the machine, then, essentially acts as an "intelligent assistant", whose role is to support operators in the control of physical processes. This notion requires that the *human* and the *machine* must be modelled in equivalent terms and that a highly integrated coupling of the two models is essential to describe and analyse the detail of the interaction, as simply schematised in figure 1.

In addition, this standpoint is based on the premise that human cognition is an active process, that can be modified and influenced by operator's objectives and by context and situation: the

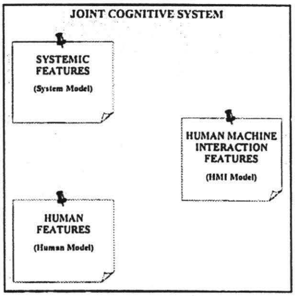


Figure 1 - The Joint Cognitive System paradigm in system modelling and description

description of the behaviour has to be seen from a global point of view and not in terms of a single human or systemic component.

THE PROPOSED FRAMEWORK

General features

The proposed framework is shown in figure 2, [Piccini, 2001]. A detailed number of available instruments (methodologies and techniques) is supplied in order to allow the construction of:

- a *System Model*, with deep knowledge about the system from structural, functional and environmental point of view;
- a *Human Model*, with selection of the most suitable cognitive and human error model/simulation and corresponding taxonomy;
- a *Human-Machine Interaction Model*, with selection of the most appropriate representation of relations between operators and supervised/controlled system.

Two final steps are recursive and complementary:

- the *Supervisory and Control System Design* phase implements the information from the previous steps integrating them with "generic" (syntactic, semantic, contextual and environmental issue) and "specific" guidelines (display and control design);
- the *Validation* phase, performed at design stage, relies upon a top-down and bottom-up assessment, and requires a complete human reliability analysis by appropriate techniques.

The Joint Cognitive System Triad approach

The first three phases of the framework must be considered strictly interconnected. Their objective is to investigate all systemic and human aspects in the definition and design of control system and human-machine interface, as well as to integrate the overall information in

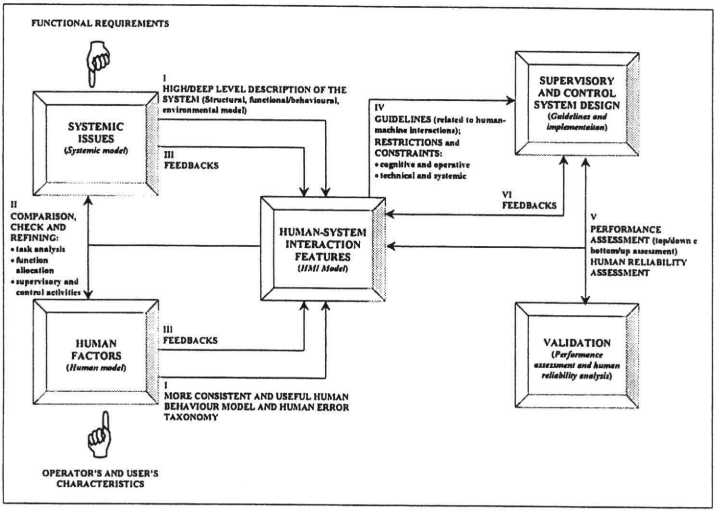


Figure 2 - Proposed methodological framework.

a comprehensive model giving a complete picture of the human-machine characteristics. This follows directly from the notion of the previously mentioned *JCS* paradigm.

From the *systemic* point of view, all analyses have to be performed taking into account the *functional requirements* of the system under development (or under upgrading/updating). These requirements, defined and established in the early phases of the design process, represent the basis for the development of:

- a structural model of the system, through high level and deep level assessments;
- a functional/behavioural model, through expression of the specific goals and functions required for different systems, subsystems and components, and a careful description of the dynamic evolution of processes;
- a contextual/environmental model, by deep inspection of the context and working environment, in order to identify the so called "context stimuli" that can potentially affect or alter human performance.

A wide range of methodologies and techniques can be found in the literature to carry out these activities: concerning the first and third steps, documentation analysis, observation and talk through, interviews and check lists are the most useful methods. As for functional analysis, FAST (Functional Analysis System Technique), SADT (Structural Analysis and Design Technique) and similar functional modelling assessments have been shown very powerful supporting means for design, [Lambert et al., 1999; Ponta, 1997].

From the *human* modelling point of view, a particular attention has to be devoted to the identification of *operator's characteristics*. According to most studies on control room design and human factors related safety, the cognitive aspects of task performance are insufficiently identified and analysed during the design phase, [Cacciabue and Hollnagel, 1998]: this can lead to unbalanced human/machine task allocations and poor or inadequate interface design. A fundamental step in a process of control systems and human-machine interface design is then represented by the selection of a reference human model, as it provides a means for the formalisation and organisation of knowledge about the "role" of the human operator. This is an essential information to determine the level of automation and the distribution of tasks between human operator and control system. In addition, it also represents a basic indicator about the kind of operator interface required in a specific situation, [Søresen, 1999].

This phase must then consist of a process of evaluation and selection of appropriate human behaviour model and error taxonomy, taking accurately into account the system functional features and human cognitive characteristics, allowing:

- to establish the weight of the human element in design process;
- to perform an identification of specific cognitive activities of reference and a related classification;
- to perform and identification process of specific error modes and types, and the related classification;
- to obtain a reference of specified human-related problems to face off in the subsequent design phases.

Literature offers several validated approaches in these fields, [Hollnagel, 1993, 1998; Cacciabue, 1998a].

The third and fundamental phase of the modelling process is represented by an activity of integration of the inputs from the previous steps and definition of the reference basis of the *human-machine interaction* for the successive design activities. The so established "triad" has to be faced in a recursive and cyclic way, where the outputs of each phase represent basis and feedback for comparison, checks and refining activities. A reference supervision model is essential to link systemic and human elements, merging them in order to give way to an evaluation and identification process of the most suitable taxonomies and models among a range of possibilities taken from the most representative and validated ones in literature, [Sheridan, 1999; Billings, 1997].

The definition of the supervision model allows to reach the following crucial results:

- establishment of the balance between human and systemic elements in the design process;
- identification of specific classes of supervisory and control activities of reference, e.g. monitoring, diagnosis, control, etc.;
- categorisation of information nature characteristics (numerical and symbolical), abstraction levels (goals and means), and temporal feedback (mean term and short term/real time).

All these elements contribute to the construction of an overall reference human-machine interaction model, supported by essential activities of allocations of functions (static and dynamic) and task analysis (classical and cognitive). Also in this case, several methods are proposed in the literature for dynamic function allocation and task analysis, particularly suited for analysing supervisory and control type tasks, [Hollnagel, 1993; Bye at al., 1999; Kirwan and Ainsworth, 1992].

Design and validation

The design and validation phases represent a further loop of the methodology, strictly linked and integrated together with the previously described modelling triad.

The actual *design phase* is a process of translation of all structured and formalised human-system interaction information, and its implementation in a control system and interface design project. This has to be done through development of generic and high level guidelines (*syntactic issues*, concerning human-computer interaction; *semantic issues*, regarding human-system interaction; *environmental issues*) and specific and low level guidelines (actual display and control design). The identified classes of cognitive activities and categories of error modes can represent inputs for evaluation of different displays project solutions; in an analogous way, function allocation and task analysis results can guide command-control design. Traditional numerical control-command systems are to be coupled with innovative approaches innovative and advanced systems specifically devoted to cognitive activities, especially "monitoring" (with related tasks of active and passive detection) and "diagnosis" (identification of causes and effects in case of abnormal running), [Riera, 2001].

The *validation phase* has to be carried out during the implementation of the human-machine system and may demand further design changes . Its contribution becomes crucial not only as a final activity of an accomplished system, but also and above all as a simultaneous assessment of an evolving system. The framework identifies three main phases:

- top-down assessment;
- bottom-up assessment; and
- human reliability assessment.

These assessments should include evaluation of qualitative and quantitative nature able to take into account cognitive aspects of human performance, and links and dependencies with the control system and human-machine interfaces, [Hollnagel, 1998; Cacciabue, 1998b; Carpignano and Piccini, 1999].

THE CASE STUDY APPLICATION

Presently, the methodological framework and its major content of information about existing and available methods has been developed and structured. The application to a real case has been started. In particular, a thermoelectric power plant has been selected in Northern Italy, where an upgrading program is under way, including the complete renewal of the control system and related human-machine interfaces. The case study concentrates on the control system and related interfaces for the turbine-alternator group of the power plant, and has a twofold objective:

- to validate the methodological framework, helping to refine the selection methods of available techniques and the links between different phases;
- to give a critical review of the existing control and HMI system.

In figure 3 the structure of the existing control system and the characteristics of the present user's interface are shown.

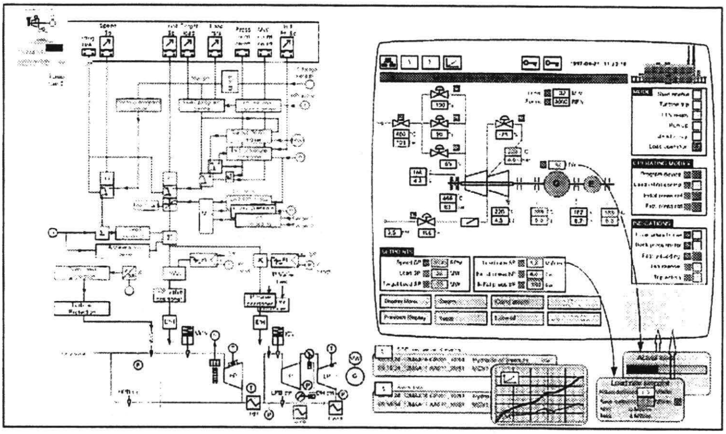
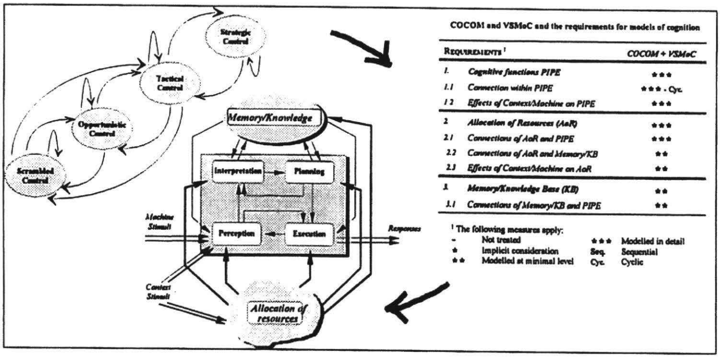


Figure 3 – Current control system and HMI of the plant.

As described in the previous chapters, each step of the framework is characterised by a number of techniques and methodologies offered by literature and by previous specific works. The selection of the most appropriate of them in relation to the particular case under study, is done merging the *user's needs* (the more cognitive and operative aspects) and the *design restrictions and constrains* (the more technical aspects). This selection and adjustment phases have been structured in selection and evaluation criteria and "aid-to-decision" matrixes, whose test and validation are among the main aims of the real case application. Figure 4 shows an example of a possible selection matrix for the cognitive human behaviour model. In this case, the selection is performed in relation to the model capability of representing the relations between the cognitive functions (Perception/Interpretation/Planning/Execution) and the cognitive processes (Memory/Knowledge Base and the Allocation of Resources) of a reference model of cognition. The model under study is in this example the "COCOM and VSMoC" by Hollnagel [1998].

A deep *functional analysis* has started, by application of techniques like FAST, SADT and a functional modelling technique specific for thermo-electrical plants, also under evaluation (figure 5), [Ponta, 1997]. This step is performed in conjunction with a deep inquiry of the environmental and contextual working conditions, by means of walk through, talks with operators and specifically compiled check lists.

A *specific human error taxonomy* is being developed, based on a proper human cognitive model and on the particular characteristics of the control operation requirements. Among the *human behaviour models*, COCOM by Hollnagel [1998], and the more classical SRK model by Rasmussen, [Cacciabue, 1998a], have been selected as the basis for the operator's behaviour characterisation. The specific human error taxonomy is being developed based on the structures of Reason's human error classification, [Reason, 1997], and CREAM and HERMES taxonomies, [Hollnagel, 1998; Cacciabue, 1994].



COCOM and VSMoC and the requirements for models of cognition

	REQUIREMENTS [1]	COCOM + VSMoC
1.	*Cognitive functions PIPE*	***
1.1	*Connection within PIPE*	*** - Cyc.
1 2	*Effects of Context/Machine on PIPE*	***
2	*Allocation of Resources (AoR)*	***
2.1	*Connections of AoR and PIPE*	***
2.2	*Connections of AoR and Memory/KB*	**
2.3	*Effects of Context/Machine on AoR*	**
3	*Memory/Knowledge Base (KB)*	**
3.1	*Connections of Memory/KB and PIPE*	**

[1] The following measures apply:

-	Not treated	***	Modelled in detail
*	Implicit consideration	Seq.	Sequential
**	Modelled at minimal level	Cyc.	Cyclic

Figure 4 – An example of selection matrix: the case of the operator's behaviour model

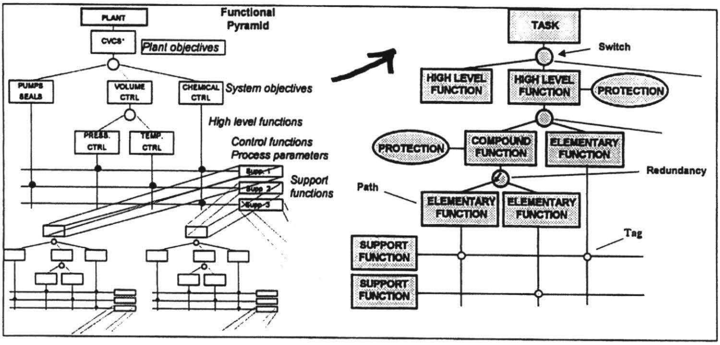
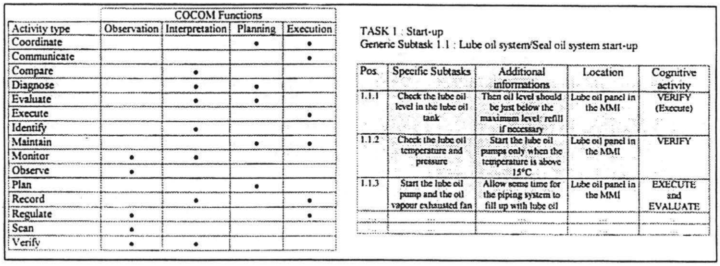


Figure 5 – The functional analysis structure

An accurate *task analysis* has finally been started, focusing on the operative aspects, as well as on the cognitive features associated with the performance requirements. A *cognitive* task analysis will then be fully applied to the system, specifying the interaction of mental procedures, factual knowledge and task objectives in the process of job performance, [Hollnagel, 1998; Cacciabue, 1998a].

This approach (figure 6) is based on the identification of the main task, of the generic and specific subtasks, and on building of a specific "cognitive demands profile" with reference to a list of critical cognitive activities and "generic cognitive activity by cognitive demand matrix", [Hollnagel, 1998].



	COCOM Functions			
Activity type	Observation	Interpretation	Planning	Execution
Coordinate			•	•
Communicate				•
Compare		•		
Diagnose		•	•	
Evaluate		•	•	
Execute				•
Identify		•		
Maintain			•	•
Monitor	•	•		
Observe	•			
Plan			•	
Record		•		•
Regulate	•			•
Scan	•			
Verify	•	•		

TASK 1 : Start-up
Generic Subtask 1.1 : Lube oil system/Seal oil system start-up

Pos.	Specific Subtasks	Additional informations	Location	Cognitive activity
1.1.1	Check the lube oil level in the lube oil tank	Then oil level should be just below the maximum level: refill if necessary	Lube oil panel in the MMI	VERIFY (Execute)
1.1.2	Check the lube oil temperature and pressure	Start the lube oil pumps only when the temperature is above 15°C	Lube oil panel in the MMI	VERIFY
1.1.3	Start the lube oil pump and the oil vapour exhausted fan	Allow some time for the piping system to fill up with lube oil	Lube oil panel in the MMI	EXECUTE and EVALUATE
...	...	...	...	...
...	...	...	...	...

Figure 6 – The cognitive task analysis approach

CONCLUSIONS

The purpose of the work presented in this paper is the development of a support methodological framework for designers of human interactions. By means of a number of articulated and formalised phases, it aims at supplying the designer with a structured guide through all "theoretical" aspects and more "practical" issues arising during a design activity. The framework offers a guideline and a compendium of existing techniques for designing and developing HMIs, that merge in the exploitation of the concepts of the *Joint Cognitive System* paradigm.

Presently, the methodological framework and its major content of information about existing and available methods has been developed and structured, and the application to a real case in under development. The outcome of this exercise will hopefully be dual. On the one hand the methodological framework and its recursive procedure can be validated with respect to the concepts of JCS-triad and to its practicality to a real case. On the other hand, it will be possible to appreciate the amount of modifications and amendments that are required in the reality order to comply with the requirements a cognitive approach, with respect to a planned process of renewal of a control system based on more classical engineering methods.

REFERENCES

Billings, C.E. (1997). *Aviation Automation.* Erlbaum, Mahwah, New Jersey, USA.

Bye, A., Hollnagel, E. and Brendeford T.S. (1999). Human-machine function allocation: a functional modelling approach. *Reliability Engineering and System Safety*, **64**, pp. 291-300.

Cacciabue, P.C. (1994). *Affidabilità dinamica e fattori umani in sistemi nucleari.* European Commission, JRC Report. EUR-15988 IT, EC Bruxelles, (in Italian).

Cacciabue, P.C. (1998a). *Modelling and Simulation of Human Behaviour in System Control.* Springer-Verlag, London, UK.

Cacciabue, P.C. (1998b). Modelling and simulation of human behaviour for safety analysis and control of complex systems. *Safety Science*, **28** (2), pp. 97-110.

Cacciabue, P.C. and E. Hollnagel (1998). Human factors methods for safety assessment in highly automated environments: task analysis, modelling and data. In: *Probabilistic Safety Assessment and Management* (A. Mosleh and R.A. Bari Eds.), pp. 741-745. Springer-Verlag, London, UK.

Carpignano, A. and Piccini M. (1999). Cognitive theories and engineering approaches for safety assessment and design of automated systems: a case study of a power plant. *Cognition Technology and Work*, **1** (1), pp. 47-61.

Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control*. Academic Press, London, UK.

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method*. Elsevier, Oxford, UK.

Hollnagel, E. and P.C. Cacciabue (1999). Cognition, technology and work: an introduction. *Cognition Technology and Work*, **1** (1), pp. 1-6.

Kirwan, B. and L.K. Ainsworth (1992). *A guide to task analysis*. Taylor & Francis, London, UK.

Lambert, M., Riera B. and G. Martel (1999). Application of functional analysis techniques to supervisory systems. *Reliability Engineering and System Safety*, **64**, pp. 209-224.

Mitchell, C.M. (1999). Model-based design of human interaction with complex systems. In: *Handbook of Systems Engineering and Management* (A.P. Sage and W.B. Rouse Eds.), pp. 745-810. John Wiley and Sons, New York, USA.

Piccini, M. (2001). Supporting Design of Advanced Supervisory and Control Systems: an Integrated and Human-Centred Approach. In: *Human Decision Making and Manual Control* (M. Lind, Ed.), pp.113-121. Technical University of Denmark, ISBN 87-87950-89-8.

Ponta, A. (1997). *Modelli avanzati e tecniche per l'analisi di affidabilità a supporto della manutenzione di sistemi complessi*. PhD thesis. Politecnico di Torino, Italy, (in Italian).

Reasons, J. (1997). *Managing the risks of organizational accidents*. Ashgate, Aldershot, UK.

Riera, B. (2001). Specification, design and evaluation of an advanced human-adapted supervisory system. *Cognition Technology and Work*, **3** (1), pp. 53-65.

Sheridan, T.B. (1999). Human supervisory control. In: *Handbook of Systems Engineering and Management* (A.P. Sage and W.B. Rouse Eds.), pp. 591-628. J. Wiley & Sons, New York, USA.

Søresen, M.U. (1999). Application of functional modelling in the design of industrial control systems. *Reliability Engineering and System Safety*, **64**, pp. 301-315.

Woods, D.D. (1993). Cognitive systems in context. In: *Cognitive Engineering in Aerospace Applications: Pilot Interaction with Cockpit Automation* (N.B Sarter and D.D. Woods Eds.), pp. 2-9. NASA Contractor Report 177617, Moffet Field, California, USA.

Comparative Assessment of Human Error in Autonomous Production Cells

Ralph Reuth, *Christopher Schlick, Holger Luczak
Institute of Industrial Engineering and Ergonomics, Aachen University of Technology
*Forschungsgesellschaft für angewandte Naturwissenschaften, D-53343 Wachtberg
<r.reuth, h.luczak>@iaw.rwth-aachen.de

ABSTRACT

Using dynamic simulation, an approach to assess the impact of human errors on overall system performance in Autonomous Production Cells (APC) is presented. The method of timed colored Petri Nets is used to model and to simulate the developed task models. Within an expert inquiry four task models were parameterized with respect to 3D-laser welding and 5-axis milling. To validate the chosen methodological approach one reference manufacturing order for each technology was executed and monitored by video analysis techniques on the shop floor. Then comparative hypothesis regarding total time on task and expertise requirements were formulated. Several simulation experiments were conducted to test these hypothesis. The results from inferential statistics show, that single operator APC have an efficiency of more than 37 % on total time on task without any significant influence of human errors.

INTRODUCTION

The development of flexible decentralized production units offers the potential to shorten the order processing time. Being still in a state of basic research, Autonomous Production Cells are based on effective interaction of both highly skilled operators and complex CNC-based machinery such as 5-axis milling and 3D-laser welding on the shop floor [APC, Eversheim et al., 1998]. From the human factors research's point of view it is important to design and assess operator's work processes already in the stage of research of the APC lifecycle [Schlick et al., 2001]. Therefore, task analysis is a crucial step in evaluating and designing systems. It provides insights into how users and systems interact and ensures that either the task is adapted to the person or, if feasible, measures can be derived to fit the person to the task [Luczak, 1997]. Thus, flexible manufacturing systems need human interventions in monitoring, maintaining and planing. As soon as the variability in human performance is recognized as inevitable, then it is easy to understand that when humans are involved, consequently errors will be made, regardless of the level of training, experience, or skill [Park, 1997]. The common sources of potential system errors that may occur in the flexible manufacturing systems include [Majchrzak, 1988]: (1) operator errors, (2) errors in the cell or shop level computer, (3) errors in the part or program that runs the specific manufacturing process at the machine level, and

(4) faults in the electronic logic that relays the mechanical movements of the machine. This is the reason why it is necessary to investigate the dynamic behavior of work systems. Therefore, this contribution deals with the comparative dynamic simulation of operator's work processes in APC and conventional manufacturing with respect to human reliability analysis aspects and the resulting shifts in knowledge and skill requirements. To validate the chosen methodological simulation framework one reference manufacturing order for the technology of 5-axis milling and one for the technology of 3D-laser welding was executed and monitored by video analysis techniques on the shop floor.

METHODOLOGY

The *external mode of malfunction* [Rasmussen, 1987] resp. observable error in the operating environment has its origin in a malfunction of human behavior typically initiated by some occurrence of an external event. With respect to the evaluation of the efficiency of work systems we will focus on their consequences. The effects of these observable errors are inappropriate states of work objects and work system. Some examples in the field of manufacturing are: loss of time, reduction of quality, production disruption and accidents [Zimolong, 1990]. Thus, the recovery of human errors can be divided in task and object related categories. Task related consequences of human errors can lead to simple task repetition or the repetition of complete task sequences [Reason, 1990]. With regard to manufacturing systems object related consequences mean reversible or irreversible quality reduction to work objects. Consequently, the operator's task spectrum has to be classified to tasks with material changing abilities (e.g. job execution) and non material changing abilities (e.g. operation planning). Both everyday observation and common sense have shown that there are probably only three ways in which people's errors are brought to their attention: (1) self-monitoring, (2) environmental error cueing and (3) by another person. The higher the *level* is on which a person is operating, the more error prone he is [Reason, 1990]. Making a postslip attentional check does not ensure the detection of error per se. Detection must also depend on available clues signaling the departure of action from current intention. Detection of errors on the skill-based level of performance where any discrepancy between the current and desired state is, is in principle at least, fairly easy to determine. But this is clearly not the case at the rule-based and knowledge-based levels [Reason, 1990]. One of the most detailed investigations of error detecting during problem solving was carried out by Allwood [1984], University of Gotheborg. In addition, Rizzo and his collaborators [Rizzo et al., 1986; Bagnara et al., 1987] examined in two separate studies the relationship between Reason's error types and self-monitoring detection processes. Averaging these three studies the overall relative detection rates are 86.1 per cent for skill-based, 73.2 per cent for rule-based and, 70.5 per cent for knowledge-based errors [Reason, 1990]. According to a dynamic prospective evaluation of work systems the term Human Error Probability (HEP) is of final interest. The vast majority of techniques and tools currently available for human reliability analysis is essentially based on research undertaken in high-risk sectors, such as nuclear and petrochemical industries. However, the same principles can obviously be applied successfully in other (with lower safety risk) industrial contexts. The *Human Factors Reliability Benchmark Exercise – HF-RBE* [Poucet, 1989] tested different HRA-techniques in terms of a specific task by different expert teams. Results show that calculated HEPs particularly sizeable deviate. As a study carried out by Kirwan [1997], HRA-methods reach good conformity by evaluation of skill-based and rule-based tasks. With respect to our former research activities we could demonstrate a significant shift within cognitive control towards rule- and knowledge-based behavior especially for APC operators in comparison to conventional operators [Schlick et al., 2001]. This proves an often only qualitative proposition of Sheridan's [1997] supervisory control paradigm. To calculate HEP, the HRA-method ESAT [Brauser, 1992] was chosen. ESAT is a relative new technique for task taxonomy, although external validity has not been proved so far. The major disadvantage of uncertainty in calculating valid HEPs in ESAT is leveled by a relative comparison of alternatives in our approach. This relative comparison of alternatives is widespread in the dynamic simulation of flexible manufacturing systems, because an accurate absolute prediction of characteristics is difficult to achieve due to the complexity of component interrelations. Objective of ESAT is the classification of *any* task in a man-machine system according to human reliability of task execution. The adjustment of Performance Shaping Factors (PSF), resp. task type, task characteristic, personnel, environment and system allow to calculate a HEP for *each* task treated [Brauser, 1992].

With respect to our dynamic simulation these considerations lead to the following model algorithm for HE, which is a simplified representation of error commitment, detection and recovery. This model algorithm was used on each task in simulation progress:

1. Check whether the task is already faulty (e.g. errors made by prior execution).
2. Calculate HEP (if task is faulty: HEP = 1) and decide by random number generation, if an error occurs.
3. Decide, in dependency of the level of cognitive control and random number generation, whether an error has been detected. If this is not the case, the task will be marked as faulty. Additionally an error counter is introduced, which will be raised by one.
4. Calculate additional time consumption in dependency of error detection.
5. If the task has material changing abilities and the error counter is larger than zero, decide whether the manufactured workpiece is reversible or not.

To rate human occupational requirements in manufacturing it is of great importance to know the operators' task structure on the shop floor. With regard to the intended dynamic simulation four task models were build on the methodological basis of Colored Petri Nets [Jensen, 1997]. Two task models represent typical work processes in APC (3D-laser welding and 5-axis milling) whereas the others represent conventional manufacturing. These task models include all work stages from process- and operation planning, machine adaptation, job execution, etc. In order to allow repetition of complete task sequences in case of erroneous actions or system failures a hybrid approach was chosen, because ESAT is not applicable. Additionally, from the psychological point of view of human errors there are doubts, whether distinct actions, which will lead in industry to errors, can be separated by so called normal actions. To optimize work activities it is of great concern for the operators' functional skills to try different techniques like "try and error" to achieve new problem solutions [Zimolong et al., 1992]. Thus, these cases have to be modeled explicitly into operator's task structure by branching and looping.

RESULTS AND DISCUSSIONS

The starting basis for task network modeling has been provided by our former research activities [Schlick et al., 2001, Schlick et al., 2000]. Existing task networks for conventional

3D-laser welding and 5-axis milling have been adjusted to recent technological and organizational developments. All task networks were tested regarding consistency and logical errors. To develop the net graphs of APC, task structures have to be acquired empirically, because this manufacturing system is still in a state of basic research, and a formal description is lacking. Therefore, a participatory task analysis and modeling approach with the developers of APC was preferred. Each meetings lasted 2 hours (see table 1). Results are shown in table 2.

Table 1: Description of expert meetings

Technology	Work System	Number of Meetings	Number of Experts	Average Occupational Experience
3D-laser welding	Conventional	3	5	6
5-axis milling	Conventional	3	7	15
3D-laser welding	APC	5	5	3.5
5-axis milling	APC	5	6	3.5

Table 2: Resulting task structures on basis of Colored Petri Nets

Technology	Work System	Places	Transitions	Arcs
3D-laser welding	Conventional	22	33	66
5-axis milling	Conventional	28	31	77
3D-laser welding	APC	28	43	91
5-axis milling	APC	33	42	89

Acquisition of task attributes

The values of task elements for all four task models were acquired with support of a questionnaire, which referred to an underlying 3D-laser welding order respective 5-axis milling order. The lot size of each manufacturing order was five parts. Two groups of raters participated in each technology. The first group consisted of APC experts. These experts were only competent in narrow domains of the whole work process. Therefore, the APC experts only rated the subset of APC task elements that were compatible with their competency profile. The second group included experienced industrial workers in terms of conventional manufacturing technologies. These experts rated the values of attributes of the whole set of task elements of the conventional manufacturing system (see table 3).

Table 3: Description of Aquisition of Task Attributes

Technology	Work System	Number of Experts	Average Occupational Experience	Data Collection Time [h]
3D-laser welding	Conventional	5	6	16.2
5-axis milling	Conventional	6	15	25.2
3D-laser welding	APC	7	3	9.2
5-axis milling	APC	8	3	7.3

To specify each task in task models the following aspects were rated: (1) time consumption, (2) task complexity, (3) task difficulty, (4) the mode of object manipulation, (5) probabilities for looping and branching, (6) the characteristic of expertise for a successful task execution and, (7) additional time consumption in case of observable errors within the task itself.

Validation of Conventional Task Models

Before experiments with task models can be executed, models have to be checked to see whether they are valid representations of the systems that are being studied. Afterwards the model can be used to identify causes and effects [Daalen et al., 1999]. With respect to our investigation this can only be done for conventional manufacturing, because APC are still in basic research. In order to validate estimations for conventional manufacturing, the reference milling and 3D-welding orders were each produced in reality by an additional experienced industrial worker. Due to high costs this could only be done once. According to this, the logical structure of task models and the reliability of expert ratings in terms of time consumption had to be verified. The experimental design for laboratory evaluation of conventional task models was based on three methods: (1) observation and time measurement [REFA, 1991], (2) recording by means of two video cameras, whereby one camera was fixed to the machine and the other mobile near the operator, and, (3) *thinking aloud* to comment essential work steps.
The reference 3D-laser welding order was manufactured on a 6-axis 3D-laser welding machine (TLC105 - Trumph Laser Cell) by an experienced operator with an occupational experience of three years. The observed processing time for a lot size of 5 finished parts was 5,39 h including 5 defective workpieces. The reference 5-axis milling order was manufactured on a 5-axis milling machine (Maho MH700S) by an experienced operator with an occupational experience of 43 years. The observed processing time for a lot size of 5 finished parts was 5,83 h. Subsequently a video analysis came to operation, which ended in one separate task network based on Colored Petri Nets each technology, either. Thereby the derived experimental task network of conventional 3D-laser welding includes a total of 32 places and 44 transitions. The control flow is represented by 90 arcs. The experimental task network of conventional 5-axis milling includes a total of 31 places and 37 transitions. The control flow is represented by 74 arcs. Afterwards, each experimental task network was rated by the relevant operator in terms of time consumption, task difficulty, etc. as presented before. Besides, measured time was given the operator as temporal basis for assessment. The mode of object manipulation, task complexity and observable state transition probabilities were acquired by post video analysis. Thus, the basis for a comparative validation via simulation was given.

Table 4: Summary of inferential statistics for validation of conventional manufacturing task structures ($\alpha = 0.25$).

Null Hypothesis	Technology	Computed Mean / Standard Deviation	Computed t-statistic	Test Result H_0
H_{02}: $TT_{expert} = TT_{model}$	3D-laser welding	$TT_{expert} = 5.54$ h / s = 0.78 h $TT_{model} = 5.58$ h / s = 0.94 h	t(df=451) = 0.561	Not Rejected
H_{03}: $TT_{expert} = TT_{model}$	5-axis milling	$TT_{expert} = 5.92$ h / s = 0.37 h $TT_{model} = 4.22$ h / s = 0.83 h	t(df=326) = -31,44	Rejected

Next the Poses++ Petri Net simulator (http://www.gpc.de) was used to compute the different conventional work processes including our affiliated model of HE. With the help of t-statistics for independent samples it was to prove, that there are no significant differences between both task networks (experimental/non experimental) for each technology. The probability of erroneously rejecting the null hypothesis is usually called the beta error probability. As alpha and beta error probabilities behave inverse proportional, we chose $\alpha = 0.25$ to minimize type-2 error. We took 234 simulation replications for 3D-laser welding and 163 simulation replications for 5-axis milling [Goldmann et al., 1998].

The results in Table 4 demonstrate a good match of total time on task according to the task models of 3D-laser welding. The corresponding t-value does not allow rejection of the first null hypothesis H_{02}. Thus the non experimental task model of 3D-laser welding seems to be useful and appropriate for further analysis, as it produces the required output within its domain of application. The second null hypotheses H_{03} shows in contrast significant differences with respect to its corresponding t-value, so that adjustment of the non experimental task model of conventional 5-axis milling is required. Thus, a detailed investigation of estimated versus measured task time consumption for each task was carried out with the help of t-statistics (α = 0.25).

Table 5: Extract of inferential statistics for task time consumption in 5-axis milling for H_{0i}: $T_{\text{Expertrating}} = T_{\text{Observation}}$ ($\alpha = 0.25$).

Task Name	N	Computed Mean / Standard Deviation	Observed Time Value	Computed t-statistic	Test Result H_0
NC-Programming	6	T_{expert} = 19.4 min / s = 17.85 min	1,5 h	t(df=5) = -9.67	Rejected
Control workpiece measures	6	T_{expert} = 3.5 min / s = 2 min	3.3 min	t(df=5) = 0.207	Not Rejected
Check Tooldata	6	T_{expert} = 3.1 min / s = 2.2 min	9.1 min	t(df=5) = -6,593	Rejected

All executed tasks were tested (see Table 5), whereas 8 significant differences in task time consumption could be found. The most obvious time difference of 1,3 h was detected within the task *NC-Programming*. The basic reason for this was, that the milling operator had to deal with a new programmable numerical control and therefore has had a lack of experience. Additionally, expert estimations were based on the availability of predefined program routines for wholes, etc, whereas in experiment the NC-program was coded each line sequentially. Consequently, all tested tasks with significant differences were substituted in the non experimental model by real time observation. Thus, the adjusted model was also tested with the help of t-statistics ($\alpha = 0.25$) by additional 173 independent simulation replications.

Table 6: Inferential statistics for adjusted task structure of conventional milling ($\alpha = 0.25$).

Null Hypothesis	Technology	Computed Mean / Standard Deviation	Computed t-statistic	Test Result H_0
H_{04}: $TT_{\text{expert}} = TT_{\text{model}}$	5-axis milling	TT_{expert} = 5.94 h / s = 0.41 h $TT_{\text{modelcorr}}$ = 5.98 h / s = 1.01 h	t(df=187) = 0.469	Not Rejected

As can be seen in Table 6, the corresponding t-value does not allow rejection of the third null hypothesis. Thus, the non experimental adjusted task model of 5-axis milling seems to be useful and appropriate for further analysis. Like already mentioned above, task time consumption for NC-programming can not be seen as a representative value. Therefore, this value was put back to expert estimations. Furthermore, task complexities and the mode of object manipulation were substituted in both non experimental task models by observed values. It has to be stressed, that these models are only valid with respect to our reference orders. Therefore, two validated reference models as starting basis for further investigation were given.

Comparative Assessment and Dynamic Simulation

The following four hypothesis were formulated to a comparative assessment of single – operator manufacturing and were later tested via inferential statistics.

H_5: 'The total time on task is significantly lower for APC'. The dependent variable *total time on task* (TT) can be computed directly from the dynamic simulations of the respective task sequences (H_{05}: $TT_{APC} >= TT_{Conv}$).

H_6: 'The portion of human error is significantly lower for APC'. The dependent variable *portion of human error* (HE) is defined as total time on task with provision for HE in relation to the total time on tasks without HE (H_{06}: $HE_{APC} => HE_{Conv}$).

H_7: 'The number of defective work pieces is significantly lower for APC'. The dependent variable *number of defective work pieces* (DP) can be computed directly from the dynamic simulations of the respective task sequences (H_{07}: $DP_{APC} => DP_{Conv}$).

H_8: 'The portion of knowledge and skills is significantly higher for APC'. The dependent variable *portion of KS* (KS) is defined as the cumulative product of execution times of task elements and relevant KS characteristics in relation to the total time on tasks (H_{08}: $KS_{APC} <= KS_{Conv}$).

The Petri Net simulator was used to investigate these different comparative hypotheses by the use of a three way analysis of variance (MANOVA). Independent variables were the technology (milling/3D-welding), the work system (APC/conventional manufacturing) and the notice of human error (with HE/without HE). Afterwards, for each level of independent variables 232 simulation repetitions were computed [Goldmann et al., 1998].

For the computed simulation repetitions of APC, the mean total time on task $TT_{APC.milling}$ is 3.05 h (s = 0.27 h) for the technology of 5-axis milling. The mean total time on task for 3D-laser welding is $TT_{APC.welding}$ = 3.52 h (s = 0.51 h). In contrast, the mean total time on task for conventional milling is $TT_{Conv.milling}$ = 4.97 h (s = 0.6 h) and $TT_{Conv.welding}$ = 5.59 h (s = 0.97 h). The corresponding main effect *work system* results in a very significant difference F(1,1848) = 5012.323, whereupon the null hypothesis H_{05}: $TT_{APC} >= TT_{Conv}$ is rejected. By comparison, the results from inferential statistics show that single-operator APC have an efficiency of more than 37 % on total time on task. This is no surprising result as flexibility in scaffolding for machining centers is beyond controversy and hardly to surpass by any other machine tools. Both time-consuming clamping and machine adjustment of several machines and transport processes have been minimized [Boetz, 2001]. Especially planning is influenced by the key idea of task integration for well skilled workers by the use of highly developed computer-aided tools. Their experience is asked for alteration planning in order to ensure a quick and independent reaction in case of disturbances [Eversheim, 1998].

When the portion of human errors is simulated a significant influence of the effect *human error* on the total time on task is calculated. The median portion of human errors for 3D-laser welding in APC has been computed to $HE_{APC,welding} = 3.36\ \%$ and $HE_{APC,milling} = 1.23\ \%$. The median portion of human errors for conventional 3D-laser welding was calculated to $HE_{Conv,welding} = 3.8\ \%$ and $HE_{Conv,milling} = 1.39\ \%$ for conventional milling. However, double interaction between effects of *human error* and *work system* is not significant $F(1,1848) = 1.074$. Consequently, the null hypothesis H_{06}: $HE_{APC} => HE_{Conv}$ can not be rejected. On the one hand, it seems that the operators' tasks may become simplified since they are merely a supervisory controller of a highly automated manufacturing cell [Sheridan, 1997]. On the other hand, cognitive functions are much more complex and therefore tend to be more error prone than they were before computer-supported process planning came into play. Concerning the chosen approach for human error assessment the influence of system factors like information feedback, quality of user interface, etc. were not regarded in our study and must be derived by additional comparative investigations of developed APC-planing tools vs. conventional user-interfaces. Consequently additional simulation experiments have to be performed to prove these thesis in the near future. The median portion of human errors for 3D-laser welding both in APC and conventional manufacturing is obviously exceeding the ones in 5-axis milling almost three times. Referring to results of inferential statistics the basic reason can be seen in significant differences in estimations for difficulties of reference manufacturing orders and resulting task complexities, which are expressed by significant differences in all three main effects. Indeed, experts sensed the difficulty of reference order in 3D-laser welding almost 2.75 times harder than in 5-axis milling. Besides, the between subject effects of *technology/system* and *technology/human error* show significant differences with respect to task difficulty and task complexity. Thereby, a stronger relief of operator's task spectrum in favor of APC with respect to 5-axis milling could be stated. But we have to keep in mind, that the automated time share of 5-axis milling is significantly larger as pure processing time in 5-axis milling exceeds processing time in 3D-laser welding 18 times.

Although no significant influence of human error on total time on task in favor of APC vs. conventional manufacturing could be stated, a significant difference in terms of the number of defective work pieces is shown $F(1,1848) = 2127.635$. Thereby, the statistic mean of defective work pieces could be reduced from $DP_{Conv,welding} = 3.28$ ($s = 1.58$) and $DP_{Conv,milling} = 0.21$ ($s = 0.13$) to $DP_{APZ} = 0.03$ both 3D-laser welding and 5-axis milling. Consequently the null hypothesis H_{07}: $DP_{APC} => DP_{Conv}$ is rejected. The significant difference in double interaction of *technology* and *work system* $F(1,1848) = 1863.276$ can be attributed to the technological level of research and development as 3D-laser welding is a relatively modern technology and therefore optimization potential is larger. This leads to the assumption of fault-tolerant systems, which are one core postulation of APC in order to achieve a long term and trouble-free manufacturing process. The fundamental idea of fault-tolerant systems is neither to reduce resp. to avoid human errors nor to tolerate them, but to prevent their consequences [Seifert, 1992]. By doing so, all three levels of operator assistance by Rouse & Morris (1987) were taken into account, whereby the operator will be directly informed on the lowest level about momentary consequences of actual activities, e.g. syntactical control in parametering 3D-laser welding sensors. The next level informs the operator about future consequences e.g. graphical simulation of NC-programming in terms of possible clamping collisions. The third level deals with *intelligent* control, identification and recovery of possible system disturbances as there is e.g. avoidance of rattle in 5-axis milling. Thus, on the more consciously controlled rule-following level, development of know-how and rules-of-thumb is depending upon a basic variability and opportunity for experiments to find shortcuts and identify convenient and reliable signs which make it possible to recognize recurrent conditions without analytical diagnoses [Rasmussen, 1990]. According to our results, APCs seem to support this paradigm, which obviously manifests in the development of automation assistance with operators' larger scope of decision.

Surprisingly, the median portions of knowledge and skills in terms of expertise have been significantly decreased both in APC 3D-laser welding and APC 5-axis milling. Consequently, the null hypothesis H_{08}: $KS_{APC} <= KS_{Conv}$ can not be rejected. This seems to be a quantitative confirmation for the *irony of automation* [Bainbridge, 1987] of highly computerized manufacturing systems. On the one hand, it seems that the operator's role may have become simplified since they are merely assigned the task of controlling a highly developed technical system running within a well-adapted organizational context, and for which they have received adequate training. On the other hand, it is also true that cognitive and social aspects of their roles are far more complex than previously when deviations to the normal system functioning occur, because of the new performance pressures. Also an ordinal interaction between the effects of technology and work system can be observed. Reasons for this can be traced back to the facts of different levels of order difficulties and the stage of technological research and development with respect to 3D-laser welding.

OUTLOOK

Our future research will focus on the simulation of autonomous manufacturing crews with respect to different APC layouts. In the long-run, these facets should aim at a simulation tool for the human centered design and management of autonomous and cellular manufacturing systems. Independently future research should aim to develop suitable HRA-techniques especially for industrial manufacturing. The fully valid mathematical formulation of HEPs in terms of the different cognitive control levels will be the major difficulty, which however is very important for prospective evaluation. In particular, the provision of human variability especially in knowledge-based tasks (e.g. operator is using short-cuts) is missing. Currently, available techniques are only qualified for first coarse estimations.

ACKNOWLEDGEMENTS

The research is funded by the German Research Foundation according to the collaborative research center no. 368 *Autonomous Production Cells*.

REFERENCES

Allwood, C. M.: Error detection proceses in statistical problem solving. Cognitive Sience, 8, 1984.

Bagnara, S.; Stablum, F.; Rizzo, A.; Fonatana, A.; Ruo, M.: Error detection and correction: A study on human-computer interaction in a hot strip mill production planning and control system. In: Preprints of the first European Meeting on Cognitive Science Approaches to Process Control, France, 1987.

Bainbridge, L.: Ironies of Automation. In Rasmussen, J.; Duncan, K.; Leplat, J. (Eds.): New Technology and Human Error. New York: John Wiley & Sons, 1987.

Boetz, V.: Kompettbearbeitung auf CNC-Bearbeitungszentren – Technik der Zukunft?. VDI-Z 143 Nr. ½, 2001.

Brauser, K.: ESAT – Ein neues Verfahren zur Abschätzung der Menschlichen Zuverlässigkeit. In Gärtner, K.-P. (Ed.): Menschliche Zuverlässigkeit. Tagung VDI und DLR: München, 1992.

Daalen, C.; Thissen, A.; Verbraeck, A.: Methods for the Modeling and Analysis of Alternatives. In Sage, A. P.; Rouse, W. B. (Eds.): Handbook of Systems Engineering and Management. Wiley & Sons, 1999.

Eversheim, W.; Luczak, H.; Haufe, A.; Klöwer, S.; Schlick, C.: Human-oriented Manu-facturing in Autonomous Production Cells. In: Fink, P.; Koningsveld, E.; Dhondt, S. (Eds.): Human Factors in Organizational Design and Management – VI. Amsterdam: Elsevier 1998.

Goldmann, D.; Nelson, B. L.: Comparing Systems via Simulation. In: Banks, J. (Ed.): Handbook of Simulation. New York (NY): John Wiley & Sons, 1998.

Jensen, K.: Coloured Petri Nets. Basic Concepts, Analysis Methods and Practical Use. Volu-me 1. 2nd edition, Berlin: Springer-Verlag 1997.

Kirwan, B.; Kennedy, R.; Taylor-Adams, S. & Lambert, B.: The Validation of Three Human Reliability Quantification Techniques – THERP, HEART and JHEDI: Part2 – Results of Validation Exercise. Applied Ergonomics, 28(1), 1997.

Luczak, H. : Task Analysis. John Wiley & Sons, New York 1997.

Majchrzak, A.: The Human Side of Factory Automation. San Francisco: Jossey-Bass Publishers, 1988.

Park, S. P.: Human Error. In G. Salvendy (Ed.): Handbook of Human Factors and Ergonomics. New York: J. Wiley, 1997.

Poucet, A. (Ed.): HF-RBE Human Factors Reliability Benchmark Exercise – Synthesis Report (EUR 12222 EN). Luxembourg: Commission of the European Communities, 1989.

Rasmussen, J.: The Definition of Human Error and a Taxonomy for technical System Design. In Rasmussen, J. et al. (Eds.): New Technology and Human Error. Wiley & Sons Ltd., New York, 1987.

Reason, J: Human Error. Cambridge u. a.: University Press, 1990.

REFA (Ed): Methodenlehre der Planung und Steuerung. Teil 3., München: Carl Hanser, 1991.

Rizzo, A. Bagnara, S.; Visciola, M.: Human Error detection processes. In Mancini, G.; Woods, D.; Hollnagel, E. (Eds.): Cognitive Engineering in Dynamic Worlds. Ispra, Italy: CEC Joint Research Centre, 1986.

Rouse, W. B.; Morris, M. N.: Conceptual Design of a Human Error Tolerant Interface for Complex Engineering Systems. Automatica, Vol. 23, No.2, 1987.

Schlick, C. ; Reuth, R. ; Luczak, H. : ActiveUI - Gestaltung einer multimodalen Benutzungsschnittstelle der virtuellen Realität. In: 46. Arbeitswissenschaftlichen Kongress, GfA Press,Dortmund 2000.

Schlick, C.; Reuth, R. Luczak, H.: Comparative Simulation Study of Work Processes in Autonomous Production Cells. Human Factors in Manufacturing. John Wiley & Sons, 2001 (in print).

Seifert, R.: Maßnahmen zur Erhöhung der Zuverlässigkeit. In Bubb, H. (Ed.): Menschliche Zuverlässigkeit. 1 Auflage. Landsberg: ecomed-Fachverlag, 1992.

Sheridan, T.B.: Supervisory Control. In G. Salvendy (Ed.): Handbook of Human Factors and Ergonomics. New York: J. Wiley, 1997.

Zimolong, B.: Fehler und Zuverlässigkeit. In Hoyos, C.; Zimolong, B.(Eds.): Ingenieurspsychologie, 1990.

Zimolong, B.; Dörfel, G.: Ursachen und Klassifizierung menschlicher Arbeitsfehler. In Bubb, H. (Ed.): Menschliche Zuverlässigkeit. 1 Auflage. Landsberg: ecomed-Fachverlag, 1992.

ADAPTIVE PILOT MODELING FOR COCKPIT CREW ASSISTANCE: CONCEPT, REALISATION AND RESULTS

P. Stütz [(1)] and R. Onken [(2)]

[(1)] ESG -GmbH
Einsteinstr. 174
81675 München, Germany
phone: +49-(0)89-9216-2604
fax: +49-(0)89-9216-2632
e-mail: pstuetz@esg-gmbh.de

[(2)] Universität der Bundeswehr
Werner-Heisenberg-Weg 39
85579 Neubiberg, Germany
phone: +49-(0)89-6004-3452
fax: +49-(0)89-6004-2082
e-mail: reiner.onken@
unibw-muenchen.de

ABSTRACT

Monitoring of crew behaviour is a vital part of the internal situation assessment process performed within Crew Assistant Systems. A prerequisite for this is the availability of a behaviour models..

The paper describes the concept, realisation and evaluation of an adaptive behaviour model focused on rule based pilot activity. It autonomously acquires and reproduces situation-dependent *individual* pilot behaviour during plan execution.

The model concept provides a hybrid architecture. Normative behaviour, describing deterministic pilot behaviour, is implemented by use of *Petri Nets*. A learning module utilises a *case-based reasoning* mechanism to customise the transition behaviour within the Petri Net structure, thereby achieving online model adaptivity. For this purpose it accesses observed pilot action events stored in a case base.

The second part of the paper goes into detail on the experimental investigation performed. The results proved the basic functionality of the adaptive pilot model as well its capability to predict situation-dependent pilot actions.

INTRODUCTION

A variety of authors [2][4] investigate on the reasons for human error in aviation. The results gained so far can be divided in the following categories:

- Loss of situational awareness

  The sudden appearance of unusual situations force the pilot to change from his observing role to an active one. It is evident that in a lot of accidents the crews where suffering from insufficient awareness concerning aircraft state (*flight path awareness*, *terrain awareness*, *energy awareness*) and aircraft systems (*mode awareness*) leading to erroneous crew actions. Often even the basic necessity to act is not clear (*controlled flight into terrain*).

- Loss of piloting skills

  Today's cockpit avionics favour the loss of sensomotorical and cognitive skills by the crew. Training in simulators only provides limited compensation.

- Loss of self-criticism

  The appropriate time for the pilot to disengage from malfunctioning avionics systems is not recognised. Instead even more resources are wasted finding an explanation for the unintelligible behaviour of the system.

So called *cognitive assistant systems* aim to eliminate these deficiencies by providing hints, warnings and situation-adapted problem solutions. A prerequisite for such intelligent support is a thorough and comprehensive understanding of the overall situation. This includes the aircraft and the environment as well as the crew.

The principal goal is to ensure that the crew is aware of the currently most urgent task and to provide relevant automation to carry out this task in cases when the crew is over-taxed. Mutual understanding of objectives and resources both on the machine and on the human side seems to be a prerequisite for such a symbiotic man-machine relationship, in order to prevent pilot error and to enhance mission success.

The Cockpit Assistant System CAMA (*Crew Assistant Military Aircraft*) is a functional prototype developed according to principals. This knowledge-based system is designed to support military transport crews performing logistic and tactical missions. Financed by the German Ministry of Defence, CAMA was integrated in a flight simulator and a test aircraft [7]. The following investigations on behavioural modelling were carried out during the years 1994 till 1998.

BEHAVIOUR MODELING

As indicated before, the electronic assistant has to go through the same process of situation assessment as the crew in order to provide efficient support. Moreover the system must also be able to derive conclusions on behavioural aspects relating to its human counterpart. In other words, it has to develop an understanding of what the crews has to do and must not do in specific situations.

Monitoring of crew behaviour can be accomplished by comparison between expected and actual pilot actions. Whenever erroneous behaviour is detected, the pilot is warned and the appropriate actions are recommended to the pilot. In addition information on expected behaviour can be used internally to anticipate upcoming phases of high workload and to deduce inherent pilot intents. It is obvious that for the generation of such expected pilot behaviour an ample and reliable knowledge base is needed as reference. In the next paragraphs, the concept and the realisation of a respective *behaviour model* will be shown.

Concept

Other engineering disciplines (e.g. system design or construction) typically require overall models concerning bio-mechanical and cognitive behaviour. These models represent a whole range of behavioural aspects through averaging. In contrary a model suitable for individual pilot assistance also has to take into account individual aspects of behaviour, neglecting this would lead to sub-optimal results.

Fundamental for the concept of the proposed model is the assumption that normative regulations and procedures provide the guidelines for pilot behaviour, and can be described as deterministic pilot behaviour documented in pilot handbooks and air traffic regulations. This normative behaviour then is steadily amended and adapted by the individual pilot within certain tolerances in order to suit his needs and preferences.

In order to imitate this iterative refinement a process within the technical system had to be created allowing a continuous adaptation of a predefined normative core knowledge base by learning from observed pilot activities, thereby establishing the *adaptive pilot model*. This customised knowledge base than allows individually correct statements on upcoming pilot actions regarding the actual situation (Figure 1).

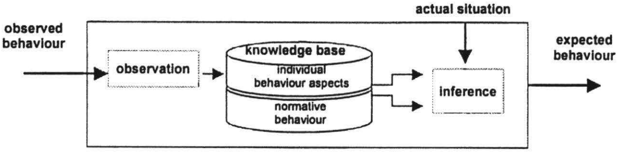


Figure 1: Hybrid concept for a adaptive pilot model combining normative behaviour with individual aspects

Realisation

Realisation of the adaptive model was performed in two major steps. At first the normative model had to be implemented. In a subsequent step an extension providing the adaptive capability was added.

Normative model

Pilot behaviour can be separated into situation assessment and action processing components. Behaviour modelling is performed for all flight segments (taxi, takeoff, departure, IFR/cruise, tactical flight, drop, approach, landing) and concerns the following tasks:

a) situation assessment

- recognition of current flight segment and process of plan execution related to flight plan and procedures

b) pilot actions:

- primary flight guidance (altitude, course, airspeed, power setting, ...)
- operation of flaps, gear, speed brakes, radios

In his analyses [5] showed that *Petri Nets* are the most suitable representation for this mainly rule-based behaviour because of their ability to formulate concurrent, discrete event driven procedures.

Petri Nets:

Petri Nets can be described as a graphical representation of a net graph based on a mathematical theory which enables the analytical verification of system properties. A typical Petri Net consists mainly of the following net primitives:

- *Places*:

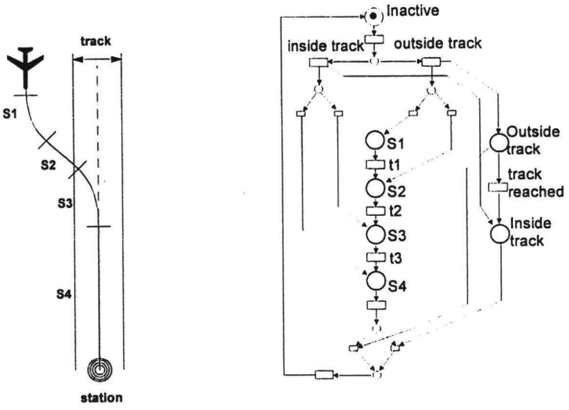
  Discrete states are represented by places. Examples are flight segments ("final approach"), conditions for subsequent actions ("turn right after passing A") and states of discrete aircraft systems ("flaps 20 degree").
- *Transitions*:

  Transitions are used to represent situation state transitions, e.g. between flight segments ("final approach → landing") and discrete aircraft systems ("landing gear up → down"). These transitions are typically evoked by fulfilled conditions ("Altitude higher 5000 ft")
- *Tokens*

  Tokens symbolise the current net state as marks on the relevant places.

Example:

An *interception* is carried out to reach a given (magnetic) course, which leads to a target station (e.g. a radio navaid). This can be required within published departure or approach procedures or can be commanded by air traffic control. In the general case, an interception covers 4 sections (see fig. 2): turning to intercept heading (S1), maintaining on intercept heading (S2), turning to the given course (S3), tracking of the given course (S4). Sections are skipped if the aircraft fulfils the characteristics of a following section, e.g. if the aircraft is already on intercept heading at the time the procedure is started, section S1 is skipped.



a) intercept procedure b) Petri Net

Figure 2

Adaptive Extension

The second step was to find an appropriate mechanism to automatically adapt the normative statements of the petri net structure mentioned above to individual habits. Several algorithms for *machine learning* and *example based learning* were evaluated. Finally the method of *case based reasoning* was chosen and implemented in a case learning module complementing the normative model core [6][3]. In this module observed pilot action events are stored in a case base and attributed according to their coherence with state transitions in the Petri Net structure. On demand the Petri Net interpreter is able to recall these cases using similarity considerations during runtime, in order to refine its state and transition parameters within given tolerances. This functionality ensures full online adaptivity, but simultaneously considers the primacy of the normative model.

Figure 3 shows modules and functions of the adaptive piloting model. The functions are explained in the following.

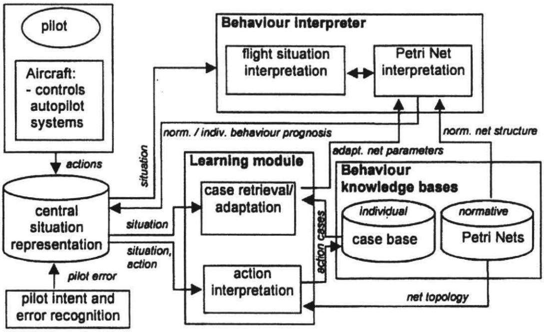


Figure 3: Modules and embedded functions of the adaptive pilot model

Behaviour interpreter

This module represents the real time processor of the Petri Net system. It loads the normative behaviour model stored in a Petri Net description language during system startup. The *Petri Net interpretation* function manages Petri Net states and transitions. State and transition parameters subjected to adaptive modelling are received on request during runtime from the learning module. The transition parameters, either purely normative or adapted are then handed over to the *flight situation interpretation* function to be monitored. When a condition is fulfilled, the Petri Net interpretation is notified and the transition takes place.

Action interpretation

This function supplies the case base with a constant stream of discrete pilot action events. Detected actions are assigned to appropriate state transitions in the Petri Net structure. Therefore this function uses information on net topology. In order to make these action cases useable for a later case retrieval, additional situation attributes under which the action took place are recorded.

Table 1: Case data stored for a course change manoeuvre

Action description		**Situational attributes**			
		required		supplementary	
pre (heading):	337°	*dist_track* (cross track):	0.66 nm	*alt*:	12180 ft
post (heading):	300°	*dist_basis* (distance to wpt.):	10.21 nm	...	
type:	Intercept	α (angle bet. current and next leg):	-56°		
		β (angle bet. planned track and acft. Hdg.)	14°		
		ias (speed):	200 kts		
		flight phase:	Enroute		

Table 2: Case data stored for flap setting action

Action description		**Situational attributes**			
		required		supplementary	
pre (setting):	14°	*ias* (speed):	132 kts	*alt*:	1631 ft
post (setting):	1°	*flight phase*:	Departure	...	

Case base

Table 1 and Table 2 show excerpts of data stored for example action cases. A commercial database tool is used for case storage management. It provides parallel read/write access to the database for case storage and retrieval, thereby enabling the continuous refinement of the knowledge base during runtime. SQL (structured query language) is used for communication.

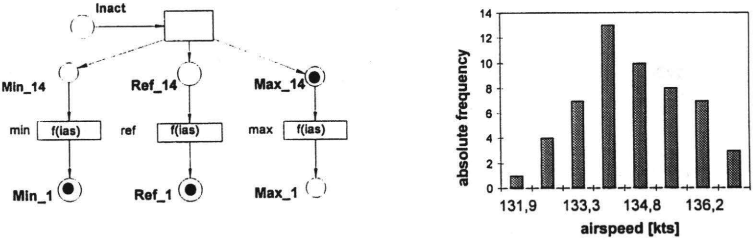
Case Retrieval and Adaptation

This function provides the on-line case base access for the Petri Net system in a way to preserve the overall normative task sequence, but also to take into account individual, admissible deviations. For illustration, it is considered that the Petri Net preconditions of a state-transition (that means a place is occupied by a token) are fulfilled.. The transition condition can now be acquired during runtime from the examples in the case base given for the individual pilot. It is thereby making use of observations which were just recently collected.

Referring to Table 1 and Table 2, the 'transition problem' is described by the pre and post state of the transition. This is passed along with other net status information to the *retrieval stage*. Finding relevant cases can be either *similarity based* or *trivial*.

Trivial case usage:

Trivial usage typically can be found for certain system setting inquiries (e.g. flaps, gear). In this case the simple check for pre and post state identity suffices to retrieve one or more example cases. An example is shown in Figure 4. The net *flaps_departure* models the flap setting behaviour during the departure phase. The net represents a range of possible settings through the run of 3 independent tracks. They symbolise the earliest possible (*Max_x*), the typical (*Ref_x*) and the latest possible transition (*Min_x*), dependent on the actual airspeed. In only normative mode this would correspond to the prescribed flap settings according to stall and maximum flap extension speed. The adaptive model replaces this for the earliest, typical and latest flap setting action considering its experience with an specific pilot. In order to do so, a statistical evaluation is done on all known flap reductions from 14° to 1° for this particular pilot (Figure 4b). An airspeed higher then the 15% percentil (132.5 kts) would allow the *Min*-Transition to fire. Firing of the *Ref*-and the *Max*-transition respectively is triggered by the median (134 kts) and the 85% percentil (135.6 kts) value. Figure 4a shows the Petri Net state at an actual airspeed of 135 kts.



a) fragment from net *flaps_departure* modelling the retraction from 14° to 1°

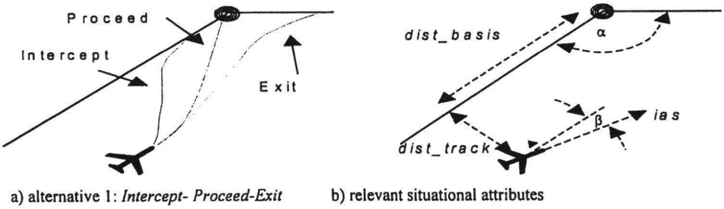
b) histogram of airspeed values when flap reduction form 14° to 1° was observed during departure (54 cases)

Figure 4: Petri Net and frequency histogram for departure flap setting

Similarity based case usage:

The above given example for trivial case usage only took airspeed into consideration as transition parameter. However, a variety of state transitions depend on the occurrence of more complex situations. It is a characteristic for these *multivariate* transitions that the set of parameters which influence the state transition is known; their absolute value and their functional relation, however, has to be estimated. In this case *similarity based* case usage has to be conducted.

Figure 6a shows an example for a more complex decision task. Given a lateral deviation from the current leg of the planned track, the pilot has to decide whether to intercept the current leg (*Intercept*), to directly proceed to the next waypoint (*Proceed*) or to disregard the current leg and to steer towards the following leg (*Exit*). Relevant situational information (Figure 6b) for this decision is considered to the aircraft's relative position (*pos* to the current flight plan leg), defined by the cartesian values *dist_basis* and *dist_track*, the aircraft's speed (*ias*), the aircraft's current heading relative to the actual track (angle *β*) and the geometrical constellation between the current and the following leg represented by the angle *α*.



a) alternative 1: *Intercept- Proceed-Exit*

b) relevant situational attributes

Figure 5: decision alternatives after considerable cross track deviation

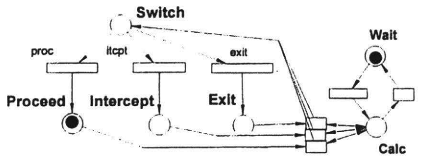


Figure 6: Petri Net fragment from net *Basic_Tracking_Off_Track*

Figure 6 shows the respective Petri Net indicating that the pilot is most likely to proceed directly to the next waypoint. Finding a normative functional relation between the relevant parameters and the final guidance decision is hard to achieve, as objective rules and recommendations do not exist. This freedom typically favours the formation of individual pilot customs and a rigid normative model looses its validity.

The way to tackle this problem in the sense of case-based reasoning is to see in which situation (decribed by the above mentioned parameters) a pilot made these guidance decisions and to conclude that he will come to the same decision again in a similar situation.

The question arises how situational similarity can be determined. In general the situational description of an observed pilot action can contain *metrical* (e.g. position, altitude), *ordinal* (e.g. flight phase) or *nominal* attributes (e.g. waypoint idents). In the given example metrical attributes prevail. For this group of attributes similarity considerations traditionally utilise *distance* measurements, the *euclidian distance* being the most well known measure. Formula 1a shows the relation between distance d and similarity *sim* for two attribute values a and b, whereby d_{max} is the maximum allowable distance. If distance d decreases to zero, similarity raises to 1 and identity between a and b is reached. More information on the determination of similarity especially for attributes other than metrical can be found in [1].

$$sim(a,b) = 1 - \frac{d(a,b)}{d_{\max}}$$

a) relation between distance and similarity

$$SIM() = \frac{w_{pos} sim_{pos}() + w_{\alpha} sim_{\alpha}() + w_{ias} sim_{ias}() + w_{\beta} sim_{\beta}() + w_{fph} sim_{fph}()}{w_{pos} + w_{\alpha} + w_{ias} + w_{\beta} + w_{fph}}$$

b) local similarities *sim* contributing to global similarity *SIM: w: local similarity weight*

Formula 1: similarity measurements

In order to use this approach to compare stored case attributes with the current situation, local similarity values are computed individually for all attributes. After that these values are combined to a global similarity *SIM* through a weighted combination. For the given example the global similarity is made up from the local similarities for position, alpha, airspeed, flight phase and waypoint ident. Note that the last attribute *(fph)* is not metrical (Formula 1b).

Assuming that the aircraft for some reason deviates from the planned track violating a certain threshold. Here the Petri Net interpreter invokes the aforementioned net *Basic_Tracking_Off_Track* and issues a request to the retrieval and adaptation function together with a description of the actual situation in order to conclude for the most likely pilot reaction. This function now tries to retrieve suitable course manoeuvre cases and sorts them according to their situational similarity. The action type of the most similar case (e.g. Intercept) is than passed back to the interpreter and the respective transition is allowed to fire.

EXPERIMENTAL INVESTIAGTION

The adaptive model was implemented in a workstation environment and validated in simulator trials. The subjects were 5 professional air transport pilots with 800 to 2800 hours flight experience. After familiarisation with the fixed base research simulator was assured, a variety of experiments were conducted in order to verify and validate the functions of the adaptive pilot model. The following chapters only focus on the results found on the topic of similarity-based case usage.

Scenario and tasks

In order to assess the prognostic capabilities of the adaptive model, the pilots lateral aircraft guidance behaviour was investigated. A scenario was set up which allowed to repeatedly provided the pilot with off track situations as described in Figure 6. Each pilot had therefore to conduct four IFR flights, two from Frankfurt (EDDF), the other two from Friedrichshafen (EDNY) to Stuttgart (EDDS). Duration of each flight was about 40 minutes. Once the cruise altitude (FL 120) was reached, the pilots were assigned *radar vectors ("Lucky07, turn left (right) heading xxx")*, typically given through ATC in order to maintain air traffic separation or to avoid bad weather areas. After these radar vectors forced the pilots to deviate considerably from the pre-planned track, they were requested to disregard the assigned heading and to follow the original flight plan again (*"Lucky07, proceed as filed" or "Lucky07, resume own navigation"*). In this situation, the pilots had to decide whether to intercept the closest flight plan leg, to proceed directly to the respective waypoint or to steer towards the following leg corresponding to Figure 6a and b.

The pilots were able to control the aircraft via sidestick and autopilot. A glass cockpit like primary flight display was used to indicate airspeed, altitude, vertical speed and aircraft altitude. On a navigational display the flight plan, radio navaids and navigational instrument indications were depicted in moving map format. On a third display the flight plan was shown in a alphanumerical flight log format. All pilot actions were recorded with their attributes by the *action interpretation function* as explained above.

Results

Figure 7 shows as an example two tracks flown by Pilot 4 from Friedrichshafen (EDNV) to Stuttgart (EDDS), They deviate from the pre-planned track due to the radar vectors given by (simulated) ATC. The diamonds depict the course manoeuvres detected by the *action interpretation* function. These pilot actions were automatically classified according to characteristic features (e.g. start and end heading) as being *Proceed*, *Intercept* or *Exit* actions. An appropriate decision table was used.

The results of such classification can be seen in Figure 8a. The diagram plots all action cases in a common co-ordinate system defined by the attributes *dist_basis* and *dist_track* (see Figure 6d). This representation supports an easy comprehension of the positional relation of the actions recorded. The accumulation of dedicated action types in certain areas suggest that the aircraft's relative position has a significant influence on the pilots decision. A demerger of the remaining overlapping zones can be assumed in other dimensions of situation space (e.g. *ias*, α).

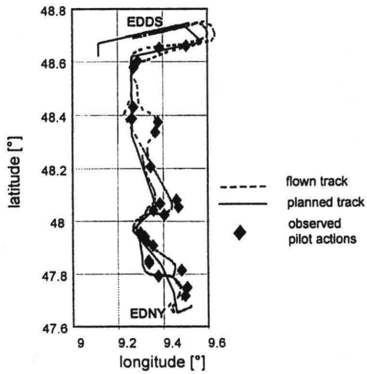


Figure 7: pilot actions (diamonds) plotted according to their geographical position during observation

General similarity computation

In order to validate the similarity based case usage mechanism, clinical test requests were sent to the *retrieval and adaptation* function after the case acquisition phase, effectively asking the function "which action case out of the case base is most similar to the test situation and which action type is associated with it".

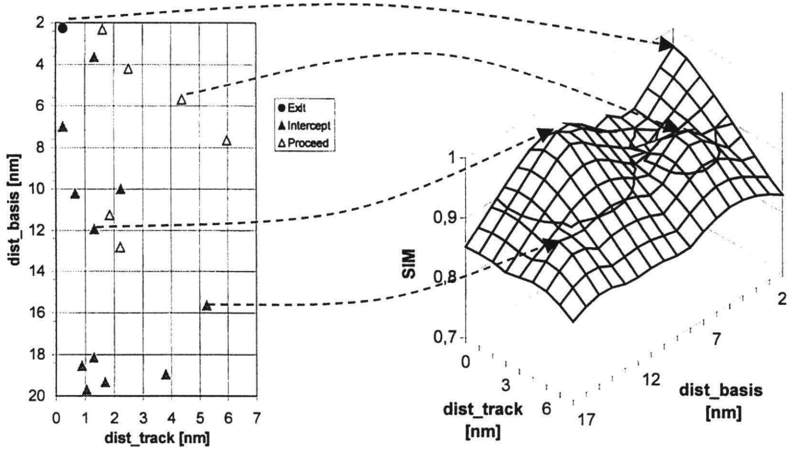
As designed, the *retrieval and adaptation function* calculated and sorted the global similarity of all known *Exit*, *Intercept* and *Proceed* action cases to each provided test situation. Using Formula 1b the global similarity *SIM* measures were computed by a weighted combination of the local similarities *sim* as given in Table 3.

Table 3: weight and variation range of attribute values

attribute	**value**	**weight**	**d$_{max}$**	***sim***
dist._track *dist_basis*	0 – 8 nm 2 – 24 nm	10	26 nm	$sim_{pos} = 1 - \frac{\sqrt{(dist\_track_C - dist\_track_S)^2 + (dist\_basis_C - dist\_basis_S)^2}}{d_{max}}$
α	-60° – 30°	5	180°	$sim_\alpha = 1 - \frac{\alpha_C - \alpha_S}{d_{max}}$
β	0°	5	180°	$sim_\beta = 1 - \frac{\beta_C - \beta_S}{d_{max}}$
ias	200 kts	1	100 kts	$sim_{ias} = 1 - \frac{ias_C - ias_S}{d_{max}}$
fph	-	*0*	-	-

Figure 8b shows the *similarity surface* gained for array of test requests applied on the case base of pilot 5, where the situation was described by varying *dist_track* and *dist_basis* values in steps of 1 nm. However, α was fixed to -45°, β to 0° and *ias* to 200 kts for all requests.

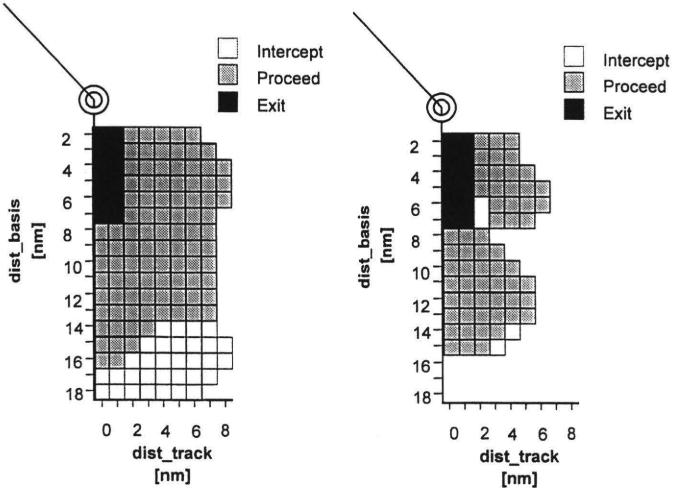
The *similarity surface* represents the global similarity values of the most similar stored action cases compared to the given situation descriptions. Cone-like structures show the strong affinity of certain situational regions to specific stored action cases. These dominant cases (marked by arrows connecting Figure 8a and b) graphically reside directly 'under' the tip of the cones in *dist_track – dist_basis*-situation space. These cases yield a strong positional similarity d_{pos} close to 1. The maximum elevation of each cone finally depends on the similarity values gained for the other, basically static situational attributes.



a) classified pilot actions plotted according to their relative position to the planned track (Pilot 5)

b) Global similarity surface; arrows depict the influence of dominant action cases

Figure 8: general similarity measurement for an array of test requests with 0 nm < *dist_track* < 8 nm, 2 nm < *dist_basis* < 18 nm, α = -45°, β = *0*° and *ias* = 200 kts



a) $SIM_{min} = 0.85$

b) $SIM_{min} = 0.90$

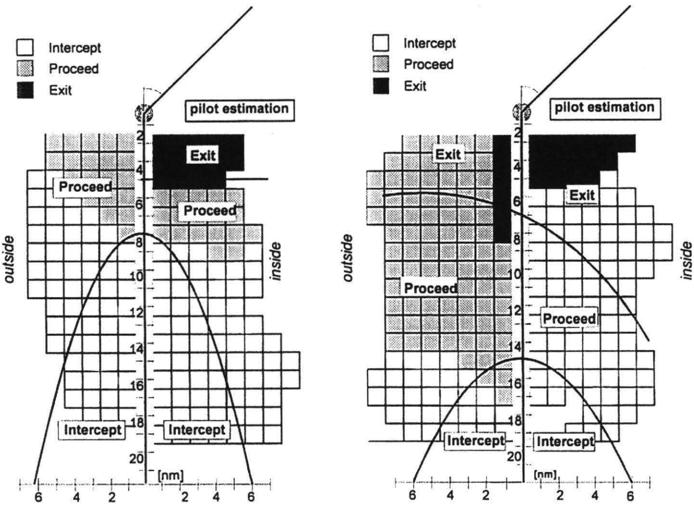
Figure 9: action type array derived for test request array (Pilot 5)

It may be observed that the computed similarity values in Figure 8b decrease significantly in test situations with large *dist_track* and *dist_basis* attributes. This can explained by the absence of action cases with positional relevance in these areas. It is obvious that a minimum similarity value SIM_{min} is required in order to regard a stored case as relevant for a given situation.

Figure 9a and b denote the action types associated with the most similar cases derived for the test requests. In Figure 9a a minimum similarity SIM_{min} of 0.85 was required, in Figure 9a SIM_{min} was at 0.9. White spaces can be understood as areas in situation space where the pilot model refrains from providing a solution due to a lack of relevant case knowledge.

Individual behaviour differences

Figure 10a and b clearly show strong differences among the action type solutions produced by the adaptive model for specific pilots. Obviously the model indicates a much stronger tendency for Pilot 5 to directly proceed to the waypoint compared to Pilot 4 when positions on the *outside* are considered. Even at quite large *dist_basis* values the model does not favour the *intercept*-option. Another peculiarity is that the model does not foresee at all the *Exit* manoeuvre for pilot 4 on the *outside*. To get more insight on the validity of these model statements, the pilots were asked to prepare a rough subjective drawing indicating the areas of their manouveral preferences. Overlaid on the results of the pilot model, the drawings indeed confirm the effects mentioned. Pilot 5 admits himself a quite large area where he would rather choose the *proceed* option than to intercept the planned track, but the tendency of Pilot 4 to prefer the intercept option up to a distance of about 8 miles to the waypoint. This threshold value was stated identically by the pilot model, considering small *dist_basis* values. Likewise Pilot 4 ruled out the *Exit*-option for his course behaviour on the *outside* and the pilot model was able to reproduce the *Exit*-area for Pilot 4 on the *inside* almost identically.



a) Pilot 4

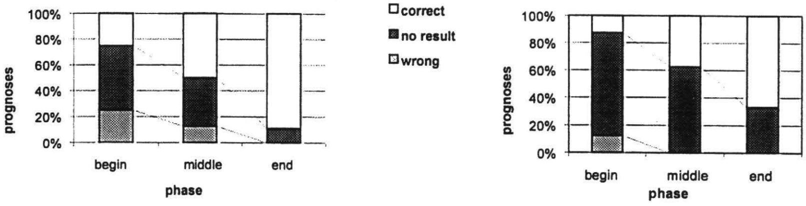
b) Pilot 5

Figure 10: action type array with $|\alpha| = -45°$, $\beta = 0°$, $ias = 200$ kts, $SIM_{min} = 0.85$; areas separated by curved lines depict the pilot's subjective estimate of own manouveral preferences; *inside* and *outside* are used as positional reference considering the turn direction of the subsequent planned leg.

Learning progress and prognostic performance

To further gain objective insight into the prognostic capability of the adaptive pilot model, the retrieval and adaptation function was requested to predict the action types of each action case A_n in the case base by only having access to cases A_1 to A_{n-1}, where n denotes the time order of case recording. Within the three phases *Begin*, *Middle* and *End* the prognosis results were evaluated as *correct*, *wrong* or *no result*. As described before the yield of *no result* depends on the setting of SIM_{min}. Figure 11a and b give the respective results. During the *begin*-phase of case acquisition the probability of finding good matching cases is quite low. Therefore *wrong* or *no results* are in the majority. In this phase a raised level of SIM_{min} biases the system towards *no result* as seen in

Figure 11b. In the following phases correct model prediction become prevalent, finally reaching a score of 90 % for correct predictions in the *end*-phase in Figure 11a.



a) $SIM_{min} = 0.85$

b) $SIM_{min} = 0.90$

Figure 11: Learning progress and prognostic performance during case acquisition phases; Pilot 5

CONCLUSION

The paper started with a description of the demand for pilot behaviour modelling within cockpit crew assistant systems. The need not only for normative models describing the general prescribed behaviour but also for individual models providing information on subjective preferences and customs was emphasised.

In order to realise an appropriate model, a hybrid concept was introduced. The concept uses Petri Nets as a representation of rule-based pilot behaviour in the area of plan execution. Transition behaviour within the Petri Net system then is customised to individual pilot preferences during runtime, using previously observed behaviour examples. Theoretical background for this example-based adaptation process is provided through the paradigm of *case based reasoning*. For closer investigation a prototype was implemented and experiments were conducted with professional pilot.

The pilots had to perform basic decision tasks concerning lateral aircraft guidance within an IFR scenario. The results verified the basic functionality of the adaptive model. Furthermore significant differences in behaviour characteristics were observed and the model's ability to predict this correctly was validated. Further investigation should be carried out expanding the areas of pilot tasks covered by the adaptive models.

REFERENCES

[1] Althoff, K.-D.; Auriol, E.; Barletta, R.; Manago, M.: *A review of industrial case-based reasoning tools*, AI Intelligence, Oxford, 1995

[2] Billings, C. E.: *Aviation Automation - the search for a human-centered approach.* Erlbaum Verlag, Mahwah NJ, 1997

[3] Kolodner, J: *Case-based reasoning*, Morgan Kaufmann, San Mateo, 1993

[4] Onken, R.: *Basic Requirements concerning Interactions in Combat Aircraft*, Workshop on Human Factors/ Future Combat Aircraft, Ottobrunn, 1994

[5] Ruckdeschel, W.; Onken, R.: *Modeling of Pilot Behaviour using Petri Nets*, 16th International Conference on Application and Theory of Petri Nets, Torino, 1995

[6] Stütz, P.; Onken, R.: *Adaptive Pilot Modelling within Cockpit Crew Assistance,* Cognitive Systems Engineering in Process Control (CSEPC), Kyoto, 1996

[7] Stütz, P.; Schulte A.: *Evaluation of the cockpit assistant military aircraft (CAMA) in flight trials*; 3rd International Conference on Engineering Psychology and Cognitive Ergonomics, Edinburgh, 2000

Support of Situation Recognition by Intention Driven Adaptive Sensing

Makoto Takahashi, Hiroki Tani, Tadashi Miyazaki, Djoko Hari Nugroho and
Masaharu Kitamura

Dept. of Quantum Science and Energy Eng., Graduate School of Eng., Tohoku University
Aramaki-Aoba 01,Sendai,Miyagi,980-8579 JAPAN
makoto.takahashi@qse.tohoku.ac.jp

SUMMARY

A framework of dynamic location of sensors, called intention driven adaptive sensing (IDAS), has been proposed with the emphasis on the integration of an inference mechanism with the active sensing module. In the present study, the validity of the proposed framework of intension driven adaptive sensing has been evaluated using the experimental setup, in which the anomalies can be simulated in real systems and the actually moving robot with exchangeable sensors can conduct dynamic measurements. Through a set of experiments using experimental setup, it has been confirmed that the causes of anomalies have been successfully identified by combined use of fixed sensors and mobile sensors.

1. INTRODUCTION

The appropriate situation recognition is quite important to achieve higher level of safety in large scale process systems, such as a nuclear power plant. Although many research efforts have been conducted to develop a diagnosis system to help operators by effectively and quickly identifying a cause of failure[1] , few of them have been applied to the practical operational situations and accepted by operators preferably. This is partly because that most of the diagnosis system have been designed under the assumption that signals required for diagnostic inference are always available and the signal characteristics are predetermined. In real situations, however, the availability of all of the required information is not always guaranteed and the contents and the grain size of referred information strongly depend on the current goal of the operators. In this regard, the authors believe that the sensing process itself should be improved in relation to the operators' diagnostic behaviors to realize a human-machine cooperative working environments. In other words, a mechanism of dynamic location of sensors according to the current diagnostic context should be provided for dealing with dynamically changing situations. The framework of dynamic location of sensors is called intention driven adaptive sensing (IDAS), with the emphasis on the integration of an inference mechanism with the active sensing module. In this framework of IDAS, the sensing and diagnostic inference are not independent any more but inevitably interact each other. The validity of the proposed framework of intension driven adaptive sensing has been evaluated using the experimental setup, in which the anomalies can be simulated in real systems and the robot with exchangeable sensors can make dynamic measurements.

2. METHODS

In this chapter, the basic concept of IDAS is described in detail, followed by the brief summary of the Bayesian network as the basis for the system model representing the cause-consequence relationships.

2.1 Intension driven adaptive sensing

The conventional AI methods generally assume that signals required for diagnostic inference are always available and the signal characteristics are predetermined. In practical situations, however, when expert operators face emergency situations, they do not wait to make decisions until all of the required symptom information become available. Instead, they try to find out tentative actions to take based on the available information and then proceed to make further measurements for obtaining additional information. In case of a nuclear power plant, for example, additional information can be obtained by accessing the CRT windows and/or by visiting the chart recorder. In some cases, the inquiry for the on-site personnel through in-house telephone is required for retrieving the important information. If sensors were installed at all points possibly required by the operators and all information were displayed on the instrumentation panel in a operation room, operators would have less difficulties in accessing the data required for decision making. However, this is an unrealistic situation in two aspects. First, it is impossible to install sensors at all possible points because all of the possible anomaly situations cannot be foreseen in advance. Second, a cost of measurements becomes huge if all sensors are installed as required by the situation assessments. One possible solution for this sensor allocation problem is to introduce a mobile sensing mechanism, by which the sensing points and the sensing modality can be altered according to a diagnostic context. Although the mobile sensing mechanism manually operated by operators themselves may be useful and effective for obtaining required information, an increase of operator workload is inevitable especially in time critical conditions. In the proposed framework of IDAS, the selection of sensing modality and sensing location is controlled by the intelligent agent, considering the current diagnostic context and the operators' intension. Three forms of interactions between IDAS and operators, shown as follows, are assumed.

a. Auto mode

In this auto mode, the failure identification process is performed fully in automatic manner. After initiating diagnosis process by detecting any symptom such as deviation from normal range, the failure identification is performed based on the hypothesis generation and test strategy. The IDAS tries to identify a cause of failure primarily by the information from fixed sensors. If the cause is successfully identified at this stage, no additional sensing by mobile sensing mechanism is required. If the cause cannot be identified, in other words, several failures hypotheses remains on the possible cause list, then mobile sensing is activated. As the additional sensing is to be performed step-by-step manner in IDAS, the next step sensing point and modality must be determined. In this auto mode, they are determined based on the pre-defined knowledge of the objective system, which is represented by Bayesian network. It should be noted here that it is allowed for human operator to intervene at any time to change the focus of attention.

b. Supervisory mode

In supervisory mode, the specific failure hypotheses are determined by the operators themselves, reflecting their diagnostic strategy. In auto mode, the next step sensing point is determined based only on the value of information of the additional measurement. In practical situations, however, the priority of the diagnosis process is determined not only by the effectiveness of the identification process, but also by taking into account other aspects such as criticality of the specific failure, the possible consequences of the failure and the cost of the additional measurements, etc. Although these factors can be represented as the diagnostic strategy in the intelligent agents, the supervisory mode has been provided in the present system to reflect operators' diagnostic intension on the failure identification process. Operators can select one specific failure hypothesis

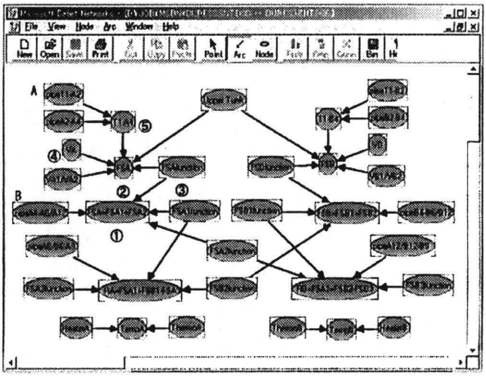


Fig.1 Bayesian Network of the objective system represented by MSBN

with higher priority to test and then the system selects the next step sensing point and modality by using the Bayesian network.

c. Manual mode

Although the additional measurements by fully manual control may increase operator workload, the manual control should be provided for the situations when operators are willing to perform exploratory search of the failure in the incipient phase of an anomaly. In this manual mode, operators can control all aspects of mobile sensing, including sensor selection and sensor positioning. The Bayesian Network (BN) is supposed to be utilized to provide support information of decision making.

2.2 Bayesian network

The Bayesian Network (BN) [2] has been utilized to represent probabilistic cause-consequence relationships of the objective system. The reasons for adopting BN are twofold. First, as the BN can be a basis for probabilistic inference, the certainty of information can be treated explicitly, which is effective in taking sensors' reliability into account. Second, by using the BN, it is possible to deal with situations when some of the symptoms are not available in the course of diagnosis. Along with the evolution of the events and incoming symptoms, the probability of each failure hypothesis in the BN can be updated incrementally. Fig.1 shows the BN for the objective system represented by using "Microsoft Belief Network" (MSBN) (Developed by the Decision Theory Group, Microsoft Research). The probability propagation has been calculated by using "MATLAB Bayes Net Toolbox 2.0".

Assignments of a priori probabilities, which are necessary for calculation, have been performed based on the empirical knowledge of the system reliability. It has been confirmed through preliminary experiments that the developed system is capable of performing the diagnosis by the interactive selection of additional information guided by the BN, starting with the limited information only from fixed sensors [3].

3. EXPERIMENTAL SETUP

3.1 DURESS

As for the experimental test bed, the configuration of Dual Reservoir System Simulation(DURESS)[4], which is a kind of standard system simulation model for an analysis of human cognitive behavior in human factor research, has been adopted for its considerable complexity in terms of an identification of a cause of failure. The piping and instrumentation diagram (P&ID) of DURESS is shown in Fig.2. DURESS is composed of two redundant feedwater lines connecting to the reservoirs. In each reservoir, a heater is installed to control the temperature of the water in it. The operator is required to maintain water level and temperature in each reservoir. Table1 summarizes the anomalies scenarios, which can be emulated in DURESS. As our DURESS is a real world system, a variety of failure modes beyond the scope of the numerical simulation are successfully incorporated. The BN representing the DURESS, as shown in Fig.1, has been utilized to control the IDAS. Table2 summarizes the meaning of the nodes in BN for DURESS. Although the BN has been configured based on the mass conservation of the water flow in the piping and the energy conservation in the reservoirs, the qualitative behavior and relationships among specific parameters have been utilized as the index of the failure because of the insufficient accuracy of the actual measurements.

3.2 Mobile sensing mechanism

In addition to the sensors equipped permanently, as shown in Fig.2, the sensor modules attached to the robot manipulator can be utilized to make flexible measurements for extensive locations in DURESS. A laser vibration sensor, a CCD camera and an infrared temperature sensor are prepared as candidates of an exchangeable sensing module attached to the manipulator hand. The sensor module exchange procedure is performed automatically without any support by human operators. As the DURESS has been built to occupy a horizontal area of 2m x 0.5m, the robot manipulator must move on the rail to cover required sensing locations. Fig.3 shows the robot

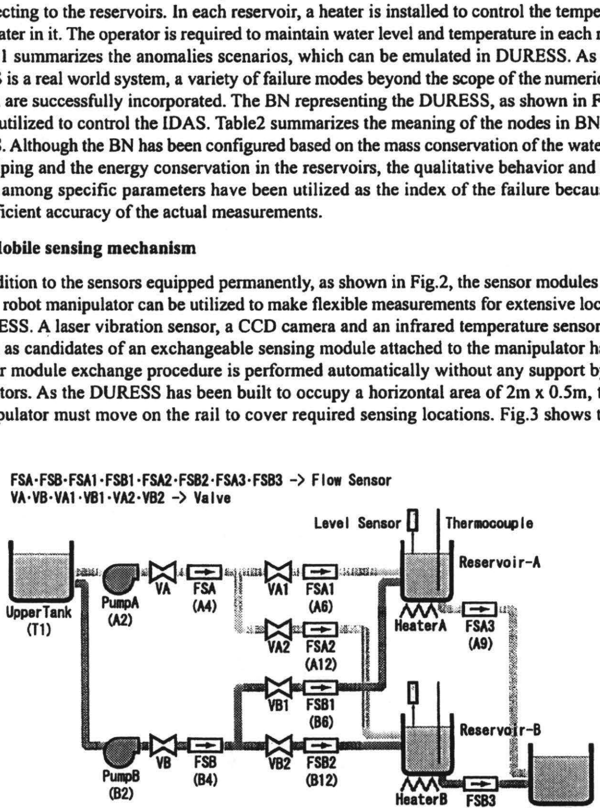


Fig.2 Piping and instrumentation diagram (P&ID) of DURESS

Table.1 Failure scenarios emulated on DURESS

Single Failure Scenario
SA1: Blockage between Upper Tank and Pump A
SA2: Leak between Pump B and FSB
SA3: Mulfunction of VA
SA4: Leak between FSA and FSA1/FSA2
SA5: Leak between FSA2/FSB2 and FSB3
SA6: Mulfunction of thermocouple A
SA7: Mulfunction of heater B
SA8: Upper Tank water level low

Multiple Failure Scenario
MA1: Air between Upper Tank and Pump A <AND> SA4
MA2: SA4 <AND> SA5
MA3: Air between Upper Tank and Pump A <AND> SA4 <AND> Leak between FAS1/FSB1 and FSA3

Table 2 Meaning of the nodes in BN

Node	Normal/Abnormal Conditions based on measurements
FSA	IF fsa = const THEN Normal IF fsa < const THEN Abnormal
FSB	IF fsb = const THEN Normal IF fsb < const THEN Abnormal
FSA=FSA1 +FSA2	IF fsa = fsa1 + fsa2 THEN Normal IF fsa < fsa1 + fsa2 or fsa > fsa1 + fsa2 THEN Abormal
FSB=FSB1 +FSB2	IF fsb = fsb1 + fsb2 THEN Normal IF fsb < fsb1 + fsb2 or fsb > fsb1 + fsb2 THEN Abormal
RA=FSA1 +FSB1−FSA3	IF (fsa1 + fsb1 > fsa3 AND ⊿la > 0) OR (fsa1 + fsb1 < fsa3 AND ⊿la < 0) OR (fsa1 + fsb1 = fsa3 AND ⊿la = 0) THEN Normal ELSE Abnormal
RB=FSA2 +FSB2−FSB3	IF (fsa2 + fsb2 > fsa3 AND ⊿lb > 0) OR (fsa2 + fsb2 < fsa3 AND ⊿lb < 0) OR (fsa2 + fsb2 = fsa3 AND ⊿lb = 0) THEN Normal ELSE Abnormal
Temp A	IF (⊿ha > 0 AND ⊿ta > 0) OR (⊿ha < 0 AND ⊿ta < 0) THEN Normal ELSE Abnormal
Temp B	IF (⊿hb > 0 AND ⊿tb > 0) OR (⊿hb < 0 AND ⊿tb < 0) THEN Normal ELSE Abnormal

manipulator on the linear guide mechanism. The horizontal movement is controlled according to the required sensing locations, which is dynamically identified by the BN.

4. EXPERIMENTAL RESULTS

The validity of the proposed IDAS framework has been demonstrated through the series of experiments, in which the anomalies were actually occurred in DURESS and the auto mode IDAS diagnosed the cause of the failure. The SA3 case in Table 1 is taken as an example to show how IDAS performs incremental measurements.

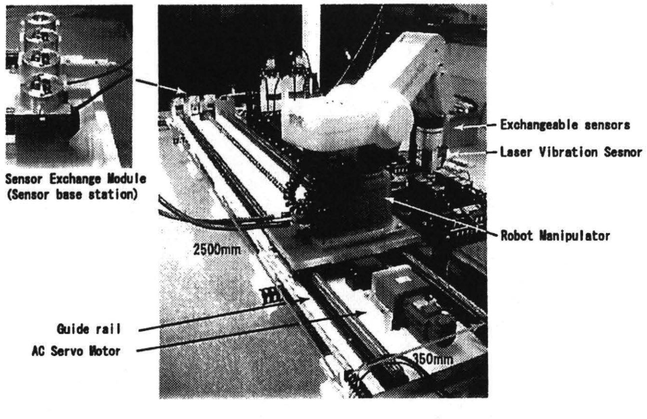


Fig.3 Robot manipulator on the linear guide mechanism. with sensor exchange module

In this anomaly case, the valve-A opens only partially, even though the operator sends the full-OPEN command. Fig.4 shows an example of the resultant window of BN inference, where the bars indicate the probabilities of anomaly for each node. Table 3 show the transition of probabilities for related hypotheses in each diagnosis step. In step 1, when only the measurements from fixed sensors were available, the probabilities for these hypotheses remained high, which means the lack of information to reach the cause of failure. The additional sensing was performed to check the hypothesis with highest probability. In this case, the pump vibration was measured by laser vibration sensor to evaluate the state of node "PipeT1-A4" more specifically. The resultant updated probabilities after confirming that the pump was functioning normally are listed in the step 2 of Table 3. The hypothesis "PipeA2-A4" was then selected as the next step focus of attention among the hypotheses with higher probabilities. ("PipeA2-A4" was selected based on the heuristics that the more upstream a hypothesis is, the higher a priority is.) The sensor was then exchanged to CCD camera to examine the existence of leak by visual inspection. The values in the step 3 of Table 3 are probabilities after confirming that no leak is present. The hypothesis "VA" (Mulfunction of Valve-A) was then selected as the next focus of attention. The valve position was then actually measured to find out that the valve opens partially contrary to the full-OPEN command. The diagnosis process ended by confirming that the hypothesis "VA1/VA2" was normal based on the additional measurement of visual inspection through CCD camera.

It has been confirmed that the proposed system could diagnose thecause of failure for all of the failure scenarios in Table 1. Although the mechanism of reflecting operators intension into the diagnosis process requires futher eleboration, the present results successfully showed the advantage of IDAS framwork in a realistic situations.

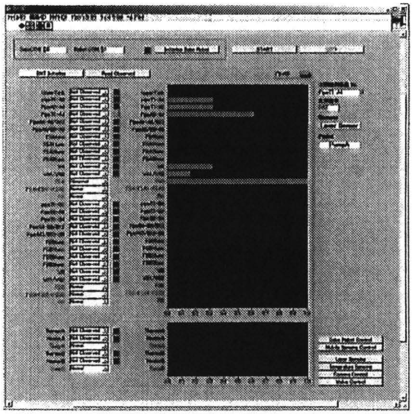


Fig.4 Example image of resultant window of BN inference

Table.3 Transition of probabilities for related hypotheses for each diagnosis step

Node	Step1 →	Step2 →	Step3 →	Step4 →	Final Results
PiepeT1-A2	0.325	Normal	Normal	Normal	Normal
PipeA2-A4	0.325	★0.433	Normal	Normal	Normal
PipeT1-A4	★0.617	0.433	Normal	Normal	Normal
VA	0.325	0.433	★0.689	Abnormal	Abnormal
VA1/VA2	0.162	0.216	0.344	★0.050	Normal
Sensor used	Laser	Camera	Valve Position	Camera	
Sens . Location	PumpA	PipeA2-A4	VA	VA1/VA2	

5. CONCLUSION

The validity of the dynamic failure identification based on the proposed framework of intention-driven sensing has been examined under realistic experimental conditions. The effectiveness and validity of the proposed framework of intension driven adaptive sensing have been confirmed through a set of experiments, where the causes of anomalies have been successfully identified by combined use of fixed sensors and mobile sensors.

It should be noted here that the intention driven sensing is analogous to a human cognitive behaviors in everyday life, where movements of eyes are required to obtain three dimensional perception and additional information searches possibly with other sensing modalities are carried out to validate a certain hypothesis. The dynamic interaction between the process of information acquisition and decision making is one of typical characteristics of expert facing a large scale systems, which is pointed out with regard to the research topic of dynamic environmental supervision[5] and naturalistic decision making[6]. The authors believe that the present framework of intention driven sensing can be a basis for advanced decision support for operators of large scale process systems.

ACKNOWLEDGEMENTS

This work was supported by Grant-in-aid for COE Research-11CE2003.

REFERENCES

[1] Kitamura,M.(Ed), 1998. Evolving Techniques of Diagnosis Toward Establishment of New Paradigm for Human-Machine Cooperation, J.At. Energy Soc.JAPAN, 40(9), pp.652-683. (In Japanese)

[2]Pearl,J.,1988. Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann Publisher,Inc.

[3]Takahashi,M.,2000. INTELLIGENT MOBILE SENSING FOR DYNAMIC FAILURE IDENTIFICATION, Proc. International Topical Meeting on Nuclear Plant Instrumentation Controls, and Human -Machine Interface Technologies (Washington, DC, U.S.A.), pp.572 - 580.

[4]Vicente,K.J.et.al., 1995.Supporting Operator Problem Solving Through Ecological Interface Design,IEEE Trans. Syst.Man Cyber.,25(4),pp.529-545.

[5]Hoc, J.M. et.al., 1995. Expertise and Technology, Lawrence Erlbaum Associates Publishers, Hillsdale, NJ., pp.19-42.

[6]Zsambok,C.E.,1996. Naturalistic Decision Making , Lawrence Erlbaum Associates, Publishers, Mahwah, NJ.

INFORMATION GATHERING MANAGEMENT AND PLANNING : THE CASE OF ANESTHESIA

H. Thuilliez[1]*, F. Anceaux[1], M.C. Beuscart-Zéphir[2], J.-M. Hoc[3]

[1] CNRS – UVHC, LAMIH - PERCOTEC, Le Mont Houy, F-59313 Valenciennes Cedex 9.
Herve.Thuilliez@univ-valenciennes.fr
Francoise.Anceaux@univ-valenciennes.fr

[2] CERIM, Faculté de médecine, Université de Lille 2, Place Verdun, F-59041 Lille Cedex.
mcbeuscart@univ-lille2.fr

[3] CNRS – IRCCyN, PsyCoTec, 1 rue de la Noë, BP 92101, F- 44321 Nantes Cédex,
Jean-Michel.Hoc@irccyn.ec-nantes.fr

ABSTRACT

This paper describes the information gathering of anesthetist and anesthetic nurse on the anesthetic file at the beginning of the preoperative part of anesthesia. Several parameters were analyzed. The role of the field of anesthetic file were categorized in term of planning and pre-conditions. The results show that the anesthetist read less field than nurse. Anesthetists search to select the appropriate plan for the anesthesia and the nurse search data to act. We think that the anesthetists only partially plan the anesthesia and take into account a limited number of data to select a plan. Anesthetists specify only the components that are immediately useful or that allow a good anticipation of risks. Conversely, nurses consider all the information. but they insist more on planning particularization components and on action precondition components.

KEYWORDS

Anesthesia, Planning, Information Gathering, Dynamic Process.

INTRODUCTION

In the research area studying the cognitive processes underlying the activities of process control, the anesthesia is often referred to [de Keyser & Nyssen, 1993 ; Hoc, 1996 : Hoc & Amalberti, 1995]. Besides this research context, financial and legal problems in healthcare domain as well as new medical trends have constituted strong motives to introduce software applications in the hospitals. But as soon as these new tools deal with medical activity and imply physician-machine cooperation, their integration in the daily working environment is difficult or result in a failure. The breakdowns often result from the inability of these application to deal properly with the users' cognitive processes. In the present study we analyze the information selected and memorized by the anesthetists in the per-operative setting. just before they induce the unconscious state. The analysis of this activity of information gathering management is performed considering the role and status of each piece of information in the planning of the anesthetic process.

BACKGROUND

Like the aircraft process [Amalberti & Deblon, 1992 ; Xiao, Milgram, & Doyle, 1997a]. anesthesia involves time pressure constraints which force the anesthetists to produce planning and anticipating activities permitting to deal successfully with the situation. Planning allows the operator to solve the problems by anticipating the needs for preparing and coordinating the actions. According to Gaba [Gaba, 1994], anesthetists' plans rely on a representation including different components: (i) the actual patient's condition; (ii) the goal of the surgery; (iii) the necessary and available resources (mental and physical human resources, material and technical resources). The plan guides the anesthetists decision by compromising with the goals and the constraints. At first, the plan is general, and it is specified during the course of the anesthetic process. Xiao [Xiao, 1994 ;Xiao *et al.*, 1997 ; Xiao, Milgram, & Doyle, 1997b] has listed the main characteristics of the anesthesia planning, which is grounded on so-called "Points for Consideration"; planning is always fragmentary, and is based upon a combination of mental preparation and physical / technical accommodations. The anesthetists tend to focus more on the problems identification than on their solving.

In France, relevant medical information are collected during the pre-operative anesthetic consultation one week before the actual surgery. The anesthetist who performs this consultation and the anesthetist in charge of the anesthesia during the surgery are usually not the same person. Then the planning of the per-operative anesthetic process relies on the anesthetic consultation file, which ensures the transmission of relevant information between the different phases of the process and between the different operators. In such a context, it is important to study how is collected the medical information (pre-operative phase), but it is also crucial to study the way this information is gone over in the per-operative setting. Our previous researches [Anceaux & Beuscart-Zéphir, in press ; Anceaux, Beuscart-Zéphir, & Sockeel. 1999 ; Beuscart-Zéphir & Anceaux, 2000 ; Beuscart-Zéphir, Anceaux, & Renard, 2000 ; Beuscart-Zéphir, Renard, & Anceaux, 1999] on information gathering during the anesthetic consultation demonstrated that the anesthetists may use alternately three different strategies to collect information, and pointed out the characteristics of the selected information as well as the role of planning in the actual anesthesia running. Therefore, we now focus on the way the anesthetist in charge of the per-operative phase and the anesthetic nurse helping him scan this information and deal with it just before they start the induction phase [Thuilliez, 2000].

MATERIAL AND METHODS

4 anesthetists (An) and 4 anesthetic nurses (Nu), all of them professional experts working in the pediatric department of the University Hospital of Lille (CHRUL). were presented 8 consultation files. Those files were real cases for which the difficulty and the number of data had been controlled. Each subject went through the files in a different random order. The files were presented on a computer with a specific software which automatically recorded several parameters such as the overall time for information reading and gathering. the time spent on each area or field of the file, the number of fields read and the order of fields considering. After each file, subjects were requested to write down the data they remembered: the number of fields concerned and the number of semantic units recalled in each field were computed. Relying on two interviews with expert anesthetists, we categorized and characterized each field according to its data's role in the per-operative activity. This classification provides us with 3 categories of information for planning (elaboration. modification, particularization) and 2 categories of information for preconditions (local preconditions for actions, general preconditions for running the whole anesthesia).

RESULTS AND DISCUSSION

The number of fields considered varies significantly for the two groups of subjects (An = 16,5 out of 29; Nu = 23; $p<.01$). Anesthetists always read the subject's name first, in order to check the concordance between the file and the patient. Then they identify which anesthetist performed the consultation, in order to "know where to get the relevant information". After that, they seek critical information allowing them to select the appropriate plan for the anesthesia. On the contrary, the nurses begin with fields containing data allowing immediate actions. The overall time for information reading is longer for the nurses than for the anesthetists (An = 126,4 sec. ; Nu =172,2 sec. ; $p<.01$). But the two groups don't differ regarding the average field reading time (An = 2,45 sec.; Nu = 2,48 sec.) The results of the recall test show the same tendency: Anesthetists recall 12 fields on the average, versus 15.5 fields for the Nurses.

If we focus on the classification of the fields according to their function in the upcoming activity (Table 1), we see that the patterns of fields reading and recalling do not differ for the two groups of subjects (reading time, number of fields scanned, number of fields recalled). However, the proportion of fields considered in each category is not the same: the Nurses scan quite exhaustively all the fields belonging to both categories (planning and preconditions) while the Anesthetists consider mainly the fields allowing them to select or elaborate a plan. As for the average reading time, all the subjects spend more time scanning the fields referring to planning modification or particularization. For the recall test, Anesthetists as well as Nurses tend to recall more fields referring to the planning elaboration.

Table 1: Characteristics of the fields' recall according to their category (the standard deviations appear between brackets)

	Plan selection	Plan Modification	Plan particularization	Global Précondition	Local Précondition
Average proportions of pointed out fields (compared to the number of fields in the category)	0,79 (0,10)	0,54 (0 ,09)	0,52 (0,08)	0,13 (0,04)	0.52 (0.08)
Average row of recall	5,10 (1,02)	6,35 (1,45)	6,78 (0,78)	6,11 (0,57)	10.02 (4.13)

Those results seem to support Xiao's interpretation of anesthetists' "fragmentary planning". Moreover, information management differs slightly for the two groups of subjects, depending on the operational function of the apprehended and recalled information. It seems that the anesthetists only partially plan the anesthesia; they take into account a limited number of data, in order to select a plan, and they specify the only components that are immediately useful or that allow a good anticipation of the risks. Conversely, the nurses seem to consider quite all the information, but they insist more on planning particularization components and on action precondition components.

CONCLUSIONS

We focused here on the beginning of the per-operative phase, for which the anesthetists rely on a partially specified plan. The plan will be refined, adapted and re-adapted all along the per-operative process. In order to study the various information management strategies related to the consultation file, it is necessary to analyze this activity in the whole per-operative context. Future studies should involve the cooperative aspects of the planning which appears to be asynchronous for the anesthetists between the phases and synchronous with the other operators within a given phase; these studies could lead to a better understanding of the mechanisms allowing the production and maintenance of the COmmon Frame Of Reference and of the actual function of this COFOR in cooperative planning. Furthermore, this study and the previous ones focusing on the anesthetic consultation led to set some recommendations for the implementation of a computerized anesthetic record in the CHRUL. The results also grounded the assessment methodology to be used for on-site trials of such computerized anesthetic patients records.

REFERENCES

Amalberti, R. & Deblon, F. (1992). Cognitive modelling of fighter aircraft process control : a step towards an intelligent on-board assistance system. *The International journal of Man-Machine Studies. 36,* 639-671.

Anceaux, F., & Beuscart-Zéphir, M.C. (in press). La consultation préopératoire en anesthésie : gestion de la prise d'information et role des données retenues dans la planification du processus d'anesthésie. *Le Travail Humain.*

Anceaux, F., Beuscart-Zéphir, M.C., & Sockeel, P. (1999). Human Machine Cooperation in the Anesthetic Consultation : importance of planning activities for information gathering. *In Proceedings of the 7th European Conference on Cognitive Science Approaches to Process Control - CSAPC'99* (pp. 15-20). Valenciennes, France: Presses Universitaires de Valenciennes.

Beuscart-Zéphir, M.C., & Anceaux, F. (2000). Méthode d'évaluation des logiciels d'anesthésie: nécessité de la prise en compte de l'activité des utilisateurs. Communication présentée à *SFIM@R 2000*. Lille. France, Oct.

Beuscart-Zéphir, M.C., Anceaux, F., & Renard, J.M. (2000). Integrating user's activity analysis in the design and assessment of medical software applications: the example of anesthesia. In A. Hasman, B. Blobel, J. Dudeck, R. Engelbrecht, G. Gell , & H.U. Prokosch (Eds.), *"Medical Infobahn for Europe", Proceedings of MIE 2000 and GMDS 2000* (pp. 234-238). Amsterdam : IOS press. Health. Technology and Informatics.

Beuscart-Zéphir, M.C., Renard, J.M., & Anceaux, F. (1999). Activity analysis for designing an anaesthesia record system. In J.M. Fouke, R.M. Nerem, S.M. Blanchard, A.P. Yoganathan, (Eds.). *Proceedings of the 1st joint Meeting of BMES & EMBS, 21st Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp.1221). Atlanta, Georgia, Oct.

de Keyser, V., & Nyssen, A.S. (1993). Les erreurs humaines en anesthésie. *Le Travail Humain. 56* , 243-266.

Gaba, D.M. (1994). Human Error in Dynamic Medical Domains. In M.S. Bogner (Ed.), *Human Error in Medicine* (pp.197-224). Hillsdale, NJ: LEA.

Hoc, J.-M. (1996). *Supervision et contrôle de processus – la cognition en situation dynamique*. Grenoble. France : Presses Universitaires de Grenoble.

Hoc, J.M., & Amalberti, R. (1995). Diagnosis: some theoretical questions raised by applied research. *Current Psychology of Cognition, 14,* 73-100.

Thuilliez, H. (2000). *La planification en situation dynamique : L'utilisation du dossier d'anesthésie comme support à l'élaboration et à l'utilisation du plan anesthésique*. Mémoire de DEA de psychologie des processus cognitifs. Université de Paris 8, Paris.

Xiao, Y. (1994). *Interacting with complex work environments : a field study and a planning model*. Thèse de doctorat de philosophie. Université de Toronto.

Xiao, Y., Milgram, P., & Doyle, D.J. (1997a). Planning behavior and its functional role in interactions with complex systems. *IEEE transactions on systems, man, and cybernetics – Part A: systems and humans, 27,* 313-324.

Xiao, Y., Milgram, P., & Doyle, D.J. (1997b). Capturing and modelling planning expertise in anaesthesiology: results of a field study. In C. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp.197-205). Mahawah, NJ: LEA.

Activity Analysis during Cognitive Training of Operational Personnel

G.E. Zhuravlyov
Russian Academy of Sciences
E-mail: gennady@rinet.ru
http//www.rinet.ru/~gennady

The report outlines necessity of the activity analysis during computerized cognitive assessment and training of a power plant operational personnel.
A conlusion is made that the computerized analyses of a subject's activity should be developed and included in the programs of computerized psychological assessment and training.

1. Introduction

At the time of working life, the psychological attributes are slowly changing according to the age and the tasks performed. It is a natural and inevitable process. Movements lose their speed and accuracy. The memory volume is curtailing. The scope of attention is narrowing. Many other professionally important psychological characteristics are changing in consequence of either being employed or idle. Managers must mention these alterations and take proper measures in a timely manner.

Our preliminary investigation demonstrated that imaginative thinking of operational personnel was developed much better than their logical thinking. For some operators, the difference between the sides of their minds was so great that they could be called advanced gifted persons in their imaginative abilities while being retarded in their logical thinking. We can suggest that operators exploit predominantly one side of their brain while the logical one is turned off and atrophies slowly without exercise. For such operators, the low logical ability reduces understanding of procedures, deteriorates performance according to safety rules, and so on [1].

An operator can make many mistakes before supervisors realized that he has changed. Partially, experience may replace mental functions. Earlier or later psychological deficiencies reveal themselves in extraordinary situations.

An useful but insufficient source of psychological data are computerized psychological assessments. A superficial simplicity and validity of the computerized testing made it very popular and promoted it broad distribution. In reality, any application of the computerized testing ought to be *предварено* thorough investigation of the activity personnel involved in. Moreover, activity analysis should be involved in conducting testing and in an interpretation of results.

Psychological data are universal. It was shown repeatedly that there is an evident direct correlation between psychological attributes and professional performance. The main obstacle stems from a potential discrepancy between psychological tasks and professional operations. This problem could and should be solved by detailed analysis of the operational personnel activity and appropriate selection of the set of psychological methods. The main merit of any psychological investigation is low cost and rapidity in comparison with the price of a default.

Regulative handling assumes the idea of supportive measures that prevent degradation of psychological functions. This idea implies a permanent correction. Also, psychological assessment and training should be conducted after any mishap or accident regardless of causes. Taking into account psychological analysis of actions we can better understand the fault and develop preventive measures.

Psychological training was implemented extensively in team training and usually directed on processes of emotional and social regulation. Their influence on operational activity was mediated by cognitive processes. In our work, *we decided to improve the main cognitive processes involved in professional performance.*

2. Psychological analysis of personnel activity

Decision making is the principal function of power union operators. Decisions are made during two main phases of work. The first one is observation. while the second one is operation. Long periods of observations alternate with series of operations aimed to establish a desirable state of equipment.

From the psychological point of view, operational personnel activity could be described on the whole by the following structure (*Figure 1*):

1. Analysis of an operational situation, including
    1.1. Scanning controls and indicators
    1.2. Prediction of potential values of parameters
    1.3. Mental construction of the system image (or of an assemblage of interdependent images)
2. Setting operational goals. As a rule, a hierarchy of the goals is installed and maintained. On the lower level, the nearest goals are like turning on a feeding pump. On the top, such goals include equipment protection and personal safety.
3. Determination of present resources, including
    3.1. Condition of equipment
    3.2. Completeness and reliability of information
4. Determination of criteria (specific demands) and limitations (restrictions) for operations under current conditions
5. Mental analysis (dynamic modeling) of different versions of the technological process development under influence of available operations. For instance, an operator must recognize dangerous tendencies in equipment behavior and predict an action which could produce a beneficial output.
6. Decision making in a narrow sense of the term, i.e. planning of a consequence of operations
7. Performing operations under strict control of the results

3. Structure of Activity

Human Factors/Ergonomics problems constitute a scientific and practical discipline dealing with human activity. We need a kind of a short description of the processes regulating this activity.

To expound the personnel activity investigation we need some new concepts connected to a content of activity, i.e. to its projection on the basic model (Fig.1). The concept of a content of activity is a mediating chain between technological and psychological concepts.

A psychological structure represents the subject's regulative processes. It is constructed from basic psychological processes like perception, attention, memory, thinking, emotions and others (see Fig. 2). Psychological structures are independent of the material inside which they are realized (human body, computer, Noosphere). That is why we implement identical terms (like memory) to different constructions: memory of the global nature, computer's memory...

A technological structure of human activity describes transformations within the object in regards to the main goal of the technological system output.

The contents structure of activity describes the activity in a whole as it is processing among three main components: subject, tools, object.

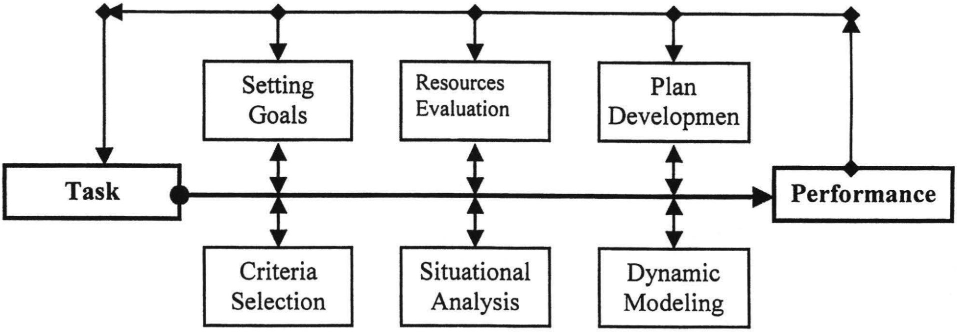


Figure 1. A Simplified Contents Scheme of Activity

All elements of the proposed structure are included in loops of direct and reverse connections. At any given moment, an activity of an operator could start from a contingent element followed by a suitable consequence of other elements of the structure.

The described operational personnel activity is realized generally by the following cognitive processes: perception, memory, thinking, and prediction (see *Figure 2*). As a matter of fact, the process of thinking plays a leading and rather managerial role.

Cognitive (informational) processes are under general regulation of the next three groups of processes: emotions, motivations, and consciousness.

As it is shown on Figure 2, there is a noospheric regulation of psychological processes by technology, social norms, culture, and nature.

The psychological work with the power personnel is based on an accurate determination of the psychological attributes which are important for the efficient professional activity. According to the structure of psychological processes, we selected the following set of cognitive attributes:

- Coordination of fine movements
- Reaction time

- Attention
- Memory
- Imaginative thinking
- Logical thinking
- Orientation in a complex situation.

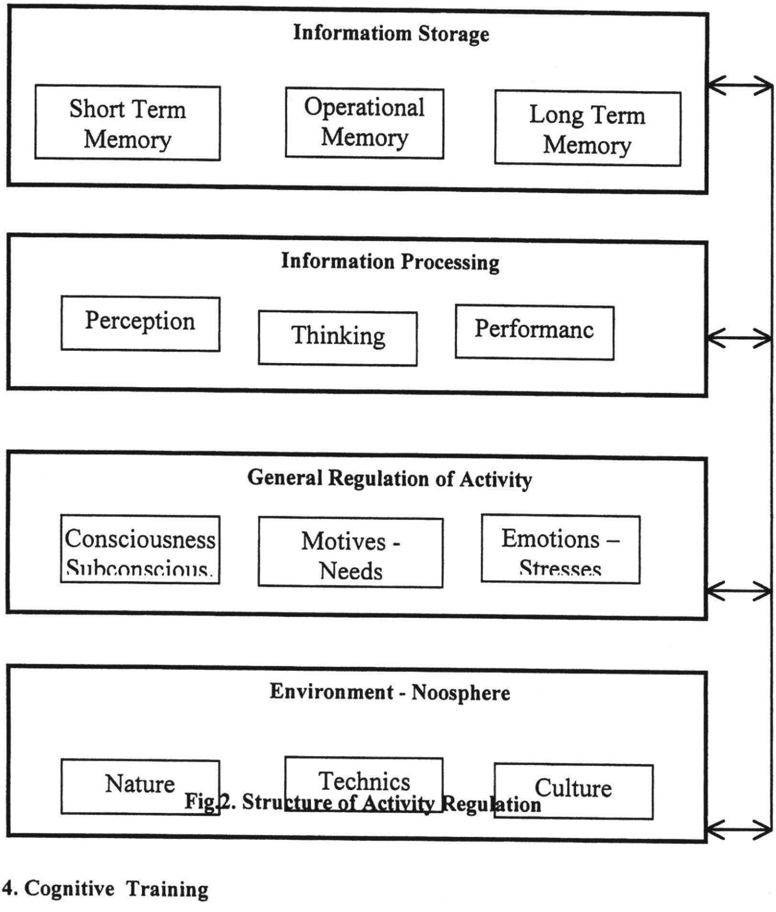


Fig.2. Structure of Activity Regulation

4. Cognitive Training

The first psychological task was to create motivation to the training. As a matter of fact, it was not a simple task. Personnel were initially prejudge to the investigation. They feared the results may negatively influence their salaries and positions. The task of motivation was solved both by experience and persuasion. For 13 years of testing, people saw that testing was not followed by any negative consequences. More over, an evident presentations of results allowed to observe positive effects of training.

Nevertheless, some people preserved negative attitude to training. The evident correspondence was revealed: the worse were results of assessment the lower was motivation. It is interesting to note that managers were covered by this rule, too. Their results were among the lowest and most of them were trying to avoid psychological training.

We administer cognitive training individually on the personnel working places: in control rooms just at the control board, in assistant operators' waiting rooms, in a room of a shift supervisor. Intervention of others was forbidden. Before starting either training, a psychologist explained the task and conditions, in general terms. Then, a subject (an operator) studied an instruction. When he felt ready, he began performance.

During training sessions the psychologist sat beside. He prevented failures induced by eventual misunderstandings. The main undeclared goal of the psychologist was to reveal elements of the subject's activity: grip of the situation, organization of actions, algorithms of operations, heuristics in use, type of errors, emotional reactions, and motivation to training. This information was highly important in correcting skills and optimizing cognitive processes.

Computerized cognitive training presupposes a significant mental load for the psychologist. It is much greater than during traditional testing sessions. The training demands permanent attention to the subject's operations and errors. A small feature of behavior could provide valuable information about the underlying psychological processes. For the psychologist, the main goal was to construct a dynamic model of the subject's mental mechanisms.

This model should permit reproducing of alternative algorithms of the subject's activity. After that stage, it was possible to evaluate to what extent a suggested algorithm differed from the optimal one. The revealed differences must be translated into recommendations comprehensible by the subject. Sometimes, it was necessary to present a short psychological essay as a commentary to recommendations. The frequent themes of the essays were structure of human memory, brain's hemispheric asymmetry, mental coding, interdependence of speech and recollection.

5. Methods of cognitive training

A set of the selected methods was formed to cover the whole scope of the operator's cognitive processes from perception through thinking and movements control. Every method generated tasks with randomly changed initial conditions. While working with the same method, the subject never met identical tasks. This important peculiarity prevented substitution of skill formation by getting accustomed to a task.

Below are described some methods applied and their output.

"Reaction Time". A subject had to choose one of the two potential reactions in response to one of the set of more than 6 million stimuli. Any stimulus was formed from 4 figures discriminated by shape and color. Before starting a given task, two figures were determined as crucial ones. If both were presented in a stimulus, the subject had to press the ENTER key. In other cases (either only one predetermined figure was presented or both were absent), the subject had to press the SPACE key. The advantage of the method was the complexity of the stimuli. Thanks to this fact, processes of memory, search, and recognition were trained along with the disjunctive reactions.

A regular error was observed in performance of the sensory-motor tasks. There was a tendency to work as quickly as possible disregarding accuracy. For operational personnel, that was impermissible deficiency of the psychological processes organization. The psychologist demanded to slow down performance. The subject gradually learned to work without mistakes. Only after that was he allowed to increase speed. *Figure 3* presents a sample of changes in reaction time, errors, and generalized index of performance during training sessions.

Under common input-output demands, very different psychological processes may be involved. There is a well known story about astronomers observed stars passing through Greenwich meridian. The head of a observatory mentioned that some astronomers make a systematic error. They delayed to mark the time a star was passing by. They were dismissed. But nowadays I can suggest that two groups of astronomers applied different tactics in marking the time. The first group predicted the moment when the star ought to cross a thread in a view field of a telescope. The second group waited for an actual moment of passing by and only after that they marked the time. The difference was just equal to the simple reaction time – about 0.2 sec. Our conclusion is evident: externally defined operations do not reflect real psychological processes.

"Switch of Attention". During the tasks a consequence of square matrices was presented. 36 cells of each matrix were **randomly** filled with numbers from 11 trough 46. A subject had to point to numbers from two sequences alternately. The first sequence was made of odd numbers in the ascending order from 11 to 45, while the other one was made of even numbers in the descending order from 46 to 12. The subject had to reproduce the mixed sequence of the numbers **11**, *46*, **13**, *44*, **15**, *42*,...,**45**,*12*. Psychologically, the method simulates the frequent situations when an operator adjusts simultaneously two or more installations (processes) of his power unit.

In the tasks on Switch of Attention a subject had to store in memory a pair of numbers, taken from both odd and even sequences. Transfer to the next step demanded accomplishment of **two** operations: to add 2 to the odd number and to subtract 2 from the even number. The two operations interfered, of course. The main failure of subjects was inability to arrange mental processes according to the installed goal.

Application of methods on attention revealed an important fact. The best performers were the same in both types of tasks. Correspondingly, the worst ones were common in the two cases. Speed of performance and accuracy for the least proficient operators differ significantly from the best ones.

Skills of information search were trained in tasks on attention, too. It was mentioned that decrements in performance sometimes appeared. A subject might spend more than 20 seconds for search of a next number instead of the regular 2-4 seconds. Analyzing microoperations we concluded that such subjects utilized the spontaneous way of searching. They were incapable of organizing an ordered scan of a matrix presented. In these cases, we told the operator about distribution of functions between right and left hemispheres of the human brain. After that, we slowly taught him a combined way of matrix scanning: simultaneous perception and consecutive scanning in groups of cells. As a result, decrements disappeared.

Unceasing load on short term memory is an important feature of an operator of any profession. We trained it by using an adequate method.

"Short term memory". The human memory is considered as the most stable and conservative cognitive process. Nevertheless, we tried to improve memory. It was shown that the volume of memory could be increased significantly. The effective output of memory is determined by two factors: the state of brain cells and cognitive algorithms of remembering and recalling. While the first factor is practically closed for psychological influence the second one can be identified and improved.

A traditional psychological method was reproduced in the computerized form. A sequence of randomly generated sets of numbers was presented. The length of a set increased gradually (starting from 3 numbers in a set) if a subject successfully reproduced on his keyboard previous two sets until a first error appeared. Since that moment three other sets of the same length were presented. The final result was compiled as an average of the last five sets.

The following failure in subject's operations was observed frequently. After catching a new set of numbers, the subject was hurrying to reproduce it as soon as possible. Before completing the input he unexpectedly mentioned that the tail of the set faded in memory. In this case, the natural cognitive process of remembering was disrupted. We recommended that he complete the internal muttering of the set, and only afterwards to start the process of recalling. The recommendation was clear and easy to follow. Results were mostly positive. In some cases we saw that a subject could not restrain himself from the intention to answer immediately. The solution of the problem was very simple. We compelled the subject to read numbers loudly and call them out while pressing the computer's keyboard. The positive results were more persuasive than words of the trainer.

The second well-known way of improving the short term memory is by arranging numbers in blocks. For example, the set 5 2 7 9 3 0 1 3 2 could be remembered much easier by being presented in the form of three subsets of three: 527 930 132. This recommendation was acquired smoothly and resulting figures improved immediately.

"Mental rotation." Some methods were aimed on bettering the imaginative thinking. It occupies a central position in the operational activity. Deficiency in imaginative thinking might arise a question of the operator's professional fitness.

The method is a modified replica of the well-known H.J. Eysenck's intellectual tests. Six resembling figures were simultaneously presented on the screen. Five of them were generated by rotating the same figure on the plane. The angel of rotation was selected randomly. The sixth figure was a mirror reflection of the initial figure. A subject had to find and point out the sixth figure.

"Square." A subject had to reconstruct a square by putting together its parts on the computer's screen.

There were other methods on imaginative thinking. Each of them was made as a game and positively accepted by a majority of operators.

Few operators revealed a low level of the imaginative thinking. Their main failure was substitution of thinking by chaotic trials in search of a solution. It was very difficult to explain by words the idea of the mental operations. Frequently, the trainer performed the task slowly showing a type of the imaginative decisions. Improvement did not occurred every time for this minority of operators.

"Vitkin Test (Complex Situation)." At the beginning of a task, a figure (a polygon) was shown for 10 seconds. Then, a complex figure of lines appeared on the screen. The subject had to find the figure among the lines and point out its apices in 2 minutes. A possibility was provided to see shortly again the sought figure. Twelve tasks were presented in a session.

The method imitates the part of the real operational tasks when an operator has to select a scheme in the whole technological mnemonics. The method was useful in improving corresponding psychological processes. Additionally, we taught operators to control themselves and return to the beginning if necessary. The output was mainly positive. Half of the operators were poor in performance. A simple repetition of the tasks rarely produced positive effect. It seems, that additional efforts should be devoted to further development of the training programs on logical thinking.

6. Results and analysis

Psychologically, the operational activity under real conditions differs significantly from a subject's activity under a laboratory conditions. At the same time, the eternal problem of transferring psychological results from the laboratory into practice can be partially overcome while investigating activity on site.

An important aspect of the cognitive training was to watch closely the subject's mistakes. A general assumption was made that there was a correspondence between the style of mistake in training and in practical activity at the control board. We had observed that some subjects control their actions insufficiently. Simply speaking, these subjects tried to do something and later on observed what output they produced. In an emergency, such operators generate failures more often than overcoming deviations.

Training produced positive results for 68% of operators. It should be taken into account that the cognitive training is a complex psychological event. Emotions, health, mood, status of equipment, weather conditions, and many other factors cast influence on the output. Luck in searching solution plays its role. It is impossible to expect 100% efficiency of the training. We consider our figure as a satisfactory one.

7. Conclusion

Computerized psychological assessment and training demands deep analysis of a subject activity. This is an ultimate condition of success.

Computerized analyses of a trainee activity should be developed and included in the program.

8 . Bibliography

Zhuravlyov, G.E. Subconsciousness and Safety Culture. - In "Proceedings of the International Topical Meeting on Safety Culture in Nuclear Installations". Vienna, 1995, pp. 631 - 644.

Modeling I

Acquisition and Modelling of A Priori Knowledge for Decision Support Systems

Bernhard Döring, Gert Dörfel, Helmut Distelmaier
Research Institute for Communication, Information Processing, and Ergonomics (FKIE)
Neuenahrer Str 20, D-53343 Wachtberg, Germany
doering@fgan.de; doerfel@fgan.de; distelmaier@fgan.de;

Abstract: Decision support systems shall assist human operators in performing their tasks in natural dynamic situations like medical emergencies, fire-fighting, police actions, or military operations, respectively. Such systems need information about the actual external system environment and the actual internal system state as well as a priori knowledge about possible external situations and possible internal system states. The paper reports about an approach for acquiring and modelling especially the a priori knowledge about possible external situations in a military environment. The approach comprises three modelling steps: a theatrical performance model that corresponds quite well with the knowledge representation of subject matter experts (SME) and therefore can be applied for knowledge acquisition, a formal description model that specifies mathematically the a priori knowledge of SME, and an object-oriented UML model that represents the basis of modern software system development. The applicability of the modelling approach is demonstrated by means of a Navy scenario that describes the operational air environment of a frigate in a surveillance mission.

Key words:
Knowledge acquisition and modelling; decision support system; intelligent user interface.

INTRODUCTION

Decision support systems as part of intelligent human-machine interfaces are used, for instance, to assist human decision making in natural dynamic situations like medical emergencies, fire-fighting, police actions, or military operations, respectively. Such situations can be characterised, e.g., by complex dynamic changing environments, time pressure, uncertain and/or insufficient information, and high risk of false decisions [Orasanu and Conolly, 1993].

Military decision making is a process that translates data into action [Wohl, 1981]. When analysing this process different activities can be identified, for instance, information collection, hypothesis generation, action planning, and action accomplishment [Wickens, 1992]. Human operators performing these activities as elements of human-machine systems need information about the actual external system situation, i.e., the system environment, as well as the actual internal system situation, i.e., the system state. But especially for generating situation hypotheses and planning actions operators require additional a priori knowledge about possible external system environments and possible internal system states.

This knowledge has to be identified and modelled for developing intelligent human-machine interfaces that support operators in performing tasks in natural situations (Fig. 1). The concept of intelligent or knowledge-based user interfaces includes an interactive graphical or multi-

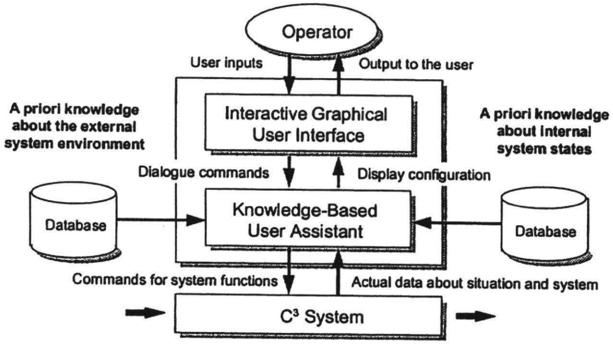


Fig. 1: Concept of a Knowledge-Based User Interface.

media user interface and a knowledge-based user assistant (KBUA) [Berheide et al., 1996]. Environmental and system status data provided by a highly automated system, e. g., a military command, control and communication (C^3) system, are processed by the KBUA and presented adaptively to the operator on the interface depending on, e.g., mission segment, tactical situation, system status, user tasks, and/or user abilities. For processing data provided by the C^3 system the assistant additionally uses the mentioned a priori knowledge. The following considerations are focused on the acquisition, mathematical specification, and object-oriented modelling of a priori knowledge about external system environments in military operations.

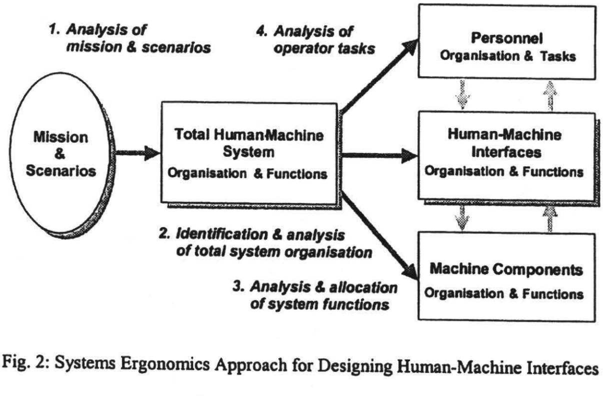


Fig. 2: Systems Ergonomics Approach for Designing Human-Machine Interfaces

Generally, to identify the knowledge needed for designing human-machine interfaces a systems ergonomics approach (Fig. 2) has to be applied. This approach derives the a priori knowledge about external system environments from missions and relevant scenarios of the system to be developed [Beevis, 1999]. Analysing missions and scenarios the organisation of the total system of interest and its subsystems can be identified. For that human-machine system whose human-machine interfaces are to be developed related system functions are determined. Then, system functions are analysed and allocated to human operators and machine components. Functions that have been assigned to the human represent operator tasks. Analysing these tasks in detail the relevant information and action requirements including the needed a priori knowledge can be determined. These requirements constitute the basis for designing human machine interfaces [Kirwan and Ainsworth, 1993].

THE MODELLING APPROACH

But with the systems ergonomics approach problems frequently arise with military systems. Subject matter experts familiar with natural military situations have normally not been trained to specify immediately the required a priori knowledge in a way that it can be mathematically formulated and finally specified using the object-oriented Unified Modelling Language (UML) which is the basis of modern software system development. To overcome this deficiency an acquisition and modelling approach has been developed that comprises three steps: a Theatrical Performance Model, a Formal Description Model, and an UML Model (Fig. 3).

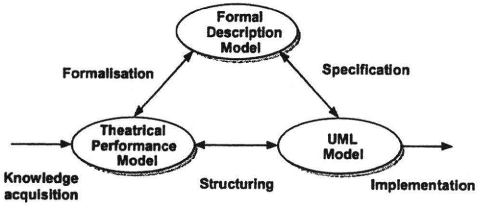


Fig. 3: Steps of the Modelling Approach

Theatrical Performance Model

The first step of the modelling approach is focused on knowledge acquisition. For that, the Theatrical Performance Model (TPM) is developed because theatrical performance can be considered as a well-known representative of naturalistic scenes. Moreover, characteristics of this performance correspond more with the internal representation of military experts concerning military situations than formalised descriptions. It is expected that by means of

TPM elements these experts are able to describe possible dynamic mission environments that may occur in actual or future system operations.

Elements of theatrical scenes (Fig. 4) are, e.g., a stage with wings, actors and supernumeraries with characteristics like roles, activities, active and passive relationships, events, directions, and a screenplay [Dörfel et al., 2000].

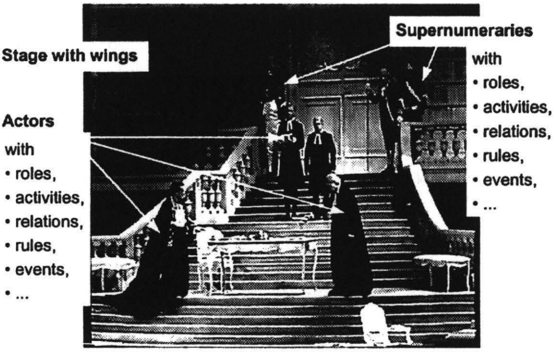


Fig. 4: Elements of a Theatrical Scene

A scene represents a situation. In this approach actors as well as supernumeraries play their individual roles with goals and activities on a stage with a defined wing. Active and passive relationships between actors, supernumeraries and the wing themselves as well as between each others have to be described. Events evoke status changes. Directions are the rules for the occurrence of events or for the course of actions and the screenplay is the script for all the actions of a scenario. In the following, the general scenario description with its elements will be explained in more detail. Structuring and modelling happens by means of theatre-specific elements which will be explained on the basis of a Navy scenario:

Stage and wing represent the scene of actions with a passive environment like area, geography, weather etc. Example: In a military scenario stage and wing are the location where the military operation takes place with all environmental conditions.

Actors are active elements of a scene with respect to the relevant problem. Example: In a military scenario they represent the different components, like aircraft, ships, and command posts, which influence this scenario actively.

Supernumeraries are active elements of an action, whose behaviour is not relevant with respect to the problem considered. Example: In a military scenario they are the different components which do not participate in the scenario actively, e.g., fishing boats.

Roles represent higher level tasks of actors and supernumeraries with respect to their activities, i.e., they are behaviour-determining requirements (goals of activities). With each role particular activities are associated. Example: In a military scenario special activities of an aircraft are associated with the role "reconnaissance". This role contains, e.g., all activities which have to be accomplished for acquiring information about enemy forces, including detecting and identifying targets, evaluating their threat potential, communicating information to own forces, etc.

Activities are decisions and/or actions by the actors and supernumeraries which have to be accomplished to reach their goals with respect to their roles. Example: see roles.

Active relationships occur between actors themselves and/or supernumeraries. They cause an exchange of material, information and/or energy. Example: In a military scenario the communication between a frigate and surrounding fishing boats represents such active relationship.

Passive relationships represent connections between scene elements without the exchange of material, information and/or energy. They occur between actors, supernumeraries, or the wing themselves or between the elements of different groups. They represent organizational or structural connections. Example: In a military scenario the organizational structure of a squadron represents such passive relationship.

A scenario event is a trigger that causes actors and supernumeraries to perform certain activities. Example: In a military scenario an event related to a frigate occurs if it detects an approaching aircraft. Then, the ship has to identify the aircraft and to evaluate its threat.

Rules are direction instructions to guide the course of a scenario. Example: In a military scenario these rules are identification regulations if the identity will not been communicated directly.

A screenplay describes the course of a scenario, i.e., a specific order of situations and actions. Contrary to predefined actions and processes, in natural settings there does not exist any fixed screenplay but the course of actions evolves analogue to an improvisational performance or on the spur of the moment.

To demonstrate the applicability of the TPM and the modelling approach to a military situation, a Navy scenario is used as an example. It is part of a military crisis reaction mission that has been specified together with experts of the German Navy. This scenario constituted the basis for the development of an intelligent human-machine interface which has been developed to support operators in a combat information centre of a frigate in identification/recognition of possible air threats [Distelmaier et al., 2000]. The scenario includes, e.g., own ship, neutral, friendly, and hostile aircraft as actors and supernumeraries, and safety zones as well as air corridors and routes as scenery elements.

Formal Description Model

The second step of the developed approach on which these considerations are mainly focused transforms TPM elements into the Formal Description Model (FDM) using formalisms of the general system theory [e.g., Klir, 1969]. This step is necessary because TPM elements are normally expressed in natural language statements that hold a high degree of semantic fuzziness. To eliminate this fuzziness TPM elements are transformed into mathematical specifications that already closely correspond to elements of the UML model used later. The modelling process starts with determining the set O of scenario objects o_i as main FDM elements:

$$O = \{ o_i : i \in I_O \} . \qquad (1)$$

An object o_i can be described by the set AKT_i of its activities akt_{ik}, the set ATT_i of its attributes att_{ik}: the value set ATT_{ik} of all values $att_{ik}(t)$ of attribute att_{ik}, and the set S_i of object states s_{ik}

$$\begin{aligned}
&AKT_i = \{ akt_{ik}: k \in I_{AKTi} \} , \\
&ATT_i = \{ att_{ik} : k \in I_{ATTi} \} , \\
&ATT_{ik} = \{ att_{ik}(t) : t \in T \} , \qquad (2) \\
&S_i = \{ s_{ik} = (att_{im}(t), att_{im+1}(t), att_{im+2}(t), \dots) : k \in I_{Si} \wedge att_{im}(t) \in ATT_{im} \\
&\qquad \wedge att_{im+1}(t) \in ATT_{im+1} \wedge att_{im+2}(t) \in ATT_{im+2} \wedge \dots \} , \\
&S \subset ATT_{im} \times ATT_{im+1} \times ATT_{im+2} \times \dots .
\end{aligned}$$

Normally, if an object o_i is in the state s_{ik} only a subset of AKT_i is activated. Therefore, a state-based behaviour SBB_i of an object o_i is defined that specifies for each object state the subset of activated activities. With $P(AKT_i)$ as power set of AKT_i it counts:

$$\begin{aligned}
SBB_i = \{ (s_{ik}, (akt_{im}, akt_{im+1}, akt_{im+2}, \dots)) : {} & s_{ik} \in S_i \wedge \\
& (akt_{im}, akt_{im+1}, akt_{im+2}, \dots) \in P(AKT_i) \wedge \\
& akt_{im}, akt_{im+1}, akt_{im+2}, \dots \text{ are activated during } s_{ik} \} , \qquad (3) \\
SBB_i \subset S_i \times P(AKT_i) . &
\end{aligned}$$

Two categories of scenario objects can be distinguished: 1. static scenario objects that correspond to TPM stage and scenery elements which posses only one state and no activity; 2. dynamic scenario objects that correspond to actors and supernumeraries of the TPM which posses different states, state transitions, and activities.

Other FDM elements correlated with TPM elements are, e.g., object roles, active and passive relationships between objects, scenario events, and behaviour rules. With SR as the set of all scenario roles sr_m and R_i the set of all roles r_{ik} of object o_i it counts:

$$\begin{aligned}
&SR = \{ sr_m : m \in I_{SR} \} , \qquad (4) \\
&R_i = \{ r_{ik} = sr_m : k \in I_{Ri} \wedge sr_m \in SR \wedge sr_m \text{ is a role of } o_i \} \subset SR .
\end{aligned}$$

With every role r_{ik} of an object o_i specific activities are associated. Therefore, a role-based behaviour RBB_i of an object o_i can be defined that specifies for each object role the subset of related activities:

$$\begin{aligned}
RBB_i = \{ (r_{ik}, (akt_{im}, akt_{im+1}, akt_{im+2}, \dots)) : {} & r_{ik} \in R_i \wedge \\
& (akt_{im}, akt_{im+1}, akt_{im+2}, \dots) \in P(AKT_i) \wedge \\
& akt_{im}, akt_{im+1}, akt_{im+2}, \dots \text{ are related to } r_{ik} \} , \qquad (5) \\
RBB_i \subset R_i \times P(AKT_i) . &
\end{aligned}$$

Active relationships specify interactions that occur between objects by exchanging matter, energy, and/or information, respectively [Haberfellner et al., 1997]. With this exchange outputs of an object o_i and inputs of another object o_k are coupled temporarily. To describe formally an active relationship $ar_{ik}(t)$ that occurs between o_i and o_k at time t the following sets are defined:

the set Y_i of all outputs y_{im} of o_i, the value set Y_{im} containing all values $y_{im}(t)$ of output y_{im}, the set X_k of all inputs x_{kn} of o_k, the value set X_{kn} containing all values $x_{kn}(t)$ of input x_{kn}., the set AR_{ik} of all active relationships $ar_{ik}(t)$ between o_i and o_k, and the set AR_i of all active relationships of object o_i:

$$\begin{aligned}
Y_i &= \{ y_{im} : m \in I_{Yi} \} , \\
Y_{im} &= \{ y_{im}(t) : t \in T \} , && (6) \\
X_k &= \{ x_{kn} : n \in I_{Xk} \} , \\
X_{kn} &= \{ x_{kn}(t) : t \in T \} , \\
& Y_i \cap X_k \neq \varnothing , \\
AR_{ik} &= \{ ar_{ik}(t) = (y_{im}(t), x_{kn}(t)) : \\
& \qquad (y_{im}(t), x_{kn}(t)) \in Y_{im} \times X_{kn} \wedge y_{im} = x_{kn} \wedge t \in T \} , \\
AR_i &= \cup AR_{ik} .
\end{aligned}$$

With passive relationships between objects we discriminate between an aggregation (is-part-of relation) which describes a decomposition hierarchy and a generalisation (is-a-kind-of relation) which describes a heritage hierarchy [Rumbaugh et al., 1991]. If an object o_k is part of another object o_i so exists an aggregation agg_{ik} between these objects. Then, the set AGG_i contains all aggregations agg_{ik} of object o_i:

$$AGG_i = \{ agg_{ik} = (o_i, o_k) : o_i, o_k \in O \wedge o_k \text{ is part of } o_i \} \subset O \times O . \qquad (7)$$

If correspondingly, an object o_k is a kind of another object o_i there exists the generalisation gen_{ik} between these objects. The set GEN_i contains all generalisations gen_{ik} of object o_i:

$$GEN_i = \{ gen_{ik} = (o_i, o_k) : o_i, o_k \in O \wedge o_k \text{ is a kind of } o_i \} \subset O \times O . \qquad (8)$$

A scenario event is defined by an output generated by an object during a certain state and sent to another object as result of an active relationship between both. Is E_i the set of all events e_{ik} produced by an object o_i, S_i the set of all states s_{im} of o_i, and Y_{in} the value set containing all values $y_{in}(t)$ of output y_{in} of o_i, E the set of all scenario events e_{ik} it counts:

$$\begin{aligned}
& E_i = \{ e_{ik} = (s_{im}, y_{in}(t)) : k \in I_{Ei} \wedge s_{im} \in S_i \wedge y_{in}(t) \in Y_{in} \} , && (9) \\
& E_i \subset S_i \times \cup Y_{in} , \\
& E = \cup E_i .
\end{aligned}$$

Finally, behaviour rules describing object activities in the case of an event are considered. In the following this is called the event-based behaviour EBB_h of object o_h. EBB_h can be described graphically by a state transition diagram. With the set S_h of all states of object o_h it counts:

$$\begin{aligned}
EBB_h &= \{ (e_{ik}, s_{hm}, s_{hm+1}) : e_{ik} \in E \wedge s_{hm}, s_{hm+1} \in S_h \wedge \\
& \qquad \text{transition } (s_{hm}, s_{hm+1}) \text{ is caused by } e_{ik} \} , && (10) \\
EBB_h &\subset E \times S_h \times S_h .
\end{aligned}$$

With equation (3) the subset of all activities activated during a state s_{hm+1} has already been defined.

As an example, the air route identified in the mentioned Navy scenario and specified in the TPM as a scenery element is considered. It consists of two segments with different reference points and directions. In the FDM the air route represents a static scenario object o_1 which contains the objects o_2 (segment 1) and o_3 (segment 2) as parts. Therefore, between these three objects there are two aggregations $agg_{12} = (o_1, o_2)$ and $agg_{13} = (o_1, o_3)$. The attributes of object o_1 are those of objects o_2 and o_3, plus an attribute att_{11} = route identifier. Attributes of object 2 (segment 1) and their related values are:

$$
\begin{aligned}
&att_{21} = \text{object identifier}, && att_{21}(t) = \text{segment } 1, \\
&att_{22} = \text{segment reference point}, && att_{22}(t) = (\text{latitude: } a/^\circ, \text{ longitude: } b/^\circ, \text{ altitude: } c/\text{ft}), \\
&att_{23} = \text{segment width}, && att_{23}(t) = d/\text{nm}, \\
&att_{24} = \text{segment length}, && att_{24}(t) = e/\text{nm}, \\
&att_{25} = \text{segment height}, && att_{25}(t) = f/\text{ft}, \\
&att_{26} = \text{segment direction}, && att_{26}(t) = g/^\circ, \\
&att_{27} = \text{segment speed}, && att_{27}(t) = h/\text{kn}, \\
&att_{28} = \text{segment flight level } 1, && att_{28}(t) = i/\text{ft}, \\
&att_{29} = \text{segment flight level } 2, && att_{29}(t) = j/\text{ft}.
\end{aligned}
\tag{11}
$$

Correspondingly, attributes of object 3 (segment 2) and their values have been defined. Because o_1 represents a static scenario object the above specified values are constant.

As another example an aircraft inside the air route is considered. In the TPM this aircraft represents an actor element. In the FDM it appears as a dynamic scenario object o_4. Some possible roles of an aircraft (AC) in the considered Navy scenario are: commercial AC, fighter AC, reconnaissance AC. The object o_4 takes the role of a commercial AC. Therefore, the set of roles R_4 contains the element r_{41} = commercial AC (Equ. (4)). o_4 is specified by attributes, their values, and activities. Attributes and related values are:

$$
\begin{aligned}
&att_{41} = \text{object identifier}, && att_{41}(t) = \text{name}, \\
&att_{42} = \text{position}, && att_{42}(t) = (\text{latitude, longitude : inside air route}), \\
&att_{43} = \text{altitude}, && att_{43}(t) \approx \text{on air route flight level}, \\
&att_{44} = \text{altitude change}, && att_{44}(t) \approx 0 \text{ ft/min}, \\
&att_{45} = \text{course}, && att_{45}(t) \approx \text{air route direction}, \\
&att_{46} = \text{course change}, && att_{46}(t) \approx 0^\circ, \\
&att_{47} = \text{speed}, && att_{47}(t) \approx \text{air route speed}, \\
&att_{48} = \text{speed difference}, && att_{48}(t) \approx 0 \text{ kn}, \\
&att_{49} = \text{identity}, && att_{49}(t) = \text{neutral}.
\end{aligned}
\tag{12}
$$

Some activities of object o_4 are:

$$
\begin{aligned}
&akt_{41} = \text{Communicate identity}, \\
&akt_{42} = \text{Operate in accordance with air route}, \\
&akt_{43} = \text{Operate not in accordance with air route}.
\end{aligned}
\tag{13}
$$

In the regarded scenario there are other aircraft which approach, cross, and leave the air route. Using their manoeuvres with regard to the air route as classification characteristics different possible aircraft states can be distinguished, for instance:

$$\begin{aligned} s_{41} &= \text{aircraft inside air route}\ , \\ s_{42} &= \text{aircraft approaching air route}\ , \\ s_{43} &= \text{aircraft crossing air route}\ , \\ s_{42} &= \text{aircraft leaving air route}\ . \end{aligned} \tag{14}$$

As an example, the object o_4 is further considered in some detail. It possesses the actual state s_{41} with the actual attribute values described above. The actual state-based behaviour belonging to the state s_{41} (Equ. (3)) is defined by the activities of equation (13). Some of these activities, that means, akt_{41} = "Communicate identity" and akt_{42} = "Operate in accordance with air route" correspond to the role r_{41} = commercial AC. Therefore, they constitute the role-based behaviour of object o_4 (Equ. (5)) whereas the activity akt_{43} = "Operate not in accordance with air route" would be rather the constituent of a fighter AC role. Further aircraft described in the TPM as scenario elements represent corresponding dynamic scenario objects of the FDM and are specified in the same way.

UML Model

The last step of the developed modelling approach transforms the FDM into the UML Model (UMLM) which is the basis for implementing the a priori knowledge into the intelligent human-machine interface. Generally, basic elements of an UMLM are instances and classes [Balzert, 1999, Rumbaugh et al., 1991]. Instances represent real and/or abstract entities of the problem domain and are described by their states and behaviour. In this meaning the above described scenario objects represent already UML instances. A class describes a group of instances with similar properties (attributes), common behaviour (operation), common relationships to other instances, and common semantics. All identified air route segments represent, e.g., instances of the class "Air route segments", all identified air route approach sectors are instances of the class "Air route approach sectors". In the same way, all aircraft identified in the scenario are instances of the class "Aircraft".

According to the specified static and dynamic scenario objects of the FDM the classes "Static knowledge elements" and "Dynamic knowledge elements" are defined for the UMLM. The above introduced classes "Air route segments" and "Air route approach sectors" with their actual instances are a kind of the class "Static knowledge elements". The above mentioned class "Aircraft" with all aircraft in the states "AC inside air route", "AC approaching air route", "AC crossing air route", and "AC leaving air route" as instances is a kind of the class "Dynamic knowledge elements" (Fig. 4). On the other side the classes "Static knowledge elements" and "Dynamic knowledge elements" are a kind of the more general class "Knowledge elements".

Essential constituents of the UMLM are knowledge patterns that consists of described knowledge elements. Analogously to design patterns [Gamma et al., 1995] the knowledge patterns describe expert knowledge that concentrates on the essential part of a problem solution expressed in a plain way. The problem of the scenario regarded is the identification/recognition of possible air threats during a military crisis reaction mission of a frigate. To determine the intent and derived from that the possible identity of aircraft static knowledge about, e.g., air route segments and air route approach sectors are combined with dynamic knowledge about the actual state and behaviour of aircraft. Therefore, a class "Air route patterns" has been defined, for instance, that comprises static knowledge elements like "Air route segments" and "Air route approach sectors" as well as dynamic knowledge elements like "Aircraft" (Fig. 5). That means with the identification process: IF a contact detected by the sensors of the ship has the specified attribute values of an aircraft with the position inside the air route THEN its identity is assumed to be neutral.

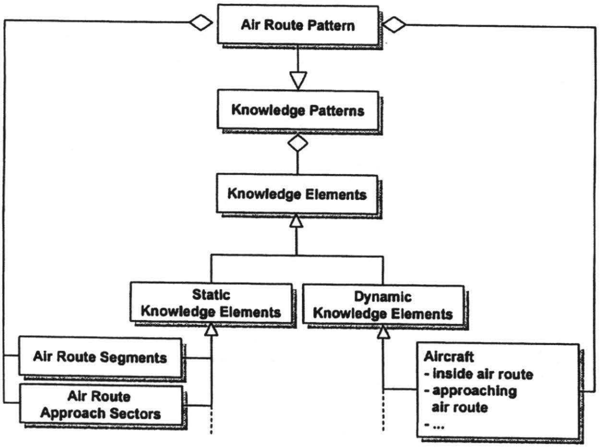


Fig. 5: UML Structure of Knowledge Elements and Patterns

Applying the described modelling process step by step the scenario situations have to be identified and compiled in the TPM by the military matter experts at first. Then, TPM elements are transformed and specified formally in the FDM. This formal specification constitutes the basis for establishing finally the UML Model. For describing needed classes and instances of this model FDM elements have to be restructured into knowledge patterns. Those elements of the UMLM constitutes the basis for the design and implementation of the object-oriented software programme. But this modelling process does not only proceed strictly in this direction. If during the specification of the FDM and/or the UMLM knowledge gaps are identified a backward directed analysis and synthesis can be carried out to fill up these gaps. That means, the modelling process proceeds in a bi-directional manner (Fig 3.).

It is planned to validate the acquisition and modelling approach developed together with subject matter experts of the German Navy for situation awareness, identification/recognition, and threat evaluation processes in future anticipated anti-air warfare situations. In a new project the approach will be used for identifying knowledge patterns that constitute the basis for designing operator decision support for the above mentioned activities. The knowledge-based assistant system (Fig. 1) will apply these patterns with a case-based reasoning mechanism where the described knowledge patterns can act as relevant cases for generating solutions for the mentioned activities.

References

Beevis, D. (Ed.) (1999): Analysis Techniques for Human-Machine System Design. Wright-Patterson Air Force Base, Ohio: Crew Systems Ergonomics/ Human Systems Technology Information Analysis Center (CSERIAC).

Balzert, H. (1999): Lehrbuch der Objektmodellierung. Heidelberg, Berlin: Spektrum Akademischer Verlag.

Berheide W., Distelmaier, H., Döring, B. (1996): Adaptive Function Allocation for Situation Assessment and Action Planing in C3 Systems. In: D. Beevis, P. Essens and H. Schuffel (Eds.), Improving Function Allocation for Integrated Systems Design. SOAR 96-1. Wright-Patterson Air Force Base, Ohio: CSERIAC, pp 159-177

Distelmaier, H., Dörfel, G., Döring, B. (2000). A Concept for Knowledge-Based User Support in Naval Environments. RTO Meeting Proceedings 57 on Usability of Information in Battle Management Operations. Neuilly-sur-Seine Cedex, France: NATO Research and Technology Organization; pp 17.1-17.8

Dörfel, G., Döring, B., Distelmaier, H. (2000): A Concept for Modeling Natural Settings. Proceedings of the KES'2000 (Knowledge-Based Intelligent Engineering Systems 2000), 30 August - 1 September 2000, Brighton, UK, pp 257-260.

Gamma E., Helm, R., Johnson, R., Vlissides J. (1995): Design Pattern. New York: Addison-Wesley Publishing Company.

Haberfellner, R., Nagel, P., Becker, M., Büchel, A., von Massow, H. (Ed. W. F. Daenzer, F. Huber) (1997): Systems Engineering. 9. Auflage. Zürich: Verlag Industrielle Organisation.

Kirwan, B. and Ainsworth, L. K. (1993): A Guide to Task Analysis. London: Taylor & Francis.

Klir, G. J. (1969): An Approach to a General System Theory. New York, London: Von Nostrand Reinhold Company.

Orasanu, J. and Connolly, T. (1993): The reinvention of decision making. In: G. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.) Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corp. pp 3-20

Rumbaugh J., Blaha, M., Premerlani, W., Eddy, F., Lorensen, W. (1991): Object-Oriented Modeling and Design. Englewood Cliffs, NJ: Prentice-Hall Inc.

Wickens, C. D. (1992): Engineering Psychology and Human Performance. 2nd Edition. New York: HarperCollins Publisher.

Wohl, J.G. (1981): Force Management Decision Requirements for Air Force Tactical Command and Control. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-11, No. 9, Sept. 1981, pp 618-639.

ANALYZING MARITIME WORK DOMAINS

Johannes Petersen[1*] and Morten Nielsen[2]

Center for Human-Machine Interaction
[1]Oersted • DTU, Automation, Technical University of Denmark, Denmark
jop@oersted.dtu.dk
[2]Computer Science Department, Aarhus University, Denmark
mnielsen@imv.au.dk

Abstract

This paper focuses on problems related to the analysis of maritime work domains using the abstraction hierarchy proposed by Rasmussen (1986). It is argued that there is need for a causal basis of the means-end relations of the abstraction hierarchy. The main reason for this is that an account of the possibilities offered by physical objects, denoted as means, depend not only on the intrinsic properties of such objects but also on the properties of causally related objects in their immediate environment. This is particularly obvious in the maritime domain where the forces produced by control devices (e.g. main propeller, rudder and thruster), and disturbances (e.g. wind and current) are highly dependent on external properties of the water, hull, etc. Through simple causal models of control devices it is illustrated that objects typically denoted as means corresponds to an individual physical object having specific causal propensities in the context of an action involving additional objects (having other causal propensities). When adopting a means-end view such additional objects would normally be backgrounded.

Introduction

The movements of a vessel are determined by the forces acting on it. Some of these forces are controllable, while others are not. The controllable forces originate from rudder, main propeller, thrusters, anchors, hawsers and tugs, whereas the uncontrollable forces are due to wind, current, sea and hydrodynamic forces acting on the hull (squat and bank effects).

An important part of maritime operations on large container carriers is to control the movements of the vessel – often referred to as *maneuvering*. In restricted waters and harbor areas the autopilot is switched off and the crew has to perform the maneuvering of the vessel manually by controlling the forces produced by control devices (or actuators) such as rudder, main propeller, and thrusters. The increasing size of vessels and traffic-intensity has led to increasing demands on the crew's ability to maneuver the vessel in a safe and efficient manner.

For an actor performing maneuvering tasks it is crucial to have knowledge of the mechanisms underlying the production of forces (controllable and uncontrollable) acting on the vessel. He or she must bear in mind that the production of controllable forces acting on the vessel is not only a result of the properties of the control devices in isolation. Rather it is a result of control devices interacting with other things in their immediate environment. For instance, transverse force produced by the rudder is determined, among other things, by the rudder angle, the area of the rudder, its lift-coefficient *and* the velocity of water inflow to the rudder. Likewise, the production of transverse force using the bow thrusters is determined, among other things, by the pitch of the propeller blades (assuming that the rate of revolutions is constant) *and* the velocity of the water flowing along the hull.

It is therefore typically the case that the interventions of an actor (changing rudder angle, changing the rate of revolutions of the main propeller, changing the pitch of the thrusters, etc.) influences only a single *player* in the force production that he or she actually wishes to control (not to mention the side-effects which are also occurring). In order to make decisions about appropriate maneuvering operations it is important that the crew is capable of performing valid assessments of the actual possibilities (and hence impossibilities) of force production offered by the different control devices in specific situations.

Although crew-members obtains extensive knowledge of how the changing environment influences the force production of control devices through training and experience, there are no dedicated information systems on the command bridge assisting the crew in performing on-line assessments of the actual possibilities of force production offered by the different control devices. It is believed that the insufficient information provided by today's bridge display systems is problematic and has to be dealt with in order to improve both efficiency and safety in maritime maneuvering operations.

A first step in improving information systems on the command bridge is to analyze the work domain. According to the Cognitive Systems Engineering approach (see e.g. (Rasmussen, 1986) and (Vicente, 1999)), analysis and modeling of the work domain itself is a precursor for the design of information systems supporting human decision-making. The overall purpose is to obtain deep knowledge of the *possibilities* of action and the *constraints* on action imposed by the work domain.

This paper deals with problems related to the modeling of maritime work domains, focusing explicitly on those parts of the work domain relevant for maneuvering. On the basis of a representation of the work domain using Rasmussen's well-known abstraction hierarchy we discuss fundamental limitations of the means-end relations implied by the abstraction hierarchy, in relation to the modeling of maritime work domains. The focus is on the lack of an explicit account of the causal role of the *immediate environment* of components at the lower levels of abstraction – a shortcoming preventing an adequate account of the actual possibilities of force production offered by the control devices fitted on a vessel.

An Abstraction Hierarchy of the Maritime Work Domain

For the purpose of designing information systems supporting human decision processes in complex work settings it is important to come to terms with the substance matter of the work domain. This involves a model of the work domain describing the possibilities of action and constraints on action given by the work domain itself. According to Rasmussen (1991) the representation of the work domain is supposed constitute the space of *possibilities* from which the *actualities* are selected in a particular situation. This is similar to a city map representing the possible routes between different locations, from which an actor can make decisions about an actual route from A to B (Vicente, 1999).

The abstraction hierarchy, proposed by Rasmussen (1986), suggests that a representation of complex work domains should be given at different levels of abstractions spanning the gap between physical form and the overall purpose and constraints imposed by the environment (the means-end dimension). Furthermore, it should be possible to consider each level of abstraction at different levels of decomposition (the part-whole dimension). The actual content of the levels of abstraction, as well as the number of levels needed, will depend on the specific domain in question (Vicente, 1999).

The maritime work domain

In Table 1, the content of the work domain relevant for maritime maneuvering is described as an abstraction hierarchy. The individual levels of abstraction reflect different types of properties of the work domain.

Table 1. An abstraction hierarchy of the maritime work domain relevant for maneuvering.

Levels	**Content**
Purpose	The location, orientation and course of the vessel.
Vessel Movement	The movements of the vessel (longitudinal velocity, transverse velocity and turn-rate).
Force Exertion	Controllable and uncontrollable force acting on the vessel.
Force Production	Force production of rudder, main propeller, thrusters, tugs, hawsers, anchors (control devices), wind, current and hull (disturbances)
Physical Form	Appearance and location of physical entities

Since the maritime work domain is very different from process plants, we have found it beneficial to deviate from the original labels of the abstraction levels proposed for process plants (Rasmussen, 1986). That is, the levels of abstraction have been given names that correspond better with the inherent structure of the maritime work domain. Mentioned in passing, deviation from the original names of the levels of abstraction when analyzing work domains that are different from process plants is in line with recommendations given by Vicente (1999).

The upper-most level (*Purpose*) describes the overall purpose of maneuvering of a vessel. The purpose is given in terms of the vessel's *position*, *orientation* (low speed) and *course* (cruising speed). These properties may be constrained by traffic, obstacles (quay, buoys etc.) and legal barriers such as traffic separation schemes and the rules of the road. The next lower level (*Vessel Movement*) is concerned with vessel movements in different dimensions, constrained by e.g. fuel saving strategies and legal speed limits. The next lower level (*Force Exertion*) is concerned with the forces acting on the vessel (*longitudinal*, *transverse* and *vertical* forces) causing changes in its state of motion. The constraints at this level reflect limitations on force exertion on physical equipment. One level down (*Force Production*) the focus is on the role of the different entities producing forces on the vessel. This includes not only ordinary control devices such as *rudder*, *main propeller*, *thrusters* and *tugs*, but also disturbances such as wind and current which are in fact occasionally used as constructive means of force production. Finally, at the bottom level of the abstraction hierarchy (*Physical Form*) we have a description of the appearance and location of physical entities.

It is important to note that a distinction is made between maneuvering and tasks having to do with navigation. While maneuvering has to do with short-term changes in position, orientation and course of a vessel (reflected in the top level of the abstraction hierarchy shown in Table 1) navigation refers to the long-term planning involved in conducting vessels from A to B. The substance matter of the work domain for navigation is not reflected in the abstraction hierarchy given above and will be discussed in this paper.

The abstraction hierarchy as a basis for the derivation of maneuvering operations

In view of the different levels of abstraction described above, the construction of concrete maneuvering operations can be understood in terms of the matching of consequences of forces production (propagating upwards in the abstraction hierarchy) and the overall purposes and constraints of the maneuvering task (propagating downwards in the abstraction hierarchy). Due to the many-to-many relations among the levels the need for decisions and iterations between the levels arises.

For instance, when a certain force interaction, which is supposed to bring about an intended vessel movement is demanded the typically thing to do is to change the balance between controllable and uncontrollable forces acting on the vessel, by changing the intensity of the controllable force being produced. In case it is not possible to bring about the desired movement of the vessel using control devices an alternative might be to inhibit the uncontrollable forces acting on the vessel or even modify the intended state of vessel motion.

Although forces stemming from wind and current are considered uncontrollable such forces are actually functions of the area of the vessel being exposed to these disturbances – the area above the water line and the area below the water line, respectively. This means that such forces can be inhibited by changing the orientation of the vessel (although in many cases this implies that the actual goal is abandoned).

Even though overall *goals* of a specific maneuvering task are always issued top-down, i.e. as desired position, orientation or course at the top level of abstraction, ultimately, the concrete *objectives* have to be derived bottom-up, in view of the actual possibilities (and constraints) of force production offered by the control devices (and disturbances) in the particular situation (see (Lind, 1999) for more details on the distinction between goals and objectives).

Limitations of means-end relations

As mentioned above, one of the purposes of the abstraction hierarchy is to capture the space of possibilities provided by the work domain. More specifically, this means that the possible means of a function at a specific level of abstraction are found at the next lower level. Things (structural entities) enter at the bottom level as *means* of functions described at the next higher level of abstraction. See for instance the abstraction hierarchy proposed above (Table 1), where the rudder is a means of bringing about a transverse force on the vessel.

Although the means-end thinking promoted by the abstraction hierarchy – identifying individual things as means of functions - serves an important pragmatic purpose it is important to emphasize that functions are not the result of the properties of an isolated thing only, but a result of individual things interacting with other things in their immediate environment. That is, in order to account for the actual possibilities of things to serve as means one needs to adopt a wider causal perspective, taking into account both the properties of the individual thing, denoted as a means, *and* the properties of things in its immediate environment with which it interacts. This is particularly needed when modeling maritime work domains.

For instance, when conceiving the rudder as a means of the production of a transverse force acting on the vessel important causal aspects are backgrounded, e.g. that the transverse force is produced by the water inflow to the rudder acting like an airfoil. That is, the transverse force is determined by properties of the rudder (area, rudder angle and lift-coefficient) *and* properties of the water inflow (velocity and density). Consequently, if there is no inflow of water the possibility of using the rudder as a means of force production vanishes.

The abstraction hierarchy is not explicit about how to describe the role of the immediate environment of the things denoted as means. Neither is many of the applications of the abstraction hierarchy, see e.g. (Vicente, 1999).

Integrating Means-end and Causal Views on Force Production

In order to control a complex system it is important to understand the possibilities of subsystems and components in relation to the production of different types of system actions or processes. Generally, an actor brings about a state change of processes in a complex systems by influencing the way its components interact, and how they interact with things in the environment of the system – roughly speaking one can say that the actions of the actor changes the actions (processes) of the system.

For instance, when the crew onboard a vessel wants to bring about a change in the turn-rate of the vessel it needs to perform appropriate actions (interventions or omissions). According to the abstraction hierarchy above (Table 1) changes in the state of motion of the vessel are achieved by means of bringing about changes in the balance of controllable and uncontrollable forces acting on the vessel, which in turn are brought about by means of changes in the forces being produced. In order to make decisions about possible interventions the crew must have knowledge of the underlying causal mechanisms determining force production. More specifically, the crew must be able to predict the possibilities of force production provided by control devices (rudder, main propeller etc.), interacting with other things in their immediate environment (water, hull, etc.).

Actually, it is important to have knowledge of the causal mechanisms underlying both the *controllable* forces (stemming from control devices) and the *uncontrollable* forces acting on the vessel (stemming from disturbances). This requires an explicit account of the causal role of environmental factors with respect to the possibilities of force production of the individual things being in focus (means), i.e. both control devices and disturbances.

Causality, action and things

In the literature on generative approaches to causality (e.g. (Bunge, 1959, 1977) and (Harré and Madden, 1975)) one can find descriptions on the link between *action* and *things* engaged in action. One of the goals of these approaches has been to anchor the concept of causality to the nature of things having *dispositions* or *causal propensities* to enter into actions. This has certain similarities with the principles of abstraction underlying the abstraction hierarchy, according to which things or structural entities are conceived as means of higher-level functions. Our goal is to take advantage of the similarity between things having causal propensities (dispositions) and things conceived as means, and propose an integration of the causal view and the means-end view of complex systems.

According to Bunge (1977) the *possibility* for an action depends on the disposition of things to enter into an action, while its *actualization* involves the actualization of thing's dispositions: "...the actualization of a potentiality of a thing x involve some thing y other than x and forming part of the environment of x. This other thing, to be called the *complement* of the first, must have a disposition matching that of the thing of interest, for the first disposition to actualize. Thus a certain key may open provided it is joined to a proper lock with the unlocking disposition. What exhibits an actual or manifest property is the whole formed by the thing of interest and its complement." (p.180-181, emphasis in original)

This complementarity of things is also found in (Harré and Madden, 1975). Here the pair of things with the potency to enter into an action is referred to as *agents* and *patients*, respectively. The agent acts upon the patient whereas the patient is being acted upon by the agent. According to this view the complementarity of dispositional properties of the things are understood in terms their degree of engagement in action, being either *active* or *passive*. The disposition of the thing being an agent is denoted by its *power* to act, whereas the disposition of the thing being a patient is denoted by its *liability* to be acted upon.

Means-end and causality

Generally, when conceiving a particular thing as a means it is presumed that it has the causal propensity to bring about a desired action or function of the system. That is, what is denoted a means is typically either a patient or an agent in a specific action context, leaving the complement in the background. For instance, when describing the means of transverse force acting on vessel the rudder (patient) would normally be in focus, while the water inflow to the rudder (agent) is backgrounded.

Hence although it serves a pragmatic function to focus on individual things as means it is also somewhat problematic as it hinders a detailed picture of the actual possibilities (and constraints) of achievement of actions or functions at the next higher level of abstraction, which are partly determined by the properties of the things forming part of the immediate environment of the thing conceived as a means. Nevertheless, in applications of the abstraction hierarchy means are often seen as individual things without an explicit focus on their immediate environment.

Examples

In this section we propose a causal account of the force production of some standard control devices fitted on a larger vessel. In agreement with the above, it is shown that force production is the result of the properties of at least two things interacting. It should be emphasized that it has not been our intention to give a complete account of the interaction between things involved in the production of forces acting on a vessel. Rather the purpose has been to illustrate a way of modeling the causal basis of means-end relations needed in order to account for the actual possibilities of force production of particular means in particular situations.

For each of the descriptions that follow the focus is on the purposeful production of a particular type of force acting on a vessel - longitudinal or transverse force. Within the scope of force production a distinction is made between things that are active and passive, referring to these as agent and patient, respectively. The relevant causal propensities of the agent (power) and the patient (liability) are described in terms of causal properties and the symbols representing their values. Furthermore, it is illustrated how changes in the value of these properties are related causally.

Main propeller as a means of longitudinal force production

Longitudinal force production of the main propeller is an action involving the main propeller (agent with power) acting upon the inflow water (patient having liability). The power of the main propeller is determined by its diameter, thrust coefficient, and its rotation, whereas the liability of the water is determined by its density and axial flow into the propeller.

Table 2. A causal account of the production of longitudinal force production of the main propeller.

Action	Longitudinal force production of the main propeller (X_P)	
Causal Roles	Agent	Patient
Things	Main propeller (fixed pitch)	Inflow water to propeller
Causal Properties	Thrust coefficient (C_T) Propeller diameter (D) Rotation (n: rate of propeller revolutions)	Density (ρ) Axial water inflow (V_A : velocity of axial water inflow)
Mathematical Expression	$X_P \cong 1/8\, \pi\, C_T\, \rho\, D^2 (V_A^2 + (0.7\, \pi\, n\, D)^2)$	
Causal Relations	X_P qprop+ n (*) X_P qprop+ V_A X_P qprop+ ρ	

* (Q1 qprop+ Q2) expresses that Q1 is *qualitative proportional* to Q2, and (Q3 qprop- Q4) expresses the Q3 is *inversely qualitative proportional* to Q4. If a quantity Q1 is qualitative proportional to another quantity Q2, it means that there is a functional relationship between Q1 and Q2, and that Q1 is increasing monotonic in its dependence on Q2 (inversely qualitative proportionalities are defined similarly, with the function being decreasing monotonic) (Forbus, 1984).

It should be noted that causal account given in Table 2 is merely an approximation that does not include all types of influences on the longitudinal force production on the rudder. For instance, the axial flow of water into the propeller (V_A), which influences the longitudinal force production, is a function of the longitudinal velocity of the vessel at the propeller *and* the so-called *effective wake fraction* which in turn is a function of water depth, propeller loading, local drift angle, rudder angle. See e.g. (Simonsen, 2000) for more details.

Furthermore, it should be noted that apart from longitudinal force the main propeller produces also a transverse force on the vessel. According to Simonsen (2000) this force can be approximated by 1 percent of the propeller thrust for forward propeller action and 10 percent for backing propeller. However, only during accelerations and deceleration (transients) the transverse force has influence on the movements of the vessel.

Rudder as a means of transverse force production

Transverse force production of the rudder is an action involving the water inflow to the rudder (agent with power) and the rudder acting as an airfoil or a wing (patient having liability). The power of the water is determined by its flow into the rudder (distinguishing between flow inside and outside the propeller slipstream), whereas the liability of the rudder is determined by its area (distinguishing between the portion of the rudder inside and outside the propeller slipstream), lift coefficient and deflection (rudder angle determining the angle of attack of the water flow).

Table 3. A causal account of the production of transverse force production using the rudder.

Action	Transverse force production of rudder (Y_P)	
Causal Roles	Agent	Patient
Things	Water	Rudder
Causal Properties	Density (ρ) Water inflow inside slipstream of propeller (V_S : velocity of water inflow inside propeller slipstream) Water inflow outside slipstream of propeller (V_A : velocity of water inflow outside propeller slipstream)	Lateral area inside propeller slipstream (A_S) Lateral area outside propeller slipstream (A_A) Lift coefficient (C_L qprop+ α) Deflection (α : rudder angle)

Mathematical Expression	$Y_R \cong \frac{1}{2} \rho C_L (A_S V_S^2 + A_A V_A^2)$
Causal Relations	Y_R qprop+ α (for $0° < \alpha < 35°$) Y_R qprop+ V_S (for $0 < V_S$) Y_R qprop+ V_A (for $0 < V_A$) Y_R qprop+ ρ

Since part of the water inflow to the rudder is generated by the propeller, it is clear that there is a tight interaction between these devices in the generation of transverse force. This is indirectly reflected in Table 3 by the role of the inflow of water to the rudder inside the slipstream of the propeller. Furthermore, the velocity of water outside the propeller slipstream (V_A) is a result of the velocity of the vessel at the propeller and the effective wake fraction. As mentioned above this implies additional influences on the force produced by the rudder from e.g. the hull. However, these are not considered in this paper.

Furthermore, as indicated in Table 3 the lift coefficient (C_L) of the rudder is largely a function of the rudder angle (Simonsen, 2000). Increasing rudder angle lead to increasing lift-coefficient. Apart from the transverse force produced by the rudder also a longitudinal force is produced. The longitudinal force increases as the rudder angle increases.

Conclusions

This paper has addressed the need to integrate the means-end concepts underlying the abstraction hierarchy with concepts originating from generative approaches to causality. The main reason for this is to enable an account of the actual possibilities offered by physical objects, ascribed the role of means. The actual possibilities of such objects depend on both their intrinsic properties *and* the properties of causally related objects in their immediate environment. It has been illustrated that this is particularly obvious for maritime work domains where the force production of control devices (e.g. main propeller, rudder and thruster), as well as disturbance sources (e.g. wind and current) are highly dependent on external properties of the water, hull, etc.

Through simple examples it has been illustrated that means-end concepts fits well into a generative causal account of force production. From a causal point of view, what is typically referred to as a means corresponds to an individual thing having a specific causal role in an action involving additional things (having other causal roles) that are typically backgrounded when adopting a means-end view.

References

Bunge, M. (1959). Causality. The place of the causal principle in modern science. Harvard University Press, Massachusetts.

Bunge, M. (1977). Treatise on basic philosophy (Vol. 3). Ontology I: The furniture of the world. D. Reidel Publishing Company. Dordrecht-Holland/Boston-USA.

Forbus, K.D. (1984). Qualitative Process Theory. Artificial Intelligence, Vol. 24, pp. 85-168.

Harré R. and Madden, E.H. (1975). Causal Powers. A theory of natural necessity. Basil Blackwell. Oxford.

Lind, M. (1999). Plant Modelling for Supervisory Control. Trans Inst MC, Vol. 21. No. 4/5, pp. 171-180.

Rasmussen, J. (1986). Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering, Amsterdam: North-Holland.

Rasmussen, J. (1991). Modelling Distributed Decision Making. In: Rasmussen, J., Brehmer, B. and Leplat, J, Distributed Decision Making: Cognitive Models for Cooperative Work. John Wiley & Sons, Inc., New York.

Simonsen, C.D. (2000). Rudder, Propeller and Hull Interaction by RANS. Ph.D. thesis. Department of Naval Architecture and Offshore Engineering, Technical University of Denmark. ISBN 87-89502-33-7.

Vicente, K.J. (1999). Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Lawrence Erlbaum Associates, Publishers, Mahwah.

Modeling II

Operationalization of non-formal theories in operator-models

Thomas Jürgensohn, Thomas B. Sheridan
3-354 MIT, Cambridge MA 02139, USA, thomasj@mit.edu; sheridan@mit.edu

Abstract

Quantitative and formal models of dynamic actions of humans in machine environments are usually based on formalisms that interrelate measurable or computable features of behaviour or set up relations of these features with time. Mainly engineers have set up such models for about 50 years as a tool for the development of traffic systems, control rooms of nuclear power or chemical plants, or aircraft. Problems in these models do not lie in the dynamics of action, but in the description of influences, which are cognitive, motivational, or emotional. This paper deals with the integration of such factors, which are normally described non-formally in psychological theories and models into quantitative and formal models of dynamic action – explained for the modelling of motivational influences and mental models. The methodological problems are pointed out for the hypothetical example of driver assistant systems.

1 An example

A typical example for formal modelling of driver behaviour is the modelling of the continuous steering of the vehicle. Such a model could be set up within the frame of developing a assisting system (e.g. obstacle assistant and predict the behaviour of the entire system). It is assumed, that the system supports the driver, but keeps him within the loop all the time. So, the driver can still be responsible for his actions.

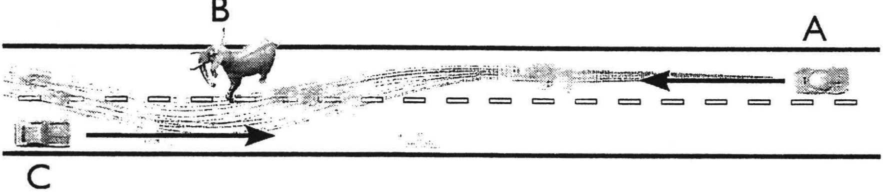


Figure. 1 Prediction of behaviour of vehicle A for sudden appearance of an obstacle B and traffic C

To evaluate the system it is crucial to know whether an accident is avoided by the system in a statistical sense, or if there is the possibility that it can be harmful, for example by disturbing the driver in his actions. Since an accident here means contact with the obstacle, answering this question is only possible if the contact can be calculated independently of the situation. This requires knowledge of the vehicle position at the moment of the obstacle appearance (figure 1). As the driver can influence the system behaviour during the support phase, a prediction of the system behaviour requires a model of the operator – i.e. the driver. Of course, the behaviour of the assisting system must also be computable.

In the above example the driver together with the vehicle and the assisting system forms a continuous dynamic system. Let us assume that the continuous actions and reactions of the driver depend on the accuracy of the driver's situation judgment, whether he already was in the same situation before (so that he can rely on learned knowledge), whether he is timid, alert, thoughtful or hectic driver, whether he is hurried, distracted or in love. Let us call these influences mental factors. The accuracy of a model that does not account for mental factors therefore depends on the extent of the mentioned influences on the action. Mental influences are normally defined non-formally. The focus of this paper is the consideration of these non-formal factors into formal-dynamic driver models. Central is the modelling of motivation influences, because the non-formal character of motives is obvious. Based on this approach the formal modelling of mental models or mental representations will be investigated.

2 Classification of Model Types

Before we embark on the central issue just pointed out, let us clarify some terminology and basic characteristics of models in general. In this paper the main distinctive feature of models is whether they are formal or non-formal.

Formal modelling is characterized by a fixation of attributes of the modelled system to free elements of formalism and at the same time changing these attributes according to the rules of the formalism. As a typical feature of formal models the behaviour of the modelled object becomes known only after computation (after several steps of the formalism). For example, a model of steering behaviour by means of differential equations needs the solution of the equations using real numbers to get an idea what really happens. Of course this solution could be done purely mentally. Therefore, the difference between formal and non-formal cannot be set equal to *computer based* versus *mental*.

Non-formal models are models that are not formal – that means one of the main features, the fixation or the change on the basis of fixed rules, is absent. However, models can be fixated and yet be non-formal. A typical example for fixated but non-formal models are descriptive models. These are made up usually of a number of non-formal sentences about a subject and together form a model of a subject. They are fixated by the medium language, but cannot be modified according to rules.

Distinctive feature of non-formal models is not necessarily the absence of formulae. Linguistic expressions translated into formulae are also non-formal, if there are no formal operations done with them. The formula $\forall \neg \exists F \mapsto r(t^{+}) \neq 0$ is non-formal, if it just means "it is true for almost all drivers, that they react immediately upon burst of the tire", and no feature of the formula is further used. On the other hand, linguistic expressions can be formal, if they are modified or combined according to rules. A set of sentences, which express implications and are logically combined, can from a formal model. This can be "mentally" computed, that means: modified according to a formalism.

Often *non-formal* is set equal to *qualitative*. However, this is not meaningful, because there can be formal qualitative models. The distinction qualitative-quantitative is related to measurable features. Qualitative Models describe measurable features by means of a small number of discrete values, on which an ordering relation is defined. Linguistic models with expressions like small, big, thin, thick, fast, slow are an example of qualitative-non-formal models. Models that use interval arithmetic or rules are examples of qualitative-formal modelling (e.g. [5]).

For the topic of this paper we must clarify what we understand by the term *dynamics model* (we distinguish here between a dynamic model and a dynamics model. The latter is a model of dynamic systems, the former a model that is dynamic itself). We use the term dynamics as it is used in system theory: "*Dynamics is the rules of changes in time*". All systems that change continuously in time need the relation to time for their description of the time variation, because causality in a continuous state space cannot be expressed by means of transition rules between discrete states. They are always dynamic systems that are described by means of their state values. Therefore many macroscopic-physical systems are dynamic systems, if they are described as being continuously time variant. If a dynamic system coupled to a non-dynamic but causal system is in one combined model, the entire model must be a dynamic system, because the time relation cannot be resolved in causal relation. Every system that influences causally a dynamic system is a dynamic system itself. Its behaviour must be described as time-dependent. In the same way two or more causal systems that cannot be coupled in a complete causal manner (sometimes stochastic causal) can be combined into a system only after transformation to a dynamic system.

3 Modelling of dynamic action during machine interaction

The modelling of dynamic action in machine environments has been investigated since the 40s of the last century. There are a number of publications that give good summaries on that issue [33, 26, 16, 5, 18]. As the focus of this paper is on non-formal models, the issue is dealt with only to make implications of the combination of non-formal and formal-dynamic models apparent.

As mentioned in the preceding section, time variant behaviour of a continuously varying system cannot be described just causally, but requires the modelling of the dynamics – that means an explicit relation to time. Of course this is also true for the system "driver of a vehicle". All approaches to model operator behaviour in dynamic machine environments describe the operator as a dynamic system. The reason is not only the fact that humans always are dynamic systems, but that they interact with dynamic systems. This assertion is explicated more in detail in the following.

Let us come back to our example: The starting point of our considerations is the set of the trajectories of the vehicle A shown in fig. 1. They are important for our issue, because their knowledge would allow a prediction of a potential accident. They are without doubt continuous in time and space, and this means that the vehicle is a dynamic system with respect to this consideration. Its behaviour can be calculated quite accurately on the basis of numerical models of the vehicle.

Since for the description of driving trajectories the vehicle forms a continuous dynamic system, and the driver affects these trajectories, the driver himself is also a dynamic system as pointed out before – his behaviour must be related to an independent time. Because the driving trajectories result from the numerical model of the vehicle, a description of driver behaviour must contain at least one causal connection with the vehicle. In our example we find influences on steering, braking and operating the accelerator pedal.

We assumed that the driver in spite of the assisting system always exerts full control and is responsible for his actions. This means that his actions are determined by the goal to minimize damage to the vehicle or harm to himself or other persons. In order to realize this goal, he needs information from the environment and the behaviour of the vehicle. Hence the driver behaviour depends on objects or states of the environment – they are causal. Using the same logic as before, when we said that a continuous driving trajectory implies that the driver is a dynamic system, we can conclude that the environmental information has to be dynamic, too. It is dynamic, even if no object in the environment is dynamic.

This somewhat tortuous deduction describes nothing but a dynamic control-loop. A model that describes this control-loop hence has to be a dynamic model. This is not self-evident. There can be models of control-loops that are not dynamic – if driving is modelled as a closed causal-loop, but none of its elements is dynamic. This could be a descriptive (linguistic) model of causality in our scenario. We can derive causal conclusions from our concept, which is formed by the model, but not related to an independent time flow – hence no dynamical ones.

4 Motives

One of the basic needs of humans is the acquisition of a coherent conception of the world. The search for rules about the "way of the world", for reasons and causes, is the most important drive of inquisitive thought. Every search for reasons has to come to an end – if only for the reason of limited time. These are the final rationale, natural laws or divine will.

Of course, our drive for fundamental research does not exclude ourselves. We search for reasons, why we behave the way we behave. Our behaviour depends to a large extent on the state of the world, on external circumstances – we adapt our behaviour to the necessities of the world. However, part of the reasons for a specific action can be found within ourselves. Here we have to differentiate between internal causes that are related with particular external present or past circumstances, and causes without this correlation. Influences on actions that have something to do with perception, evaluation and decision are cognitive determinants of an action. Contrary to that are the causal sources of behaviour that lie completely within us and are not directly related with a particular situation: the *motives*.

Closely connected to motives are *emotions* – often there is no sharp distinction between the two, because emotions can also be thought of as causes of behaviour. The co-operation of intellect and emotion/motivation is articulated in a very illustrative way by Plato (-380). He thinks of the *psyche* as a chariot: the charioteer "reason" conducts his chariot drawn by the horses "courage" and "desire". Reason corresponds to what we call today cognition; courage and desire could be compared to our motives.

Due to limited space it is not possible to recapitulate the characteristics of motives as they have been investigated in motivation psychology. Let us instead summarize the properties of motives, which can be deduced from their definition (internal causes of behaviour). For further reading please see [14,10,22,25,35].

Because motives are not related to a specific situation, one single cause cannot explain single actions, but only certain *aspects* of an action – due to the great variety in external conditions. Some aspects of an action become more probable than others. Whatever can be thought of as an aspect of an action depends on the type of action. Examples are the starting point, the vigour, or strength of an action element. Very often these aspects are high-dimensional patterns of features and difficult to describe, but possible to comprehend. The features that describe a "brave" action, for example, are too complex and mutually interdependent to be completely explicable. Motives, when conceived as origins, are always taken as one-dimensional. Their effect on actions combines usually a number of features and aspects.

Taking motives as causes of actions indicates a difference in the dynamics of motives and actions. Only longer observations reveal possible tendencies in actions. In system theoretic jargon: the more noise is in the system the more time is needed to identify system parameters. Motives as an explanation of complex behaviour therefore have to be prolonged over behaviour, but are not necessarily time-independent.

Motives describing action aspects have to be associated with actions (we include thought as a "mental action"). Depending on the motive, this association can be more or less specific to a certain class of actions. Another important consequence for motives comes from the definition as an *internal* cause: Motives as entities of our body or mind must be identifiable or measurable – or external causes must be impossible. An internal cause can be measured with ease if it is related to something perceivable like hunger. Therefore we can find a lot of motives related to entities of our body in theories of motivation. With same reason, we also find many motives combined with emotions, as fear, anguish, joy etc in the literature.

External factors as action causes can also be ruled out if different people act differently in an identical situation. For this reason dispositions or personality factors, if associated with an action are also named motives. If we just observe a limited action fragment it is not possible to distinguish between permanent and temporary behaviour dispositions. It is not possible to decide from the behaviour of a driver in a certain situation, whether he is a timid driver or whether he is just anxious at the very moment. This would require long-term observations. Yet another condition to rule out external causes for behaviour is the observation that one person behaves differently in similar situations. If driver A needs 11 minutes to drive to his office instead of the usual 15 minutes, even if all external factors remain constant, there must be an internal cause, which we could call "hurry". Of course, the hurry certainly has a reason, but this is beyond our observation scope.

What motives are and how many exist, does not depend on human behaviour but also on the benefits in the interpretation. Plato wants to describe the entire human soul and restricts himself to two motives – towards the end of the 19th century drive psychology had "discovered" several thousand motives, each of which served as an explanation in a very specific action environments. This could not prove useful, because it was not possible to decide which of the many motives was valid in a given situation. Modern motivation psychology reduced the number of elementary motives to 6-15 for a given action context (e.g. traffic behaviour, purchase, etc.).

Motives, therefore, do not only depend on our behaviour, but on the way we want to generalise behaviour in the sense of non-formal modelling or what we want to explain. Both number and choice of motives depend on our capabilities for *non-formal modelling*. Another issue is measurability. The basis of motives are attributes of action, which are based on non-formal measurements. The procedure of identification of action tendencies and action aspects can be interpreted as modelling of a *vague causal correlation*.

Because of the merely tendencial effect of motives on actions they can explain a single action only with respect to an aspect of multitude. If the number of possible motives has to be limited in a non-formal modelling process, each motive has to appear in many different actions as an aspect or tendency to cover all kinds of actions. But a multitude of actions together are always continuous are quasi-continuous. Tendencies or aspects of this bunch of actions thus have to be continuous, too. This means that motives are measures, entities with an ordinal characteristic, to which moderators like "more" or "less" can be applied. In the literature, only few motives can be found that are not conceived to be continuously measurable. All reactions of the body combined with emotions and feelings can be associated with an intensity. This is one reason that they were so often named motives.

We have outlined the motives from a viewpoint of modelling behaviour. Motives result as certain elements of non-formal modelling, base on observation and feeling. Their effect is on level above direct action. Because of this direct relationship with the observation scope they are more dependent on the objectives of the modeller than on action determinants, which are related to particular actions. What is declared a motive therefore depends strongly on what is to be explicated. For this reason motives are often named "constructs" in psychological literature. This gives them a certain subjective touch. Motives are tied both to the modelled actor and to the modeller. This means, a person can have different motives, depending on the view of the modeller.

5 Motives in quantitative-dynamical models

To translate non-formal motives into quantitative-dynamical driver models three strategies can be applied. Let us use the motives fear and hurry for our investigation.

1. Translate the non-formal constructs fear and hurry into a formal model. Then this formal model is combined with formal-dynamical aspects of the driver model in a total model.
2. Fear and hurry are also considered as states of a person, visible in action attributes. But now the non-formal modelling based on human observation is omitted. Instead only experiments with physical measurable attributes are used. By this way, fear and hurry are modelled and thereby defined in another way.
3. Instead of the motives being formally modelled, their effects on actions in the investigated action context are defined. In this case the basis of the model set up are non-formal constructs, but without explicitly modelling them. Here the mental application of a mental model of the situation is formalised, where the mental model comprises motives as elements.

By the introduction of the construct of a motive *variant 1* can be ruled out as practically unworkable: Since motives like fear or hurry are aspects of non-formal modelling they depend to a large extent on the modeller. Thus models of motives have to include a model of the non-formal modelling procedure. It is beyond doubt that this can be ruled out because of the possible influences of fear on different actions on the one hand and the principle impossibility in formalisation of a feeling like fear.

Variant 2 seems possible – the concept of a motive as a cause of behavioural aspects of a person would be translated into formal modelling. But a brief reflection shows that this is either unrealisable or does not conform to the original objective. One can think of recognize personal attributes in measurable features of behaviour (in our example steering, braking or pressing the acceleration pedal), and combinations of these could even be named motives at times, since their role in the description of behaviour is just that of "normal" motives. Because of the very limited possibility to experimentally investigate life situations as a whole, only "micro motives" could be found, and no comprehensive motives like fear and hurry.

The requirements for *variant 3* are the weakest. In a simplistic formulation this procedure means to set up a formal model that behaves in a way as "we imagine". It can be imagined that the probability of our driver choosing to brake before the object instead of accelerating and overtaking to resolve the conflict depends whether he is generally timid or bold. This prediction is possible because of our non-formal model of fear and our general world knowledge. A formal modelling according to variant 3 now is nothing but modelling of these conceptions. That means that fixated elements of a model are interpreted as motives or aspects of motives. But it does not mean that motives themselves become formal constructs.

5.1 Basic rules of formal modelling taking account of motives

In the preceding chapter some conditions have been formulated to consider motives in a quantitative dynamics model starting from a reflection on the motivation concept from a model theoretic viewpoint. The main result was the statement that both modelling of motives and the validation in formal models can only be non-formal due to the non-formal character of motives. An important consequence is that meta-personal objectivity of these models as seen in formal models is impossible

All the interpretations of structures, parameters, or variables in dynamics models collected under variant 3 (see the preceding section) as an expression of motivational influences are fixated, but due to their semantic coupling to a non-formal construct they remain non-formal and thus closely connected to the modeller. However, they are not arbitrary, because a certain consensus between the modellers can be assumed. Some basic rules for the consideration of motives in formal-dynamical models are summarized:

1. Parametric influence

The influence of motives is always in a level above the time variant state variables. In a formal application motives are therefore represented as *parameters* of a certain description of behaviour. The parametric influence can range from parameters of a description of movements, to decision thresholds or structural parameters. The parameters themselves can be time variant. Also a feedback of state variables on the motivation parameters is possible, but only as long as the dynamic of both levels remain distinct.

2. Strength

Motives are one-dimensional parameters having strength. In a formal modelling this means a mapping onto a measuring variable. This are usually real positive numbers. All approaches known to the authors, where motives exist in formal models, are realized in this way [e.g. 37,20,4,14,9,19]. Besides real measuring variables also qualitative measures are thinkable (tiny... huge) or fuzzy measures. Whether motive strength only takes values of an interval (e.g. out of [0,1]) or without limit cannot be derived from the motive conception. A limitation has the advantage that results can be compared.

3. Objectivation and Validation

In principle motives like hurry or fear can be modelled in formal dynamics models without specific psychological knowledge – they are part of common world knowledge. Some examples of general knowledge about the influence of fear on behaviour were stated. However, a model set up can only be "optimal," if the non-formal basic knowledge is optimal as well. The more elaborate the modelling conceptions about motivational (and also cognitive) influences on driver behaviour are, the more valid will be the application in a formal model. As non-formal models can only exist within the heads of the modellers, this means that formal modelling of mental influences on driver behaviour requires a corresponding education of the modeller. "Objectivation" of the models in this context is equivalent to "professionalise" the modelling by communication and knowledge exchange.

4. Formal extension

A formal extension of motives requires inclusion of measurable and thus person independent variables and indicators for motives or motive strength on the one hand, and incorporation of these indicators into the parts of motive constructs that cannot be formalized on the other hand. Formal and non-formal parts must be consistent and must not be mutually contradictory.

The extension can be done practically by finding a systematic correlation either in experiments or simulations (again validated by experiments) between indicators that can be physically validated and motives that can only be estimated in a non-formal way. However, a formal extension must not destroy the broad character of motives. This means that correspondences found between formal and non-formal elements only for a limited action domain cannot be interpreted as formal extension. An example for a meaningful correspondence is that between motives and general movement parameters, possibly limited to driving. It would not be meaningful if it were limited to overtaking situations only.

6 Mental Models

"Mental model" and similar notions like "internal model", "mental representation", "situation awareness" or "picture" – to name just a few – are presently some of the most often used terms when describing cognitive processes. Unfortunately the frequent usage comes along with an extensive diversity in conveyed contents (see for example [23,38,38]). There is no consistent distinction between "internal" and "mental" or between "model" and "representation". Because of that inconsistent use we will "internal" and "mental" use in the same sense in the following passages. Possible distinctions between model and representation will be tackled later.

In general, K. Craik (1943) is taken as the originator of the idea to apply the term "model" for the description of thought phenomena. His original objective to introduce the model conception was to describe an analogy between the structure of matter and features of thought derived from properties of neural processes. Because of this mechanical analogy he chooses the conception of model as a specific representation. A purely mathematical representation is not a model in his view. His opinions are strongly characterised by physics and the analogy to calculating machines: a central conscious processor uses mental models to generate predictions of the world. The models themselves are built of symbol systems and the processor is "aware" of them.

Parallel and mostly independent from the discussion in psychology the term "internal model" is used in engineering models of operator behaviour in dynamical systems beginning n the mid 50s. It is in a certain way very similar to Craik's conception. For example the Crossover model by McRuer is named an "implicit internal model" by himself [27]. An "internal model" is realised explicitly in adaptive models or optimal filter models, which describe the capability of operators to adapt to very dissimilar machine dynamics (e.g. [34,21]). The operator is here, contrary to the conception of mental models by Craik, not aware of these internal models. They are the result of an unconscious learning process. In these models, there is generally no distinction between model and representation. Internal models of the operator models are a priori formal models, mostly formulated as transfer functions.

The term "mental model" became known in psychology mainly by the work of Johnson-Laird (1983). His ideas are based on those of Craik, but further developed. The content of Johnson-Laird's mental models is their representation characteristics. The idea is firstly that the owner of the mental model has *beliefs* about the world. Second, a mental model is a *representation* that is similar to the world or has an *analogous structure*. Thirdly, the models are built *temporarily* for problem solving. This follows in great part the mechanistic conception of Craik. A model is a model of facts only if it models the "real structure". The specific point of Johnson-Laird's mental model is their time variance, which means they can adapt to a dynamic environment. This is contrary to the concept of von de Kleer and Brown or Gentner & Gentner [13,8], who mean by a mental model conceptions of the physical world that have been stored in long term memory.

Due to limited space it is not possible here to give a survey on the different utilisations of the term mental model and its derivatives – we have already mentioned the surveys [23,38,38]. In order to clarify the width of the concept, let us present three more examples:

The secure, efficient, and fast control of dynamical air traffic through the air traffic controllers is based on its mental representation of current and future traffic situation updated continuously. These mental representations mediated by radar screen, communication, and additional information (e.g. flight strips) are called "pictures" by the air traffic controllers themselves. Bierwagen [2] means by this term a "*mental model of the actual traffic situation with a category metric*". The term *picture* corresponds to the concept *situation awareness* and is often used synonymously in the literature [14,15].

Situation awareness is another example of mental modelling. It is mainly known in field of air traffic and denotes a mental model of a pilot about the state of the airplane, the automation system, and environment variables. Endsley [11] and Flach [12] have published two examples of non-formal theories (or models) of situational awareness. A formal model of the picture of air traffic controllers based on production systems (ACT-R) is published by Niessen & Eyferth [29]. They call their model *a model of a mental model.*

A third example is a model based on Dempster-Shafer evidence theory, which reflects the ability of a car driver to navigate in a city [3]. The authors do not distinguish between model and representation and propose a pure mathematical model of the navigator's knowledge about the environment that consists of the three components: whereabouts, topology, and landmarks.

The differences in the presented examples make it probable that the conception about the termini varies, too. This can be explained with the origin of a termini-like model, presentation, or even motives as results of non-formal modelling procedures. We try to explain the basic properties of mental models from a "meta viewpoint".

Let us come first to the term representation, that can be explained a little bit more easily. Compared to the term model, which came up in its present meaning only in the middle of the 19th century the term representation is older. Its central element is the function of "standing for". The usage of the term internal representation is very much broader than that of the term model. Beginning with Barlow [1] there is the idea of the internal representation within one single neural cell. This was put to top by Marr [24] with his famous „grandmother cell", which fires, when grandmother enters the room. Rumelhart and Norman [32] have summarised this in a few words: "*a representation is something that stands for something else*". Let us follow this "definition". It is important that for a representation always two components are present: the representandum and the representat. But in addition somebody has to fix these two components. There are no "natural" representations, only those that are taken as such by people.

This division is valid also for models. In each modelling there is a modelled, a model and a modeller. If one of these components is absent, one cannot speak of modelling. Contrary to a representation using the term modelling something "actively constructed" is involved. A modeller builds a model for a certain purpose. Johnson-Laird or Genter & Genter use the conception of a model in just this sense.

If one took this property as a necessary element of the term model, then it becomes clear that internal models of process dynamic of operators must be representations, because they are certainly not generated *intentionally* by the operator – the internal model cannot even be consciously reflected. But now one can show that models need not necessarily be consciously construed, but can be result of an interpretation process. A complete clarification of the term model is still missing, but there can be no doubt that existing systems can be used by a modeller as models of the modelled – without the modeller generating the system. For example planetary systems can be regarded as models for atoms. This means that the internal representation of process dynamics is also a mental model – model and representation are synonymous in this context. But most important is that this is not true for the operator, but for the modeller, who thinks about the operator. Of course the modeller can himself be the operator as well. For the operator as the owner of the representation the representation cannot be a model, because for him there is nothing modelled. In the same way a person who thinks of the world as the middle of the universe, circled by the sun and the stars has a mental model of "reality", which he cannot be aware of. It is a mental model and not reality only for the reason that the modeller who thinks about a mental model knows what is "real" There is a good reason that there are no "mental theories", because the terminus theory does embody the property of reduction of models: models are always thought as something "imperfect" relative to reality.

The conceptions of the modeller about the internal representation can be formal or non-formal. Accordingly, the mental model is formal or non-formal. Hence like motives the mental model depends not only on the acting person but also on the modeller. The part of the modeller in mental models becomes especially clear in so called Shared Mental Models [6,30]. This concept outlines the common part of the situation awareness of crewmembers. It is obvious that such a Shared Model cannot exist as an entity of a mental system, but only in the head of a modeller. The crewmembers themselves generally do not know about such a model. A mental model is therefore a model of a modeller in the head of one or more others. A mental model is an expression of the difference between the real world and (supposed) conceptions about the world. This difference can exist for a third party only, which observes and models both systems.

The consequence for the issue of this article is that mental models can be formal in the sense of an external modeller. It is therefore not necessary to speak of formal models of mental models – mental models can be formal themselves. If formal models of mental models are to be dealt with of, then the same is true for non-formal motives and for the translation of non-formal mental models into formal dynamics models of operator behaviour. A complete modelling of non-formal mental model is not possible, because this would again require modelling the modeller. Possible is only the interpretation of elements of formal modelling as properties of the non-formal mental model, denoted as variant 3 in chapter 5. Therefore there are no models of non-formal mental models, but only non-formal mental models. On the other hand models of formal mental models can be conceived, as in these models the influence of the modeller is not so strong.

At the end a special feature of mental modelling (that means the interpretation of mental objects as models) shall be mentioned. The way motives are related to aspects of action from the viewpoint of the modeller, mental models like all representations are related to external occurrences, thought of by an external modeller. Although the interpretation process is entirely in the responsibility of the modeller, it is improbable that he discovers correlations in completely different things. Therefore the internal model of a vehicle's dynamics is again a vehicle's dynamics, the internal model of linguistic constructs are linguistic constructs, and the internal model of motives are motives. Differences in the description between internal and external circumstances become necessary only if the modeller want to model not only similarities, but also dissimilarities. If one assumes that our driver of the example knows his car very well, it is possible to describe that in a way that his internal model of the vehicle corresponds to the latest state of the art in scientifically describing vehicle dynamics. If we assume that he knows less than the present scientific status of vehicle dynamics, one will choose a different description known to be less accurate, for example a fuzzy model. But this less accurate description depends only on the knowledge of the external modeller and his interpretation of "inaccuracy". Therefore it does not make sense to search for a "true" description of internal models, because this can be impossible.

7 References

[1] Barlow, H.B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception, 1*, 371-394.

[2] Bierwagen, T. (1999). Cognitive processes of air traffic controller during conflict recognition: modeling and implementation (ZMMS-Spektrum, Band 9). Sinzheim: Pro Universitate Verlag.

[3] Boer, E.R.; Goodrich, M.A. (1998). Mental Models in Micro Worlds: Situated Representations for the Navigationally Challenged. *IEEE Int. Conf. on Systems, Man, and Cybernetics, San Diego, CA, USA, Oct. 11-14.*

[4] Bösser, T. (1987). Dangerous and Aggressive Behaviour of the Driver during Distance Control on Highways 2. Köln: Verlag TÜV Rheinland.

[5] Cacciabue, P.C. (1998). *Modelling and Simulation of Human Behaviour in System Control.* London: Springer-Verlag

[6] Converse, S.A.; Cannon-Bowers, J.A.; Salas, E. (1991). Team Member Shared Models: A Theory and Some Methodological Issues. *Proceedings of the Human Factors Society 35th Annual Meeting,* 1417-1421.

[7] Craik, K.J.W. (1943). *The nature of explanation.* Cambridge, England: Cambridge University Press.

[8] de Kleer, J.; Brown, J.S. (1981). Mental Models of physical mechanisms and their acquisition. In Anderson, J.R. (Ed.), *Cognitive skills and their acquisition* (285-305). Hillsdale, NJ, USA: Erlbaum.

[9] Dörner, D.; Hille, K. (1995). Artificial Souls: Motivated Emotional Robots. *IEEE Conference Proceedings, International Conference on Systems, Man, and Cybernetics; Intelligent Systems for the 21st Century. Vancouver, Volume 4,* 3828-3832.

[10] Dörner, Dietrich (1994). Eine Systemtheorie der Motivation. In J. Kuhl (Ed.), *Enzyklopädie der Psychologie, Motivationspsychologie.* Göttingen: Hogrefe-Verlag

[11] Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37(1),* 32-64.

[12] Flach, J.M. (1995). Situation Awareness: Proceed with Caution. *Human Factors, 37(1),* 149-157.

[13] Gentner, D.; Gentner, D.R. (1983). Flowing waters and teeming crowds: Mental models of electricity. In Gentner, Dedre; Stevens, Albert L. (Eds.), *Mental models* (99-129). Hilsdale, NJ: Erlbaum.

[14] Irmscher, Marita (2001). *Modelling and simulation of mativational influences to the driver*, VDI-Fortschrittberichte, Reihe MMS, ZMMS-Spektrum, Band 12). Düsseldorf: VDI-Verlag

[15] Isaak, A.R. (1997). Situational Awareness in Air Traffic Control. Human Cognition and Advanced Technology. In Harris, D. (Ed.), *Engineering Psychology and Cognitive Ergonomics* (S. 185-191). Aldershot: Ashgate.

[16] Johannsen, G.; Boller, H.E.; Donges, E. & Stein, W.H. (Ed.) (1977). *Der Mensch im Regelkreis, Lineare Modelle.* München, Wien: Oldenbourg Verlag

[17] Johnson-Laird, Philip Nicolas (1983). *Mental models.* Harvard: Harvard University Press.

[18] Jürgensohn, T. (2000). User Modelling. In K.-P. Timpe; T. Jürgensohn & H. Kolrep (Ed.), *Mensch-Maschine-Systemtechnik* (107-148). Düsseldorf: Gabler-Verlag

[19] Jürgensohn, T.; Irmscher, M.; Willumeit, H.-P. (1997). A Model of Individually Acting Drivers Influenced by Motivational Factors. In Borys, B.-B.; Johannsen, G.; Wittenberg, C.; Strätz, G. (Eds.), *Proceedings of the 16th European Annual Conference on Human Decision Making and Manual Control Dec. 9.-11. 1997, Kassel, Germany* (260-267).

[20] Kageyama, I.; Pacejka, H.B. (1991). On a New Driver Model with Fuzzy Control. *Vehicle System Dynamics, Vol. 20,* 314-324.

[21] Kleinman, D.L.; Baron, S.; Levison, W.H. (1970). An Optimal Control Model of Human Response, Part I: Theory and Validation. *Automatica, Vol. 6,* 357-369.

[22] Kuhl, J. & Heckhausen, H. (Ed.) (1996). *Enzyklopädie der Psychologie.* Göttingen, Bern, Toronto, Seattle: Hogrefe Verlag für Psychologie

[23] Lloyd, Dan (1987). Mental representation from the bottom up. *Synthese, 70,* 23-78.

[24] Marr, D. (1982). *Vision.* San Francisko: Freeman and Company.

[25] McClelland, David C. (1985). *Human Motivation.* Glenview: Scott, Foresman and Company

[26] McRuer, D.T. & Krendel, E.S. (1974). Mathematical Models of Human Pilot Behavior. *AGARDograph, No. 188.* London: Technical Editing and Reproduction Ltd. Harford House

[27] McRuer, D.T. (1980). Human Dynamics in Man-Machine Systems. *Automatica, Vol. 16,* 237-253.

[28] Moray, Neville (1997). Models of Models of... Mental Models. In Sheridan, T.B.; van Lunteren, T. (Eds.), *Perspectives on the human controller* (271-285). Hillsdale: Erlbaum.

[29] Niessen, Cornelia; Eyferth, Klaus (2001). A model of the air traffic controller's picture. *Safety Science, 37,* 187-202.

[30] Orasanu, J. (1994). Shared problem models and flight crew performance. In Johnston, N.; McDonald, N.; Fuller, R. (Eds.), *Aviation psychology in practice* (255-285). Aldershot, UK: Ashgate Publishing.

[31] Platon (-380). *πολιτεία.*

[32] Rumelhart, D.E.; Norman, D.A. (1988). Representation in memory. In Attkinson, R.C.; Herrnstein, R.J.; Lindsey, G.; Luce, R.D. (Eds.), *Steven's handbook of experimental psychology* (511-587) (2. edition). New York: John Wileys & Sons Ltd.

[33] Sheridan, T.B. & Ferrell, W.R. (1974). *Man-Machine-Systems: Information, Control, and Decision Models of Human Performance.* Cambridge, Ma: MIT Press

[34] Smallwood, R.D. (1967). Internal Models and the Human Instrument Monitor. *IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 3*, 181-187.

[35] Thomae, Hans (Ed.) (1983). Theories and Forms of Motivation (Enzyklopädie der Psychologie, Themenbereich C (Theorie und Forschung), Serie IV (Motivation und Emotion). Göttingen, Bern, Toronto, Seattle: Hogrefe Verlag für Psychologie

[36] Whitefield, D.; Jackson, A. (1982). The air traffic controller's picture as an example of a mental model. *Johannsen, G.; Rijnsdorp, J.E. (Eds.), Proceedings of the IFAC Conference on Analysis, Design and Evaluation of Man-Machine Systems.* London: Pergamon Press, S. 45-52.

[37] Wiedemann, R. (1974). *Simulation of Traffic Flow*, Habilitationsschrift, Universität Karlsruhe.

[38] Wilson, J.R.; Rutherford, A. (1989). Mental Models: Theory and Application in Human Factors. In *Human Factors* (617-634).

Effort-Management: a necessary concept to understand driving

Hans-Peter Krüger, Alexandra Neukum, Volker Hargutt
Center for Traffic Sciences, University of Wuerzburg

Driver-adaptive assistance systems have to take into account both the cognitive and energetical aspects of the driving task The article deals with the regulatory processes in the energetical domain. It is empirically demonstrated that driving has direct energetical consequences. Additionally, in the context of long term driving it is shown that the drivers alter the difficulty of the driving task in order to reach or maintain an energetical state needed for safe driving. It is concluded that these processes of effort management should be included into the modeling of driving behaviour.

The Need for Driver Models

There is a growing interest in powerful models of driver behaviour. New challenges come from CAE applications where it is aimed to run models of cars or new technical solutions within a simulated environment at a very early stage of the engineering process. Up until now, controller models which have little or no connection to the human way of driving, have been used to simulate driving within these models. They do not claim to model human behaviour itself, but merely to hold the car on the road. In contrast, adaptive driver assistance systems need valid information about the individual driver behaviour to decide whether a system action should be launched or not. Thus the choice of actual system answers, must be based on a deep knowledge of drivers' normal behaviour in given situations. The first attempt to construct a system which, in contrast to controller solutions tried to model the human action itself, was proposed by [Onken, 1994] who developed a system able to learn the driver's behaviour in a very short time.

The recent developments in intelligent cars have made it increasingly necessary to combine both issues. For example, to test a new solution for a steering mechanism requires a human simulation as well as including inherent steering errors, for a steering analysis to be complete. This new class of driving models must refer directly to the components of human action. Up until now, those driver models have mostly been cognitively oriented, based on models of information processing. The first group of models start on the micro-level of single actions, looking on chains of input-processing-output-Relations (or in psychological terms: on stimulus-organism-reaction S-O-R-connections). Those models are used for example in the discussion of automated vs. controlled actions. The second group of models is more general, establishing a hierarchical system of information processing units. Most widely used is the model of [Rasmussen, 1986] with the discrimination of skill-, rule- and knowledge-based behaviour.

Unrelated to this cognitive dimension of driver behaviour, a second intensive discussion has been taking place in the last few years. It deals with the energetical processes of driving behaviour and is connected with terms like "state of the driver", "driver conditioning", and "fatigue". There is substantial scientific effort to develop systems, able to detect the actual state of the driver (e.g., "drowsiness warning") and to intervene in case of critical situations. The main finding in this area is that the drivers' state is a highly regulated process, strongly influenced by the driving task. Even more important is the finding that in order to establish a sufficient energetical state to solve the driving task, the driver makes use of compensatory actions which lead to a continuous redefinition of the driving task.

At this point, a third present-day discussion comes into focus. The potential problems represented by the fulminant development of infotainment systems, likely to be implemented into cars in the not too distant future, have until now, mainly been discussed within the framework of human-machine-interface and the interference of those systems with the driving task by using the same resources. Modelling has to take into account the fact that infotainment systems will be used during driving and, therefore, have to be integrated into a concept like the "global driving task".

Taking these considerations together, the task of the driver (which has to be modelled) has to be redefined into three subtasks:

- solve the basic driving task which can be looked at in terms of information processing,
- guarantee a sufficient activational state which allows the solution of the driving task,
- integrate other sources of informational inputs given by situational demands and infotainment systems which are not directly linked with the driving task.

Future modelling has to take into account this triple task and its management by the driver. The following contribution will concentrate on the activational dimension of driving and its interactions with the driving task. The basic intention is first to demonstrate how strong the driving state is interrelated with activational processes. Second, it will be shown how regulatory processes of the driver in order to maintain a given activational state are changing the driving task itself. The basic elements of this process of "effort management" by the driver will be demonstrated as well as the consequences on the definition of the driving task.

The Energetic Management

The workload approach

The basic assumption is that the informational elements of the driving task (i.e., every task) are strongly linked with energetic processes. A well known and widely used concept combining both views is the concept of workload. Beginning in the 1970s, the workload approach was developed in various areas, especially in the field of human factors, and now comprises a wide variety of methodological issues. The most generally agreed definition of workload is given by [O'Donnell & Eggemeier, 1986]: "The term workload refers to that portion of the operator's limited capacity actually required to perform a particular task. The objective of workload measurement is to specify the amount of expended capacity". In a comprehensive overview, they classified the assessment procedures into three major categories: (1) self-report measures, (2) performance-based measures, and (3) physiological measures. The workload approach basically assumes that workload is a characteristic of the process of working rather than an outcome. It is a multidimensional, multifaceted concept, which applies in at least three parameter domains: in quantity and quality of performance, in somatic actions and reactions, and in cognitive representations of the task (difficulty) and the worker's own state (e.g. fatigue). All three domains are valid representations of workload and are mutually interdependent. In fact, workload is more a methodological concept than a formal model. It requires (1) to apply different measurement domains and (2) to concentrate the measurement onto the working process itself. There was no attempt to develop a model of human action within this issue – surely one reason for the extended application of the concept in heterogeneous contexts.

The cognitive-energetic approach of SANDERS

The first to combine the cognitive with the energetic perspective on human performance within a formal model was [SANDERS 1983]. Figure 1 shows his model. The following considerations will make use of his terminology.

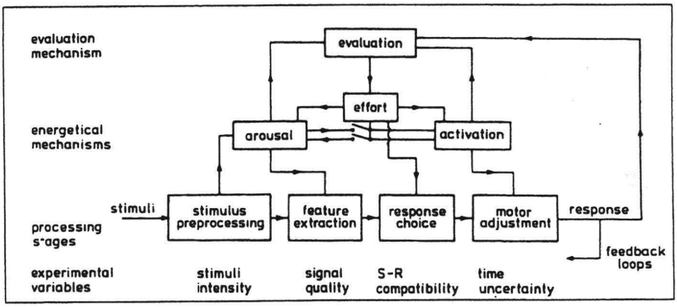


Figure 1: The model of stress and human performance [Sanders, 1983].

Sanders' issue is cognitive because on the processing level he makes use of the traditional model of information processing. The issue is energetical because the elements of the information model are directly linked with energetical systems (stemming from studies conducted by [Pribram & McGuiness, 1975]. Informational input is linked with an arousal system which is mostly activated by the stimulus situation. High input of stimuli (intensity, frequency) leads to an increased arousal by bottom up mechanism. The activation system is located at the mechanism, responsible for coordinating the output. The main task is the control of motor processes and their adaptation on the actual situation where the reaction takes place. In case of automated actions, the arousal and the activation system work in direct connection. They are primarily stimulus driven and are able to solve automated tasks within a given range of variations in the stimulus as well as in the reaction situation. Mostly this process is unconscious. In cases where automated actions are lacking or if the stimulus-response-connection gets incompatible, controlled actions must be introduced. This is done by a higher order effort system which plays a central role in the conscious regulation of activational processes. In the subjective experience this systems is represented as "being effortful", by a controlled assignment of psychophysical energy.

The basic implication of this cognitive-energetic model is that human action must be understood in a twofold manner. In case of automated reactions the system will work in a more or less bottom-up mode which is dominated by the features of the task. In case of controlled actions the control is reversed and works in a top-down mode. Here the assignment of arousal and activation is controlled by an effort mechanism leading to a controlled way of information processing. Applying these considerations to the driving task requires empirical evidence that

- driving is directly linked to energetical processes and
- regulation of the energetical state by the driver leads to changes in driving.

In some way comparable to the model of Sanders is the proposition from [Hockey, 1993] to understand control and stress regulation. He postulated two main control loops. Loop A is seen as a "routine" control system maintaining behavioural stability. This system "compares feedback from current cognitive activity with the target state, and activates familiar (automatic) adjustments which modify behaviour until it matches the target". If the discrepancy between current state and target state is too high, loop B as supervisory level takes control over the adjusting processes. Either additional resources are recruited by introducing effort as compensatory mechanism or central motivational priorities are adjusted downwards in order to meet available resources. A shift from automatic to controlled processing is initiated.

Driving as an energetical process

The relationship between workload and task difficulty

Vehicle handling and workload. To prove the relationship between driving parameters and workload measurements we selected two manoeuvres which are described in more detail in [Krüger et al. 2000]. Manoeuvre #1 was a double-lane-change manoeuvre following [ISO-3888, 1999], driven with a 2-door BMW 840 Ci, equipped with an 8-cylinder, 4.0-liter engine, manual shift, and cruise control. Manoeuvre #2 was a cornering manoeuvre with a diameter of 50 meters, driven with a BMW 318, equipped with a 4-cylinder, 1.8-liter engine, manual shift, and cruise control.. Both manoeuvres were driven on special driving grounds within restricted areas during dry weather. In both studies, subjects were unselected average drivers, males and females ranging in age between 20 and 60 years, with yearly driving of about 15,000 to 25,000 km. Manoeuvre #1 with n = 8 normal drivers and n = 2 expert drivers, Manoeuvre #2 was conducted with n = 4 subjects. As workload measurement a rating scale with five categories ("very low," "low," "medium," "high," "very high") was used where each category was subdivided either into three points (manoeuvre #2) or in 10 points (manoeuvre #1) yielding a 15-point-scale or a 50-point-scale, respectively. As a second measure, puls frequency was used.[1] The test procedure begins with an extended training of the subjects, both for the vehicle handling and for the manoeuvre itself. When subjects were familiar with the car and the course, they were intensively introduced into the scaling procedure. At least five trials per experimental block were driven. Judgments were given immediately after each trial. The manoeuvres were driven with the experimenter as passenger.

Manoeuvre #1, a double lane-change manoeuvre, forces the driver to apply severe steering inputs, as might be the case in an emergency situation. Subjects had to drive at different constant speed levels (held constant by cruise control), starting with 60 km/h and ending at the highest individual speed level that could just be mastered successfully. Within each speed

---

[1] We also looked at different performance measures which are described in [Krüger et al., 2000].

block, subjects had to perform several trials, usually between ten to twelve. Figure 2 shows at the left the results for two typical average drivers, at the right the results for the two expert drivers.

The range of speed levels starts at 60 km/h and ends for the average drivers between 85 and 100 km/h. Expert drivers reach 120 km/h. Both functions show a nonlinear trend, indicating a quadratic relationship between speed and subjective effort. The broader speed range of the expert drivers is the result of their higher skills and therefore lower difficulty of the driving task. Both groups reached their maximum of workload with the same functionality between task difficulty and subjective workload.

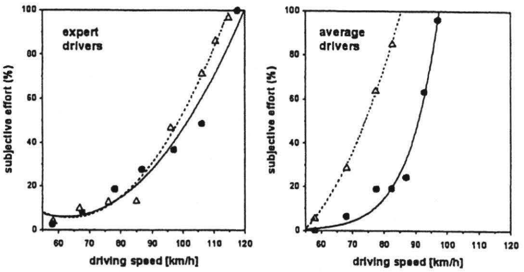


Figure 2: Functions of subjective effort according to speed levels realized in Manoeuvre #1. Left: results for expert drivers, right: results for average drivers.

An even more compelling prove of the direct relations between task difficulty and workload is given if the above experimental procedure is reversed. Instead of asking the subjects for judgments about given physical qualities a category production method was introduced. Here we instructed the subjects to adjust their actual driving to increasing levels of subjective effort (effort production method). In Manoeuvre #2 (steady state cornering) the subjects had to establish five different levels of subjective effort ("very low," "low," "medium," "high," "very high") by driving a curve of constant radius at different speeds. The appropriate driving speed level had to be chosen without feedback from the speedometer and had to be held constant over a 30-second interval. Figure 2 shows the resulting lateral accelerations for four drivers, each driving the course two times with the same instruction.[2]

The individual functions slightly vary in incline, but they generally show the same quadratic trend. Repeated measurements with different drivers yield small deviations of about 7% within each individual (measured in units of lateral acceleration). The "very low" category is linked with lateral accelerations between 2.5 and 3.5 m/sec^2, the "very high" category with 8 to 9 m/sec^2. Obviously, subjects are not only able to give reliable judgments about workload but they are also able to produce different energetical states with high precision.

---

[2] To be correct, the subjective effort as the factor varied experimentally should have to be depicted on the abscissa. For better comparison of the results we chose a consistent delineation with objective parameters on the abscissa, subjective experience on the ordinate.

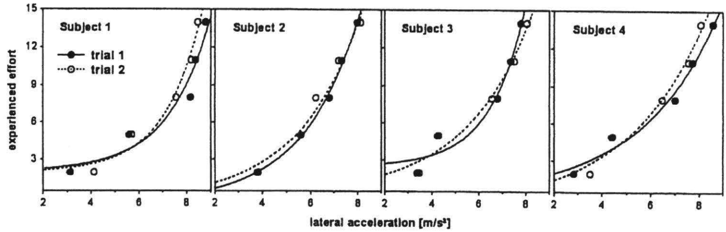


Figure 3: Lateral accelerations produced in five levels of subjective effort in the cornering Manoeuvre (Vehicle #1).

To validate these results with respect to workload, we used heart rate as a measure of global activity. Because the workload induced by the operating and stabilization task cannot be separated within the measurement of the driver's heart rate, a passenger was added, assuming that the passenger is mainly affected by the bodily impacts of driving on the organism. To control for individual differences, each subject participated as both driver and passenger in manoeuvre #2. The results are displayed in Figure 4 which shows individual heart rates as a function of lateral acceleration for four different subjects.

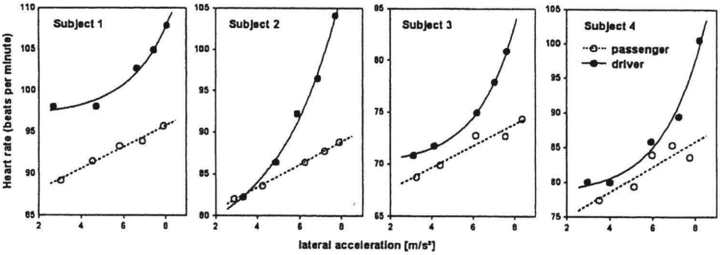


Figure 4: Individual heart rates for four subjects subjects as functions of lateral acceleration. Each subject participated as both driver and passenger (Manoeuvre #1, Vehicle #1).

As can be seen in Figure 4, the course of driver's heart rate follows a quadratic trend, whereas the course for the passenger is linear. In the passenger role, subjects show an increase of between 5 to 10 beats per minute over the full range of lateral acceleration, whereas under the driver condition, increases of 10 to 20 beats per minute can be observed over the range. Typically, the functions differ only slightly at their starting level, but diverge in a multiplicative way with increasing acceleration.

Summary. Taking these results together, we find a very strong relationship between task difficulty and workload, be described in a quadratic function. With respect to the modeling of the driver we have to take into account that the driving task will have direct consequences on the energetic state of the driver particularly in combination with time on task. On the other hand, we have to accept that the driver is highly sensitive to different stages of workload and the resulting activational state. This is a necessary precondition for regulatory processes by means of which the driver can establish a sufficient activational state to solve the driving task. If the driver wants, he or she can make use of task variations to adjust the actual activational state to the required level. The result is obvious: in the modelling of driving the driving task cannot be treated being constant but strongly depends on regulatory processes within the energetical domain.

Processes altering the activational state of the driver

Task difficulty is only one determinant of the activational state. Two additional processes are responsible for variations in this state: (1) effects of time on task and (2) circadian rhythms.

Effects of time on task and time of day. The fact that driving is an energy consuming process is best demonstrated in the study of long term driving. As effects of time on task we observed:

- Firstly, a decrement in vigilance, mainly in very easy driving situations. This decrement can best be described as a loss of stimulus driven arousal, mainly caused by processes of habituation. The bottom-up source of activation gets weaker and weaker.
- Secondly, an increase of saturation which can be described as a loss of motivation to perform the task for any longer time. Subjects have to coerce themselves to stay within the task, accompanied by very negative feelings which they try to override by introducing effort with the aim to perform as well as possible.
- Thirdly, a decrease in overall capacity, best described as an exhaustion of resources, somatic as well as cognitive. This exhaustion clearly depends on the duration of driving, but is also strongly influenced by the time of day. The decrease of overall performance is much stronger during night-time driving and is highest in the early morning. This clearly indicates that overall performance is strongly dependent on circadian rhythms.

These three processes together with the workload caused by task difficulty are the determinants influencing the actual state of the driver who has the task to continuously manage these determinants in a way that the primary task of safe driving can be fulfilled.

Regulatory processes in long term driving. Within a series of studies we determined four distinct states of the driver by using different parameters of blinking behaviour. These states typically follow each other during long-distance driving [Krüger & Reiß 1995, Hargutt & Krüger 2000a, b]. Phase 1 "Awake" is followed by phase 2 "Reduced Vigilance". Once the driver realizes this state, compensational mechanisms are initiated. This phase is followed by phase 3 "Drowsy" and finally phase 4 "Sleepy", in which the driver is already fighting against falling asleep.

In order to reveal the regulatory actions the driver introduces to maintain a sufficient activational state, we developed an experimental procedure called the "Min-Max-Method". In simulator studies N = 12 subjects had to drive a rather easy road for 3 x 5 minutes within a predefined speed interval of 90-110 km/h (low effort period). Thereafter subjects were allowed to drive for 5 minutes "as fast as possible" without making driving errors (high effort period). These blocks of 4 x 5 minutes were repeated (without pause) until the driver fell asleep. The experiment started at midnight. In the morning and in the afternoon additional trips were conducted (lasting one hour) in order to control for the effects of time on day.

The mean driven speeds in the high effort periods were then analysed depending on the fatigue level (determined by blinking parameters) in the preceding normal effort periods. As Figure 5 shows, drivers increased their speed up to phase 3 "Drowsy", demonstrating a natural attempt, during moderate fatigue levels, to try to increase their arousal by speeding up. A decrease in drivers' activation caused by habituation and saturation is followed by an active attempt of the driver to compensate for this reduction. In the limited situation of driving an easy route, the only method to of achieving an increase in the driving difficulty is to speedup.

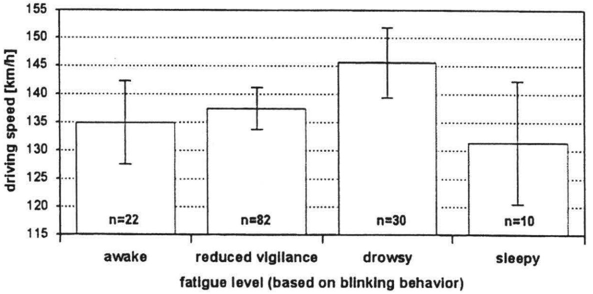


Figure 5: Mean driving speed in the high effort periods depending on the fatigue level (determined by blinking parameters) in the preceding normal effort periods.

Enhancing central arousal utilizing bottom-up mechanisms (fast stimuli input and increased task difficulty) seems to be the favoured strategy to combat drowsiness. This compensatory action finds its limits in the total capacity of the activational system which is influenced by time on task and - to a great extent - by time of day. Reaching this limit causes a change in compensation. The only way to perform even longer is to reduce the demands of driving. In the restricted context of driving this could only be done by lowering speed.

A model for energetic regulation. The aforementioned results are modelled in Figure 6. With time on task on the x-axis and performance on the y-axis, a medium level of performance at the beginning of a task is assumed.

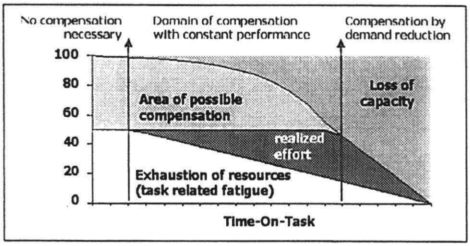


Figure 6: Illustration of effects of effort with respect to basic capacity of the human system.

With continued time on task, the resources needed for task completion will be evermore exhausted. Depending on time on task and due to circadian rhythms, the maximum capacity of the system will also decrease. By introducing effort as a compensation strategy, the detrimental effects of continued resource utilization can be compensated for thus keeping performance stable. However, if the actual demands reach the capacity limit, effort is no longer efficient. The compensation strategy changes to demand reduction – the goals are changed. Effort can be introduced in various ways. First, in a top-down manner, the driver can try "to be effortful", to concentrate more on the task. Second, he also can try to introduce bottom-up strategies by changing the task characteristics or by introducing new tasks besides driving. Nonetheless, if the capacity limit is reached, the dynamics of the system will get lost. The, only way to counteract this effect, is thus to redefine the driving task in an easier manner, neglecting all elements of the driving situation which are not absolutely necessary.

Summary and Conclusions

Efficient driver models are needed in different applications. One important field for modelling is to examine different constructive solutions either in driving dynamics or in the design of MMI. Here, the assumption of a driver who is awake and motivated to drive is very reasonable. The main emphasis of such studies is the exact measurement of the information processing by the driver. The situation changes when looking at long-term driving. In this context, it is inevitable to introduce the energetic consequences of driver actions. The energetic costs of driving and the limited capacity of the organism require an energetic management of driving which must guarantee that the driver is in an activational state sufficient to fulfill the driving task. This activational state depends on

- the difficulty of driving and the resulting task-oriented workload, from
- the time on task and the resulting state-oriented workload, and from
- the total capacity of the organism which is determined by the sum of previous workload and by processes of circadian rhythms.

To regulate the activational state to an optimal or at least tolerable level, the driver actively introduces mechanisms of effort. These mechanisms can be classified into two groups depending on the level of the energetic capacity of the organism:

- in case of sufficient total capacity, the driver tries to augment his activational state by increasing the difficulty of driving
    - either by varying the driving task (e.g., driving faster)
    - or by introducing additional tasks (e.g., infotainment).
- in case of reduced total capacity, the driver tries to decrease the activational demands of driving by simplifying the driving situation (e.g., neglecting peripheral cues, concentrating on important cues only).

These compensatory processes result in a change of the driving task either in speed or in precision. Therefore, if long-term driving is under study, driver models must include a modeling of energetic processes. The main consequence of this additional model is the fact that the driving task will be varied by the driver in order to compensate for energetic processes.

References

Hargutt, V., & Krüger, H.-P. (2000a). Compensation for drowsiness and fatigue – a driving simulation study. Paper presented at the International Conference of Traffic and Transport Psychology (ICTTP), 4.-7. September, Berne.

Hargutt, V. & Krüger, H.-P. (2000b). Eyelid movements and their predictive value for fatigue stages. Paper presented at the International Conference of Traffic and Transport Psychology (ICTTP), 4.-7. September, Berne.

Hockey, G.R.J .(1995). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In: A. Baddeley & L. Weisskrantz. Attention: selection, awareness, and control. Oxford: Clarendon Press.

ISO 3888-1 (1999). Passenger cars – Test track for a severe lane-change manoeuvre – Part1: Double lane change.

Krüger, H.-P. & Reiß, J. (1995). Einschlafen am Steuer – Ursachen, Verlauf, Folgen, Gegenmaßnahmen. Psycho, 21, 23-26.

Krüger, H.-P., Neukum, A. & Schuller, J. (2000). A workload-approach to the evaluation of vehicle handling characteristics. SAE-Paper, 2000-01-0170.

O' Donnell, R.D. & Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), Handbook of perception and human performance. Volume II: Cognitive processes and performance. New York: Wiley.

Onken, R. (1993). What should the vehicle know about the driver? In: A.M. Parkes & S. Franzén (Eds). Driving Future Vehicles. London: Taylor & Francis.

Pribram, K.H. & McGuiness, D. (1975). Arousal, activation and effort in the control of attention. Psychological Review, 82, 116-149.

Rasmussen, J. (1986). Information processing and human-machine interaction. An approach to cognitive engineering. New York: North-Holland.

Sanders, A.F. (1983). Toward a model of stress and human performance. Acta Psychologica, 52, 61-97.

COSA – A GENERIC APPROACH TOWARDS A COGNITIVE SYSTEM ARCHITECTURE

Dipl.-Inform. H. Putzer*, Prof. Dr.-Ing. R. Onken

Universität der Bundeswehr München
Institut für Systemdynamik und Flugmechanik, D-85577 Neubiberg, Germany
[henrik.putzer\reiner.onken]@unibw-muenchen.de

Abstract: *This paper presents an overview of COSA, a Cognitive System Architecture, which is a generic framework for a unified architecture for cognitive systems. – Conventional automation and similar systems lack the ability of cooperation and cognition, thus leading to serious deficiencies when acting in complex environment especially in the context of human-computer-interaction. This paper describes the Cognitive Process as the theoretical background to overcome these deficiencies. The Cognitive Process can be considered as a model of the human information processing loop whose behavior is solely driven by 'a-priori knowledge'. It is implemented on top of a symbolic AI subsystem to form the core component of COSA yielding a cooperative system with transparent behavior. The CORBA based basic layer of COSA leads to a distributed system. The 'language front end' enables a variety of knowledge formats suitable for any designer's preferences to model the system's 'a-priori knowledge'. The design and implementation with its actual state is described along with the experiences resulting from the first application based on COSA.*

Keywords: *cognitive systems, distributed system, knowledge modeling, knowledge representation, assistant systems, tutor systems, autonomous systems, human centered design, human information processing loop, implementation, unified framework, CORBA, MICO, SOAR*

1 MOTIVATION AND INTRODUCTION

The continuously increasing complexity of environments, the extension of the range of use and the situation of competition put great demands on crew and aircraft performance. Within the last several years the complexity of cockpits have been successfully reduced by introducing automation to support the crew in certain tasks. The result so far is, that efficiency and safety can be increased (see figure 1). Investigations on modern aircraft cockpits show, that a further increase in use of conventional automation will not necessarily result in increased productivity because automation itself became a complex element within the already complex environment of the cockpit. In some cases conventional automation has already become the key factor for decreased safety (e.g. "mode confusion").

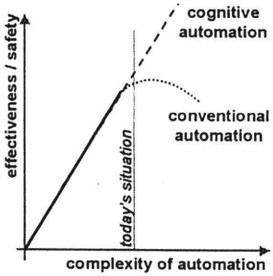


figure 1: *Today's situation*

Extensive research has been done to analyze this situation and it was found that conventional automation is in many cases too complex, acts in unpredictable ways, which means not consistent with the pilot's mental models, and provides the flight crew members with too little feedback (see [1], [2] and [3]). In most cases conventional automation is used in more or less separated systems. There is no integrated approach. Coupling of elements and the complexity of the environment has to be handled by the operator. A new *integrated* approach based on *cognitive automation* overcomes these deficiencies by making use of *cognition*, working in *cooperation* with the human operator and showing goal-consistent and *transparent behavior*.

Cognitive Automation is an extension of conventional automation and follows the human knowledge processing scheme according to Rasmussen (see [4]). As shown in figure 2, Cognitive Automation implements the whole process of building an internal comprehensive representation of the relevant parts of the external world (left column in figure 2). This can be considered as the principal basis upon which all considerations and actions of the system, especially for the crucial part of goal driven knowledge based behavior (upper row in figure 2), are created.

Another important feature of next generation's automated system is the ability of cooperating with the environment, especially with the pilot. Thereby synergic benefit is gained by supporting the strengths of both the human operator (instinct, abstraction, creativity, etc.) and the automatic system (objectivity, stress resistants, parallel processing, complex calculations, etc.). This means to implicitly and explicitly communicate with the crew in order to continuously match the technical system's knowledge about the situation against the flight mission goals and the intents and actions of the pilots.

Although the research at the Universität der Bundeswehr München is originated in the area of aeronautical applications, the approach of cognitive automation is equally appropriate in other domains with technical processes controlled by human operators like road vehicle guidance, power plant control or even business management.

2 THE COGNITIVE PROCESS

The Cognitive Process (CP) is a technical process which mimics the human information processing. This model is based on knowledge about human behavior and known design philosophies for cognitive systems. It is not designed to verify or to comply in detail with the theories about processes and features of the human brain. It is rather motivated by the need of modeling behavior that is similar to the main stream of behavioral characteristics of humans and thus comprehensible to humans.

The central component of the CP is the 'body', the oval part in figure 3, which hosts a great bunch of data. These data represents the knowledge of the system that is specific to the CP-subprocesses and to the application. The inner oval (slightly darker) contains the 'a-priori knowledge' which is fed into the CP before any processing starts. This 'a-priori knowledge' is the only origin of behavior of the application. The outer oval (light gray) contains the posteriori situational knowledge which is created during runtime. This

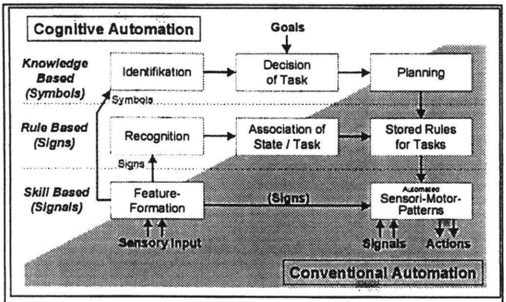


figure 2: Cognitive Automation

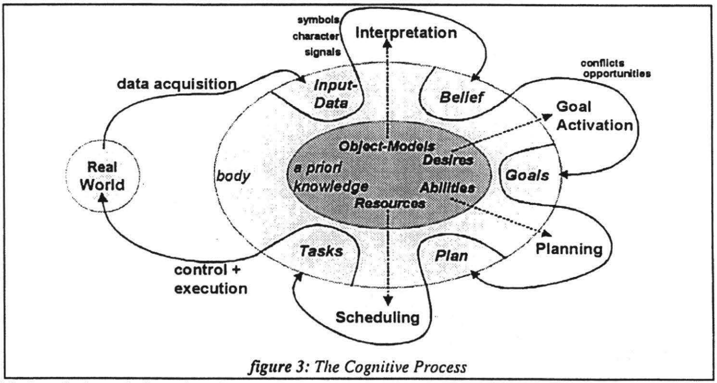


figure 3: The Cognitive Process

kind of knowledge is also called the 'cognitive yield' because it results from the operation of CP-subprocesses.

Processing of the knowledge is done by the so-called Transformators which are represented by solid lined arrows in figure 3 around the body. These Transformators can read from the whole body to retrieve their necessary input and write their results into designated areas of the body. The functions of these Transformators are designed according to the recognition-act cycle (in figure 3 clockwise around the body): *interpretation*, *goal activation*, *planning* and *scheduling* (see [9] and [5]). Additionally, two specialized Transformators establish the interaction with the outside world: via the *data acquisition* sensory data gets into the body and via a *control + execution* link external actuators and other output devices can be controlled.

The execution of all Transformators is assumed to be in parallel and event-driven. As soon as a Transformator detects a significant change at its input this change is evaluated using the available knowledge and the output is adjusted. A direct consequence is that there will be no unnecessary recalculation of the plan, if the changes of the situation stay within predicted bounds.

For a better understanding of the functions and interactions of all Transformators within the Cognitive Process a closer look at each of them is given:

- **Interpretation** – This Transformator generates the complete picture of the situation. As a simple example, one can imagine to take sensory data from a radar device and to create a structure within the area of 'beliefs' that reflects the knowledge (e.g. position, heading and speed) about other aircraft. Besides natural objects this Transformator may also establish structures reflecting relations. For example, a structure describing the binary relation of distances between the own vehicle and other objects could be established.
- **Goal Activation** – Primarily using the beliefs, this Transformator selects desires from the a-priori knowledge to activate them if the situation requires or implies it. Thus, for example in a military scenario, if there is an incoming missile, the desire of surviving by evading is activated. Activated desires are called goals.
- **Planning** – Using all planning abilities and the goals as activating elements this, Transformator generates a plan of how to achieve these goals. This plan may

contain parallel as well as sequential portions. It may also contain alternative operations. However the plan is hierarchically structured and has a time span up to the end of the current mission.

- **Scheduling** – The primary input to the scheduling Transformator is the plan. The action to be processed at each point in time is selected from the hierarchical structure of the plan and converted into simple tasks, which can be executed directly.

If necessary and if it is modeled by the 'a-priori knowledge', the behavior of these Transformators covers all three levels of human behavior (see figure 2). It is solely based on 'a-priori knowledge' specific to the Transformator. Without this knowledge the Transformators will do nothing. The basic idea about the purpose of the body is, that simply by adding knowledge all Transformators know how to use this knowledge. New behavior evolves from this by *automatically* combining the newly added knowledge with the previous knowledge.

The organization of the knowledge within the body follows the *object oriented* paradigm: it has a uniform structure in terms of models. Each model can be instantiated. Instances have data members describing their state. The model also comes with template functions which describe the behavior of all instances. This concerns the whole life cycle of instances including creation, general behavior and removal. The combination of all *micro-behaviors of all models* (instances) within the CP's body forms the *macro-behavior of the whole system*.

It can be stated that the concept of the Cognitive Process will be adequate for a wide range of cognitive system applications. It can be used independently of domains and characteristics of the application. So the Cognitive Process can serve as a core element in systems for aeronautical applications, driving, managing power plants or other domains. Possible applications are tutorial systems, assistant systems or even autonomous systems. The domain and the kind of application evolve solely from the 'a-priory knowledge' which is put into the CP. Before doing so, the concept of the Cognitive Process has to be put into a design, an architecture and an implementation of a suitable framework which we call COSA. These steps are described in the next chapters 3, 4 and 5.

3 DESIGN AND GOALS

In most cases of current research special designs are used to model certain isolated aspects of cognitive systems and solve problems of the respective application. With COSA we try to design a system which uses results from several sources and combines them to get an improved, holistic and unified architecture for cognitive systems. This section will give a more high level view on the design and the design goals leaving detailed aspects to be explained in chapter 4 in relation with the design decisions.

The first step towards a new system was a specification phase which included the analysis of former cognitive systems like CAMA (see [8], [10]) and other state-of-the-art systems. As a further step towards the design goals the requirements of three different groups of users were analyzed:

- **Programmers** add functionality or external components. For this task they need good documentation, simple interfaces for internal and external subsystems, good modularity and support to easily extend and maintain the system. Design decisions regarding these design goals are mentioned in the chapters about the 'architecture' 4 and the 'implementation' 5.

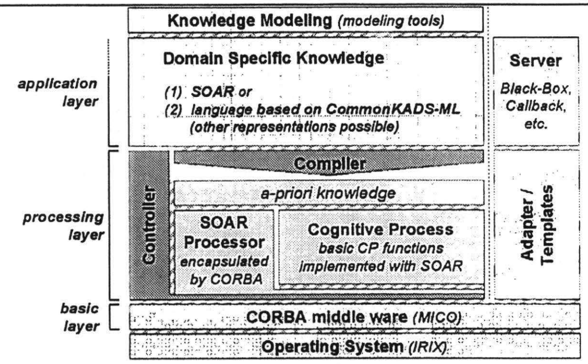


figure 4: COSA block diagram with layers

- **Knowledge engineers** design the "a-priori knowledge" to implement the systems behavior. They need an interface to potentially any modeling language and a structured concept to partition the knowledge. Ideas about that can be found in the chapters about the 'language front end' 0 and the 'controller and components' 4.2.3.
- **Users** of the resulting cognitive application put some more general requirements on the system. These are described in the following paragraphs along with other features derived from our own experience with cognition.

COSA is a framework or a unified architecture of cognitive systems. It is based on our model of human information processing, the CP, described in chapter 2. The CP as the core element of the system makes all its properties available for the resulting framework: domain and application independence through general knowledge processing, reusability, comprehensible processing to humans and support for cognitive operations (see chapter 2).

The kernel of COSA, which is the Cognitive Process, is designed to carry out high level knowledge processing. It is not designed to implement number-crunching algorithms or high frequency control loops. As it will turn out, this is no restriction for an application because there are interfaces to external components which can cope with such requirements.

Certification is not yet addressed in the first implementation of COSA. It can be identified, though, that COSA supports approaches to strengthen system integrity.

4 ARCHITECTURE

This chapter describes the architecture of COSA already taking into account aspects of the implementation. Chapter 5 is about the 'implementation' and covers more general aspects.

4.1 Layers

A layered architecture is used to ensure clear internal interfaces and decoupling of functional components (see figure 4). The functions of the layers can be described as follows:

- The ***'basic layer'*** uses the industry standard CORBA (see [13]) to build a COSA specific distribution layer. It also covers basic libraries that are used to implement COSA.

- The ***'processing layer'*** encapsulates the knowledge processing engine and implements the Cognitive Process on top of it. This layer does not contain any domain or application specific knowledge but it 'knows' how to convert knowledge into behavior. It is the organizational element that integrates all distributed components to form one system.
- The ***'application layer'*** is the only layer that contains domain specific knowledge. This layer uses the 'processing layer' as an operating system by transferring the knowledge to it.

A block diagram identifying the three layers of COSA along with some details can be found in figure 4. The architecture and functionality of all layers and their components are described one by one within the next sections.

4.2 Processing layer

The primary task of this central layer is the implementation of the Cognitive Process. This approach follows the human centered design by making the behavior of the resulting system similar to human information processing. This ensures system behavior that is comprehensible to human operators. Here both are defined: symbolic processing in general and the over all processing scheme.

The higher the abstraction of programming languages is, the more functionality can be covered by a team within a given time. So above C++ we are using SOAR to simplify the implementation of the CP that delivers its abstraction level to the knowledge engineer and application designer. This step is covered by the Cognitive Process Library (CP library) while SOAR itself is the processor. To enable the components of the system to connect to the kernel a distribution layer is implemented by using CORBA. This results in a system model where the knowledge can be distributed but is centralized at startup resulting in a central knowledge processing. The following subsections describe the components of the processing layer.

4.2.1 Processor

As the central element of the processing layer a component is needed that can process 'knowledge to yield behavior. SOAR (see [6], [7]) which is developed and maintained by the University of Michigan is a good candidate because it meets all requirements. The main reasons for selecting SOAR as the processor are the following: it is easy to learn, the developing community is very active and reacts fast on requests and last but not least: SOAR comes with portable source code and can be easily integrated in C / C++ environments.

SOAR stores its 'knowledge' about the situation in its working memory which has a uniform and symbolic structure similar to conceptual graphs. The 'behavior' is uniformly stored in rules so SOAR is a kind of production system but has a very special multi-stage processing loop so even advanced features like learning are supported.

As in other production systems SOAR offers a very fine-grained interface: the rules. With this feature, extending existing knowledge models is as easy as writing new rules and loading them into SOAR (even at runtime). In SOAR all loaded productions can fire in parallel so this supports the idea of Cognitive Automation: apply *all* knowledge at *any* situation.

It turns out that maintenance can be done as easy as extensions: the debug code is a set of rules that can be loaded into SOAR at any time. This code can trace values, set breakpoints or just print out portions of the working memory. This is supported by symbolic representation of the working memory that is understandable by human beings.

On the other hand SOAR lacks some key features. For example it has got no means of organizing knowledge in components and it does not implement the Cognitive Process, which should be the kernel of COSA. These features are added to SOAR by loading the CP library.

4.2.2 CP-Library

Building the abstract layer of the Cognitive Process is the function of the CP library. Three main features of the CP library lead to the required functionality:

- **Timers / triggers** handle simultaneous events and synchronization throughout the whole system. They are needed for internals and are not used in the actual application.
- The **OO approach** introduces object oriented design philosophies to SOAR. With this feature SOAR supports 'models' as they are described in section 2. Models can be instantiated which is a similar process to instantiating a class in C++. Instances contain data describing their individual state. Models also contain the behavior for all instances throughout the whole lifecycle: creation, general behavior and deletion.
- **Component management** enables the kernel to keep track of used components and determine dependencies and priorities. This is done in cooperation with the controller component of the processing layer.

The CP library is implemented as a set of SOAR productions in separate files. These files can be loaded into a running SOAR processor to yield a Cognitive Process scheme. This basis for cognitive applications does not include any domain specific knowledge.

4.2.3 Controller and Components

The controller can be regarded as the central server encapsulating the processor. Components register with the controller right after startup and transfer their knowledge to the controller. The controller takes the knowledge and loads it into the SOAR processor along with the registration data. During this phase dependencies of the components are checked.

This introduces a special kind of 'knowledge library' with dependency mechanisms. Thereby COSA supports manageable intelligence through dividing knowledge into thematic portions following the approach of distributed knowledge.

4.3 CORBA middle ware

COSA's basic layer is based on CORBA, an industry standard for distributed systems. It serves as a middle ware to connect the object of the controller which encapsulates the processor to other objects containing knowledge and I/O interfaces. These objects can be distributed over a network and differ in programming language, operating system and computer platform.

To use the SOAR processor within the CORBA environment the interfaces of the processor have to be mapped to the middle ware layer. The main feature of COSA's middle ware layer are the following:

- The **Client-Server-Structure** puts the controller into a central position. Components of the application register with the controller on startup so that they can be used to retrieve 'a-priori knowledge' or by accessing other interface functions.
- The **knowledge mapping** is the link between the different knowledge representations in SOAR and in the transport layer CORBA. It defines the representation CORBA uses to transport any piece of knowledge via the network from the processor to other objects and vice versa. The mapping is done by implementing a graph structure for CORBA.
- The encapsulation **of callbacks** connects I/O functions and internal callbacks of the knowledge processor to published

member functions of distributed objects. This is done by dispatching calls to registered member functions of objects.

With these features COSA can dispatch tasks to external processing units that are distributed objects connected via an adapter or implemented by a template (see right side of figure 4). This is the proper way to integrate servers, black boxes and external systems to do number crunching, implement high frequency control loops, interface data base systems or to connect any other subsystem to the application.

4.4 The Language Front End

The main purpose of the language front end is to decouple different representations of knowledge needed for knowledge modeling and the knowledge processor. This allows the knowledge engineer to use a modeling tool or environment which is best suited for his problems and saves him the learning of the format the processor is programmed in. This way object oriented or procedural approaches for any standard of knowledge modeling can be supported or even mixed to yield a joint behavior.

The central component of the language front end is the compiler which is a the triangular component of the processing layer in figure 4. This compiler takes knowledge of the representation that is received from a connected object and converts it to run on top of the CP library. The main problem consequently is to convert knowledge with any format and philosophy to a form complying with the requirements of the Cognitive Process.

To avoid that kind of automatic reformulation of knowledge and to give an ideal interface to communicate knowledge into the Cognitive Process a future goal is the design of a specialized 'Cognitive Process Language' or specialized languages for assistant systems, tutorial systems or autonomous systems.

```
KNOWLEDGE-MODEL vehicle;
  CONCEPT uav;
    SUB-TYPE-OF: aircraft;
    ATTRIBUTES:
      speed      : INTEGER =
0;
      heading    : INTEGER;
      altitude   : INTEGER;
      longitude  : REAL := 0;
      latitude   : REAL := 0;
      ...
    BEHAVIOR:
      ...
  END CONCEPT uav;
END KNOWLEDGE-MODEL vehicle;
```

figure 5: Sample Knowledge Model of an UAV

So far, besides using native SOAR representations there is another language that is based on the CommonKADS Markup Language (CML, see [18]). This language supports the idea of active models (objects) within the body of the CP. These models can be formally described as it is shown in figure 5. This piece of code defines the model '**uav**' which is derived from the model '**aircraft**' within the knowledge portion called '**vehicles**'. Each model is specified with its inheritance information, its attributes and behavior. The attributes are defined by their name, type and an optional default value. The behavior consists of SOAR productions.

5 IMPLEMENTATION

The operating system of our implementation is IRIX, a UNIX version from SGI. As programming environment the IRIX native tools are used along with the latest versions of the packages listed below. These are distributed with source code and are portable to Windows and to many UNIX derivatives:

- **STL** the Standard Template Library providing basic types and containers. It is part of the IRIX native C++ environment but is also distributed freely for other platforms (see [11]).

- **MICO** which is a free and very good implementation of the CORBA standard (see [14]).
- **SOAR** as the implementation of the Unified Theory of Cognition (see [6] and [7])
- **QT** library is planned to be used for the (graphical) user interface (see [12])

The documentation for most parts is done with DOXYGEN (see [15]) which generates the documentation from the C and C++ structures by using code comments of a special format. Design tools are also used for some aspects of the implementation, such as Rational Rose (see [16]) and the CommonKADS IDE (see [18]) which are based on the unified modeling language UML (see [17]).

A first and simple application is implemented on the basis of COSA and is already running. This first application is integrated with the flight-simulator and models a UAV during the takeoff phase.

The focus of this implementation is to verify core features of COSA's architecture to some extent especially concerning the implementation of the Cognitive Process. The language front end has not got many features yet. So far, besides using knowledge written in native SOAR, a second language can be used which is based on the CommonKADS Markup Language (CML). With the level of abstraction this language provides, we obtained promising results.

For the near future, improvements of COSA are planned in many implementation details, especially concerning the CP library. Although the performance of the system was no design goal, the system at its actual implementation is capable of running in real environments. There is still potential for some improvements, however, especially within the interface between MICO and SOAR. A research related goal is the improvement of the language front end.

6 CONCLUSION AND PERSPECTIVES

Analyzing the current situation of automation and similar systems in complex environments serious deficiencies were detected. To overcome these deficiencies the concept of the Cognitive Process was developed. This concept takes into account the research results gained in our and other projects on cognitive systems. The Cognitive Process is used as the basis for the design of COSA which is a uniform architecture for cognitive systems.

Meeting the requirements of several user groups COSA became a highly flexible and usable framework with its implementation based on free libraries. Especially the implementation of the Cognitive Process as the kernel for the uniform architecture for cognitive systems as a SOAR library fits very well. This is very plausible because SOAR implements the Unified Theory of Cognition (see [6]) which should be a good basis for the unified architecture of cognitive systems, namely COSA.

A first application is implemented with a simple approach proving that COSA is a valid framework to go on with. There is still a great amount of research needed and in progress to further verify the concept of the Cognitive Process and the design of COSA. Future activities concentrate on improving the performance of COSA and on implementing a more complex system which will be called COSY^{flight}, a Cognitive autonomous System for the flight domain that is based on COSA, the Cognitive System Architecture.

7 REFERENCES

[1] *E. L. Wiener:* "Human Factors in Aviation", *Academic Press, San Diego, California, 1993.*

[2] *N. B. Sarter, and D. D. Woods*: "Strong, silent, and 'out-of-the-loop' properties of advanced (cockpit) automation and their impact on human-automation interaction", *Technical Report 95-TR-01, Ohio State University Cognitive Systems Engineering Laboratory, February 1995.*

[3] *C. E. Billings:* "Human centered automation: a concept and guidelines", *NASA Technical Memorandum 103885, Moffett Field, CA, August 1991*

[4] *J. Rassmussen:* "Skills, Rules and Knowledge; signals, signs, ad symbols, and other distinctions in human performance models", *In IEEE Transaction on Systems, Man and Cybernetics, volume SMC-13, pages 257-266, 1983.*

[5] *Jennings, Nicholas R.; Wooldridge, Michael J.:* „Agent Technology, Foundations, Applications and Markets", *Springer Verlag, Berlin Heidelberg, New York 1998,*

ISBN 3-540-63591-2

[6] *A. Newell:* "Unified Theories of Cognition", *Harvard University Press. Cambridge, Massachusetts, 1990*

[7] *University of Michigan, Ann Arbor:* "Soar Home Page", *http://bigfoot.eecs.umich.edu/~soar*

[8] *A. Walsdorf and R. Onken:* "Intelligent crew assistant for military transport aircraft", *In NATO RTO System Concepts and Integration Panel Symposium. Sensor Data Fusion and Integration of the Human Element, Ottawa, Canada, 13.-17. September 1998.*

[9] *A. Walsdorf, H. Putzer, and R. Onken:* The cognitive process and its application within cockpit assistant systems. *In IEEE / IEEEJ / JSAI International Conference on Intelligent Transportation Systems, Tokyo, Japan, 5.-8. October 1999.*

[10] *A. Schulte; P. Stütz:* "Evaluation of the Cockpit Assistant Military Aircraft CAMA", *NATO System Concepts and Integration Panel Symposium. Sensor Data Fusion and Integration of the Human Element. Ottawa, Canada, 14-17 September 1998.*

[11] *SGI AG:* "Standard Template Library Programmer's Guide", *http://www.sgi.com/tech/stl, 2001*

[12] *Trolltech Software Company:* "Qt Reference Documentation", *http://doc.trolltech.no, 2001*

[13] *Object Management Group (OMG):* "The OMG's CORBA Website", *http://www.corba.org, 2001*

[14] *K. Römer; A. Puder; F. Pilhofer:* "MiCO – Mico Is COrba", *http//www.mico.org, 2001*

[15] *D. van Heesch:* "DOXYGEN homepage", *http://www.stack.nl/~dimitri/doxygen, 2001*

[16] *Rational Software GmbH:* "Rational Rose", *http://www.rational.com/products/rose/index.jsp, 2001*

[17] *M. Page-Jones:* "Fundamentals of Object-oriented Design in UML", 2^{nd} *Edition, 1999,*

ISBN 020169946X

[18] *"CommonKADS, Engineering and Managing Knowledge", http://www.commonkads.usa.nl, 2001*

Postersession II

Open a Window to the Cognitive Work Process!

Pointillist Analysis of Man-Machine Interaction

Frank O. Flemisch
NASA Langley Research Center
Crew / Vehicle Integration Branch, MS 152
Hampton, VA 23681, USA
f.o.flemisch@larc.nasa.gov

Reiner Onken
Universität der Bundeswehr München
Institut für Systemdynmik und Flugmechanik
Werner-Heissenberg-Weg 39, D 85577 Neubiberg / Germany
reiner.onken@unibw-muenchen.de

Summary

Technological progress keeps on changing our lives and work. We can only improve this mutual influence between man and technology if we understand the impact of technology on the work process, and if we are also able to share this understanding with other people in the design process. Quantitative assessment methods like workload offer at least a keyhole to the work process, but only an appropriate combination with qualitative methods can open this keyhole into a wider window. The pointillist approach uses a powerful metaphor, borrowed from arts, to systematically combine quantitative with qualitative aspects. Based on this approach, a tool for human factors analysis was built and used to analyze the interaction between a pilot and an assistant system in a simulator. Three examples illustrate the possible interplay between quantitative and qualitative aspects of the approach.

Why do we need a window at all?

The tremendous scientific and technological progress of the last centuries has changed our lives, more for the good, in some respects for the bad, but at least significantly. Now, at the beginning of a century, new powerful information technology promises even more, but we can already anticipate that this "more" will not automatically provide benefits for us humans with our natural and, in the context of this conference, cognitive characteristics.

In aviation, which is sometimes seen as a technological forerunner, there were tremendous efforts during the last twenty years through cockpit automation like the introduction of flight management systems or full computerized "glass cockpits". With all undoubted benefits of this development: the relative safety remained surprisingly constant. Considering the anticipated future increases of air traffic this will bring along unbearable losses of aircraft and lives [Billings 1997][FAA 1996].

Cognitive automation with its realization in assistant systems, as opposed to conventional automation, offers a technical answer to these problems. More about this in [Onken 1999] and in other papers of this conference.

Another answer was the demand for a better understanding of the work process and for an increased use of human factors knowledge. One of the most widely used concepts of human factors is workload. Indeed, a major intent of conventional automation was a reduction of workload, and this reduction was clearly shown by the corresponding human factors work. The effect was so enormous that, as a result, we now face a "pilot-out-of-the-loop" problem, e.g. [Endsley &Kiris 1995]; we speak about the "ironies of automation" [Bainbridge 1987]; and operators complain of "99% boredom, 1% panic" [Kraiss 1994]. One of the conclusions with regard to our understanding of the work process is that the concept of "workload", as it is operationalized, is still too weak. It opened a keyhole, but not a transparent window.

This irony could be more fundamental: Human achievements in science and technology based on rational thinking, which increasingly raise questions about its influence on the human and the work process, might not be equally helpful to answer these questions in a satisfying manner. As a consequence, scientifically weakly founded statements like the following might find more applause than challenge or opposition:

> *"The relationship to the world that the modern science fostered and shaped now appears to have exhausted its potential. It is increasingly clear that, strangely, the relationship is missing something. It fails to connect with the most intrinsic nature of reality and with natural human experience. It is now more of a source of disintegration and doubt than a source of integration and meaning. It produces what amounts to a state of schizophrenia:* ***Man as an observer becomes completely alienated from himself as being... The abyss between rational and spiritual, the external and internal, the objective and subjective, the technical and moral, the universal and the unique, constantly grows deeper."***
>
> [Havel 1994]

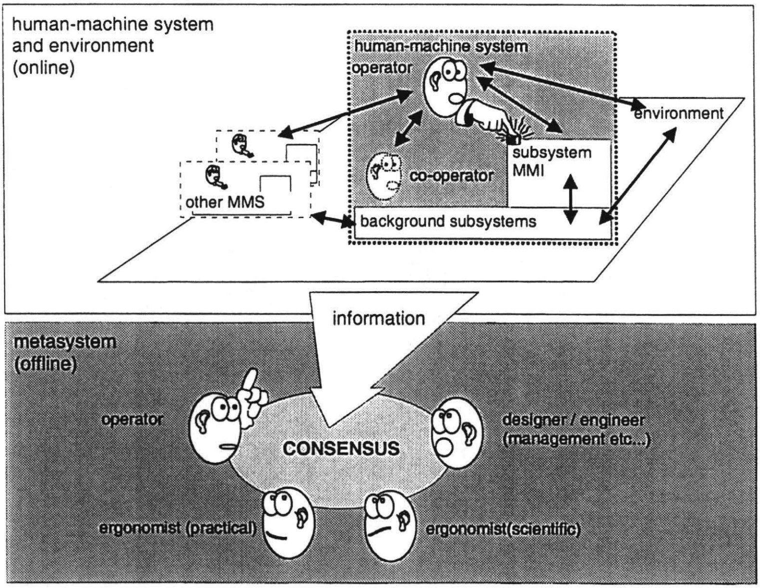
As regards our domain, humans and machines, it is therefore imperative to widely open a window into the work process in order to understand in depth the interaction phenomena and thereby to uncover hidden weaknesses before they cause problems. The limitations of e.g. the workload concept do not mean that the concept is not valuable, but we have to supplement it with additional concepts like situation awareness, e.g. [Endsley 1995] and usability, e.g. [Nielsen 1993]. As we already know a great deal about these concepts, it is now important to make techniques available to exploit these concepts in a qualitative and quantitative manner, and, like Havel himself, to use every inspiration our rich culture can offer to bridge potential abysses. This is the main objective of the work described in the following.

Analyzing the details of the window

Before we concentrate on the window itself, it is worthwhile to start from a solid base, and identify in terms of systems engineering, e.g. [Blanchard1991] or [Chapanis 1996], the parts of our challenge and a potential location of the window.

The work process we are talking about in the context of this conference needs at least a human-machine system with its operator(s), technical subsystems with their emerging part "the interface" and the environment. The key element for the analytical understanding of the work process itself is the communication between these subsystems themselves and of the subsystems with the environment.

In order to understand not only the work process but also the window to the work process, we have to step back one more step and look at ourselves analyzing this process:



Picture 1: Human-machine system and metasystem

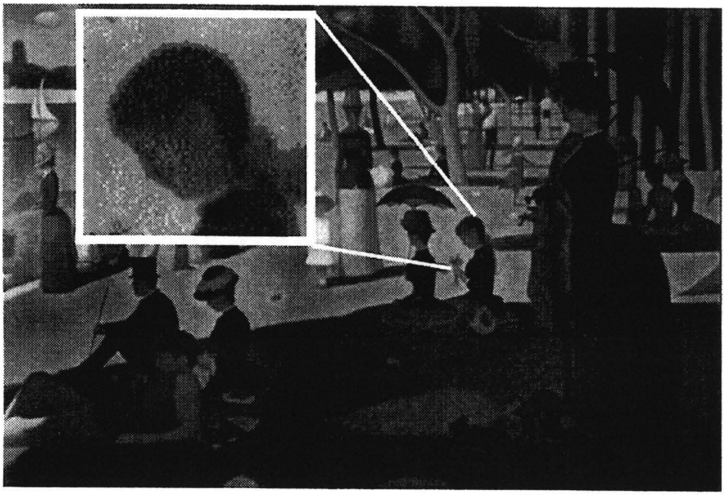
Human factors analysis is mostly performed "offline" in a "meta system" [Haberfellner et.al. 1992]. The "subsystems" here are mainly people, e.g. operators/users, designers/engineers and ergonomists/human factors engineers. They have to find a consensus about a certain problem, a solution, a common behavior regarding how to react to problems, which [Weick 1979] calls "interlocked behavior". This interlocked behavior is often hard to reach because of the "equivocation" between the partners, which corresponds with Havel's abysses. There is often a deep equivocation between the subjective impressions vs. the need for objective data, quantitative vs. qualitative methods, scientific vs. practical needs, or between the global concept of a technology and the tiny, hidden detail, which can cause "nasty surprises" ([Meister 1987]) like the failure of the complete human-machine system.

The key for bridging these abysses is the transfer of information about the communication in the human-machine systems into the metasystem. This information window has to satisfy all different needs of the different people and aspects in an ergonomical situation, which can only be achieved with a high bandwidth and in an "ecological" manner [Rassmussen et.al. 1994].

It is essential at this point of the discussion to understand that the analytical approach described here, that is, the "breaking down into details", can help to identify the abysses and the location of the window, but it cannot bridge the abysses itself. It cannot bring the details together again in an intuitively understandable manner. To bridge the abysses with a window, we have to use more than analysis.

From details to a meaningful picture: Pointillist Approach

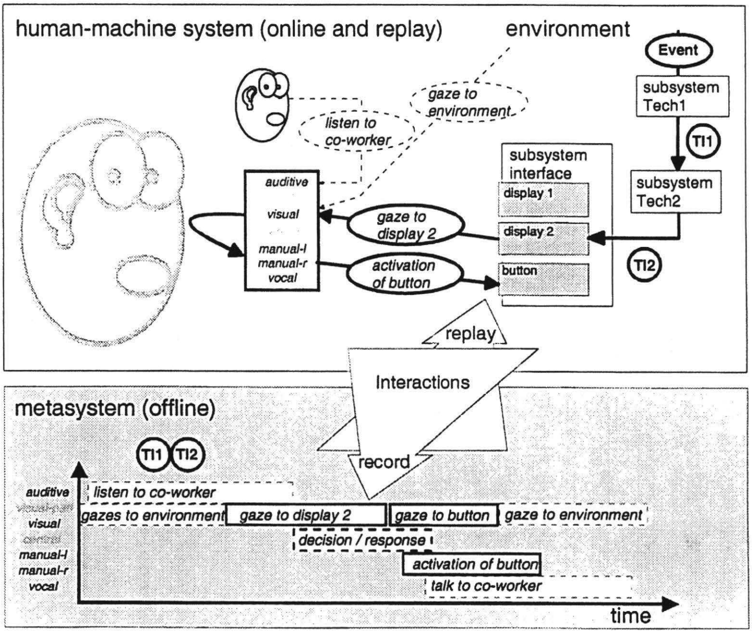
A powerful metaphor for bridging abysses can be found in the arts: The dominating painting style in Europe in the second half of the last century was impressionism, which was, oversimplified, able to force the subjective impression of scene, but was not theoretically well founded. George-Pierre Seurat (1859-91) succeeded in combining theoretical background (theory of divisionism) with high subjective impression. He invented a painting style with small "points" of unblended color (pointillism), which, from a distance, created a similar impression like the original scene. And "finally the world shined through the picture" [McLuhan 1964].



Picture 2: Grande-jatte by Georges-Pierre Seurat (1859 -91)

Science often needs to measure, focus on details, and abstract the continuous world into discreet elements. These abstractions bring about a loss of quality, but allow us to understand something about the underlying structure. They also allow us to match this model of the world with quantitative data from the original world. But if we could succeed in recombining these objective details in a manner that reflects also the original quality, the original work process as a whole could shine through the data like in Seurats pictures. Wouldn't this be worth some effort?

Picture 3 shows a simple example of how this metaphor can be used as a "window" to the work process. The exemplary chain of interaction starts with an event in the environment, which is here detected by a technical sensor Tech1. The information is handed over to a subsystem Tech2 and displayed. We can assess the smallest discrete elements of this information transfer, these interactions between technical subsystems. We can transfer, store and analyze them later on in the metasystem, as has been done for years in the testing of complex technical systems.



Picture 3: pointillist analysis of the work process

But the chain is not complete with the technical interaction alone. The operator has to perceive the information, even if he/she is already occupied with for example looking at the environment, and he/she has to select a response and activate this properly, and so on. We can measure most of this interaction, e.g. with eye tracking and keystroke recording, and can dissect this into smallest elements, which we call "basic interactions". They are the link between the interaction resources of the operator, described e.g. by [Wickens 1992], and the technical subsystems or the environment. Examples for basic interactions are e.g. a gaze to a certain area of interest on a display or the activation of a certain hardware button.

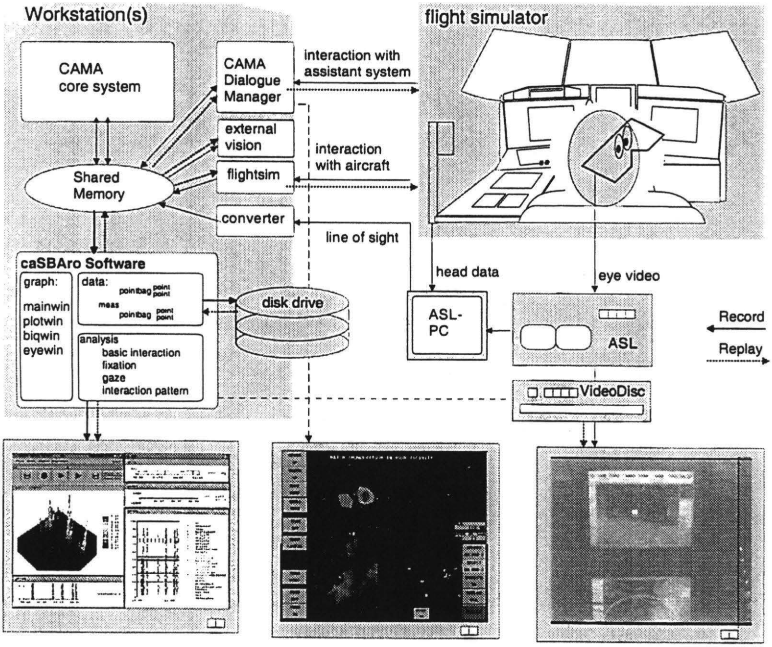
We can analyze these objective data in a quantitative way. We can measure the length of these basic interactions: how long it takes to get a certain information out of a display, how long it takes to find the right button. We can detect when certain interactions are left out or made unnecessarily, which can give us hints for existing usability problems. If we gather data about the visual interaction, e.g. measure the point of foveal visual attention, we cannot necessarily prove existing situation awareness. But if we know that for a certain task foveal visual attention is necessary to build up situation awareness, and measure missing foveal attention, we have a strong hint for missing situation awareness.

This analysis of data is what we as ergonomist already do for our understanding of the work process. But this might be not enough to make the other partners in an ergonomical situation understand and react properly. But we can recombine the same interactions in a more intuitively understandable manner. There are many ways to do that, starting with a simple graphical representation on a timeline, plots averaging over time, operators or workspace, etc.

One of the most intuitive representations of these objective data arises when we feed it back into a replay system similar to the original system. If the replay system subjectively behaves like the original system, we can sit or stand inside the replay system and get the same impression as the operator had in the original event. We are still based on objective data and guided by rational analysis, but we as normally "alienated observers" (Havel) can get involved not just with our rationality, but also with the rest of our senses. Our window is not only clear, but also wide and open enough.

caSBAro: Pointillist analysis for the assistant system CAMA

Based on this concept and fruitful experience with a replay in [Schulte 1996], caSBAro was developed for the evaluation of the cockpit assistant system CAMA (Crew Assistant Military Aircraft), which itself was developed together with DASA/EADS (European Aeronautic Defence and Space Company), the German Aerospace Center DLR, the Elektroniksystem- / Logistik-GmbH ESG and the University of Armed Forces Munich. CAMA was tested in a flight simulator in 1998 and flight tested in 2000, e.g. [Onken 1999].



visualization of interaction controls

visualization of environment / situation

scene video with eye point

Picture 4: Hardware and software of caSBAro

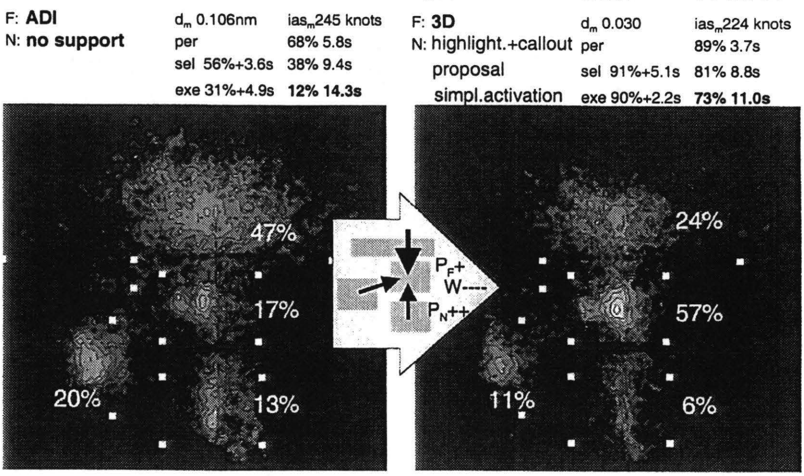
caSBAro is based on a generic flight simulator with collimated external vision and uses the same soft- and hardware environment as the assistant system (Silicon Graphics, Unix, C/C++). caSBAro integrates additional user interfaces (X/Motif, visualization library XRT/PDF), a commercial eye tracking system (ASL4000, cornea-reflex and video based pupil measurement, magnetic head tracker) and a videodisc recording system (Sony).

This is how caSBAro brings the pointillist concept into reality: each information exchange between the more than 40 modules of the assistant system (via a Unix Shared Memory), and every interaction between pilot and assistant system/aircraft, is picked up by caSBAro. A time stamp is added, the interactions are visualized and stored to a hard disk. The data can be analyzed online or offline, for example in respect to the spatial and temporal distribution of attention or the timeliness of specific interaction pattern. The data can also be replayed in the simulator, correlated with the eye tracking video. This replay can be operated as simple as a VCR.

Examples for the use of caSBAro and the pointillist approach

caSBAro was used for the evaluation of CAMA with pilots in a flight simulator. Without going into too much detail, the following examples of this evaluation can illustrate the spectrum opened by a pointillist approach, and its connection to "classical" concepts.

Example 1 connects the pointillist approach to "classical" concepts of quantitative analysis. A series of 15 generic simulator experiments focused on the influence of technical support on subjective workload (SWAT with conjoint scaling), objective performance (mean distance to track) and objective visual attention. The work process here is typical for vehicles and combines a higher frequency-guiding task "head up" (supported e.g. with a 3D-display) with a navigation task "head down" (supported e.g. with the highlighting of new conflicts or an automated proposal for conflict resolution).



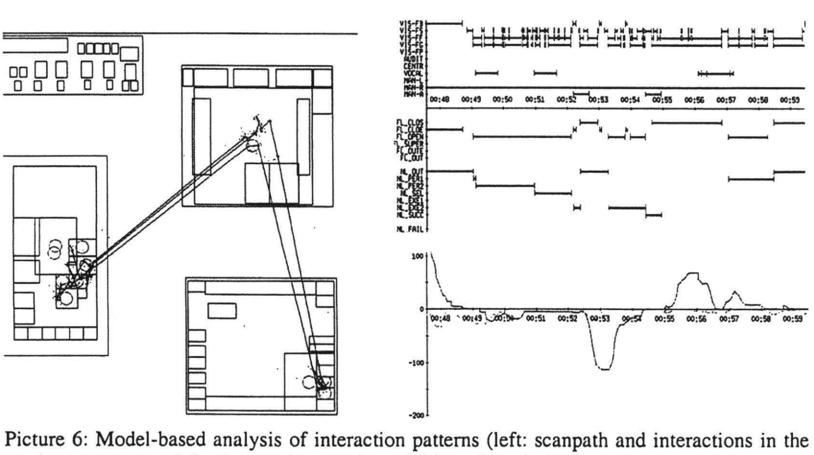
Picture 5: Effects of technology on workload, performance and visual attention, explanation in the text, [Flemisch & Onken 2000]

The manual and visual interactions are evaluated quantitatively with respect to reaction time and success rate. The visual interactions are used to derive the classical percentage number of the attention distribution, and to generate a 2D or 3D graphics (mountains) of visual attention, which can also be interpreted intuitively.

Picture 5 shows a comparison between 2 (of 15) different configurations of technical support. E_15, the base line with a classical ADI display and no additional support for the Navigation task, is compared with E_14, the configuration with a 3D-display for the flight task, and full support for the navigation task (highlighting, callout, automated proposal and a simplified activation "roger do it"). Subjective Workload W and mean distance to track d_m decrease significantly, with a light decrease of average indicated air speed ias_m. The performance for the navigation subtask, P_N, is improved on all stages of information processing (perception, selection and response execution), the rate for successfully solved conflicts increase dramatically with faster solution time.

The distribution of visual attention also shows the other side of the technological coin: The visual attention is concentrated into the 3D display (from 17% to 57% of the flight time), and withdrawn from other sources of information like the external vision (47% to 24%). The chance for building up appropriate situation awareness about vital information in the environment that is not displayed on the 3D-display (e.g. other aircraft), is hereby clearly decreased. More details, configurations and a critical experiment about that concentration effect can be found in [Flemisch & Onken 2000].

Example 2 illustrates the potential of the pointillist approach for a quantitative, model-based analysis of interaction patterns. The interaction patterns (sequence of interactions) of two subtasks, "manual flying" and "navigation", are modeled with state machines / Markov-chains and quantified with data from the above-mentioned experiments. These simple models combine visual with non-visual interaction and include interaction failures. The interaction pattern of manual flying for example is modeled according to the states of the control loop, which can be *closed* (pilot looks at the flight guidance display and moves the stick), *open* (pilot does not look, but moves the stick), *supervisory* (just looks with no control input) and *out*.



Picture 6: Model-based analysis of interaction patterns (left: scanpath and interactions in the workspace, upper right: interactions and model based analysis on a timeline, lower right: sidestick action. Direct snapshots from caSBAro, explanation in the text)

Picture 6 shows the interplay between navigation (left and lower display) and flying (upper display) with the output of the model-based analysis (upper right time-line) and the corresponding stick actions (lower right time-line). In data too large to analyze manually, model-based analyses are often the only way to detect hidden usability problems. These "tiny" problems can hook the operator's attention so extremely, that he/she can loose control over his/her vehicle and crash. Precursors to a potential hooking effect of a new, not yet modeled technology can be found by analyzing the impact on a known and modeled base task (like in this instance manual flying) [Flemisch 2001].

Example 3 focuses on the qualitative extreme and evaluates the replay. Pilots flew with a newly developed electronic checklist, which also included configurable automates. After the flight, they rated these features with a conventional questionnaire, then with a replay of the original flight in the simulator, including the eye tracking video, and finally rated this type of evaluation itself. While the most critical points were already mentioned in the conventional questionnaire, a lot of other important points were mentioned only in the replay session. The more complex problems with the automation in particular could be discussed in an environment where each of the partners, pilot, engineer and ergonomist, had his own appropriate kind of information and representation. Engineers/designers especially liked the chance to look very deep into the technical system if necessary and to have an objective reference in case of subjective complaints about certain features. Pilots clearly rated that the dialogue with the developer becomes more efficient, more objective and that the approach also holds a high potential for the improvement of their training [Flemisch & Onken 1999], [Flemisch 2001].

These three examples also illustrate the possible interplay between the different perspectives. The more global approach of example 1 might show a certain percentage of visual attention in a certain display, which might, together with a higher workload, indicate a hidden usability problem between different subtasks and their interaction patterns. The model based approach of example 2 can go into more detail and can isolate within larger data the specific cases where this problem could be severe. An intense replay session of these cases, with changes between an ergonomical, technical and operational focus, can not only reveal the technical and ergonomical details of that specific problem, but also bring developer, operator and human factors engineer together in a manner, that a solution is possible.

Conclusion

These are just a few examples what can be done with caSBAro, which is, shaped by its counterpart CAMA and technological/organizational constraints, only one instance of a pointillist approach. The pointillist approach itself, even if it was instantiated here concretely, is not so much a strict concept but a metaphor to show a potential direction of future human factors analysis.

We shall not let ourselves be discouraged by Havel's abysses. If we find ways to combine quantitative with qualitative, objective with subjective and analytical with intuitive approaches, we might reap the benefits from both worlds, which are one anyway. The scientific and technological progress will continue even faster and will change not only our technology, but also our lives and us. We are not doomed to simply accept this. If, and only if we have a clear window to what happens with us and technology, we can influence this process and to some respect choose the role, which we want to play with the technology of tomorrow. This one world is not going to stop and wait for us: Time's up.

Acknowledgement

This work was only possible with the help of many people. Representative for these stand Marius Morawski for his work with the eye tracking system, Kurt Kagleder for his programming in caSBAro, and Dr. Jennifer Burt for the revision of this text. Thank you!

Literature

[Bainbridge 1987] Bainbridge, L.: Ironies of Automation; In: New Technology and Human Error; Ed. Rasmussen, Duncan, Leplat; Wiley, 1987

[Billings 1997] Billings, Charles E.: Aviation Automation - The Search for a Human-Centered Approach; Lawrence Erlbaum, Mahwah New Jersey, 1997

[Blanchard 1991] Blanchard, B.S.: System Engineering Management;Wiley, NY, 1991

[Chapanis 1996] Chapanis, A.: Human Factors in System Engineering; Wiley, NY, 1996

[Endsley 1995] Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems; Human Factors, 37(1), p32-64; 1995

[Endsley & Kiris 1995] Endley, M.; Kiris, E.: The Out-of-the-Loop Performance Problem and Level of Control in Automation; Human Factors, 37(2), 381-394; 1995

[FAA 1996] Federal Aviation Administration Human Factors Team Report on: The Interfaces Between Flightcrews and Modern Flight Deck Systems; USA / Europe, 1996

[Flemisch & Onken 1999] Flemisch, F.O.; Onken, R: caSBAro, a Computer Aided ErgonomicS tool to detect and solve problems in human-computer interaction; *International Conference on Computer-Aided Ergonomics and Safety CAES*, Barcelona, Spain, 1999

[Flemisch & Onken 2000] Flemisch, F.O.; Onken, R: Detecting usability problems with eye tracking in airborne battle management support; NATO RTO HFM Symposium on „Usability of Information in Battle Management Operations"; Oslo; Norway; 2000

[Flemisch 2001] Flemisch, F.: Pointillistische Analyse der visuellen und nicht-visuellen Interaktionsressourcen am Beispiel Pilot- Assistenzsystem; Dissertation; Universität der Bundeswehr München, 2001

[Haberfellner et.al. 1992] Haberfellner, R.; Becker, M.; Büchel, A.; von Massow, H.; Nagel, P.; Hrsg. Daenzer, W.F.; Huber, F.: Systems Engineering; 7. Auflage, Verlag Industrielle Organisation, Zürich, 1992

[Havel 1994] Havel, V.: The Need for Transcendence in the Postmodern World; Speech in the Independence Hall; Philadelphia, USA; July 4, 1994

[Kraiss 1994] Kraiss, K.-F.: „99% Langeweile und 1% panische Angst" –über die Schwierigkeiten beim Umgang mit hochautomatisierten Systemen; In: M. Kerner (Hrsg.) Technik und Angst. Zur Zukunft der industriellen Zivilisation. Aachen, 1994

[McLuhan 1964] McLuhan, M.: Understanding Media; 1964; MIT Press; Massachusetts, USA; 1999

[Meister 1987] Meister, D.: Systems Design, Development, and Testing; In: Ed. Salvendy, G.: Handbook of Human Factors; Wiley New York; 1987

[Morawski 1999] Morawski, M.: Blickbewegungsanalyse von Pilotenverhalten; ISF, UniBwM, 1999

[Nielsen 1993] Nielsen, Jakob: Usability Engineering; AP Professional, Cambridge, MA, USA, 1993

[Norman 1988] Norman, D.A.: The Psychology of everyday Things; Basic Books, USA, 1988

[Onken 1998] Onken, R.: Die zwei Gesichter der Automation; Workshop Cockpit Automation, München 1998

[Onken 1999] Onken, R.: The Cognitive Cockpit Assistant Systems CASSY/CAMA; World Aviation Conference; San Francisco, USA; 1999

[Rasmussen et.al. 1994] Rasmussen, J.; Pejtersen, A.M.; Goodstein, L.P.: Cognitive Systems Engineering ;Wiley Interscience, NY, 1994

[Schulte 1996] Schulte, A.: Visuelle Routinen von Piloten; VDI Reihe 12 Nr. 278; Germany 1996

[Weick 1979] Weick, K.E.: The Social Psychology of Organizing; McGraw-Hill, 1979

[Wickens 1992] Wickens, C. D.: Engineering Psychology&Human Performance; Harper Collins 1992

Application and further development of CREAM exemplified by operation and control of a chemical process

Hans-Gerhard Giesa & Tanja Köhler
Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft,
Fachgebiet Mensch-Maschine-Systeme, Sekr. J 2-1, Jebensstr. 1, D-10623 Berlin,
E-mail: giesa@zmms.tu-berlin.de

The method CREAM provides a set of cognitive activities to build cognitive demands profiles that are used for human error prediction. The intention of this study was to improve the practicability of building cognitive demands profiles and to increase its interrater reliability by a revision of the definitions of the cognitive activities. The first step was a review of the definitions by 14 German experts by means of a questionnaire. The revised definitions are presented in this paper. A further evaluation of the revised definitions was performed with the support of five HRA-experts who were asked to build cognitive demands profiles for an exemplified CREAM-analysis twice: first with the original definitions and second with the revised definitions. The concordance between the HRA-experts was already surprising high in building a cognitive demands profile using the original definitions and could not be improved by the revision of the definitions. The HRA-experts assessed the support provided by the revised definitions as noticeably better than by the original definitions although the scope of interpretation was not reduced.

1 Analysis of Human Reliability in Human-Machine-Systems

1.1 HRA-Methods

In the area of reliability engineering, "reliability" is considered to be a probabilistic parameter for the quantitative assessment of a system. The reliability is expressed by the probability of not failing during a predefined time under given conditions. Numerous methods have been developed to express human reliability as well in quantitative probabilities. According to the German guiding principle VDI 4006, such Human Reliability Analysis Methods (HRA) should be able

- to analyse human actions in a qualitative way,
- to identify likely erroneous actions,
- to identify the weak points of the system for the generation of remedial measures, and
- to quantify human reliability for an evaluation of the efficacy of possible interventions.

In recent years, existing methods for analysing and assessing human reliability and their underlying methodology have been criticised especially by psychologists. (Giesa & Timpe, 2000; Timpe & Giesa, 1998). New methods have been developed to predict cognitive failures which become more and more important in human-machine-systems.

1.2 CREAM

The HRA-method CREAM (Cognitive Reliability and Error Analysis Method, Hollnagel, 1998) is constructed as a bi-directional method. CREAM is feasible both for prospective analysis (performance prediction) and for retrospective incident analysis (e.g. accident analysis). The prospective estimation of failure probabilities is based on the estimation of a control modus (basic method) and the building of cognitive demands profiles for specific tasks (extended method). Based on a task analysis, dominant cognitive activities are assigned to each identified sub-task. Each of these cognitive activities is described in terms of four cognitive functions (observation, interpretation, planning, execution). On the basis of the frequency of the cognitive functions, the cognitive demands profile of a task or sub-task is built. This profile is then used to identify cognitive function failures.

For experimental studies in which the dependability of new communication interfaces for the Airbus A 340 were investigated (Fricke et al., 2000), cognitive demands profiles were built for selected task-sequences using categories from the extended method of CREAM (Giesa, Müller & Anders, 2001; Müller, Giesa & Hauß, 2000). Two independent analyses were performed using video-recordings from these experiments.

There were some inconsistencies between the two analyses in assigning the cognitive activities to specific sub-tasks. These inconsistencies were ascribed to the scope of interpretation for the given definitions of activities. These experiences were the starting point for a revision of the definitions of cognitive activities used in CREAM.

2 Application and further development of CREAM

It was assumed that a redefinition and a more detailed description of the activities would improve the practicability of CREAM and would increase the reliability of building cognitive demands profiles within the extended method of CREAM for performance prediction. The first step of the revision was a review of the definitions by 14 German experts by means of a questionnaire. The further evaluation of the revised definitions was performed with the support of HRA-experts. They were asked to assign the cognitive activities to the documented sub-tasks in an exemplified CREAM-analysis twice: first with the definitions given by Hollnagel and second with the revised definitions.

2.1 Revision of the definitions of the cognitive activities

14 German experts (eight psychologists and six engineers) in human reliability reviewed the definitions of the cognitive activities by filling in a 27-sided questionnaire. The concept of using cognitive activities and a cognitive demands profile for HRA and their meaning and use in CREAM was explained in the questionnaire. For each activity, the experts were asked whether they accept the given definitions. Furthermore, the experts were requested to provide two typical examples from human-machine-systems for each activity. In addition they were asked whether the given set of cognitive activities is comprehensive or whether activities should be added. Table 1 shows the results concerning the acceptance of the original definitions of the cognitive activities.

Table 1: Acceptance of the definitions of the activities

Activity	n	Acceptance				Rejection	
		agree	generally agree	total	total in %	disagree	in %
Scan	14	11	3	14	100		
Execute	14	8	6	14	100		
Verify	14	9	5	14	100		
Plan	14	9	4	13	92,9	1	7,1
Regulate	14	7	6	13	92,9	1	7,1
Observe	12	8	2	10	83,3	2	16,6
Evaluate	12	5	5	10	83,3	2	16,6
Compare	14	7	4	11	78,6	3	21,4
Diagnose	14	5	6	11	78,6	3	21,4
Monitor	13	8	2	10	76,9	3	23.1
Record	13	5	4	9	69,2	4	30,8
Identify	12	4	4	8	66,6	4	33,3
Communicate	14	7	2	9	64,3	5	35,7
Co-ordinate	14	4	3	9	64,3	5	35,7
Maintain	13	5	3	8	61,5	5	38,5

The categories "agree" and "generally agree" are summarised as one category "total" which represents a general acceptance. Only the degree of agreement to the definitions of "record", "identify", "communicate", "co-ordinate" and "maintain" is less than 75 %. The different numbers of "n" arise from missing responses. Table 1 shows that most of the definitions are accepted generally. Added comments often referred to the given assignment of the four cognitive functions to the activities and to the complexity of the definitions.

The revision of the definitions mainly results from the outcomes of the questionnaire and from our own experiences from the CREAM-Analyses in the Airbus A340-project. Table 2 shows the revised definitions of the cognitive activities. Each definition is supplemented with two examples selected from the experts to illustrate the meaning of the definitions.

The original set of definitions is complemented by the activities "supervise" and "decide". "Supervise" was requested by some experts who missed the consideration of human-human-interface in the given activities: the activity "monitor" is seen as too technical. Though the activity "decide" is implicitly included in other cognitive activities, it was missed as an separate activity by some experts.

2.2 Evaluation of the revised definitions by HRA-experts

For the evaluation of the revised definitions of the cognitive activities, a CREAM-analysis of the operation and control of a chemical process was prepared. The evaluation was performed with the support of five HRA-experts. All had previous experience with the application of CREAM.

Table 2: The revised definitions of the activities and two typical examples

Activity	Revised definitions and typical examples
Scan	Quick review of displays or other information source(s) to obtain a general but preferably complete impression of the state of a system or sub-system. (Examples: Quick review of the essential instruments in an aircraft cockpit. Quick review of the operating console in a control room.)
Observe	Look for or read one or more specific measurement values or system indications.(Examples: Read the values of temperature, pressure, and power for the protocol. Read a system message like ‘high engine vibrations”.)
Monitor	Follow the development of one or of a set of specific parameters or keep track of system states over time. (Examples: Watching the development of specific parameters like temperature, pressure and power of a reactor. Watching the altitude of an aircraft during the descent.)
Compare	Examine the qualities and/or quantities of two or more entities (measurements) with the aim of discovering similarities or differences. (Examples: Nominal/actual value comparison of the pressure of a reactor. Comparison of the indicated values of analogue/redundant panel meters.)
Verify	Confirm the correctness of a system condition or measurement or confirm the success of a taken action. (Examples: Confirm the successful start of a pump after checking that it reached the correct mass flow. Confirm the correctness of an alarm after inspecting the related displays and instruments.)
Identify	Establish the identity of a plant state or sub-system or component state. (Examples: Determine the identity of an engine according to power. Define the stability of a sub-system after the performance of a survey.)
Diagnose	Recognise or determine the nature or cause of a condition. (Examples: What is the reason for a loss in oil pressure? Diagnosis of an alarm in the control room, i.e. to which component the sign belongs.)
Evaluate	Appraise or assess an actual or hypothetical situation based on available information without requiring special operations with regard to one or more criteria. (Examples: Assessing whether the level of the temperature increases fast enough to reach a given goal in time. Appraising whether the planned point of landing is approachable with the remaining fuel.)
Plan	Formulate a set of actions by which a goal will be successfully achieved. Plans may be short-term or long-term. (Examples: Developing a flight plan. Developing a check list.)
Execute	Perform a previously specified action. (Examples: Open or close a valve. Switch off or on a button, machine, pump etc..)
Regulate	Alter speed or direction of a control (system) manually with an immediate effect in order to attain a goal. Adjust or position components or subsystems manually with an immediate effect to reach a target state. (Examples: Regulating manually parameters like speed and course in a car. Regulating manually altitude in an aircraft during the descent without using the autopilot.)
Maintain	Sustain a specific operational state. (Examples: Keep the level of the temperature in a rectification system within two limits (e.g. 35 and 38C) by manually switching on and off a heating band. Sustain the course of a ship due to weather by manually correcting the position of the rudder.)
Co-ordinate	Bring two or more independent system components into a specific relation at the same time. (Examples: Switch on two different buttons at the same time. One is for regulating the temperature, the other is for regulating the power. Open two different valves in a specific relation at the same time.)
Record	Write down or log system events, measurements, etc. (Examples: Noting observed measurements with the aid of a given list. To document a manufacturing process into a log-book.)
Communicate	Pass on or receive verbal person-to-person information needed for system operation. (Examples: Communication via VHF in an aircraft cockpit between crew and ATC. Tell another operator face-to-face, what further job he has to do.)
Decide	Goal-directed process of assessment of two or more possible options followed by the choice for one option. (Examples: To settle for the execution of a flight plan change or not. To select which of the given alternatives for a flight plan will be taken.)
Supervise	Observe and control the operations of one ore more involved operators. (Examples: To keep an eye on the activities of another operator in a control room. The Pilot Non-Flying watches the activities of the Pilot Flying in an aircraft and vice versa.)

2.2.1 Object of the analysis: a rectification system at the Technical University Berlin

The rectification system used for the exemplified analysis is installed at the Technical University Berlin. This package column for a three-phase-rectification has not yet been implemented in industry. It was chosen due to the easy access to a package column and to the operators. To operate the rectification system, a wide range of typical chemical engineering tasks is required to start the package column, continuous operation, liquid and vapour sample withdrawal, an analysis of the samples, and to stop the package column. Though the operation of the package column is not particularly dangerous, it is expected that the results can be transferred to safety relevant processes. The actions differ only in the degree of the consequences resulting from erroneous actions.

2.2.2 Design of the study

The five HRA-experts received a 60-page documentation of the rectification system. This documentation included a general description of the structure of the plant, a site plan, a process diagram of the plant, and extensive photographic documentation. The documentation also included the results of a given task analysis which was performed using observation interviews during the operation and control of the package column. 11 main tasks were identified which consist of 146 sub-tasks.

The HRA-experts were asked to assign the activities to the documented sub-tasks using the given definitions. Afterwards, the experts were asked to fill in a questionnaire which includes several items concerning the practicability of the given definitions. The experts had to assign the activities and to answer the questionnaire twice: first using the original definitions and – seven months later – using the revised definitions of the cognitive activities.

2.2.3 Results

Some selected results concerning the practicability and the interrater reliability are presesented as follows. A more detailed analysis of the results can be found in Köhler (2001). An important criterion for the evaluation of the practicability of a method is the effectiveness of support attributed to the method itself. Therefore, the HRA-experts had to estimate the effectiveness of support from the list of definitions for building the cognitive demands profile in the questionnaire. The range of the scale was 1 (no effective support) to 5 (effective support). Table 3 shows the ratings for both analyses. There is a noticeably better estimation for the revised definitions.

Table 3: Effectiveness of the support of the list of activities

	Original Definitions	Revised Definitions
Expert 1	4	5
Expert 2	4	5
Expert 3	1	4
Expert 4	4	5
Expert 5	2	4
Mean	3	4,6

To estimate whether the scope of interpretation of the definitions was reduced by the revision of the definitions, the HRA-experts had to rate this scope of interpretation. The range of the scale was 1 (no scope of interpretation) to 5 (very large scope of interpretation). There was only a slight tendency in reduction of the scope of interpretation from the original to the revised definitions. Table 4 shows considerable individual differences between the ratings of the experts.

Table 4: Scope of Interpretation

	Original Definitions	Revised Definitions
Expert 1	4	4
Expert 2	2	1
Expert 3	5	5
Expert 4	2	1
Expert 5	4	3
Mean	3,4	2,8

The experts were asked to specify the time they needed for the examination of the documentation and for assigning the activities to the given task-steps. The results are shown in Table 5. There are relatively high individual differences between the experts independent from the definitions used. When compared to the original definitions, the time for building the profiles was reduced for all experts with the revised definitions. The reduction in time is in a range from 20 to 85 %. It has to be taken into account that a considerable part of this time reduction should be attributed to the familiarity with the process when performing the second analysis with the revised definitions. In the presented results, the reduced time could be supported by the improved practicability of the definitions. However, the design of this study does not allow a clear statement at this point.

Table 5: Time needed for assigning cognitive activities to task-steps

	Original Definitions		Revised Definitions	
	Time for documentation	Time for assignment of activities	Time for documentation	Time for assignment of activities
Expert 1	240 min.	180 min.	no response	no response
Expert 2	180 min.	300 min.	60 min.	180 min.
Expert 3	150 min.	390 min.	60 min.	60 min.
Expert 4	60 min.	120 min.	20 min.	40 min.
Expert 5	300 min.	600 min.	120 min.	480 min.
Mean (Exp. 2-5)	172,5 min.	352,5 min.	65 min.	190 min.

Against the expectation there are not much differences between the cognitive demands profiles of the first and second analysis. Figure 1 shows the resulting cognitive demands profiles for the operation and control of the rectification process built with the original definitions (1. cognitive profile) and with the revised definitions (2. cognitive profile). The presented figure represents an averaged profile which is based on the activities with the highest match between the five experts.

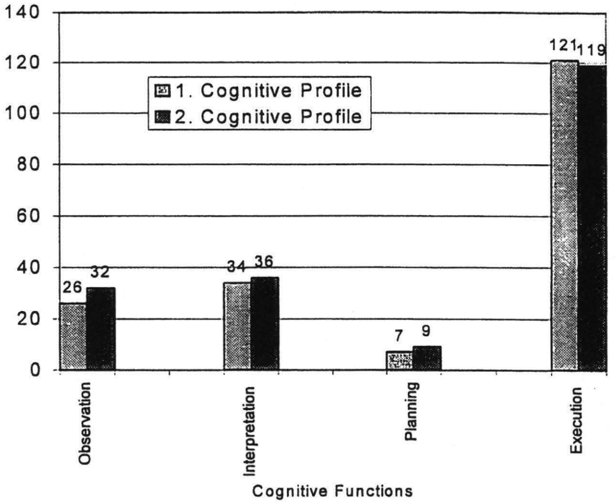


Figure 1: Cognitive demands profiles for the operation and control of a rectification process

Overall, the revision of the activities had relatively little influence on the resulting profiles. There was already a high interrater reliability of 70 % in building the cognitive demands profile with the original definitions. This high concordance could not be improved by the revision of the activities.

3 Discussion

The results indicate that practicability could be improved by means of the revision of the definitions of the cognitive activities. The HRA-experts assessed support provided by the revised definitions as noticeable better than by the original definitions. The scope of interpretation was reduced negligibly. The clearly reduced time for building the cognitive demands profile can be attributed in a significant way to familiarity with the task resulting from the first analysis. But, this reduction is probably enhanced by the improved practicability of the definitions.

Agreement between the HRA-experts was already surprising high in building the cognitive demands profile using the original definitions and could not be improved by the revision of the definitions. This result was not expected due to our experience in building the cognitive demands profiles in the Airbus A340-project. One possibility that explains this result is a supposed difference between novices and experts: the analysis in the cockpit of the A340 was performed by HRA-novices and the analysis for the operations and control of a rectification system by HRA-experts. Maybe the experts possess a similar representation of the activities due to their experience which is independent from the given definitions.

Finally, it should be pointed out that the assignments of the cognitive functions to the activities were much more criticised by the 14 experts who participated in the review of the definitions than the definitions themselves. A revision of these assignments seems to be reasonable based on these results.

4 Literature

Fricke, M., Dehn, D. & Müller, T. (2000). Co-operative Air Traffic Management - Human Factor Aspects of Air/Ground Data Link and Multi-Sector Planning, Fortschritt-Berichte VDI, Reihe 22 Nr. 4. Düsseldorf: VDI Verlag.

Giesa, H.-G., Müller, T. & Anders, G. (2001). The influence of airborne Data Link on system dependability. In D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics, Volumes 5 & 6. Aldershot: Ashgate.

Giesa, H.-G. & Timpe, K.-P. (2000). Technisches Versagen und menschliche Zuverlässigkeit: Bewertung der Verläßlichkeit in Mensch-Maschine-Systemen. In K.-P. Timpe, T. Jürgensohn & H. Kolrep, Mensch-Maschine-Systemtechnik (S. 63-106). Düsseldorf: Symposion Publishing.

Hollnagel, E. (1998). Cognitive Reliability and Error Analysis Method. CREAM. Oxford: Elsevier.

Köhler, T. (2001). Anwendung und Weiterentwicklung von CREAM am Beispiel der Bedienung und Überwachung einer verfahrenstechnischen Anlage. Diplomarbeit am Institut für Psychologie und Arbeitswissenschaft der TU Berlin.

Müller, T., Giesa, H.-G. & Hauß, Y. (2000). Bewertung eines kooperativen ATM-Konzepts in Simulatorstudien. In K.-P. Timpe, H.-P. Willumeit & H. Kolrep (Hrsg.), Bewertung von Mensch-Maschine-Systemen (S. 115-135). Düsseldorf: VDI Verlag.

Timpe, K.-P. & Giesa, H.-G. (1998). Man-Machine-Systems Engineering. In H. Schmidt-Traub (Hrsg.), Computer Application in Process and Plant Engineering (DECHEMA Monographs, Bd. 135, S. 331-345). Weinheim: Wiley-VCH.

VDI 4006-2. Menschliche Zuverlässigkeit – Methoden zur quantitativen Bewertung menschlicher Zuverlässigkeit (Entwurf), April 1999.

The Inexorable Link between System Control and Designed Performance

Iain S MacLeod

Aerosystems international
West Hendford
Yeovil
Somerset BA20 2AL
UK
Tel: + 1935 443000
FAX: +1935 443111
email: iain.macleod@aeroint.com

Summary

This paper argues that regardless of the structure of the system, system control and designed performance are inexorably linked. With the advances in Information and Communications technologies, systems are becoming larger and more diverse than hitherto. The architecture, processes, and form of functions within systems are evolving with the use and handling of knowledge becoming increasingly important. This latter issue requires that both human and machine must share knowledge-related tasks in order that interoperability between sub systems will be actioned. However, such interoperability will not necessarily depend on the co-location of system assets but will be strongly influenced by the culture, organisation, situations, and contexts under which the system has to operate. The System of Systems concept is introduced as being capable of supporting many diverse and geographically distributed sub systems. Control of the performance of such a system is argued to be more effective if the structure of command and control is flat rather then hierarchical.

Introduction

Engineered control is inexorably linked to the engineered architecture of the system, and directs the system's achievable levels of functional performance and operational capability. However, quality design of operator controls and system feedback can enable the system to exceed designed performance through evoking system functions early in anticipation of future events and often, through anticipation, circumnavigation restrictions on system usage such as speed or processing constraints. Good human control inputs to the engineered components of a system can provide the necessary bonding in a system to enable its operating performance to be greater than the sum of the performances of its parts.

Good human control is assisted by human skills, through expertise developed in the control of the particular system, by effective use of technology in systems design, teamwork, and through associated designed levels of sensible automation of system functions. Importantly, human control of a system is actioned through heuristics but, traditionally, relies on the support of engineered algorithmic-based control mechanisms. In the future, more control systems will be engineered through a combination of software heuristic and algorithmic techniques. Table One shows examples of some modern technologies, some heuristic based, and their uses.

Table One. Some Modern Technologies and Their Uses

TECHNOLOGIES	EXAMPLE APPLICATIONS
ESTABLISHED TECHNOLOGIES	FLIGHT MANAGEMENT SYSTEMS
KNOWLEDGE BASED SYSTEMS (KBS)	TACTICAL DECISION AIDS
CASE BASED REASONING	ENGINE MAINTENANCE AIDS
NEURAL NETS	FUEL MANAGEMENT
BAYESIAN LOGIC	ENGINE VIBRATION ANALYSIS
FUZZY LOGIC	TRANSPORT SCHEDULING
INTEGRATED TECHNOLOGIES	INDUSTRIAL AUTOMATION

But technology can only be optimised for human control if it is designed to support the operator cognitive functions required for the direction of system performance towards the achievement of system goals. Cognitive functions refer to those functions of a human-machine cognitive system (Hollnagel & Woods, 1983) that encompass the system's use of knowledge, its awareness and use of assessments of the environment, an anticipation of future events, the management of system resources, supervision of system performance, judgement, and the dynamic direction of system operation towards goal achievement. Within the constraints of cost and capability limitations of engineering systems using new technologies, system related cognitive functions may be conceived to be specified to reside in man, machine or be shared by both (MacLeod, 2000).

The fundamental problem that must be addressed by control is that of the appropriate timing of the control. Traditionally, engineered control has been bounded by a clocked machine iteration to equate hysteresis within the control loop. This is functioned through a set of specific engineered rules. In contrast, human control is functioned under an individual's awareness of the control needs and the time available to perform control tasks, this appreciation strongly dependent on the individual's repertoire of skill based heuristics and that individual's expertise in the particular domain and environment. Further, certain system controls are cognitive in nature and traditionally reside with the human operator(s) of a system. This working state is especially relevant to the dynamic replanning and application of tactics within the military environment.

This paper will consider changes to the nature and purpose of systems, addressing issues related to system architecture (Rechtin & Maier, 1997) and the control of system functions and performance. Rapid advances in Communication and Information Technologies support these changes and promote developments in socio-technical systems where the advances in technology are encouraging a levelling of hierarchical team structures. In parallel, there is a promotion of the alternative flatter team structures as improved means of allowing the communication of information between teams (see Allusi, 1992 on the 7 'C' of teamwork) and the complementation of human knowledge and skills towards the achievement of goals.

Different Properties of Data, Information, Advice, and Knowledge

It is important to understand the nature of what is communicated within systems as confusion in this area has possible affects on systems interoperability and safety. To understand the different properties of data, information, advice, and knowledge. Data in itself has no meaning and requires an application of appropriate rules to convert it into information. Information can exist in many forms, one taxonomy considering information as:

- Collateral. Collateral is historic information that assists the planning and execution of a task. Examples of this type of information are mission planning information or a geographical map.
- Ephemeral. Ephemeral information is information where its use only has a short term validity. Examples of this type of information are the speedometer readout whilst a car is in motion or the height information on an aircraft during descent.
- Individual. Individual information concerns properties of an individual human or system component that is only pertinent to that individual and cannot be aggregated with other system related information. An example would be the information contained in a personality profile of an individual.
- Real World. Real World information is information concerning the detection, classification, and status of real world objects. An example would be the details on aircraft type and specific identity.

Advice is a result of the application of rules to combine information and knowledge as a means of predicting future events and making explicit associated recommendations for direction of future activity (MacLeod, 1998). Knowledge is a collection of information and the rules for the appropriate application of that information with relation to the perceived environment and work context.

Concept of Agent Related to Interoperability and Cognitive Functions

The concept of an Agent will be used to assist the argument of the paper. By its nature an Agent has to be able to use data, information, and knowledge but may not necessarily proffer advice. An Agent can be considered to be any system-related entity that has a delegated responsibility to act in a particular fashion within its competency and its environment(s). This responsibility requires that, within given remits, an Agent co-operates with other system Agents to contribute to the control and optimisation of the host system performance.

Furthermore, the concepts of interoperability between systems and system components will be examined as an avenue to a discussion on the roles of system Agents as mediators on the utilisation of designed system functions and processes. Such mediation concerned primarily with the degree of a system's use and handling of imbued knowledge as allowed by the architecture and functionality of the designed system. It will be argued that system control and performance are inexorably linked but that the quality of that link depends ultimately on the specification of system related cognitive functions; the system functions related to the adaptability of the system's use and handling of knowledge.

Changing Considerations on the Nature of Systems

The rapid advances in Information Technologies and Communication Technologies have promoted a consideration of systems that is beyond the traditional concept of a single product or system engineered for a specific purpose and physically co-located within a specified environment. This has produced misunderstandings about the similarities and differences between requirements, traditions, and benefits of the 'old' and the 'new' (Moore, 1995). The 'Systems of Systems' concept considers a system as a diverse range of sub systems working towards a common high level goal, fitting into a common overall architecture sustained by interacting machines, processes, information flows, organisations, and human-machine teams. This concept is unlike the relatively simple system that has traditionally been considered to involve one human through one interface interacting with a combination of engineered sub systems or equipment. The System of Systems concept involves teams of humans and teams of engineered systems interlocked through a top level architecture of the whole and also a series of processes and information flows providing the glue to allow the disparate parts to perform effectively as one to address dynamically changing goals.

The traditional problems with engineered control, namely the direction of system control through human cognitive functions and by human inputs to the system, remain regardless of the situation of the control, its form, or the technology used in the physical design of the system. However, with a System of Systems the idea of interoperability takes on an expanded meaning. Interoperability remains in part concerned with the standardisation of interfaces, protocols, and practices. However, it also encompasses a careful consideration and implementation of derived system processes and architectures involved with the rapid transfer and adoption of knowledge between many diverse, and distributed, system components. These components are effected by intrinsic System of Systems societal properties related to human culture, involved organisations, and participating teams. A System of Systems is a socio-technical system. Careful understanding of the above is necessary to meet the challenges involved with the real time adaptation of the overall performance and goals of a System of System necessary to meet the dynamics of changing challenges emanating from perceived artefacts distributed geographically over many locations and contributing sub systems. Figure One illustrates some of the constituents of a System of Systems.

Agents and Functions Related to the Use and Handling of Knowledge

Agents can be animate, inanimate, or combinations of both. There are many forms of Agents depending on their defined roles or capabilities. By considering role we refer to the system related functions and duties performed by the Agent during the life of a system. There are many detailed definitions of Agents. Arguably, some common properties of Agents are that:

- The Agent produces a timely response to changes within its environment;
- The Agent responses to external changes are goal orientated;

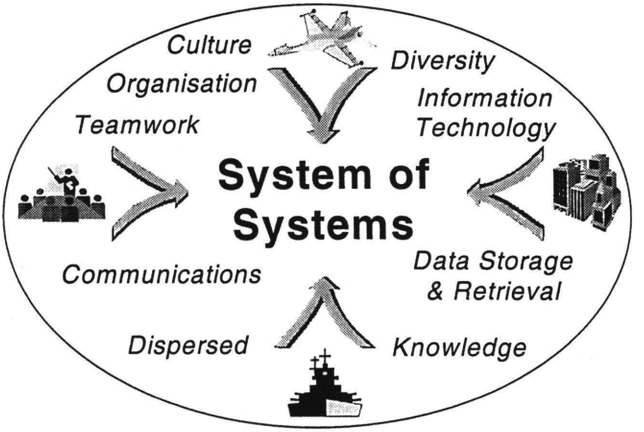


Figure One: Some of the Constituents of a System of Systems

- The Agent has some form of autonomy over the control of its activities;
- The Agent is continuously on if its host system is activated;
- The Agent communicates as required with other Agents of the system.

Inanimate engineered Agents act primarily on pre programmed tasks aided by communication with, and engineered assessments of, the needs or actions of other Agents in a system. In contrast, Humans, as sentient Agents, bring intrinsic properties of awareness and assessment to their roles. Consciousness is the basic ingredient of awareness and is the property that links the object that has consciousness with its environment and community (Dennett, 1996). The quality of conscious behaviour depends on intelligence and its appropriate application. Intelligence has many definitions but can be described as an ability to adapt to different and changing environments. The quality of that adaptation can be considered under terms of effectiveness.

Cognition implies some form of 'engine' concerned with the use and handling of knowledge. Cognition encompasses the processes, analysis, and knowledge/experience basis, currently assumed to primarily reside in the processes of the human mind, and that are sustained and supported by intelligence and consciousness. The communication, reception, and acknowledgement of explicit state information and intentions between diverse co-operating Agents are manifest as activities representing the effective facets of cognition used by a system. In the case of the human Agent this effectiveness is manifest through awareness and operator expertise at work performance. Figure Two illustrates some of the factors influencing the performance of a human Agent

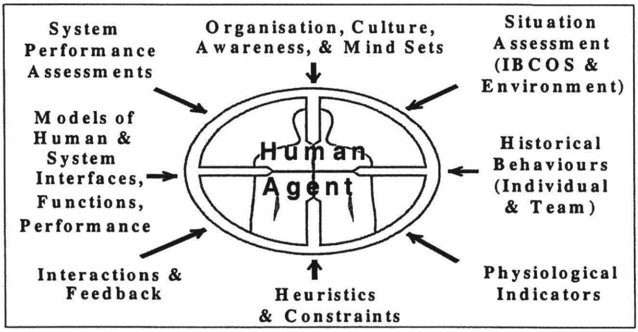


Figure Two: Some Factors Effecting Human Agent Performance

However, purposive use and handling of knowledge by engineered Agents can be argued to require forms of an in-built awareness related to the Agent's environment and goals. Organisation and culture are peripheral to a simple system, but are an intrinsic part of a System of Systems.

Currently, the fusion of the diverse areas encompassed by system related cognition requires the application of human expertise in the workplace. The machine Agent can accurately assess situations and conditions within its designed remit, and can perform quality and timely work. Moreover, a machine Agent can be argued to have a 'consciousness' and intelligence if it has in-built awareness of the existence, purpose, and capabilities of other machine Agents (Haig, 1998). Nevertheless, the machine on its own has no consciousness in the human sense, awareness of its inbuilt knowledge or of its inbuilt processes, or of the reasons supporting its performance goals. It must be expected that in the near future the ultimate responsibility with the management, control, and direction of the machine will reside with trained human Agents filling roles of operators, supervisors, or managers.

Furthermore, the role of an Agent must be planned to fit with the system performance requirements driving the designed amalgam of the functionality, architecture, and processes of the overall system throughout the system design and development life cycle. As such, the system specification must address issues related to system architecture and process, issues frequently not addressed by current specification methods. In addition, the specification must logically address all forms of system functionality, including that covering the use, handling, communication and control of system related knowledge essential to maintain system performance towards the achievement of system goals i.e. functions related to cognition. At the present, cognitive functions are poorly addressed or totally absent from system specifications. They often are considered later in the physical phase of system design through various means of implementation of derived functionality e.g. human computer interface modifications, retrospective automation of tasks previously found onerous by the human, or 'patches' to aspects of design implementation. Figure Three is a simple schematic of the fundamental phases of a system design life cycle.

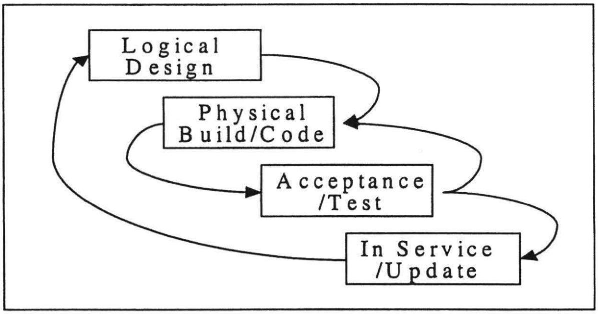


Figure Three: Fundamental Phases of System Design Life Cycle
(Derived from IEE P1220)

System Functions – Considerations on Type and Form

At a high level of description, system functions can be considered as properties of the system. At a more detailed description of system functionality, functions can be specified as system requirements accompanied by an associated performance. For example, *Provide fire extinguishing system capable of producing *** of *** at the locus of the fire within *** seconds of fire detection.*

When functions are used as a basis for the specification of the requirements of a system, they are defined through consideration of the capability requirements and the user requirements for the system. Often accompanying such a specification is a delineation of constraints and assumptions placed on the system and its operation. However, the customary use of the logical definition of system functions is flawed, this partly caused by a lack of consideration on system control and use of knowledge and partly because of the engineering based form of specification interpretation generally used the system design community. Moreover, a performance specification may define system functions and their associated performance but usually fails to define any requirements for system utility and ease of use or 'fitness for purpose'.

One example of a fitness for purpose form of requirement is:

Radar equipment controls will be provided that allow the Radar operator to retain the focus of their visual attention on the Radar display (Guidance contained in Standard xyz) and receive timely and appropriate feedback on the control operation through that display (e.g. between 0.2 to 0.5 second depending on the function).

The above problems in generating secure requirements are exacerbated by fixed price contracts that do not allow incremental or evolutionary approaches to design and development. These latter approaches allowing room for reappraisal of system functionality and the associated derivation of any necessary additional system functionality missing from the initial system functional and performance specification.

Some of the potential problem areas in system control that are normally poorly addressed by initial specification of requirements are:

Problem Areas and Some Related Issues in Control

1. Timing (e.g. poor timing can place control out of phase and prevent appreciation of control inputs).
2. Inappropriate form of control (e.g. the form of control does not allow the required degree of course or fine control).
3. Poor positioning of control (e.g. the operator finds the control hard to reach or the location of the control detracts from attention to main task performance).
4. Poor range of control (e.g. the range over which the control can be applied is inadequate to support its function).
5. Difficulties in activation of control (e.g. the control can only be activated by a command embedded in the depths of a hierarchical chain of commands).
6. Convoluted control (e.g. the functioning of many interactive controls is required to achieve the required result).
7. Control thresholds, gain, and device laws inappropriate to the form of control (e.g. control device insensitive with relation to specific control task).
8. Poor form of feedback (e.g. feedback of status of a continuous variable given by a digital rather than an analogue presentation; over memorisation of system states required of the operator).
9. Poor location of feedback (e.g. feedback position inappropriate to its appreciation with relation to form and activation of control).
10. Poor presentation of system status information (e.g. system control is difficult if the status of system activation or serviceability is hard to determine).

Some Design Implications

1. Poor design of system's engineered control requires undue attention to control by the operator at expense of other operator activities.
2. Under and over engineered functionality in the system requiring undue operator attention to system control.
3. High system noise obfuscating requirements for operator control (e.g. too much information, presentation of unnecessary information, presentation of inaccurate information, presentation of historical information as ephemeral information).
4. Poorly conceived automation of the system leading to poor system management, supervision and hence system control.

5. Inadvertent activation of controls.

6. Finally, performance of system cannot be dynamically adjusted to suit changing task requirements (e.g. the system consistently fails to meet its goals).

The above considerations illustrate some of the issues and implications related to the control of a simple system with relation to the associated designed performance of the system. There are associated issues and implications related to the use and handling of system knowledge within all systems, especially those related to the cognitive functions necessary for system direction and control. These functions are of high importance within the socio-technical remit of a System of Systems. Furthermore, if they are not carefully considered by design and development processes, their absence is partly the cause of perceived differences between specified system requirements and the system acceptance requirements related to perceived system utility and ease of use (Lane & MacLeod, 2000).

Judgement in System of Systems Direction and Control - Knowledge and Appreciation of Environment and Context

The direction of a system towards the satisfaction of high level goals involves the use of guidance from strategic planning. In contrast, the real time reactive control of a system is concerned with the timely fulfilment according to critical task related goals. Thus, both direction and control of a system involve exercises of judgement by Agents within the System of Systems. For example:

- ➢ Direction: The changing of the tasking of an asset by a military commander (human Agent) requires an assessment of the risk placed on the planned strategy related to mission effectiveness;
- ➢ Control: Flight changes en route by an aircraft pilot, required to avoid a flight hazard, necessitating an increase of aircraft speed in order to achieve a mandatory time on target.

Such judgement involves the appraisal of communicated information and knowledge, throughout the overall system, or the retrieval of information and knowledge from forms of easily accessible storage. A major problem in such retrieval and communication is to find the most pertinent information and knowledge for an intended purpose, and communicate it in a timely fashion to the correct recipient. Such activity can be performed automatically by the system provided the information or knowledge can be identified by known location and type. If the information has to be searched for, advice based on knowledge of the states of the System of System could be used, this considering the System of Systems sensed influences of the current environment, context, and the locations of similar previous retrievals.

However, a System of Systems is a socio-technical system where the form of command and team structures must have an important influence on the effectiveness of control and communications. It is important that all command, control, and team structures share a clear set of strategic goals and can co-operate to co-ordinate the communication of the expected outcomes with the planned recipients. The sharing of knowledge in any dispersed system must rely on System of System nodes of teamed Agents being responsible for particular tasks, the control of work, and the node's possession of relevant forms of information and knowledge. The promulgation of the outcomes of the activities of any node is likely to be in clear forms of information, knowledge, and advice that depend on situation and context.

Thus the team structure, relationships, direction of communication flows between the nodes, and the leadership of teamed nodes, will all have to be flexible to cater for changes in demand dictated by external influences. This suggests that the overall management and communication structure of a Systems of Systems will be flat rather than hierarchical as compromises will have to be continually strived for to maintain a balance within the workings of the system. Thus the relationship between control and performance will be close but may have to be constrained to prevent facets of the System of Systems becoming unmanageable and out of control. For these latter reasons, a careful but dynamic choice and amalgam of all forms of Agents must be possible depending on the influences of situation, context, and goal. The actioning of such choice must depend on the effectiveness of the System of Systems architecture, processes, and the reliability of communications (i.e. as secured by the completion of transmission, reception, and acknowledgement of the understanding of the contents of a communicated message).

End Notes

It has been argued that advances in Information and Communication technologies are changing the structure of systems toward a form more suitable for the use and handling of knowledge and the formation of larger collectives of systems termed System of Systems. These changes involve alterations within the requirements of interoperability between sub systems and also to the command, control, and team structures associated with these changes. Regardless of the argued changes to system structures, system control and designed performance are inexorably linked through the utility of the system and the ease of use of its assets.

References

Alluisi, E.A. (1992) in the Foreward to the book, R. Swezey, & E. Salas (Eds), *Teams: Their training and performance*, Ablex, Norwood, NJ.

Dennett, D.C., (1996), *Kinds of Minds*, Weidenfeld & Nicolson, London.

Hollnagel, E. and D.D. Woods, (1983) Cognitive systems engineering: New wine in new bottles, International Journal of Man-Machine Studies, Vol. 18, pp: 583-600.

IEEP1220 , (1994.), *Standard for the Application and Management of the SE Process*, IEEE Standards Department, Piscataway, NJ.,

Lane, K. & MacLeod, I. S. (2000), The Process of Certification: Issues for New Technologies, in *Proceedings of the 3rd International Conference on Engineering Psychology and Cognitive Ergonomics*, 25th to 27th October, Edinburgh, UK.

MacLeod, I.S. (1998), Information and Advice: Considerations on their Forms and Differences in Future Aircraft Systems, in Scott, P.A., Bridger, R.S., and Charteris, J (Eds), *Global Ergonomics*, Elsevier Science, Oxford, pps 667-672.

MacLeod, I.S. (2000), A Case for the Consideration of System Related Cognitive Function Throughout Design and Development, *Systems Engineering*, Vol.3, No.3, Wiley.

Moore, G.A. (1995), *Crossing the Chasm.* Harper Collins Publishers, New York.

Haig, N. (1998), The possible value of consciousness to guided weapons and other unmanned platforms, *Journal of Defence Science*, Farnborough, UK, Vol. 3, pp: 451-454

Rechtin, E. and Maier, M.W., (1997), *The Art of Systems Architecting*, CRC: Florida.

A socio-technical approach to safety assessment

Patrizia Marti, Antonio Rizzo, Petra Scrivani

Department of Communication Science
University of Siena
Via dei Termini 6,
53100 Siena, Italy

[marti/scrivani@media.unisi.it, rizzo@unisi.it]

Abstract

This paper proposes a new approach for the safety assessment and the proactive evaluation of the impact that new technological tools can have on the safety related work context such as Air Traffic Control. The approach considers all the resources that contribute to making control activity tightly interactive and context dependent. For this reason, the evaluation of safety critical system requires an in-depth analysis of the socio-technical context of work in order to assess the role that each component plays in the process. The Critical Interaction Analysis (CRIA) method, designed for this approach, has been applied to the evaluation of MIDAS prototype, a simulator of ground movements in Fiumicino airport.

Introduction

The evaluation of safety critical systems is a composite and articulated activity. Thanks to the availability of advanced technological tools, operators can demand routine tasks to the system and concentrate on higher level mental operations. Therefore the activity of the operators evolves towards a flexible and context dependent process, where the knowledge that is daily produced is used to face new, incoming situations. Indeed critical situations are not due to the availability of certain information in the execution of a procedure but to the way in which different components in the process (software applications, organisational and cultural aspects, the physical layout, human operators) are balanced and interact to avoid or provoke breakdowns in the activity.

For this reason, the evaluation of complex safety critical systems requires an in-depth analysis of the socio-technical context of the work in order to assess the role that each component plays in the process.

In this paper we present an approach to the evaluation of safety critical systems that systematically assesses the distribution of different resources among process components and their interactions. Each of these resources is a stakeholder in the knowledge needed to carry out the process. However, this knowledge distribution must not be considered static: any process can be executed with a different allocation of resources. Dynamic and complex system environments.require flexible allocation of knowledge for process execution in order to deal with breakdowns resulting from the wrong distribution of resources among components.

The introduction of new technology that is going to be used to achieve better process performance, change work practices and may impact on social practises and influence the user's knowledge and behaviour. For that reason the design of new technological tools has to consider the work environment as an interwoven and interdependent system, as a socio-technical system, where any change of system resources impacts on the remaining ones.

In the paper we present a method, called Critical Interaction Analysis (CRIA), that permits the proactive assessment of which aspects of the system may impair or enhance safety after the introduction of new artefacts in the work setting.

The ShelAtc method

This method is inspired by the well known conceptual model, named SHEL, developed by Elwyn Edwards [Edwards, 1972]. The model describes the behaviour of interactive systems with special regard to human factors. SHEL is an acronym for Software, Hardware, Environment, and Liveware. Software refers not just to computer software but to the rules, procedures, practices that define the way in which the different components of the system interact among themselves and with the external environment. Hardware is used to refer to any physical and non-human component of the system such as vehicles, tools, manuals, signs and so on. Liveware refers to any human components of the system in the relational and communicational aspects. Environment refers to the socio-cultural and organisational environment in which the different components of the process interact.

The SHEL model concentrates on the interfaces between people and all system components including other liveware resources. The important point about SHEL is that it offers a systemic view where humans are not considered as isolated from the other system components. This view is consistent with recent theoretical work in cognitive psychology including Distributed Cognition [Hutchins, 1995] and Activity Theory [Nardi, 1996] but is grounded in simple concepts that can be understood by system designers without this theoretical background. In particular, Activity Theory assumes that human behaviour is not a set of disembodied cognitive acts (e.g. decision-making, classification, remembering). Rather, conscious activity takes place in everyday practices and it is inextricably embedded in a social matrix of which every person is an organic part. In this respect, the unit of analysis is wide and articulated. It consists of a subject (individual or group), an object or motive, artefacts (or tools) and socio-cultural rules and norms. Hence human activity should be considered as a socially and culturally organised ensemble where artefacts play a critical role in mediating human activity.

The SHEL model fits well with this theoretical framework since it considers any specific process as a combination of these three resources. This combination changes as soon as the process evolves and it is not exclusive in the sense that many combinations may occur during the process.

The method developed at the University of Siena has been applied in different contexts of safety critical systems, from requirement specification to incident analysis [Rizzo, et al., 2000], [Sujan, et al., 1999].

In this paper we describe a specific application of the method, that is, the safety assessment and the pro-active evaluation of the impact that new technological tools can have in the context of Air Traffic Control. The case study is an evaluation of MIDAS, a simulator of ground movements in Fiumicino airport, developed by Alenia Marconi Systems. MIDAS simulates new technological tools that may be operational in the Control Tower in the near future. These include a data link system, a stripless environment, a short-term conflict detection tool and monitoring aids.

Application of CRIA method

In the proposed approach, the first step is an in-depth analysis of the work system where the development of a new support system is intended. The activity analysis uses different counterbalancing techniques of data gathering ranging from observations, interviews, story telling, etc. It focuses on the weak points in the current system thereby eliciting safety issues and problems in the actual distribution of resources within the Shel components as well as their interaction.

The CRIA method develops in the following phases:

Preparation of test material:

- Identification of the basic Software (S), Hardware (H), Liveware (L) components.
  As stated above, our approach is based on a systemic and dynamic model: just a little change in one of the components of a working activity could significantly affect the mutual relationship between the components, their distribution, and even the process itself.
  For this reason, the first, fundamental step is to identify the basic SHL components that may affect the use of MIDAS: the ATC standard procedures (S) that the MIDAS prototype enables to carry out, and, for each of them, the new tools (H) and the professional figures (L) involved.
- Identification of safety issues.
  These are the issues that impact on safety in the context of the Fiumicino airport. Six main safety issues were identified: reliability, consistency and integration of the information needed to perform the activity; short and medium term planning; controller-pilot communication; management of the hand-over procedure; feedback on the accomplishment of tasks in progress.
- Scenario building.
  A set of scenarios was selected representing critical interactions between H, S, L components. Indeed, the evaluation does not aim to sequentially test each single procedure as standing alone, but to create a simulated realistic operational context, in which non-linear interactions among components could emerge.
  Scenarios were built through the following steps:

1- matching safety issues and MIDAS tools. We verified that the selected scenarios matched the identified safety issues and highlighted which MIDAS tools could impact on these safety issues.
2- implementing scenarios on the simulator. Our aim was to provide the evaluator with a situation as closer as possible to the original scenario observed in the Control Tower.
3- identifying SHEL components (H: all the tools implemented in MIDAS simulator needed to execute a task or a procedure; S: all procedures and the action sequences needed to the process development; L: the actors involved in the simulation, mainly Tower and Ground Controllers, plus the pilot and different pseudo-pilots).
4- envisioning interactions among components. For each scenario we tried to identify which interactions between the operator and the other system components (L, S, H) could be safety critical using the MIDAS tools.
5- structuring scenario for test sessions to plan a complete and meaningful test. Elements of the structure included: rationale, estimated temporal duration, actors, goal (the objective of the scenario that the evaluators had to reach), initial condition (status of the interface), operational context (meteorological and traffic conditions), MIDAS tools involved, other external supports available to the controllers.

- Preparation of the CRIA Question Table (SQT)
  The SQT is a list of questions related to the coupling of H, S, L components of the selected scenarios.
  For each components of the selected scenario we asked about the possible impact on the other components. The aim of the SQT is not to deeply analyse a specific component. On the contrary, the aim is to extend the analysis horizontally to the other systemic components and only after, if should be necessary, going into detail. In the early stage of the evaluation of new technology, focusing on the detail of a particular components could lose the view perspective of the interaction with the other component and, moreover, the understanding of the general process.
  Examples of CRIA questions are: L-L: " Does data link support communication between the Ground Controller and a pilot who has little familiarity with airport layout?"; L-S " Does the system support the Tower Controller in managing emergency procedures?"; L-H " Does the system provide all information that in the operational context come from different sources (windows, strips, radar)?".

Run the test

The simulation session were aimed at highlighting the strengths and weaknesses of some specific procedure carried out in a realistic tests as close as possible to the real operational situation. The simulation allowed us to control some variables otherwise difficult to evaluate, providing: feedback to refine the analysis; contextual elements that were not noticed during the analysis; empirical evidence of how the actors' mental models, consolidated from years of experience in managing ATC processes could be easily or with difficulty used to carry out the same processes using different tools, ways and procedures.

Each test was developed in three phases:

- Warming up

The session started with a brief explanation of the evaluation session, its objectives, and the schedule. Before starting the evaluation, the controllers were asked to familiarise themselves with MIDAS.

- Scenario execution

  The controllers received the scenario objective on paper. They were requested to think aloud in case the simulation involved only one controller. In test sessions involving two controllers, they could communicate with each other. In this way we collected data about: how a single controller interacted with the MIDAS interface; how are the interactions between controllers and the support that the MIDAS tools could provide in these situations.
- Post test: retrospective comments, focus group with the user

  After each scenario, the controllers and the designers were involved in a debriefing session based on the video recording of the test. The controllers were asked to freely comment their performance even if the designers drove the discussion towards the assessment of safety issues. At the end of the debriefing the SHEL Question Table was completed.

The data analysis

The data, which emerged from test sessions, controllers' retrospective comments and focus group were analysed to evaluate the impact of MIDAS tools on safety issues and to proactively evaluate the possible occurrence of new safety issues due to the introduction of these tools.
The findings of the analysis be provided in three different formats:

- Answers to the SHEL Question Table

  As stated above the SQT has been built up on the base of all possible matches between different components. This allows us to focus at every step on each of the tree components and their possible interactions. The SQT table constrains us to detect all the tree components at the same level.
- Assessment of safety issues

  This is the qualitative assessment of the Safety Issues detected during the activity analysis phase. They could be covered satisfactorily by MIDAS solutions, not covered at all or covered but requiring evaluation and eventual clearance during the later design stages.
- Synoptic view of critical interactions among SHEL components and pro-active safety assessment.

  It consists on a proactively evaluation of the possible emergency of new Safety Issues, since changing tools and procedure may cause new safety criticality, not predictable in the original context

Results

In the following we provide two examples of criticality we discovered applying the method. The examples show both real and estimated critical interactions related to the management of the landing and the hand-over procedures.

The landing procedure is one of the most critical processes in the control tower. During the test it was clear that in the execution of this procedure, a critical interaction occurs between the MIDAS simulator and the Tower controller. The controller has trouble in monitoring too many different information provided by the system. All the effort to integrate them is in charge of the controller who receives poor support from the system during this activity.
The transcription of the retrospective comment of the tower controller can help to clarify the problem.

"I didn't see in the approach window the third landing a/c since I focused my attention on the other a/c. I cannot monitor the approach window because it is separated from the rest of the view and because it displays arriving a/c too late. Even in the operational context I usually do not monitor the radar too much otherwise I cannot see other thing through the windows. In the standard approach procedure the pilot calls me saying "On final" so I immediately start managing the calling a/c. In this system I have to look for information about arriving a/c rather than waiting for this information as in the operational context. If there is high traffic in the apron, there is no way to monitor the approach window.
If we project this critical interaction to the development of the entire procedure, we can discover that MIDAS, in the current implementation, could put at risk also the interaction between the pilot and the tower (since the controller doesn't contact the pilot at the right time) and the interaction among the two controllers, since the hand-over procedure cannot be executed in safety conditions.

The second example is related to the hand-over procedures. This is implemented in MIDAS through the exchange of incoming or outgoing messages on the screen display of the Ground and Tower controllers. In the real operational context the same procedure is executed through a strip transfer between the two controllers.
The following is a description of what happened during the test.
The tower controller doesn't issue the take off to AZA 886 because he is managing a landing a/c. He doesn't notice that the Ground controller has already released the AZA 886, therefore this aircraft is under his responsibility. For this reason, the AZA 886 waited for two minutes before receiving *the take-off clearance. This delay had a negative impact on the whole departure sequence.*
During the debriefing session the controller commented as follows:
...I delayed the take off because I expected a "visible hand-over" from the Ground controller. I didn't check the incoming message window because I was busy...

The scenario and the controller's comments clearly show a problematic distribution of the information resources in the MIDAS simulator. The hand-over procedure is not executed under safety conditions because the system does not support a "visible" hand-over. In the operational context this procedure is executed through a strip transfer (both paper and electronic strips). These are dedicated tools that convey a specific meaning for the activity. In MIDAS there are no dedicated tools representing the hand-over: the transfer of responsibility is carried out by monitoring the incoming message window, the same are that is used for other procedures. This way to represent the hand-over is not effective, resulting in a lapse of attention by the controller. Furthermore, if we systematically project this critical interaction between L-H components (that is between the Ground controller and MIDAS) onto the other L-L components, we can foresee potential critical interactions also among the Ground and Tower

Controller and the pilot and Ground controller, even if this critical interaction has not been directly observed during the test.

The table below summarises the findings we obtained from the application of the CRIA method. Indeed the table show the safety problems relating to the introduction of the MIDAS tools in the current operational context and can be used to proactively assess the impact of these tools on the other components of the process that were not directly observed during the test scenario.

The table 1 is organised as follows:
in the first column we inserted the hardware and the liveware we identified in MIDAS, that is the pilots and the ground and tower controllers;
in the second row we put the software component, that is the procedures that, from the domain analysis, resulted in being safety critical;
the black lines represent the critical interactions we observed during the scenario simulation;
the dashed lines are the "estimated" critical interactions that can be inferred as a result of the method.

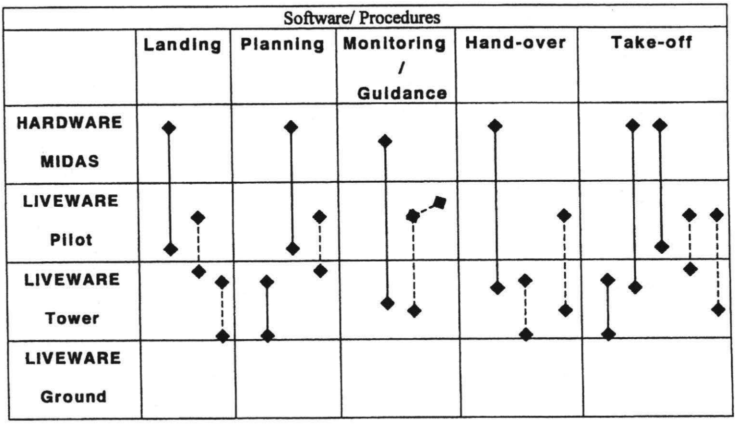


Table 1: Findings of the application of the CRIA method.

The table allows us not simply to assess existing safety problems, but to elicit the possible weakest interactions in the work setting. It allows and constrains us to foresee possible design solutions in term of these emergent situations. This mean that possible redesign solutions have to take account of the complex inter-dependencies between resources and the opportunities offered by their correct integration.

Conclusion

In conclusion, the methodology we presented was successfully tried out in different contexts of safety critical applications [Wimmer et al, 1999] [Rizzo et al., in press]. The particular application described in this paper enlarges its potential to the safety assessment of such systems including a pro-active evaluation of the impact of new technological tools in real operational settings. In particular the method offers the following advantages:

- it allows critical interactions about system components and to infer new ones to be detected;
- it allows limitation of scenarios that represent categories of single events to be overcome;
- it provides the knowledge necessary to specify requirements and re-design defects.

Indeed, the method clearly detects at what level the problem occurs and which interactions between system components should be redesigned to solve them.

References

Edwards, E., 1972, Man and machine: Systems for safety, Proceedings of British Airline Pilots Associations Technical Symposium (British Airline Pilots Associations, London), pp. 21-36.

Hutchins, E. (1995). *Cognition in the Wild.* MIT Press.

Nardi, B. (Ed) (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction* Cambridge: MIT Press.

Rizzo, A. Save L. (in press) SHELFS: A proactive method for managing safety issues. D. Embrey (Ed.) In Human Reliability. Paris: NATO Press.

Rizzo, A., Pasquini A., Di Nucci, P., Bagnara, S. (2000) SHELFS: Managing critical issues through experience feedback in railways. Human Factors and Ergonomics in Manufacturing, 10, 83-98.

Sujan, M. A., Pasquini, A. Rizzo, A. Scrivani, P. Wimmer, M. (1999). Activity Theory as a Framework for Considering Human Affect in the Design. Proceedings. of the IEEE Systems, Man, and Cybernetics Conference, Los Alamitos, Ca: IEEE Press, pp. 726-731.

Wimmer, M. Rizzo, A., Sujan. M.A. (1999). A holistic design concept to improve safety related control systems. Proceedings of the Safecomp '99. London: Springer & Verlag, pp. 297-309.

G.P.'S ACTIVITY ANALYSIS : A RETROSPECTIVE STUDY OF PATIENTS' FOLLOW-UP VIA PATIENTS' MEDICAL RECORDS ANALYSIS

B. Rajaonah*[1], F. Anceaux[1], M.C. Beuscart-Zéphir[2]

[1] CNRS-UVHC, LAMIH- PERCOTEC, Le Mont Houy, F- 59313 Valenciennes Cedex 9
bakorajaonah@wanadoo.fr
Francoise.Anceaux@univ-valenciennes.fr
[2] CERIM, Faculté de médecine, Université de Lille 1, Place Verdun, F- 59041 Lille Cedex
mcbeuscart@univ-lille2.fr

ABSTRACT

This research aims to analyze the General Practitioner's (G.P.) activity within the framework of dynamic situations : we consider the patient as a dynamic physiological process. We analyzed the contents of 6 patient's medical record followed for more than ten years, as well as the mails exchanged with the specialists. We coded theses contents according to the formalism in predicate and arguments by helping us of the DSM model [Hoc & Amalberti, 1995]. We shown that the activity of the G.P. aims at maintaining within acceptable limits various sub-processes evolving in interaction, and that this objective is carried out by 2 types of activity: the follow-up and the management of incidents. We shown that the current representation during the consultation was constituted by biomedical knowledge and practical experience. metaknowledge as well as medical and contextual data on antecedents and present situation.

KEYWORDS

Dynamic situations' management, medicine, current representation, analysis of activity

INTRODUCTION

In the research area of the management of dynamic situations, emergency medicine as well as anesthesiology are activities classically studied. As a dynamic situation, the General Practitioner (G.P.) activity has generated relatively few studies. In this paper, we propose to analyze the general practitioners'activity relying on a general cognitive architecture of dynamic situations management [DSM, Hoc & Amalberti, 1995, 2000]. We carry out a retrospective study of patients' follow-up through an analysis of the information contained in the patient's medical record. Part of the results were used for the development of an inter-mediation platform allowing bi-directional electronic communications between the G.P's Electronic Patient Record (EPR) and the Hospital Information System, but the results can also provide interesting recommendations for the design of usable and efficient EPR.

BACKGROUND

Standard researches on medical diagnosis were primarily focused on the study of medical reasoning and on the analysis of the knowledge underlying the diagnosis [Elstein, Shulman. & Sprafka, 1978; Larkin, McDermott, Simon, & Simon, 1980; Patel & Groen, 1986; Elstein, 1988; Boshuizen & Schmidt, 1992]. Most of this researches relies on academic case studies. Some more recent studies regard this medical activity as a process supervision activity in which the diagnosis is considered as a component of a broader decision-making activity [Boreham, Foster &, Mawer, 1992; Boreham, Mawer, & Foster, 1996; Hoc & Amalberti. 1995]. Within this framework, the G.P.'s objective is to maintain the parameters of the patient's health process and sub-processes within acceptable limits. In order to perform this task, the G.P. relies on a representation of the patient's health condition. This representation is particularized according to several types of information: previous medical data, current data and anticipated data. The diagnosis thus relies on a temporal window integrating past, present and future; this temporal window is permanently moving towards the future [Hoc, 1996]. Then. the medical diagnosis can precede the actual decision-making, but preservative actions can be engaged before any diagnosis has been set.

A previous work [Anceaux, Beuscart-Zéphir & Houzé, 1999] pointed out the role of the patient's record in the elaboration of the current representation of the patient's medical condition. This study outlined the general structure of the events and activities and issued a categorization of the situations the G.P. is confronted with. In the present research [Rajaonah. 2000], we focus on the episodes of care and on the cognitive activities implemented during these episodes.

MATERIAL AND METHODS

We worked on the medical records provided by Two GPs "professional experts" , with practicing experience of 10 years. Each of them, provided us with a copy of three anonymous patients' records. Each patient record represents a follow-up of 10 years and comprises the following: patient file, lab results, liaison, consultation and discharge letters. The patient file is partly a paper record (from 1989 to 1997) and partly electronic (GPs were equipped with a CEPR two years before the study). Those data were completed with interviews of the two GPs focusing on their representations of those patients.

The information contained on the patients' files was coded following the predicate and arguments structure suggested by Hoc and Amalberti [1999, 2000]. This method leads to identify the three main classes of cognitive activity described in the DSM model [Hoc & Amalberti, 1995, 2000], which are information elaboration, diagnosis and prognosis, decision-making (integrating feedback evaluation).

RESULTS AND DISCUSSION

The analysis of patients' files and interviews demonstrates that the current representation is constituted of two classes of knowledge: a knowledge base, including academic, procedural and metacognitive components; a fact base, specific for each patient including medical as well as contextual information.

We performed a longitudinal analysis of the patients care and follow-up. Then relying on the categorization of situations suggested by Anceaux *& al.* [1999], we can describe an alternation of two types of episodes:

- **Follow-up episodes** during which the patient health process is balanced. During these episodes, the representation of the patient's medical condition is stabilized and can be last over a more or less long duration punctuated by consultations. The recorded classes of activity are part of the follow-up process : (i) there are few diagnosis activities, (ii) information elaboration activities are restrained mainly to the gathering of specific information devoted to standard follow-up (blood pressure, auscultation, biological analysis results) and (iii) the actions set and performed aim mostly at the maintenance of the health process balance (e.g. via the prescription of the usual treatment). During these types of episode, the activity seems to be regulated on two levels: (1) by the short term loop of automatic control consisting in the activation of schemas like information gathering scenario or symptomatic treatment prescription, and (2) by the medium term loop which initiates an adjustment of the current representation allowing a better understanding of the situation. This adjustment of the representation may rely for example on a prescription of a complementary exam to be performed by a specialist physician in order to confirm an hypothesis. These two loops may function in parallel.
- **Incidents management episodes** indicating a breakdown of the balance, which is to be restored by the physician. Along those types of episodes, information elaboration activities are performed, that aim at understanding the incident and which relies on the elaboration of a new representation of the patient's health process. Diagnosis activities (identification, hypothesis generation and testing) appear along with non routines actions (e.g. information requested from specialists or prescription of a new treatment). During these episodes, the long term control loop is activated, in parallel with the others control loops. It allows the G.P. to seek for new elements of comprehension in his knowledge base, leading to a deep modification of the current representation of the actual patient's medical condition.

CONCLUSIONS

This study identified two types of episodes in the GP's patient care activity : (a) standard follow-up episodes relying on stabilized representation of the patients conditions. During these episodes, information gathering and elaboration activities aim at determining the necessity of amending the health process balance and (b) incidents management episodes during which the representation needs to be adjusted, or rebuilt. Along these episodes, information elaboration aims at setting a diagnosis and restoring the health process balance.

A second significant finding concerns the critical role of the cooperative activities involved in the G.P.'s patient care activity. At sometimes during the standard follow-up episodes. and quite systematically during the incident management episodes, the GP calls in the specialists aid in order to check the balance of physiological parameters or to get an help to set the diagnosis and prescribe a new treatment.

Part of these results and findings have already been used to support the design of an intermediation platform which supports medical exchanges between health professionals. In the long term, the patient could be integrated in this network and participate in the communication process concerning his own healthcare.

REFERENCES

Anceaux, F., Beuscart-Zéphir, M.C., & Houzé, P. (1999). Coopération Homme-Homme médiatisée par la machine : coopération asynchrone entre les médecins généralistes et les différents partenaires de santé. Communication présentée au *Séminaire Psy Ergo Hainaut*. Valenciennes, France. Déc.

Boreham, N.C., Forster, R.W. & Mawer, G.E. (1992). Strategies and knowledge in the control of the symptoms of a chronic illness. *Le Travail Humain*, 55, 15-34.

Boreham, N.C., Mawer G.E., & Foster, R.W. (1996). Medical diagnosis from circumstantial evidence. *Le Travail Humain, 59*, 69-96.

Boshuizen, P.A. & Schmidt, H.G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science* 16, 153-184.

Elstein, A.S. (1988). Cognitive processes in clinical inference and decision making. In D.C. Turk. & P. Salovey (Eds). *Reasoning, inference & judgement in clinical psychology*. New-York : Free Press / Macmillan.

Elstein, A.S., Shulman, L.S., & Sprafka, S.A. (1978*). Medical Problem Solving. An analysis of clinical reasoning*. Cambridge, MA : Harvard University Press.

Groen, G.J., & Patel, V.L. (1988). The relationship between comprehension and reasoning in medical expertise. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise* (pp. 287-310). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hoc, J.M. (1996). *Supervision et contrôle de processus: la cognition en situation dynamique*. Grenoble. France: Presses Universitaires de Grenoble.

Hoc, J.M., & Amalberti, R. (1995). Diagnosis: some theoretical questions raised by applied research. *Current Psychology of Cognition, 14*, 73-100.

Hoc J.M., & Amalberti. R. (1999). Analyse des activités cognitives en situation dynamique: d'un cadre théorique à une méthode. *Le Travail Humain, 62*, 97-129.

Hoc J.M., & Amalberti. R. (2000). Modelling NDM cognitive activities in dynamic situations: the rôle of a coding scheme. *Proceedings of the NDM5 meeting*. Tammsvik, Sweden, May. CDROM.

Larkin, J., McDermott, J. , Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335-1342.

Patel, V.L., & Groen, G.J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science, 10*, 91-116.

Rajaonah, B. (2000). *Analyse de l'activité du médecin généraliste*. Mémoire de maîtrise. Université Charles de Gaulle, Lille.

Raufaste, E., Eyrolle, H., & Mariné, C. (1998). Pertinence generation in radiological diagnosis: spreading activation and the nature of expertise. *Cognitive Science, 22*, 517-546.

A methodology for analyzing the dynamic collective organization of nuclear power plant operators in simulated accidental situations

Jacques Theureau
Centre National de la Recherche Scientifique / Université de Technologie de Compiègne,
Compiègne, France : jacques.theureau@hds.utc.fr)
Geneviève Filippi, Geneviève Saliou
Electricité De France / Recherche & Développement / Etudes de Sûreté & de Fiabilité / Groupe
Facteurs Humains, Clamart, France : genevieve.filippi@edf.fr, genevieve.saliou@edf.fr
Pierre Vermersch
Centre National de la Recherche Scientifique, Paris, France : pvermers@es-conseil.fr

This research deals with modelling of emergency operation of a nuclear power plant involving hard-copy instructions distributed among different operators. Emergency operation is seen here as *experienced, embodied, dynamically situated (including socially), indissolubly individual and collective, and cultural.* Based on preliminary modelling of significant elementary units (SEU), sequences and macrosequences of the courses of experience of various agents, and on the collective interaction of those SEUs and sequences, two sorts of progress were made : the first, which lies just slightly lower than the SEU, concerns modelling of the *dynamics of attention windows*; the second concerns the *dynamics of openings*, i.e. the diachronic and synchronic relationships between the courses of experience of the different agents.

Keywords : Situated action, Modelling individual & collective action, Emergency operation, Computer Supported Cooperative Work.

Introduction

This ongoing research is part of a long-term dialectic between ergonomics research and ergonomics practice undertaken some time ago by the two teams (see, for example, [2] and [3], and [11]). It looks at emergency operation of a nuclear power plant in which hard-copy instructions are shared out among the reactor and water-steam operators, the supervisor, the operations manager, and the safety engineer, and is based on full-scale simulator tests. The five agents—who may be joined by auxiliary operators—work in a space divided into functional zones, written instructions in hand. Audio-video recordings of all control-room activity and of very short self-confrontations of the reactor operator and supervisor were systematically analysed for two 150-minute tests chosen for their common complexity and their differences. The objective of the research is to make progress in modelling of this activity.

1. Theoretical and methodological framework

The theoretical and methodological framework of this research is that of *cognition seen as experienced embodied, dynamically situated (including socially), indissolubly individual and collective, and cultural* :

* *experienced cognition* : this is what some authors call "the explanatory gap of cognitive sciences", i.e. absence from the mainstream—so far—of a description and explanation of actors' access to emotional, sensory, and cognitive processes, which obviously does not imply that such access is total or even correct;

* *embodied cognition* : cognition, decision, action, communication, but also emotion, attention, etc.—taking these terms in their common meanings—are inseparable;

* *dynamically situated cognition (including socially)* : cognitive phenomena are not situated merely in the head, or even merely in the body, but also in the relationship between corporal dynamics and situational dynamics;

* *indissolubly individual & collective cognition* : individual cognition includes a relationship with other agents, including through the traces they leave in the situation, and collective cognition cannot be described and explained without considering its relationship with individual cognition;

* *cultural cognition* : this culture is more or less shared by the agents.

This theoretical and methodological framework can be seen as a development of what was implied in the notion of situated action. It may be remembered that the theme of situated action was brought into the public domain in 1987, with the book by Lucy Suchman [5]. Since then a variety of authors (see [13] in particular) have reduced it to i) a methodological innovation, the study of cognition in the field, and ii) a theoretical criticism of the notion of plan and of the reduction of man-machine interaction to carrying out procedures or instructions. Now, if indeed there is any methodological innovation, it is very relative, since methodologies for studying cognition in the field were developed a long time ago, particularly in the work analysis developed by the French-language ergonomics movement. And if indeed there is any theoretical criticism, it is far from being the essence of the theoretical proposals made. For as an alternative to cognitivism, i.e. to the paradigm of "man as an information-processing system" as formulated in [4] for example, Suchman proposes considering "lively, moment-by-moment assessment of the significance of particular circumstances" and "to explore the relation of knowledge and action to the particular circumstances in which knowing and action invariably occur" ([5], p. 178). This involves not just describing situated action, but also proposing situated action/cognition as a strong theoretical hypothesis. This strong hypothesis has critical and constructive aspects. The well-known theoretical critical aspect is that human actions are far from being generated and controlled from start to finish by plans, or in other words, by internal representations specifying in full the different steps in the performance of human actions, whether these internal representations are produced directly by the agents or are the result of their reading procedures or instructions. The theoretical constructive aspect is that plans and procedures or instructions are just some of the resources available for carrying out actions. The fundamental point to be considered is not so much the problem of the plan or of the procedure or instruction as the idea that action/cognition calls on other resources, i.e. the material, social, and cultural characteristics of the environment in which events occur and which constitute the situation of the agent(s). As these characteristics can change at any time, to adapt to them individuals adjust their actions to the new environmental circumstances. This adjustment is done on an improvised *ad hoc* basis. Seeing situated action/cognition as a "strong hypothesis" thus implies that cognition is "situated", including in laboratory experiments, and that here too, research procedures taking account of this situated nature must be developed.

While one of the major contributions of the situated action/cognition movement is that it highlighted the opportunistic and improvised nature of human action and co-operation, together with its material and socio-cultural anchoring, this contribution has been limited to a scientific practice with sparse interest in systematic modelling of the phenomena studied, and it has thematized essentially material anchoring of cognition, to the detriment of corporeal and cultural anchoring. In addition, while calling on the accounts of the agents concerned, it has not thematized the fact that the agents too experienced personally these accounts and the corresponding actions carried out. To go beyond these limits and, in so doing, raise things to the level of scientific requirement formulated and practised by Herbert Simon in connection with an entirely different paradigm (refer in particular to the introduction to [4]), it is necessary to develop a "harder" phenomenology, i.e. one directed towards a systematic modelling of the greatest possible wealth of empirical data. This amounts to : defining an adequate paradigm for human cognition (versus the paradigm of "man as an information-processing system") and the corresponding theoretical object(s); defining an observatory for these theoretical objects at the level of the "minimum theory" of data collecting on human cognition, as specified by Ericcson & Simon in "Protocol analysis" [1]; defining a phenomenology or analytical model of situated action/cognition, i.e. a theoretically coherent set of descriptive categories of human experience, of experienced cognition, or, in other words, a generic analytical model of it (as compared to the descriptive categories of "problem-solving graphs", those of "information status" and of "information-processing operators"); defining a generic synthetic model of situated action/cognition (as compared to today's conventional "cognitive modelling" complying with the paradigm of "man as an information-processing system"); defining a method for designing practical models, i.e. models intended to guide the designers of technical spaces and tools in the control room, of instructions and their supports, and of the training and organisation of agents.

2. Principles of the modelling carried out

This research takes as granted the paradigm of living systems defined by Maturana & Varela (see [12], for example), along with the theoretical objects and observatories defined by [7], [9], and [14], and is restricted firstly to construction of an empirical analytical model dealing with the activity of agents as they experience it at any time, and more precisely with what we call here *collective interaction of courses of experience*. The research makes it possible to extend, enrich, and more effectively validate the analytical and synthetical comments that can be made : empirical comments on the production of agents' experience at any time, based on the characteristics of their status (corporeal anchoring), their situation (material and social anchoring), and their culture (cultural anchoring), pending analytical and synthetic modelling of that production; ergonomics comments on the possible consequences in terms of the design of instructions, interfaces, and organisation.

The research picks up from work done previously by the EDF team (see [2] and [3], and, more remotely, [9] and [10]) : construction of recording, observation, and verbalization data from emergency-operation tests (verbalization = self-confrontation of the main operators); transcription of each test (lasting between 120 and 150 minutes); special transcription of data concerning the reactor operator and the supervisor (which gives rise to different theoretically-based decisions, in particular with respect to the description of the actions performed); reconstitution of the tracking of instructions based on pages of the hard-copy emergency-operation instructions used; salient events noticed by observers; preliminary analytical modelling in terms of significant elementary units (SUE), sequences, and macrosequences of each agent, and in terms of the collective interaction of those significant elementary units, sequences, and macrosequences; preliminary series of empirical and ergonomics comments arising out of the preliminary modelling.

The essential principle of the preliminary analytical modelling and its extension under the present research project is to determine—retrospectively—, from a given moment in the activity of agents, which periods of that activity are perceived as a unit by the agents, and what sort of relationships the agents establish between those periods. The units and relationships set up at a given moment may last throughout the rest of the activity, or they may be called into question. The relationships between units build units of a higher order than the previous ones. These higher-order units can be continuous or discontinuous.

Some of these units and relationships can be elicited by the agents themselves, through their interviews and self-confrontation. Others can be inferred by analysts, by comparing with other parts of the test transcripts (which include continuous observation and recording of the behaviour of agents and various forms of verbalization by those agents, particularly self-confrontation), but also—subject to some precautions—by means of the experience of analysts and their knowledge of the training the agents are given, the tasks they perform, and the tools they use. This modelling produces a graph of the concatenation, imbrication, and partial overlapping and embedding of units. The units are themselves categorised in terms of significant structures, by considering the different sorts of relationships between the significant lower-order units. The significant structures that the previous studies determined by comparing a theoretical approach and an inductive approach are different kinds of elementary structures, diachronic structures (sequences of different orders, series of different orders), and synchronic structures, i.e. structures that are indifferent to temporal succession (synchronous structures of different orders), the definition of which can be found in [9].

The starting point for modelling the collective interaction of the courses of experience of certain agents with an essential role in a given collective situation is the models for each of the agents concerned, i.e. the graphs of concatenation, imbrication, partial overlapping and embedding of significant units of their activity. The modelling results in several parallel graphs and allows the relationships between the units of which the graphs are made to be determined.

The heuristic value of this modelling of courses of experience and of their collective interaction lies in : its utility in terms of the intelligibility of the activity of the agents in the units and relationships determined; the possibility it offers for close comparison between different transcripts and fragments of transcripts; the possibility if offers for inductive construction and specification of generic categories with respect to the activity of agents; the effective constraints in the activity of agents that can be detected from it, together with the categories of constraints that can be defined, these constraints concerning the current state of agents, their material situation (particularly the content and presentation of instructions, the type and layout of indicators in the control room, etc.), organisational situation (the roles of different agents, their possibilities for communication, etc.), and cultural situation (commonness and differences in training, experience, etc.).

The first modelling of SEUs, sequences and macrosequences, and their collective interaction makes it possible to specify : (1) the string and embedding of dynamic contexts of situated interpretation of instructions by the supervisor and reactor operator, and their constraints and relationships with the demands of the process; (2) the co-operation between the supervisor and the reactor operator and between them and the other agents, its constraints and effects; (3) information concerning the competencies of the supervisor and reactor operator, and the effective, in-situation mobilisation of those competencies, together with the constraints, effects, and compatibility with the process of that effective mobilisation.

To extend this first modelling, the SEUs determined from the following criteria are taken as granted : continuum of perception-action of situated interpretation/following of instructions; period of stoppage on an instruction or set of instructions; continuum of communication with other agents regarding a given topic; periods with varied contents in which the agent departs from the instructions. On the other hand, we consider the resulting sequences and macrosequences as being simply a first step in the revelation of more complex structures, both diachronic and synchronic.

This research thus makes it possible to make two kinds of modelling progress : one—just slightly below the level of the SEU—concerns modelling of the *dynamics of attention windows*; the second concerns the *dynamics of openings*, revealing the diachronic and synchronic relationships between SEUs, sequences, and macrosequences of the courses of experience of different agents. This progress in modelling is developed iteratively, as was done in the case of the previous modelling : determination of the different kinds of links each SEU has with others, i.e. the openings of different types; determination of attention windows and their role in the local development of courses of experience; analysis of the series of different types constituted by the links between SEUs; analysis of the different kinds of synchronic links between these series, as seen from the point of view of each agent; formulation of the empirical and technological gains obtained by means of this modelling. Such a process of modelling is not monotonic : at each step, one has to look back to the previous steps to detail and refine their results, or even to call how they were done into question (which, with respect to the construction of data, can obviously only have an effect on the next study or stage of research).

3. Score reading, disturbance, and dynamics of attention windows

Let us look first at the dynamics of attention windows. Whenever information is acquired by reading signs (as opposed to simply identifying presence/absence or threshold overruns, which are a matter of indices or signals), attention has to be focused. In other words, there is a moment when the agent can pay attention to a single thing, when he momentarily excludes other information from his field of conscience, when he sets his mind to taking in the meaning of the information read, something he can do only if he does only that. It is therefore reasonable to assume that there is a strong relationship between reading activity (reading of hard-copy, but also reading of screen displays or plots) and a temporary mind-set in which the field of attention is focused on a single thing, temporarily inhibiting and excluding everything else. Now it is precisely this sort of reading activity which dominates emergency-operation activity driven by procedures. In general terms (with substantial variations depending on their role), agents read text, move to a different location in the control room, adjust controls, communicate, and wait. But first of all, they read and read : main procedures, auxiliary procedures, etc. This reading can in fact be called "score reading" : each item or set of information read (an individual instruction, a test, etc.) on hard copy corresponds to an action to be carried out (go get information to document an instruction) or a test to be done (change documents and open another, communicate information, phone another agent, carry out a control action, make an adjustment, etc.). The notion of "score" is that of sheet music, where each sign is meant to produce a determined action : play a certain note, for a certain time, with certain alterations, with a certain expression, etc. Taking the analogy a little further, a music score is the transcription of the result to be achieved, the notation of the sound expected. However, this notation is but a poor reflection of what the player has to do to get that result, for in order to play the notes correctly, he has to interpret the written music, correct it, and complement it to make it an effectively adapted guide for action. A beginner is incapable of knowing exactly how to play the music on the paper in front of him once it starts getting complex, which hand plays what, for example (on piano). Similarly, if operation instructions are the end result of a whole process of capitalization of knowledge and experience, agents have to add a good dose of knowledge of their "scores" in order to apply them correctly. It is a matter of expertise, constitution of a procedure-reading habitus which appears to be broadly underestimated, yet which is foremost among the preoccupations of simulator-training supervisors who see it as a central requirement.

What must be stressed is that there are constant changes in focus/mind-set. A line is read, an instruction is taken in. To do this, the agent has to discern precisely what he perceives, so he restricts his field of visual perception. In most cases this reading leads both semantically and spatially to another instruction, but also to movement of the agent to another point in the control room, and to another kind of reading—as occurs when the agent reads a value off a display—, or it can take him to another document which must be extracted from its classification system and thumbed through until the right sheet is found. There is then a new focus of attention, etc., one characteristic of which is that sooner or later the agent will go back to the main document that he set down previously, and pick up again precisely where he left off, so as to ensure the imperative of continuity of his sequential reading. To these changes in focus which must be managed by the agents' working memory are added interruptions which can cause them to lose the thread they are following. These interruptions can be diversions of occupation (and therefore of focus) : while the agent is proceeding with an adjustment, something extrinsic to that activity interrupts him and requires him to suspend his current occupation and turn to something else. More locally, these interruptions can be changes in focus : while the agent is reading off a series of values on a screen, the phone rings, or another agent needs an answer to a question; the agent responds quickly to the solicitation and immediately returns to his instructions. For example, just as the reactor operator starts implementing the Orientation and Stabilization Document (OSD), the operations manager interrupts to ask if he has called the safety engineer. "No. Do it, will you, Colin", says the reactor operator. The reactor operator has been momentarily interrupted, but it is clear that he does not need to reflect on the matter or take in new information to be able to answer the question, but nevertheless he does more than just answer since he delegates performance of the task to the person who asked the question. Which implies that he gets an answer back since his own answer contains a question. There is indeed an interruption, but all the conditions point to there being no change in occupation, merely a momentary 'blip' authorised by the fact that the question is oral and consistent with the reading in progress, and because the activity required to answer does not require a new occupation that would compete with that already engaged. In addition, it can be assumed that the OSD is sufficiently fragmentary for it to accommodate simple interruptions. Equally, it can easily be imagined that in certain activities requiring closer attention, the simple fact of being addressed in this manner could more or less seriously affect one's state of concentration and compromise the efficacy of the activity in progress.

These interruptions, and particularly interruptions to occupations, are potentially sources of errors when agents return to pick up an activity where they left off. For instance, a phone rings during a basic-cycle phase, just when the reactor operator is documenting a test from the readout on a screen. The agent decides to interrupt this phase of work in progress, i.e. without completing it and mentally "bookmarking" his instructions. When he comes back to it, he picks up at the phase of work interrupted, but at the wrong place in the procedure sheet. Interruptions such as this require agents to perform additional marking and verification tasks in order to ensure the continuity of their activity. For example, during the same phase, the supervisor asks for information while the reactor operator is reading off values from plotters. The operator does not reply immediately; first, he finishes his readings, then goes back to his procedure, and finds that he has to go to a new page; he turns to that page and only then turns to the supervisor to answer. He did not take the risk of interrupting the continuity of application of the instructions before reaching a stable and easily identifiable point. Conversely, agents whose activity requires them to interrupt the activity of other agents develop an additional activity of following the other agents' activity and controlling the interruptions they have to provoke. For example, the reactor operator tells the supervisor the conclusions reached from implementation of the OSD, as he is required to do by instructions. He gets the ECP1 (Reactor Control Status 1) instructions ready and, without opening them, says to the supervisor "OK, you can run your loop". In the ensuing period, he sits back, holding the instructions closed under his arm, says nothing, and makes no verbal or non-verbal communication other than a general sign of withdrawal. When the supervisor has reached his own conclusions—the same as the operator's: follow the ECT1 procedure (the one corresponding for him to ECP1)—he confirms "Yeah", and, in a different register, "Here we go" to signal that he is starting to implement the instructions. These two sorts of additional activity dovetail together.

4. Control with distributed instructions and synchronic management of openings

Another property of the course of action of each agent is its opening to a more or less indeterminate future. This is why we detail the common-sense notion of occupation by introducing that of *open action*, or, put more simply, *opening*. An elementary action can be fully completed : thus, looking at the simplest example, the agent makes phone call, gets hold of the right person, and gives his message : "You're wanted in the control room". In this case, once he has hung up, the operation has been carried out and completed. On the contrary, if the agent makes a call and cannot speak the right person, he leaves a message asking to be called back. In this case, when he hangs up, he creates an opening or, in other words, an action which has not been completed, which remains open to a future end. The same can apply in the first case too if there are other contingencies accompanying that of arrival of the person called, such as briefing him on the situation.

In fact, this notion of opening is a very general one. Its relevance extends well beyond cases like this. As soon as a test begins, an opening is created for each operator and experienced by him : operation under normal circumstances that will be turned into emergency operation, in one way or another. As soon as any operator gets involved in an emergency procedure, an opening is created : the situated following/situated interpretation of the instructions, until it has been successfully accomplished or until the evolution of the process means he changes procedure. The notion of opening even brings us to a set of even more general hypotheses on human activity, synthesised by the notion of *hexadic sign* (see [8]), which cannot be detailed here.

Openings (occupations) and attention focus should be considered jointly. Let us look at an opening. The moment at which an opening is created, progresses, or closes may fit into the sequence of action managed coherently by the agent; it may also occur out of the blue at any moment, while an attention focus is going on. Consequently—to take up the previous example again—the moment the person calls back, having been given the message, may be precisely when the operator is recording information to document the response to an instruction, causing him to interrupt his reading to answer the phone. There is a break in the focus of attention, and once the opening has been closed (or has progressed)—i.e. the phone has been answered—, the agent has to pick up again exactly where he left off, or just afterwards, having often moved several metres from the panel to the telephone. The structure of an opening therefore leads towards potentially inopportune closure—or progression—; in all cases, whatever happens, or when it happens, is beyond the immediate control of the agent who initiated the opening. The graph of the reactor operator's and supervisor's openings during a test allows for analysis of the interruptions in focus of attention and of their diachronic and synchronic management of openings. It also serves as a base map for closer local analysis of the dynamics of attention windows generally.

The openings in the activity of an agent at any time build up. The time between creation and closure of an opening can engender a more or less nagging sense of worry for the agent who continues to manage the rest of what the instructions require him to do. The number of openings in place simultaneously and the agent's skill at managing them synchronically are part of conditions which can diminish vigilance, induce moments of confusion, and lead to distraction. None of these conditions alone can do this, but together they may. For example, in the reactor operator's and supervisor's analysis graph of openings for one of the tests systematically analysed, about an hour after the start the reactor operator is simultaneously managing situated following/interpretation of the ECP2 instruction, in the knowledge that the emergency continuum will eventually require the next set of instructions up, ECP3 (following a previous message from the supervisor), while carrying out the instructions of an auxiliary instruction sheet (RFLE58), and at the same time the supervisor is simultaneously managing situated following/interpretation of the corresponding ECT2 instruction, in the knowledge that the emergency continuum will eventually require the next set of instructions up, ECT3 (following a previous message from the operations manager), is looking through the instructions regarding the criteria for changing to ECT3, is examining and carrying out actions concerning the state of a particular system (ASG Auxiliary Feedwater), and is also awaiting the results of an *in situ* inspection of the system by an auxiliary operator.

While modelling of the changes in focus of attention and opening, filling, and closure of openings concerns the course of action of individual agents, i.e. that part of their individual-social action which is significant for them, it in fact points towards the collective activity of the entire control team, for these events are the result not only of the specific activity of each agent, but also of the relationships each agent has with activities of other agents. Each agent is supposed to follow his instructions (ECP for the reactor operator, ECT for the supervisor) without giving a thought to the others, except at special predetermined times when either he has to give some of the others information about the system controlled or the point he is at in his instructions, or he and the others have to change instructions. We saw an example of the latter case above, concerning the change from OSD to ECP1. It can be seen that in fact the agents interact with each other well beyond these particular moments. They observe each other, organise their access to resources (operations sheets and logs, for example) and their respective instructions to auxiliary agents, express their feelings, co-ordinate with each other, wait for each other *extra*-instructions, exchange information and diagnoses/prognoses on the dynamics of the system, criticise each other's actions and movement through the instructions, and sometimes even put their minds together to collectively solve problems. Most of these events are not "meta-functional", i.e. they do not concern just the construction and maintenance of the particular social relationship of emergency operation, but comply with the constraints of operation with distributed instructions. For example, through the instructions he follows, the supervisor checks the operations carried out by the reactor and water-steam operators. If he is ahead of them, nothing prevents him reading even further ahead, but he will have to go back over the instructions to actually carry out the checks. If he is too far behind them, his verification will be too late. So to the above description of the simultaneous openings of the reactor operator and supervisor respectively one hour after the start of the test must be added openings concerning what each of them thinks of the work of the other : one of the reactor operator's opening is to not disturb the supervisor during his laborious search for the criteria for changing to ECT3, and to get through ECP2 as quickly as possible and find a way to ECP3 that would overcome the problem; one of the supervisor's opening is to wait for the reactor operator to change to ECP3. Parallel modelling of the activities of the reactor operator and supervisor and of the interactions between them thus gives a vision of the collective activity as it is co-constructed by the two agents and their particular dynamic situations, including other agents. The procedures—including auxiliary procedures—sheets, logs, and the entire control room are tools for capitalising on and managing the knowledge which contributes more or less successfully to this co-construction.

Conclusion

This progress in analytical modelling helps detail and create a better foundation for the empirical and ergonomics comments made previously, and also engenders new empirical and ergonomics comments. Some of these comments arise directly out of the model. Others arise indirectly, and are the starting points for subsequent development of the model. In particular : (1) this research reveals the *real requirements for competence* of agents, something that goes well beyond simply following written instructions, particularly those requirements concerning *extra*-instruction communication with other agents, including those concerning adjustment of the timing of actions by different agents and those concerning the precautions to be taken with respect to disturbing other agents and how to deal with being disturbed by them; (2) it allows for analysis of *the genesis of errors and their contributing factors*, in the written style of the instructions and in variations in that style, in the quality and multiplicity of simultaneous openings, in interruptions in focus, and not just in "human weaknesses"; (3) it introduces questions about *the roles allocated to the different agents*, and in particular to the supervisor who must, on one hand, follow instructions for checking the actions of the reactor and water-steam operators, and, on the other, in dealing with the operations manager, step back from the instructions and consider the evolution of the process, and who must also carry out certain operations detailed in auxiliary procedures that the two operators are not able to carry out within a reasonable time. A further appreciable aspect of the modelling carried out is the analysis graph which enriches the available *resources for dialogue with designers.*

In sum, research such as this reveals the interest of a long-term dialectic between ergonomic research and ergonomics practice : the former makes for development of the latter, and the latter provides a starting analysis and questions for the former, the two being performed at once by official researchers and official practitioners.

References

[1] Ericsson K.A., Simon H. (1984) *Protocol Analysis. Verbal reports as data,* MIT Press, Cambridge.

[2] Filippi G. & Saliou G. (1997) *Essais de mise en situation recréée : orientations méthodologiques pour des analyses approfondies,* EDF/DER/ESF, HT - 54/97/006/A.

[3] Filippi G. & Saliou G. (1998) Data collection and analysis to assess operators cognitive activity in simulated control rooms, 7th IFAC/IFIP/IFORS/IEA symposium on "Analysis, design and evaluation of man-machine systems", 16-18 Sept., Kyoto, Japan.

[4] Newell A. & Simon H.A. (1972) *Human problem solving,* Prentice Hall, Englewood Cliffs.

[5] Suchman L. (1987a) *Plans and situated actions: the problem of human/machine communication,* Cambridge University Press.

[6] Suchman L. (1987b) Critique, in Understanding computers and cognition : Four book reviews and a response, *Artificial Intelligence,* 4-3702/87, 213-261.

[7] Theureau J. (2000) Anthropologie cognitive & analyse des compétences, in J.M. Barbier, Y. Clot, F. Dubet, O. Galatanu, M. Legrand, J. Leplat, M. Maillebouis, J.L. Petit, L. Quéré, J. Theureau, L. Thévenot, P. Vermersch, *L'analyse de la singularité de l'action,* collection Education & Formation, PUF, Paris, 171-211.

[8] Theureau J. (1992) *Le cours d'action : analyse sémio-logique,* Peter Lang, Berne, Suisse.

[9] Theureau J., Jeffroy F. & coll. (1994) *Ergonomie des situations informatisées : la conception centrée sur le cours d'action,* Octares, Toulouse.

[10] Theureau J., Filippi G. (2000) Analysing cooperative work in an urban traffic control room for the design of a coordination support system, chapter 4, in P. Luff, J. Hindmarsh & C. Heath eds., *Workplace studies,* Cambridge Univ. Press, 68-91.

[11] Theureau J., Jeffroy F., Vermersch P. (2000) Controling a nuclear reactor in accidental situations with symptom-based computerized procedures : a semiological & phenomenological analysis, CSEPC 2000, 22-25 Nov., Taejon, Korea.

[12] Varela F. (1996) *Invitation aux sciences cognitives,* Seuil, Paris.

[13] Vera A.H & Simon H.A. (1993) Situated action : a symbolic interpretation, *Cognitive Science,* 17, 7-48.

[14] Vermersch P. (1994) *L'entretien d'explicitation,* Editions ESF, Paris.

User Interface

An Experimental Study on an Adaptive CAI System for Training of Diagnosing Nuclear Power Plant Anomalies

Takahisa OZAWA[1]*, Hiroshi SHIMODA[1], Hidekazu YOSHIKAWA[1]
Yuji NIWA[2] and Erik HOLLNAGEL[3]
1: Graduate School of Energy Science, Kyoto University
Gokasho, Uji-shi, Kyoto 611-0011, JAPAN
2: Institute of Nuclear Safety System, Incorporated
64 Sata, Mikata-cho, Fukui 919-12, JAPAN
3: Department of Computer and Information Science, University of Linköpings
SE-581, 83, Linköping, SWEDEN
ozawa@uji.energy.kyoto-u.ac.jp

Abstract

The training system for diagnosing nuclear power plant anomaly, named Instruction Simulation System (ISS), has been developed. The ISS is a specially developed experimental environment which aims at deriving effective educational strategy utilized in the full-computerized adaptive CAI system. The ISS has several sub-systems and Eye-Sensing Head-Mounted Display is one of major sub-systems, which can detect trainee's eye gaze point in real time. An experiment has been conducted by using ISS with the participation of one instructor and several students, where the expert on nuclear power plant operation was actually employed as the instructor, to detect error patterns of trainees and to observe his way of instruction to reduce effective teaching strategy for adaptive CAI.

1. Introduction

Large and complex machine systems, such as nuclear power plant and airplane, must keep safety and stability from the social and economical points of view. In such machine systems, the role of human is very important in the both operational and maintenance periods. And, the people, who are engaged in such systems, are required to be highly trained. However, training of skilled professionals needs high cost with long training period. Addressing this issue, the authors have constructed an adaptive computer-assisted instruction (CAI) system which aims at the efficient and effective training of knowledge and skill necessary for diagnosing nuclear power plant anomalies.

The adaptive CAI system will have the following functional features.

(1) On-line measurements of human information activities: The system can detect a trainee's eye gaze point and recognize the trainee's think aloud in real time.
(2) The system can estimate the trainee's thinking process (reasoning path) in real time.
(3) Adaptive interaction to individual trainee in accordance with the estimated internal states: The system can provide educational critique by detecting the trainee's errors in real time.

By possessing these features, the system will cope with various states and behaviors of trainees (e.g. skill level, knowledge level, thinking process).

The authors have designed the adaptive CAI system equipped with these characteristics. Figure 1 shows the block diagram of the system. In this system, a PWR plant simulator simulates the dynamics of the plant abnormal events in real time. A trainee will be asked to find out the primary cause of the events by monitoring the information presented on Eye-Sensing Head-Mounted Display (ES-HMD) [1]. This HMD was developed by the authors and it not only provides virtual

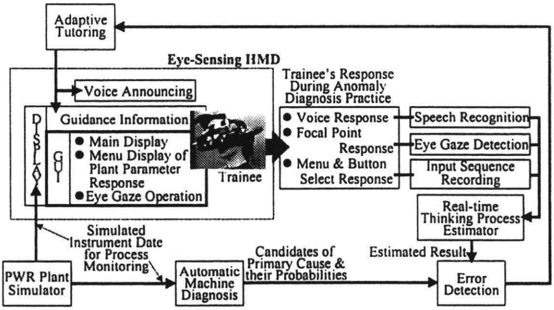


Figure 1: Block diagram of the adaptive CAI system

environment, but also monitors the trainee's eyes by CCD cameras and calculates ocular information such as eye gaze point, pupil diameter and eye blinks in real-time. The system also has a module which estimates the trainee's thinking process in real-time, utilizing two types of information, think aloud protocol and eye gaze point. At the same time, the machine diagnosis module automatically calculates candidates of primary cause of the current event and their probabilities. By comparing the result of thinking process estimation and the result of machine diagnosis, real-time feedback can be provided to the trainee.

However, the proposed CAI system still lacks two crucial blocks Figure 1: Error Detection and Adaptive Tutoring. For implementing the both, two kinds of knowledge are needed: knowledge about how to detect the trainee's mistakes and how to correct the mistakes. In this study, an experimental approach was taken to obtain their knowledge. For this purpose, an experimental environment was constructed to observe the real interactive behavior of instructor and trainee in CAI. In the developed experimental setup, an expert, who is the expert of operation and diagnosis of nuclear power plant, can monitor the diagnostic behavior of a novice trainee from remote place and can provide educational guidance. Therefore, effective educational strategy will be derived from analyzing the subject's behavior recorded in this laboratory experiment.

The experimental system, named Instruction Simulation System (ISS), was developed and a laboratory experiment was conducted using the ISS system. In this paper, the detail of ISS and its sub-systems will be first described, and then the detail of the experiment and experimental results will be presented.

2. Development of Instruction Simulation System (ISS)

As mentioned above, the purpose of the developed Instruction Simulation System (ISS) is to prepare an experimental environment to detect error patterns of trainees, to detect types of instruction and to derive some kinds of educational strategy utilized in the adaptive CAI system.

Figure 2 shows the overall configuration of ISS. The configuration of ISS is basically the same as that of the adaptive CAI system shown as in Figure 1. Both the adaptive CAI system and ISS consist of many computers connected with each other by Local Area Network. The salient deference between the both systems is that a human expert plays the role of function to detect trainee's error and correct the error in ISS. In ISS, all kinds of information about trainee's diagnostic behaviors which will be utilized in the adaptive CAI system will be presented to the human instructor, while the instructor's educational messages will be provided to the trainee in real time. All kinds of

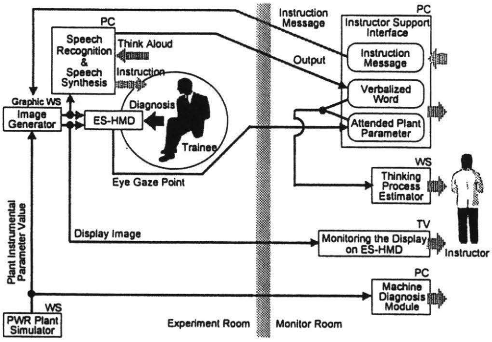


Figure 2: overall configuration of Instruction Simulation System

information presented to the instructor are (1) Trainee's attended plant parameter, (2) Trainee's think aloud protocol, (3) Trainee's view on ES-HMD, (4) Result of thinking process estimation (supplemental), and (5) Result of machine diagnosis (supplemental).

The features of ISS can be summarized as follows.

(1) Because the instructor is in the other room, the trainee at ISS can feel as if he/she studies by him/herself using CAI system.

(2) Eye Gaze Point and Think Aloud Protocol are monitored and recorded in ISS, and so those recorded data can be utilized for analyzing the trainee's diagnostic process.

(3) The dynamic behavior of the instructor is also recorded in ISS, and this record can be utilized for analyzing instructor's teaching way and for reducing effective critiquing method for adaptive CAI.

In the next chapter, details of the major subsystems of ISS will be described.

3. Subsystems of ISS

3.1. Eye-Sensing HMD and Eye Gaze Point Detection

Figure 3 shows a snapshot of ES-HMD. The user can be provided with a virtual environment to both of the user's eyes via small CRTs, and two infra-red CCD cameras equipped in ES-HMD can monitor both of his/her eyes. Then, the eye images are processed to calculate ocular in-formation such as pupil diameter, eye blinks and pupil center position in real time [1].

Figure 4 shows the outline of eye gaze point detection by using ES-HMD. The method of eye gaze point detection consists of two steps as follows;

Step(1) Detection of pupil center position in the original image of CCD camera.

Step(2) Detection of eye gaze point on CRT display by transforming eye gaze point in CCD image to that in CRT image.

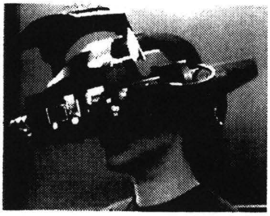


Figure 3: Snapshot of ES-HMD

In Step (1), the pupil contour data is approximated to the fittest circle, and then the center of the circle is interpreted as the pupil center position. In Step (2), the rotation angle of the eyeball is calculated from the pupil center position obtained in Step (1), and then the eye gaze point on the display of ES-HMD is detected in accordance with the rotation angle.

A real time software system called EGOS (Eye Gaze Operation System) was developed to process the above both steps in real time and to make the eye gaze point as input function. Error rate of eye gaze point detection by EGOS is about 1.5 [deg.] [2]. And so, EGOS is good enough to be used in ISS.

In the next section, the user interface for the trainee in ISS will be described. This interface has eye gaze input function by using EGOS.

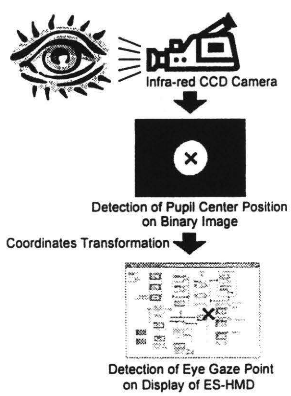


Figure 4: Outline of eye gaze point detection

3.2. User Interface for Trainee

The user interface for trainee in ISS is presented on the display of ES-HMD. He/she will be asked to diagnose abnormal events on plant simulator by monitoring the information presented on this interface.

This interface is a Graphical User Interface and consists of 16 windows. Figure 5 shows an example of the snapshot of the user interface. A mimic diagram of the plant subsystem and/or trend graph of the important plant parameters are provided by the same way as in the CRT display of actual plant operational console, and the trainee can select sub-windows by pressing window change buttons. The effectiveness of this window design was reviewed by the expert of plant operation and diagnosis.

The special function of this interface is that the pointing cursor automatically appears on the the interface display to locate the eye gaze point. This function is realized by applying the eye gaze point detection of EGOS. Due to this function, the trainee can easily pinpoint the window change buttons only by gazing them. Moreover, the interface can detect the trainee's attended plant parameters in real time from the location of the pointing cursor.

3.3. Thinking Process Estimation

The authors have already developed a real-time estimator to estimate users' thinking process [3]. The major features of this estimator are as follows;

(i) It uses the Causal Relation Diagram (CRD, Figure 6) as standard model of the user's knowledge. The CRD represents the physical and functional cause-consequence relationship between the major PWR plant parameters.

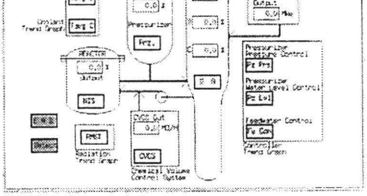


Figure 5: Snapshot of the user interface

(ii) By combining the CRD with two types of information (think aloud protocol and attended plant parameters), the estimator can compute the Confidence Score (degree of users' confidence in judging primary cause of disturbance) for each plant parameter on CRD.

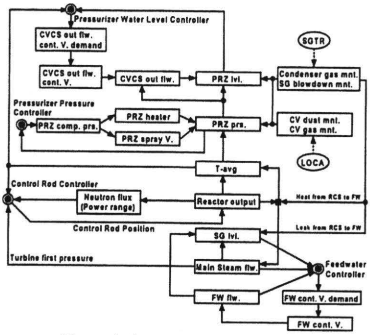


Figure 6: Causal Relation Diagram

To calculate the Confidence Score, the estimator determines the root of the abnormal propagation from the network of CRD, by using the above-mentioned two types of information. The meaning of Confidence Score is the possibility of each parameter to be primary cause rather than the confidence on a specific hypothesis the trainee holds in mind. Since the estimator can dynamically trace the transition of the Confidence Score for each parameter, it can in a sense be said to reflect the user's thinking process.

By the way, the speech recognition subsystem recognizes the trainee's think aloud in ISS. The total number of words, which can be recognized by this subsystem, is restricted to 23 for easy memorization by the trainee on one hand, while for proper recognition on the other hand. His/her thought on his attended plant parameters on ES-HMD is inferred from the verbalized words.

The result of thinking process estimation will be presented to the instructor as a supplemental information about the trainee's diagnostic process.

3.4. Machine Diagnosis

The purpose of this subsystem is to construct a model of correct behavior for comparison with the trainee diagnostic behavior [4]. The method of diagnosis is summarized as follows.

Step(1) This subsystem finds all the parameters deviated from the normal condition of the plant system. Then, it picks up the root parameters in the abnormal propagation network of CRD. This root will be defined as a candidate of primary cause. The search method for the root is almost the same way as in thinking process estimation in section 3.3.

Step(2) This subsystem generates the abnormal propagation course derived from these candidates on CRD.

Step(3) The subsystem decides the degree of deviation from normal condition for all the parameters on the course obtained in Step (2). The degree of deviation is calculated by the membership function for each parameter which is obtained from the subjective report by human in previously conducted experiment. The feature of this step is that the subsystem has the same criterion of decision as human.

Step(4) The subsystem interprets the degree of deviation as the individual evidence supporting each candidate, and then calculates probabilities of the candidates by using Dempster-Shafer theory.

The result of machine diagnosis will be presented to the instructor as a supplemental information about the situation of abnormal events which the trainee is facing.

3.5. Instructor Support Interface

The Instructor Support Interface will be primarily used by the instructor in ISS. Originally, the instructor will be asked to monitor several kinds of information about trainee's diagnostic behavior.

So, if all kinds of information are merely displayed to him/her in random manner, a considerable workload will be imposed on him/her. Consequently, the authors have developed the interface which can present several information of trainee with easily understandable and viewable formation and can accept educational message from the instructor for the trainee.

Figure 7 shows an example of the snapshot of this interface. This interface consists of four sub-windows as follows.

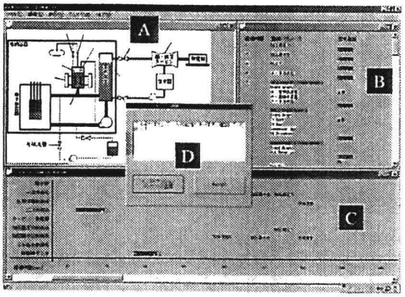


Figure 7: Snapshot of the Instructor Support Interface

(1) Mimic: Presenting trainee's attended plant parameter on a mimic diagram of plant (Figure 7 – A).

(2) Enumeration: Listing all attended parameters and think aloud protocol (Figure 7 – B).

(3) Timetable: Presenting the significant attended parameter and think aloud protocol along the time axis (Figure 7 – C).

(4) Popup: The instructor can input the educational message for the trainee by keyboard (Figure 7 – D).

The educational message given by instructor through this interface will be transferred to the trainee in real time and presented to the trainee by two modalities, one is the speech synthesis and the other is the text on a popup window which appears on the user interface for trainee.

4. Experiment and Analysis

A laboratory experiment has been conducted using ISS, in which an instructor and trainees have actually joined. The purpose of the experiment can be summarized as follows.

(1) To find out error patterns of the trainee's diagnosis and the instructor's instructions for correcting the error patterns, by analyzing both the trainee's and the instructor's behaviors.

(2) To discuss with the instructor about what kind of tutoring strategy is effective for ISS.

(3) To ask the instructor and the trainee to evaluate the usability and effectiveness of ISS.

In the following sections, the detail of the experiment and the result of the experiment will be described.

4.1. Instructor, Trainees and their Tasks

In the experiment, an expert, who had really engaged in plant operation as a chief operator, was employed as the instructor. And, three graduate students were employed as the trainee (Subject A, B and C). They were novices for diagnosing nuclear power plant anomaly.

The tasks of the trainees were as follows.

(1) Diagnosis: They were asked to find out the primary cause of abnormal events by monitoring GUI displayed on ES-HMD.

(2) Window change: They were asked to change windows for checking parameters.

(3) Think aloud: They were asked to speak words about their judgement on the value of attended plant parameters.

On the other hand, the tasks of the instructor were as follows.

(1) Monitor: He was asked to monitor four displays as shown in Figure 2.

(2) Instruction: He was asked to detect the trainee's mistakes and to give appropriate instruction from educational viewpoint.

4.2. Schedule

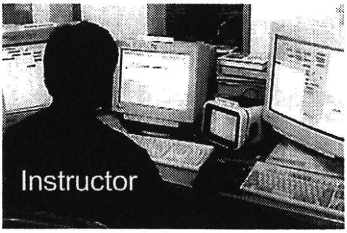


Figure 8: Snapshot of the instructor

The experiment was conducted for three days. On the first day, a lecture was given the novice trainees. The contents of the lecture were (1) fundamental knowledge about PWR power plant, (2) structure and function of the user interface, and (3) tasks in the experiment. On the second day, exercise practice of experimental task was carried out for the trainees. The contents of the practice were (1) operation of the user interface, (2) eye gaze operation by ES-HMD, and (3) diagnosing samples of abnormal events. On the last day, the experiment using ISS was actually carried out. In the experiment, each trainee was asked to diagnose four or five abnormal events and the instructor was asked to provide instruction to correct trainee's mistakes by his own will. And, the instructor designated which abnormal event should be occurred. The time limit of each trial (diagnosis for one event) was 10 minutes. But the trainee could finish a trial within the time limit if he could detect the primary cause. And, after each trial, both the instructor and the trainees were asked to make a subjective report about the process of diagnosis, cause of instruction and effect of instruction. In addition, after all trials were finished, the trainees were asked to make a subjective report about the effectiveness and usability of ISS and the instructor was asked to give a commentary report on discussion about what kind of tutoring strategy is effective for ISS. Figure 8 shows a snapshot of the instructor during the experiment.

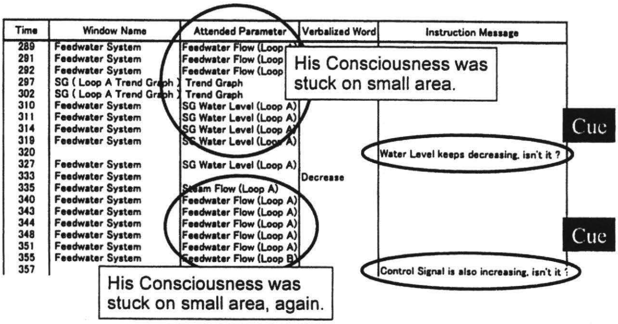
4.3. Result of Experiment

Error and Instruction

Table 1 shows the result of the diagnosis performed by trainees and the evaluation for the diagnosis by the instructor. As shown in Table 1, the instructor could detect all misdiagnosis. The authors think that it indicates that he could recognize the trainees' diagnostic behavior. On the other hand, in some trials where the trainees finally reached a correct conclusion, the instructor evaluated that the diagnosis was not sufficient. The authors think that it also indicates that he could recognize the trainees' thinking process in the details only by utilizing the information available in ISS.

Table 1: Result of Experiment

Subject	Plant Anomary	Trainee's Diagnosis	Evaluation by the instructor	Number of Instruction Message
A	LOCA	Correct	Good	3
	Pressurizer Spray Valve Fail-Open	Wrong	Deadlocked	1
	Pressurizer Spray Valve Fail-Open	Correct	By accident ?	5
	Feedwater Flow Sensor Failure (fail-high)	Wrong	No Good. This anomary may be too difficult.	5
	Pressurizer Pressure Control System Failure	Correct	By accident ?	5
B	Pressurizer Heater Failure (fail-high)	Wrong	No Good	8
	LOCA	Correct	Almost Good	1
	Feedwater Flow Control Valve Stuck	Correct	Good	6
	Pressurizer Heater Failure (fail-high)	Correct	Almost Good	5
C	LOCA	Correct	Unreliable	5
	Pressurizer Spray Valve Fail-Open	Correct	By accident ?	7
	Feedwater Flow Control Valve Stuck	Wrong	Deadlocked	15
	Pressurizer Heater Failure (fail-high)	Correct	Good	9



Time	Window Name	Attended Parameter	Verbalized Word	Instruction Message
289	Feedwater System	Feedwater Flow (Loop		
291	Feedwater System	Feedwater Flow (Loop		
292	Feedwater System	Feedwater Flow (Loop		
297	SG (Loop A Trend Graph)	Trend Graph		
302	SG (Loop A Trend Graph)	Trend Graph		
310	Feedwater System	SG Water Level (Loop A)		
311	Feedwater System	SG Water Level (Loop A)		
314	Feedwater System	SG Water Level (Loop A)		
319	Feedwater System	SG Water Level (Loop A)		
320				Water Level keeps decreasing. isn't it ?
327	Feedwater System	SG Water Level (Loop A)		
333	Feedwater System		Decrease	
335	Feedwater System	Steam Flow (Loop A)		
340	Feedwater System	Feedwater Flow (Loop A)		
343	Feedwater System	Feedwater Flow (Loop A)		
344	Feedwater System	Feedwater Flow (Loop A)		
348	Feedwater System	Feedwater Flow (Loop A)		
351	Feedwater System	Feedwater Flow (Loop A)		
355	Feedwater System	Feedwater Flow (Loop B)		
357				Control Signal is also increasing. isn't it ?

Figure 9: An example situation of (1) Deadlock

By observing the instruction and by analyzing the comments by the instructor, the trainees' errors or mistakes can be classified roughly as follows.

(1) Deadlock: Their diagnosis could not proceed, and sometimes their consciousness was stuck on small area. It is supposed that they could not recognize the current situation because of lack of knowledge or failure of applying knowledge.

(2) Insufficient situation awareness: Their diagnosis proceeded at least, but it made slow progress. They could not acquire a significant information.

(3) Delay: They sometimes paid attention to an area of plant which is not related to the abnormal event.

Figure 9 shows an example of the situation of (1) Deadlock. In this situation, in spite of an advice from the instructor, the consciousness of the trainee was stuck on one place and his reasoning did not transit. And, Figure 10 shows an example of the situation of (2) Insufficient situation awareness. In this situation, owing to an advice from the instructor, the trainee could barely notice the significance of Pressurizer Heater.

Instructions for above types of errors can be classified as follows.

(1) Cue: The instructor provides some hints, when the trainee's reasoning is deadlocked.

(2) Guidance: The instructor guides the trainee, when the trainee begins to check the instruments which have no relation to the plant anomaly or when he does not notice the important instruments.

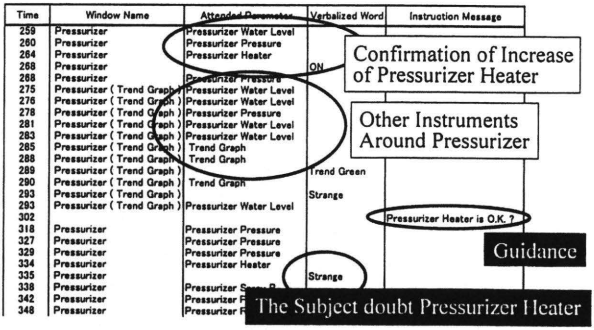
In addition, other kinds of instructions were seen in the experiment as follows, which were not aiming at correcting errors.

(3) Confirmation: The instructor asks the trainee to check other possibilities of plant anomaly.

(4) Disturbance: The instructor intentionally confuses the trainee from educational viewpoint, when the trainee can quickly detect the primary cause.

Tutoring strategy on ISS

From the discussion with the instructor after all trials were finished as well as the report from the instructor, an idea of effective educational strategy on ISS was acquired. The strategy can be summarized as follows.



Time	Window Name	Attended Parameter	Verbalized Word	Instruction Message
259	Pressurizer	Pressurizer Water Level		
260	Pressurizer	Pressurizer Pressure		
264	Pressurizer	Pressurizer Heater		
268	Pressurizer		ON	
268	Pressurizer	Pressurizer Pressure		
275	Pressurizer (Trend Graph)	Pressurizer Water Level		
276	Pressurizer (Trend Graph)	Pressurizer Water Level		
278	Pressurizer (Trend Graph)	Pressurizer Pressure		
281	Pressurizer (Trend Graph)	Pressurizer Water Level		
283	Pressurizer (Trend Graph)	Pressurizer Water Level		
285	Pressurizer (Trend Graph)	Trend Graph		
288	Pressurizer (Trend Graph)	Trend Graph		
289	Pressurizer (Trend Graph)		Trend Green	
290	Pressurizer (Trend Graph)	Trend Graph		
293	Pressurizer (Trend Graph)		Strange	
293	Pressurizer (Trend Graph)	Pressurizer Water Level		
302				Pressurizer Heater is O.K. ?
318	Pressurizer	Pressurizer Pressure		
327	Pressurizer	Pressurizer Pressure		
329	Pressurizer	Pressurizer Pressure		
334	Pressurizer	Pressurizer Heater		
335	Pressurizer		Strange	
338	Pressurizer	Pressurizer S		
342	Pressurizer	Pressurizer F		
348	Pressurizer	Pressurizer F		

Figure 10: An example situation of (2) Insufficient situation awareness

Step(1) Initial evaluation of experience and performance of the trainee: By imposing a simple malfunction, the instructor can evaluate trainee's level by observing how quickly he can identify the cause. when the trainee is qualified as novice, the training should be simple consisting of malfunctions of an element of local control systems. After the trainee become experienced to certain degree, he can move to next step.

Step(2) Evaluation of monitoring scope: The instructor observes which windows the trainee opens. Once the trainee detects the symptom/deviation, he/she will navigate to associated window. Experienced trainees can predict which malfunction has been just imposed at this stage. Instructor looks to see if trainees miss the significant information or alarm.

Step(3) Evaluation of the occurrence of coolant leak: At this stage, it is strongly requested for trainees to ensure event identification or screening of the predicted events. The monitoring of Radiation Monitoring System (RMS) is necessary whenever an accident occurs, since the role of operators is to avoid the excessive nuclear release to the environment. If a break did not occur and/or reactor trip does not follow, then trainees must make laborious identification by checking the status of on-site and off-site power, and local control system of secondary side.

Step(4) Debriefing: The instructor explains which malfunction(s) has been imposed and if there were serious mistake/errors, same practice may be made. Occasionally, the instructor enquires the reasoning process of the event identification.

Usability and Effectiveness of ISS

The evaluation for usability and effectiveness of ISS by the instructor can be summarized as follows.

(1) ISS has too many displays to monitor. It is required to integrate all kinds of information about the trainee's diagnostic behavior into a single interface.

(2) Eye gaze point and think aloud are very effective information for recognizing the trainee's diagnostic thinking. So, ISS is available for training.

(3) Thinking process estimation is available only for recognizing the focus of the trainee's consciousness. On the other hand, machine diagnosis has little utility. However, it can detect abnormal symptom earlier than human.

The evaluation by the trainees can be summarized as follows.

(1) Design and usability of the user interface are very good. The function of eye gaze operation was useful in almost cases, but it infrequently had low accuracy.

(2) The instructions provided by text and speech synthesis were easy to recognize.

From the above evaluation, the usability and effectiveness of ISS as a training system was confirmed.

5. Conclusion and Future Works

From the laboratory experiment using ISS, possible error types committed by novice students and the types of instruction for those error types were successfully obtained, and the effective educational strategy was reduced for the framework of ISS by discussing with the expert. In addition, it was confirmed that ISS is useful enough for the training from the evaluation by the expert and the trainees.

However, from the view point of developing the adaptive CAI system shown in Figure 1, the result of the experiment is insufficient, because only 13 cases for three trainees was carried out. In order to develop a tutoring strategy or rules for a computerized system, it is necessary to conduct more experiment to collect sufficient amount of data.

In addition, all the experimental data have not been analyzed. So, in the future, the authors will study several topics, for example the cognitive process of novice students by analyzing the data of eye gaze point, and so on.

6. References

[1] H. Yoshikawa, H. Shimoda, S. Fukushima, D. Morikawa and H. Fujiyama: Development of Eye-Sensing HMD and its application Experiments for Observing Human Ocular Information Processing Characteristics, Proceedings of AROB III '98, pp.7-10, (1998).

[2] T. Ozawa, Y. Aotake, H. Shimoda, S. Fukushima and H. Yoshikawa: Real-time Eye Gaze Point and Eye Blink Detection Using Eye-Sensing HMD, Transactions of the Society of Instrument and Control Engineers, (2001) (to appear, in Japanese).

[3] T. Ozawa, H. Shimoda and H. Yoshikawa: Application of Machine Speech Recognition for Real-time Estimation of Human Cognitive Process at Man-Machine Interface, Proceedings of IEEE-SMC '98, pp.4127-4132, (1998).

[4] T. Ozawa, Y. Aotake, H. Shimoda and H. Yoshikawa: A Study on Mutual Adaptive CAI System for Knowledge of Plant Anomaly Diagnosis – An Examination of Machine Diagnosis Module –, Proceedings of Annual Meeting of the Atomic Energy Society of Japan, p.680, (2000) (in Japanese).

Associates with Etiquette:
Meta-Communication to Make Human-Automation Interaction more Natural, Productive and Polite

Christopher A. Miller* and Harry B. Funk
Smart Information Flow Technologies
1272 Raymond Ave.
St. Paul, Minnesota, 55108 U.S.A.
{cmiller,hfunk}@SIFTech.com

Summary

Several different architectures for adaptive automation systems have now been demonstrated. Yet users remain ambivalent them—as our own work on the U.S. Air Force's Rotorcraft Pilot's Associate (RPA) reveals. Thus, we argue, research should be shifting away from how adaptive interfaces and automation can be made to work, and onto how they should *behave* in conjunction with human users. We find it useful to think about behavioral details and roles and responsibilities in terms of various 'etiquettes' for human-machine interactions. Results from RPA suggest that one aspect of establishing acceptable etiquette may be more and better inclusion of 'meta-communication' on the part of the system—the ability to report and accept feedback on the system's perceptions and intentions. We discuss these results in the broader context of etiquette recommendations for adaptive automation systems.

Introduction

The past several years have seen an explosion in the creation of "Adaptive" or "Intelligent" user interfaces which modify system performance, automation behavior and/or information presentation in response to some aspect of context, user, task or situation. Much of this research and growth, and much of the average user's familiarity with such systems, has been in the desktop software environment—with Microsoft™'s Office Assistant (the ubiquitous animated paperclip called "Clippit™" being the best known, if not always most loved, example. Works by Maybury [1998], Puerta [1996], and Bauer [2000] are examples of this trend. At the same time, "real world" examples of adaptive and intelligent interfaces—those which deal with a controlled system 'off the desktop' [Miller, 2000]—are moving closer to reality and fruition. Our own work on military attack/scout helicopters [Miller, Hannen, Guerlain, 1999], oil refineries, [Cochran, Miller and Bullemer, 1996] military command and control [Funk, et. al., 2000] as well as those of Bonner, et. al. [2000], Onken [Strohal and Onken, 1998], and Mitchell [1999], serve as examples.

As the number of such adaptive interaction efforts increases, we are collecting more data about what works and what doesn't. In our own experience, across multiple domains, it is clear that users want, even demand, to remain in charge of actions, yet they also want, even require, the benefits that intelligent, adaptive information and automation can provide them. Work on the Rotorcraft Pilot's Associate (RPA) Cockpit Information Manager [Miller, Hannen and Guerlain, 1999] proves instructive from several perspectives. For example, in developing RPA, we interviewed multiple pilots and designers to develop a consensus list of prioritized goals for a "good" intelligent cockpit configuration manager. Two of the top three items on the list were "Pilot remains in charge of task allocation" and "Pilot remains in charge of information presented." Thus, in spite of generally 'buying into' the need for and benefits to be provided from a sophisticated adaptive automation system, these pilots and designers remained ambivalent about how much control and authority they wanted to give it.

There are good reasons to design and use adaptive systems at the higher levels of automation. By definition [Sheridan, 1987; Parasuraman, Sheridan and Wickens, 2000], such systems share responsibility, authority and autonomy over many work behaviors with human operator(s) to accomplish their goals of reducing operator workload and information overload. While operators may wish to remain in charge, and it is critical that they do so, today's complex systems no longer permit them to be fully in control of all system operations—at least not in the same way as in earlier cockpits and workstations [cf. Miller, Pelican and Goldman, 2000; Perrow, 1999]. So, we are faced with a dilemma. Advanced and adaptive automation is necessary to allow humans to achieve the levels of performance required in today's world, yet human users frequently reject such automation when it makes them feel out of control. Worse yet may be when they fail to or cannot reject such automation—as the range of mode control errors and pilots' sense of failing to understand 'what it will do next' show all too clearly [Sarter, Woods, and Billings, 1997].

How can we build adaptive automation systems which leave the human in charge even when the human is less directly in control? The proposed solution, all too frequently, is to just make the automation 'better'—more aware of the user, more aware of the context, encompassing a broader range of behaviors, etc. The problem is that this puts the onus on the automation to know what is the right thing to do for the operator at any given time. While such a capability would certainly be nice, we believe it has proven (and will continue to prove) far too difficult to achieve with sufficient accuracy to be practical. We are not suggesting that task tracking, intent inferencing or cognitive modeling approaches [Charniak and Goldman, 1993; Hoshstrasser and Geddes, 1989; Coury, Santarelli & Mitchell, 2000] should be ignored, merely that they should be augmented. This augmentation should come through defining acceptable, desirable roles and relationships between human operators and advanced, adaptive automation systems—relationships that the automation can reliably and predictably deliver. We are finding it useful to think of such a package of defined roles and methods of relating as an 'etiquette' for human-machine interaction.

The remainder of this paper illustrates the benefits to be derived from designing adaptive automation around such an etiquette, using examples from our work on the Rotorcraft Pilot's Associate (RPA). We then go on to discuss the 'etiquette' of human-machine interactions and propose a tentative initial list of etiquette rules to stimulate future work.

Etiquette in the Rotorcraft Pilot's Associate

Description of the RPA

The US Air Force's Pilot's Associate programs were among the first efforts to implement large, adaptive interface and automation management systems [Banks and Lizza, 1991]. The particular type of automation targeted was called an 'associate' system because it was intended to provide many of the same functions and operate in the same relationship as a human associate in a single-seat fighter cockpit. 'Associates' are collections of intelligent aiding systems that, collectively, exhibit the behavior of a capable human [Riley, 1989; Miller and Riley, 1994]. They can (a) perform roughly the same breadth of activities as a human expert in the domain, (b) take initiative when necessary, but generally follow a human's lead, and (c) integrate over ongoing activities to exhibit robust, coordinated, appropriate behavior.

The US Army's Rotorcraft Pilot's Associate (RPA) program was a five year, $80 million research contract managed by the U.S. Army's Aviation Applied Technology Directorate at Ft. Eustis that built on the Pilot's Associate work, but extended it in many ways [Collucci, 1995].. The goal of RPA was to develop and demonstrate in flight an 'associate' system in a next-generation attack/scout helicopter. A critical sub-goal was to manage the information available in future helicopter operations so that human crews can attend to all and only relevant portions at a given time. RPA must accomplish this without increasing pilot workload or decreasing situation awareness. In practice, the RPA module designed to accomplish these goals was the Cockpit Information Manager (CIM) [Miller, Hannen and Guerlain, 1999]. CIM was designed to perform five major functions from the pilots' perspective:

1. *Page (or Format) Selection*—selection of a complete page or format to present on any of the aircraft's presentation devices. For example, the selection of a weapons page instead of a sensors page on the Right Multi-function Display, or the presentation of warbling tone at a specific 'location' via the 3D stereo sound system.
2. *Symbol Selection/Declutter*—turning specific symbols on or off on a selected page (e.g., include/suppress intervisibility symbology on the Tactical Situation Display).
3. *Window Placement*—control of the location for pop-up windows which would overlay some other visual imagery on the Multi-Function Displays.
4. *Pan and Zoom* –control of centering and field of view of map and sensor displays.
5. *Task Allocation*—the assigning of tasks to various pre-defined, legal combinations of the two human pilots and automation.

Each behavior was adaptive and made use of an inferred task context to determine which information should be presented or which tasks should be allocated in what way. In addition to relying solely on inferred pilot tasks, however, we also implemented several mechanisms to allow the pilots to control and interact with the adaptive behaviors described above. Pilots could, during initial configuration, control the set of options that CIM was permitted to consider, and could apply preference weights to those options. Pilots could also, during the mission, individually turn each of the options on or off. Pilots could also command any display state they desired and the CIM would respond by avoiding modification to the commanded display for a period of time. In this sense, the pilot retained control over the behaviors that the CIM could, and was likely to, perform. CIM's behavior was not entirely predictable—after all, it could still choose between authorized options on the basis of what was appropriate in context—but the parameters of its relationship vis-a-vis the pilots were well defined.

A Crew Coordination Display

The types of pilot control over CIM functions mentioned above were not new to RPA. What was new was a serious look at the 'etiquette' which existed on the helicopter flight deck and an attempt to design the CIM to 'behave well' according to that etiquette.

Prior associate programs (especially the Pilot's Associate—cf. [Banks and Lizza, 1991]) were concerned about forcing the pilot to take on additional workload if s/he had to explicitly communicate and coordinate with the associate, and therefore relied more heavily on intent recognition alone. Direct pilot interactions with the associate (e.g., about what information the pilot wished to see or what actions he was in fact engaged in) were regarded as adding to that workload since they were, after all, interactions the pilot would not have had to perform if the associate were not present. While direct interactions might be warranted, they inevitably had to justify their added cost and were, generally, not favored.

By contrast, initial interactions with helicopter crews and reviews of domain training approaches revealed that, in the attack/scout helicopter domain, as much as a third of crew members time while in the cockpit is engaged in what might be called 'meta-communication' activities—discussion of plans and intentions, allocations and affirmation of responsibilities, maintenance of situation awareness, etc. In fact, helicopter crews were currently undergoing substantial training and review of Crew Resource Management procedures [e.g., Foushee and Helmreich, 1988], which were designed to strengthen crew coordination and team situation awareness. We believed that for the RPA CIM to fail to behave in accordance with this operational 'etiquette' – that is, to fail to be able to report on its activities, its perception of the activities of others, and to take instruction about the activities it should be engaged in—would guarantee that it fail to 'play by the rules' that human crew members expected.

Our response was to create a method for simple 'meta-communication' interactions between the pilots and the associate. Based on initial designs by Dr. Stephanie Guerlain, and ultimately implemented under the supervision of Matthew Hannen at Boeing [cf. Miller, Hannen and Guerlain, 1999], this "Crew Coordination and Task Awareness" display consisted of four small LED buttons located in the upper portion of each pilot's main instrument panel. Each button was capable of displaying up to two eight-letter lines of text. The buttons were used to report, in textual form, (1) the associate's current inference about the general, high-level mission context (e.g., that we're currently engaged in an attack task rather than an evade task), (2) the associate's inference about the highest priority current pilot task, (3) the task which the associate is engaged in currently which it believes has the highest priority, and (4) the highest priority, inferred copilot task. Pressing these buttons permit either pilot or co-pilot to override CIM's current inferred tasks and assert new ones (from an automatically scrolled list of higher-level tasks from the overall task network) via a single push button input. Figure 1 shows the RPA cockpit simulation created for evaluation trials at the Boeing Company in Mesa, Arizona. The location of the Crew Coordination display is circled and an enlargement and interpretive sketch of the display is provided for clarity.

The goal of the Crew Coordination and Task Awareness display was to provide the crew with insight into 'what the associate thought was going on'—as well as a direct ability to affect it. Although such interactions would actually add workload to pilots' duties (over a perhaps unrealistic baseline where no interaction with the associate was needed or provided), we felt that the 'etiquette' of the flight deck demanded that the associate provide at least this level of meta-communication about its intents and its knowledge of the intent of others. For the associate not to have this capability might make it seem less trustworthy. We expected that the inclusion of a Crew Coordination display would facilitate user acceptance of the RPA by making it seem more like a 'team player' rather than a silent automaton, as well as allowing the crew to make task-based intent inputs which would improve the accuracy of RPA's aiding overall. Since the inclusion of a direct method for viewing and interacting with the intent estimation and task network software was a new development in the RPA cockpit (over prior associate system work), we were especially interested in how pilots would regard it.

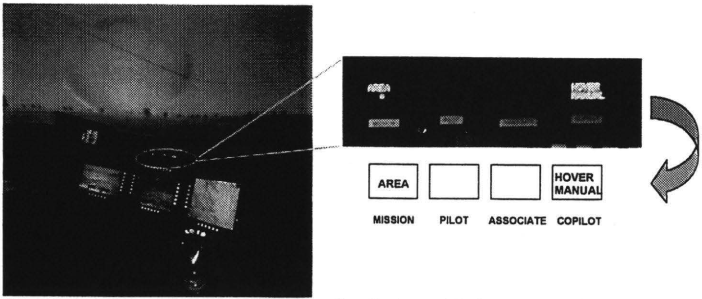


Figure 1. Location and appearance of the Crew Coordination and Task Awareness Display.

Evaluation Results—User Acceptance

Evaluation results designed to determine subjective user acceptance bear these hypotheses out. Extensive full mission simulations were performed to help evaluate the RPA CIM behaviors and implementation, as well as to assist in prioritizing issues for the flight demonstration. Four Crews (a total of eight pilots) trained and flew together, as they do in actual field operations. Crews were given realistic mission briefings and objectives and were permitted to make their own tactical decisions about how to achieve them. Tests were flown in Boeing's full mission simulator (see Figure 1) and included full fidelity RPA cockpits, dome visuals, an extensive range of passive and active threats, and human control of the aviation Tactical Operations Center, friendly artillery, and 1 to 3 wingmen. Realistic communications, including change of mission Fragmentary Operations Orders, were maintained between these players. Each pilot received an average of 10.8 hours of training in the simulator and 13.9 hours of classroom training over a two-week period.

Each crew flew 14 part-mission test scenarios, 7 with the full RPA Cognitive Decision Aiding System (CDAS), and 7 with an Advanced Mission Equipment Package (AMEP) alone. The AMEP was represented the capabilities of an advanced attack/scout helicopter platform, including an impressive array of automation and decision aids, but without the integrating support of an associate system. Each crew also flew four full-mission scenarios—two with the AMEP alone and two with the RPA CDAS in addition to the AMEP. Full-mission scenarios were designed to be highly realistic and crews were given free reign to pursue their commander's objectives via whatever methods they thought appropriate. Crews flew the two AMEP or CDAS full missions in sequence and then switched technology conditions and flew the remaining two missions with the other set of technologies. The sequence in which crew interacted with the different technology packages was counterbalanced to minimize training effects. The simulation test segments and missions were constructed to include numerous examples of the CIM page selection, window location, pan & zoom, and symbol selection behaviors in a variety of tactical mission contexts. (CIM's task allocation behaviors were not implemented in the simulation due to time and budget constraints.)

In order to obtain crew acceptance data, a questionnaire was administered to the pilot and copilot after each of the final AMEP and CDAS full-mission test trials. All of the questionnaire responses utilized complete verbal anchoring and a linear response scale with five equal intervals, in accordance with [Charlton, 1989]. The criteria value for satisfactory CIM behavior, was set at an average score of 3.5 or greater for each response.

The criterion was met for three of the four CIM behaviors. Figure 2 presents the average and range of pilots' ratings of the behaviors . In general, pilots found the CIM behaviors to be 'Of Use' or 'Of Considerable Use.' Figure 3 presents pilots' ratings of their perceptions of the frequency with which they had to override or correct CIM's actions. The average over the CIM behaviors fell between 'Seldom' and 'Now and Then' with symbol selection capabilities performing notably better.

Figure 4 shows pilot ratings of CIM as a whole. CIM was seen as 'Frequently' providing the right information at the right time and was seen as almost always predictable in its behaviors. Finally, Table 1 compares pilot ratings of their effectiveness over four mission types with CDAS versus with the AMEP. On average, pilots found themselves to be more than half a point more ef-

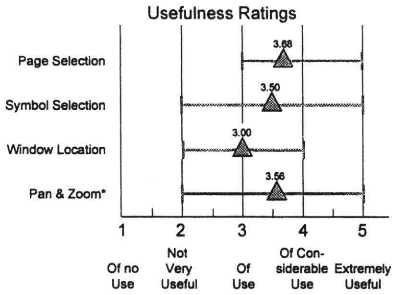
Table 1. Perceived effectiveness in different mission tasks with CDAS and AMEP alone (where 3.0= 'Fair', 4.0='Good' and 5.0='Excellent'.)

Average Rating	AMEP	CDAS
Zone Reconnaissance	3.75	3.88
Area Reconnaissance	3.75	4.25
Deliberate Attack	4.13	4.75
Change to Attack	3.63	4.63

fective (10% of the scale length) with CDAS than without.

The RPA CDAS also produced overall benefits relative in one other critical area. Using TLX measures of subjective workload collected at the end of each trial, workload scores were consistently lower for CDAS conditions than for AMEP conditions (46 points versus 57 points). This difference was significant in an Analysis of Variance [F(1,6)=11.524, p<.05]. Furthermore, separate ANOVAs were conducted for each of the six TLX subscale ratings to determine CDAS' contributions to overall workload reduction. These results are presented in Table 2.

Reduced workload with the RPA CDAS is apparent in the mental demand, physical demand, temporal demand and effort subscales. There is also a marginal finding for the frustration subscale (p=.07). Means in all cases indicate that the RPA CDAS provides a benefit to the pilot. Examination of the perceived performance ratings, however, shows no effect of configuration. This may indicate that pilots use a different subjective criteria in rating their own performance, possibly judging it based on how well they felt they should have done in the given context, which would include the cockpit configuration.



* Pan & Zoom questions were asked as pilot estimates of situation awareness with CDAS pan & zoom capabilities, using anchor terms "Terrible, Poor, Fair, Good, Excellent".

Figure 2. Pilots' average ratings and ranges for the usefulness of the four CIM behaviors.

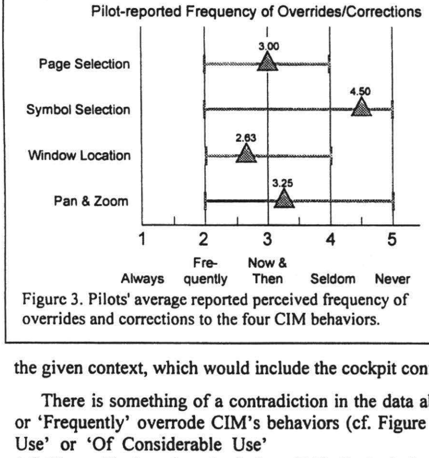


Figure 3. Pilots' average reported perceived frequency of overrides and corrections to the four CIM behaviors.

There is something of a contradiction in the data above. Pilots said they 'Now and Then' or 'Frequently' overrode CIM's behaviors (cf. Figure 3), yet they found those behaviors 'Of Use' or 'Of Considerable Use' (cf. Figure 2), they thought their effectiveness was better with CDAS (cf. Table 1), and their TLX ratings confirm that CDAS offered significant perceived workload reductions. How can such levels of perceived usefulness be achieved along with such perceived error rates?

Table 2. Analysis of the TLX subjective workload subscale ratings.

TLX subscale	AMEP mean	CDAS mean	F-Value (df: 1,6)
Mental Demand	61.77	46.25	10.487*
Physical Demand	54.48	40.31	12.042*
Temporal Demand	62.08	45.73	14.061**
Perceived Performance	35.00	42.08	2.429
Effort	62.60	48.54	20.470**
Frustration	52.81	45.63	4.961

*p<.05 ** p<.01

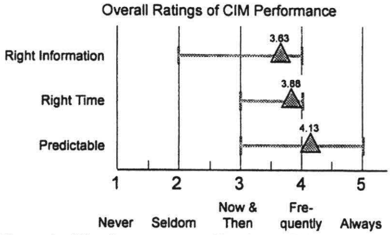


Figure 4. Pilots' average overall ratings of CIM performance.

Table 3. Perceived usefulness of the Crew Coordination and Task Awareness Display (where 4.0='Of Considerable Use' and 5.0='Extremely Useful'.

LED Button for:	Score
Mission Task	4.4
Pilot Task	4.3
Copilot Task	4.3
Associate Task	4.0

The answer may lie partly in the etiquette around which RPA was designed. As seen in Figure 4, CIM was seen as very predictable in its behaviors. It was also, generally, easy to override when it made a mistake. In addition, the inclusion of the Crew Coordination and Task Awareness display seems to have had a significant impact. Most pilots found this display 'Of Considerable Use' or 'Extremely Useful' (cf. Table 3). This provides some supporting evidence that the capability for the crew to interact directly with the associate's assumptions about active tasks was a capability that pilots welcomed—and that may have served to improve their overall impressions of CIM's capabilities and usefulness. Just like human colleagues, the RPA provides value when it is correct, and it facilitates interactions designed to manage, mitigate and correct its mistakes when it is not.

The Implications of Etiquette

Why Etiquette?

The inclusion of this novel interface type stemmed from taking seriously the 'associate' metaphor and using it to guide the behavior of sophisticated and semi-autonomous systems. 'Associates,' whether they be human or automated, must behave in ways appropriate to the established culture of the work environment and in ways that will further the goals of the team. We began with a close inspection of the 'etiquette' or culture used by team members in advanced attack/scout helicopter operations. Then, when attempting to design an 'associate system'—a highly capable, partly autonomous, dynamically adaptive aid which needed to interact with and, in a real sense, to be a part of the crew complement of this vehicle—we used our knowledge of that 'etiquette' to guide design decisions which seem to have paid off.

The success of this interface innovation has led us to think more seriously about the implications of the associate metaphor for adaptive automation in many domains, and about the role that the 'etiquette' of human-automation relationships does and should play. Etiquette seems to be a useful way of thinking about the formalizing the relationship between humans and automation systems. This thought process places the emphasis less on the construction of hardware, software or even knowledge architectures (though these must be in place) and more on the perceived behavior of the adaptive aiding system. What are the ways in which an automated associate *should* behave to support optimal human-system performance?

Etiquette rules are rarely created whole cloth by the Emily Posts or Miss Manners of the world. Instead, they attempt to observe good practices already existing in 'polite society' and then formulate them for others and/or infer from existing practices to propose etiquette for new situations. By proposing etiquette rules for adaptive automation and information management systems to follow, we should take a similar approach: observe good information exchange practices between humans and humans, or between humans and those systems that already exist, and attempt to both explicate good practices for others to follow and extend and generalize good practices to novel domains and situations.

Etiquette rules are not the same for all situations. While the mavens of etiquette may advocate a formalization of behavior for 'polite society', it is clear that there are different kinds of etiquette for different settings and domains. What may be appropriate in the boardroom would be strange in the parlor, and what might be genteel in a formal dinner would seem cold and contrived in a poker game. Furthermore, etiquette rules don't always have to be followed. More importantly for the types of domains we have been concerned with, attempts to make automated systems friendly and sociable [e.g., Reeves and Nass, 1996] may actually interfere with their ability to perform useful work in a team setting—just as human-human teams with a time critical job to do may forego many of the 'niceties' of social interaction. We maintain, however, that this type of goal-directed, team-coordinated behavior is simply another type of etiquette, and not an abandonment of all 'rules of good behavior'.

Finally, there may be times where the conscious and systematic violation of etiquette is highly useful. Nevertheless, consistent violation of rules appropriate to a domain relegates one to an undesirable position in society. We don't claim that every adaptive system should adhere to the same etiquette—just that most should try to find a good and useful notion of the 'rules of behavior' appropriate to the domain and their role in it, and then try to stick to them.

A Tentative List of Etiquette Rules

Given our experience in working on adaptive automation systems (especially, but not exclusively, the RPA work described above) and our familiarity with others in the literature, we have recently drafted a set of 12 'Etiquette Rules' for adaptive automation system behavior. This list attempts to be general; it is quite clear (given the arguments laid out above) that it should be adapted and extended to any specific domain of interaction. For example for rotorcraft pilots (and, probably, most human operators of high criticality systems), predictability in automation function is very important and should be sought after in interactions. To games players, for example, such a rule might be less important or even reversed. To some extent, this list has emphasized desirable behaviors over practicality of implementation though, as the Crew Coordination and Task Awareness display presented above illustrates, it is sometimes possible to fulfill an etiquette rule (i.e., number six below) with comparatively little interface sophistication. As discussed below, this list is evolving and changing. It is intended to provoke discussion and thought about behavioral standards more than to set those standards itself.

1. Make many, many correct interaction moves for every error made
2. Make it very, very easy to override and correct your errors
3. Know when you are wrong—the easiest way to do this is to let the human tell you—and then get out of the way.
4. Don't make the same mistake twice
5. Don't show off—Just because you can do something, doesn't mean you should.
6. Be able to talk explicitly about what you're doing and why—humans spend a lot of time in meta-communication activities facilitating coordination, especially in distributed work environments.
7. Be able to take instruction; not only will this help you adapt to the user's expectations, it may actually make you look smarter.
8. Make use of multiple modalities and information channels redundantly; understand the implications of your communications on *all* the levels on which it operates.

9. Don't assume every user is the same—be sensitive and adapt to individual, cultural, social, contextual differences
10. Be aware of what the user knows—especially if s/he knows it because you recently conveyed it (i.e., don't repeat yourself).
11. Try not to interrupt. There may be times when something you want to convey is important enough to warrant interruption, but this will usually not be the case. Err on the side of caution.
12. Be cute only to the extent that it furthers your interaction goals.

The Very Idea . . .

It is important to distinguish between the idea that there ought to be a list of etiquette rules for adaptive interface behavior and the specific list we've included above. That list is intended as the starting point for a discussion and series of inquiries whose endpoint might be a full understanding of proper adaptive interface behavior across a wide variety of domains and applications. We can, and should, argue about whether this set of etiquette rules is right or complete. We suspect that the discussion about what works and what doesn't in different contexts will be informative.

Having drafted our list independently and on the basis of our own experience, we found it intriguing to see a very similar list drafted by Dr. Erik Horvitz [in Horvitz, 1999], one of the principal designers of the Microsoft™ 'Paperclip' Office Assistant. Horvitz sees his task as striving to create assistants "with the sensitivity of an intuitive, courteous butler." That is one specific style of etiquette, it may not be right for all circumstances. In future work for our domains, we are intrigued by attempting to apply the lessons learned about intra-crew communication and resource allocation from the extensive work on Crew Resource Management [e.g., Foushee and Helmreich, 1988] to other goal-driven, high-criticality, 'real world' collaboration environments. In any event, the notion that there is an appropriate etiquette for various kinds of human-automation interaction is powerful. Now all that is needed is work designed to discover those 'rules' and to get our machines to play by them.

Acknowledgements

This work was supported by the U.S. Army AATD, contract number DAAJ02-93-C-0008. The authors would like to thank Stephanie Guerlain, Matt Hannen, Mark Ringer, Denny Rother, Erik Steinmetz, Ruben Yanas, Bernie McBryan, and Ray Higgins for their assistance in performing the work reported here.

References

Banks, S. & Lizza, C., (1991). "Pilot's Associate; A cooperative knowledge-based system application." *IEEE Expert*, June 1991, 18-29.

Bauer, M. Dengler, D., and Paul, G. (2000). Instructible Information Agents for Web Mining. In *Proceedings of the International Conference on Intelligent User Interfaces*. January 9-12; New Orleans, LA.

Bonner, M., Taylor, R. and Miller, C. (2000). Tasking Interface Manager: Affording pilot control of adaptive automation and aiding. In *Proceedings of the Annual Conference of the Ergonomics Society*, Grantham, UK, 4-6 April.

Charniak, E. and Goldman, R. (1993). A Bayesian approach to Plan Recognition. *Artificial Intelligence, 64*, 53-97.

Cochran, T., Miller, C. and Bullemer, P. (1996). "Abnormal Situation Management in Petrochemical Plants: Can a Pilot's Associate Crack Crude?" In *Proceedings of the National Aerospace and Electronics Conference (NAECON)*. Dayton, OH; May.

Charlton, S. G. (Ed.). (1989). *Operational Effectiveness Questionnaire Handbook*, Subjective Assessment Working Group, US Air Force Operational Test and Evaluation Center, Technical Paper 7.1, July.

Colucci, F. Rotorcraft pilot's associate update: The army's largest science and technology program. *Vertiflite*, (March/April 1995), 16-20.

Coury B., Santarelli, T. and Mitchell, T. (2000). Expert Modeling for Adaptive Automation and Intelligent Systems. In *Proceedings of Human Performance, Situation Assessment and Automation: User-Centered Design for the New Millennium*. Savannah, GA; October 15-19. 160-165.

Foushee, H.C. and Helmreich, R.L. (1988). Group Interaction and Flight Crew Performance. In Human Factors in Aviation. Academic Press, San Diego CA.

Funk, H., Miller, C., Johnson, C. and Richardson, J. (2000). Applying Intent-Sensitive Policy to Automated Resource Allocation: Command, Communication and Most Importantly, Control. In *Proceedings of the Conference on Human Interaction with Complex Systems*. Urbana-Champaign, Ill. May.

Horvitz, E. (1999). Principles of Mixed-Initiative User Interfaces. *Proceedings of CHI '99, ACM SIGCHI Conference on Human Factors in Computing Systems*, Pittsburgh, PA, May.

Hoshstrasser, B. and Geddes, N. (1989). OPAL operator intent inferencing for intelligent operator support systems. In V. Shalin and G. Boy, (Eds.) *Proceedings of the IJCAI-89 Workshop on Integrated human-Machine Intelligence in Aerospace Systems*. 53-70.

Maybury, M. and Wahlster, W. (1998). *Readings in Intelligent User Interfaces*. Morgan Kaufman.

Miller, C. (2000). Intelligent User Interfaces for Correspondence Domains: Moving IUIs 'Off the Desktop.' In *Proceedings of the International Conference on Intelligent User Interfaces*. January 9-12; New Orleans, LA.

Miller, C. and Riley, V. (1994). "Achieving the associate relationship; Lessons learned from 10 years of research and design." In *Proceedings of the 3rd International Conference on Human Computer Teamwork*. Cambridge, UK; September.

Miller, C., Pelican, M. and Goldman, R. (2000). "Tasking" Interfaces for Flexible Interaction with Automation: Keeping the Operator in Control. In *Proceedings of the Conference on Human Interaction with Complex Systems*. Urbana-Champaign, Ill. May.

Miller, C., Hannen, M. and Guerlain, S. (1999). The Rotorcraft Pilot's Associate Cockpit Information Manager: Acceptable Behavior from a New Crew Member. In *Proceedings of the American Helicopter Society's FORUM 55*, Montreal, Quebec, May 25-27.

Mitchell, C. (1999). Model-based design of human interaction with complex systems. In A.P. Sage and W. Rouse (Eds.) *Systems engineering and management handbook.* New York; John Wiley.

Parasuraman, R., Sheridan, T. and Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans, 30*(3). 286-296.

Perrow, C. (1999). *Normal Accidents*. Princeton University Press; Princeton, NJ.

Puerta, A. (1996). The MECANO Project: Comprehensive and Integrated Support for Model-Based Interface Development. In *Proceedings of CADUI96: Computer Aided Design of User Interfaces*. 1996. Numur, Belgium.

Reeves, B. and Nass, C. (1996). *The Media Equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press; New York.

Riley, V. A general model of mixed-initiative human-machine systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Denver CO, October 1989) 124-128.

Sarter, N.B., Woods, D.D., and Billings, C.E. (1997). Automation Surprises. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd edition) (pp. 1926-1943). New York, NY: Wiley.

Sheridan, T. (1987). Supervisory Control. In G. Salvendy (Ed.), *Handbook of Human Factors*. John Wiley & Sons, New York. 1244-1268.

Strohal, M. and Onken, R. (1998). Intent and error recognition as part of a knowledge-based cockpit assistant. In *Proceedings of SPIE, Vol. 3390: Applications and Science of Computational Intelligence*. April 13-17; Orlando, FL.

Social aspects of wearable control rooms

Charlotte Skourup

ABB AS, Corporate Research Center, Bergerveien 12, P.O. Box 90, N-1375
Billingstad, Norway, charlotte.skourup@no.abb.com

ABSTRACT

A general trend is to equip personnel with wearable computers to increase information access anywhere and anytime. Before implementing new technology, it is essential to answer the question "Why?". The balance between man and technology also needs to be balanced against the organisation. Wearable control rooms blur the boundaries between control room operators and field operators and may introduce the all-round operator. Operators make decisions in collaboration. Wearable computers have a great potential to provide operators with new functionality for communication and collaboration such as video conferencing, shared applications and digital images. There is a risk that transfer of information and knowledge may be lost in a pure virtual control room environment. Social implications of introducing wearable computers into process industry need to be further investigated.

Keywords
Wearable computers, human-machine interaction, cognitive processes, social processes.

INTRODUCTION

A general trend within the military and maintenance & repair is to equip the personnel with wearable computers to increase information access when and where the operators need it. The process control environment has been more cautious of taking this technological step partly due to safety-critical processes and cost. So far, primarily research environments and technology suppliers have initiated wearable computers to operators in the process industry as pilot studies. In this paper, a wearable computer is defined as:

> *A compact, self-contained portable full function computing device which is completely supported by a user for the input, processing and output of information.*

Wearable computers provide the operator with information any time and any place. Various interaction techniques exist to interact with the system. An example is voice control that allows the operator to perform tasks using both hands while interacting with the wearable computer. The benefits of wearable computers seem obvious, but wearable computers may also introduce new challenges and problems that are difficult to predict. First of all, there is a risk that the wearable computer itself takes the focus away from the operator's primary tasks. A beneficial balance between the operator and the technology is an essential condition for success. Not only the man-machine roles may be affected but also the roles between the operators. Today's operators normally work in two teams; The control room operator team and the field operator team. The two operator teams have access to different information sources and perform different tasks. Control room operators and field operators often complement each other and collaborate to a high degree on problem solving. Communication between control room operator and the field operator mainly takes place using two-ways radios. Such communication may cause misunderstandings and errors due to noise and poor quality of the communication. While wearable computers provide other and more reliable communication means such as text and video transfer, they may however influence the two operator roles as a consequence of increased information access. Both operator teams may be almost independent of information from the other team as they already have extensive information access. A technological challenge for wearable computers functioning as wearable control rooms is to support the transfer of individual experience and knowledge, which generally takes place in social settings that are difficult to substitute by the technology. On the other hand, the technology offers new possibilities for social interaction. The wearable control room has the potential to introduce new human relationships and networks such as support from centralised experts, experience exchange with peers from other process plants, and online maintenance & repair guidance from suppliers.

Research on wearable computers supporting tasks relevant for process operators focuses mainly on hardware and software, HCI, operator perception of the systems and system and work performance. These focus areas are absolutely required to develop a useful and beneficial system of wearable control rooms. Boing's pilot studies in the late 1990s aim at increased performance and collaboration for manufacturing workers [Billinghurst and Starner 1999]. General Electric has tested wearable computers for an assembly line for electric power generators focusing on information access [Ditlea 2000]. Researchers at CMU have tested a prototype system for maintenance and collaboration among train maintenance personnel [Siewiorek et al. 1996]. Researchers at Honeywell have applied wearable computers to refinery operators in the petrochemical and oil and gas industries [Guerlain et al. 1999]. Few of these papers include however extended discussions on collaborative operations and social aspects of wearable computers.

The wearable control rooms have the potential to solve problems concerning communication and information transferring. On the other hand, the technology may introduce new challenges in process control concerning collaboration, cognitive problem solving and social relationships. This paper highlights and discusses social aspects of wearable computers as the futures' control rooms for process control operators.

BALANCE IN MAN-TECHNOLOGY-ORGANISATION SYSTEMS

Before implementing new technology, it is essential to answer the question "Why?". Some system designers, engineers and suppliers have an interest in implementing new technology simply because of the technology itself. The answer should unambiguously be to support the operators and their tasks and to increase the organisational performance. Obviously, general wearable computers as "wearable control rooms" are beneficial in some situations whereas they have the opposite effect in other situations. The ideal wearable control room needs to be adapted completely to the operators and their tasks to avoid useless technology. In addition, today's work processes may be influenced by the introduction of wearable computers as the the technology may assist the operator to perform tasks differently and at different places. Not only will the technology affect the operators but also the entire organisation. Therefore, the man-technology system also needs to be balanced against the organisation [Bye et al. 1997]. From the design perspective, the socio-technical structure needs to be considered before going into details concerning the system consisting of operators, technology and the interaction between these components. The effect of wearable control rooms on the operators' roles and the entire organisation is further discussed in separate sections.

The respective roles of a wearable control room and an operator influence the balance between man and technology [Reigstad and Skourup 2000]. In process control, the operator needs an invisible system that provides information anytime and anywhere. The operator is not interested in focusing on the wearable control room itself but to use it as a tool such that the interaction with the wearable computer becomes a secondary task. The choice of technology (hardware and software including HCI) decides to a high degree whether the operators can continue to perform their primary tasks as usual without disturbances and interruptions from the system. Today's wearable computers provide different interaction techniques, which support a low workload from the operator, such as voice control, eye movement control and alternative keyboards [Ditlea 2000; Reigstad and Skourup 2000]. Also, the design of wearable computers varies dependent on the usage. Examples are technologies for information presentation and particular forms of the displays. Billinghurst et al. [1998] have evaluated different display styles including navigation and information presentation based on information display types and user task performance. Particularly, the results are interesting regarding human collaboration between wearable computers. For pure interaction between people, body-stabilised displays provide advantages over conventional head-stabilised displays. Guerlain et al. [1999] have explored two different wearable hardware systems based on a belt-worn wearable computer (commercial-off-the-shelf) with:

- A hard-hat-mounted monocle display device
- A hand-held unit that combines a miniaturised projection display and a three-button user input device

The system was aimed at operators in petrochemical and oil and gas industries for enhanced information access during field operations. For the specific usage, the hand-held device was selected for further investigation due to reluctance regarding having a hat-mounted system that could obstruct the operators' vision, added weight to their head and was likely to get bumped endlessly by obstructions typical of the refinery environment.

THE ALL-ROUND OPERATOR

Operators' physical location give them access to different types of information. The control room operator and the field operator therefore have different roles and perform different tasks. Control room operators control the process from the, often centralised, control room where they have access to information such as measurements, analyses, reports and historical data. Field operators operate directly in the field where they perceive various input from the process ranging from periodically instrumentation readings to sensory input such as heat and vibrations. The two operator teams exchange information to get a more correct picture of the entire process. Control room operators and field operators are experts in their respective fields. During problem solving, decision making and planning, the operators often need additional information than they have direct access to. As they provide each other with the missing information, they also involve the other part in the task. Thus, they utilise an advanced collaborative network to transfer information and individual knowledge [Skourup 1999].

Wearable control rooms change the socio-technical system and blur the boundaries between the two operator roles. Information access independent of physical locations allows the field operator to take over parts of the control room operator's tasks and vice versa. The field operator gets the opportunity to search for information without necessarily interrupting the control room operator. Similarly, the role of the control room operator also changes as the control room operator gets immediate access to information collected in the process by the field operator via the networked wearable control room computer. As a consequence, information anytime and anywhere makes it possible to merge the two roles. The ideal solution from an organisational viewpoint is that the all-round operator has access to information when and where he needs it. Many misunderstandings and errors due to communication problems will be avoided. The total number of operators may also be reduced on a short-term perspective. On the other hand, the wearable control room puts an extended workload and pressure on the all-round operators. They need to be specialists in both the traditional control room tasks and field tasks. In a long-term perspective, the question that should be answered before the introduction of wearable control rooms is which consequences such wearable control rooms have concerning the social dimension in process control [Martin 1993].

COLLABORATION

Operators frequently solve problems in collaboration, particularly between the control room and the field. While an operator requests information, the other operator automatically becomes involved in the decision making process. The individual operator's situation awareness affects collaborative decision making where situation awareness is defined as [Endsley 1995]:

> *The perception of elements in the environment, the comprehension of their meaning in terms of task goals, and the projection of their status in the near future.*

The model of situation awareness describes different individual and environmental factors that influence the operator's behaviour in operating complex systems. When a team of operators works together, team situation awareness is the sum of the situation awareness for each individual [Kaber and Endsley 1998]. Team situation awareness is independent of any overlap in situation awareness requirements among the operators. In case an operator's situation awareness includes a serious lack of, for example, environmental factors the shared situation awareness suffers. Such a situation may result in a decision so that the team goal cannot be achieved. The situation becomes less critical if the operators have the possibility to acquire missing information. Wearable computers with a focus on extending collaboration between the operators have the potential to improve team situation awareness if the individual operator gets access to information that expresses a common understanding of the environment and the process.

According to the theory of team situation awareness, not all operators need exact situation awareness of the entire process. An obvious problem within team situation awareness is if an operator cannot acquire specific information that is needed for a given task [Kaber and Endsley 1998]. A beneficial wearable computer system therefore contains the essential elements of the process required to develop common situation awareness. A wearable computer system can definitely support team situation awareness. There is a risk that wearable control rooms may exclude a significant part of the natural collaboration between operators, as they get extended information access. A critical consequence is that the team situation awareness decreases, and the failure rate due to misinterpretation and misunderstanding increases.

The extreme wearable control room combines the tasks of the control room operators and the field operators. The all-round operators input instrumentation readings in the field while the wearable computer incorporates the traditional control room instrumentation and control capabilities. This optimal wearable control room allows all operators access to basis information whereas parts of the information may be password protected for various user groups. Still, the wearable control room does not guarantee that individual knowledge is made available for the rest of the team. Implicit knowledge and individual experience exist as structures in the brain and can only be made available by special techniques such as training and special designed operator support systems [Skourup 1999]. General information access, training and increased awareness of mental models are techniques to support team situation awareness.

Wearable computers can also aid collaboration. They have a great potential to provide the operators with new functionality such as video conferencing with peers and experts, whiteboard collaboration and transfer of images from the plant to experts and databases for documentation [Billinghurst and Starner 1999; Siewiorek et al. 1996; Ditlea 2000]. Hence, larger networks of operators and experts, also outside a specific process plant, can be involved in decision making processes. Furthermore, collaborative decisions are better documented with the use of camera transfer and shared applications. Introducing technology for new collaborative means that the socio-technical system is likely to change. Therefore, it is essential to investigate these changes before designing and introducing wearable control rooms to the organisation.

SOCIAL ENVIRONMENTS

Information and knowledge transfer often takes place in informal environments. Within process control, the workload varies during a shift and in periods where the operators have less to do. They meet and have coffee and lunch together while they discuss different situations and problems in the process. In this context, individual knowledge and experience become explicit, and the operators transfer such information and knowledge. These settings are difficult to make up in the virtual control room unless the all-round operators still have a centralised control room from where parts of the control takes place. Alternatively, the all-round operators need a resting room for breaks and informal social time together. In addition, the operators will need to spend time during a shift together with colleagues in such a room. There is a risk that the transfer of information and knowledge may be lost in a pure virtual control room environment. The social implications of introducing new technology such as wearable computers are difficult to predict, but significant for a successful result [Guerlain et al. 1999; Martin 1993].

DISCUSSION

The wearable control room introduces both advantages and disadvantages. First of all, the operators can perform tasks with both hands while getting access to information anytime and anywhere. The technology also allows new forms of interaction between operators. They can collaborate in larger networks with peers and experts, also outside the plant. Information transfer and problem solving can be carried out using new media such as digital images, video conferences and shared applications. On the other hand, extended information access blurs the boundaries between the roles and tasks of the control room operator and the field operator. The introduction of wearable control rooms may result in an all-round operator. A risk is that the workload on each operator increases. The traditional and informal information and knowledge transfer between operators dissappears. The wearable control room must be adapted to the socio-technical system. Also, new informal procedures and practices within the socio-technical system must be developed and accepted to secure a successful introduction of this new medium.

Wearable computers are still used rarely in process control. This new technology and its consequences regarding the social aspects in particular needs to be investigated and tested in more depth to give a clear picture of the implications of introducing wearable control rooms into the process industry.

REFERENCES

Billinghurst, M., Bowskill, J., Dyer, N. and Morphett, J. 1998. An evaluation of wearable information spaces. In: Proceedings of Virtual Reality Annual International Symposium, pp. 20-27.

Billinghurst, M. and Starner, T. 1999. Wearable devices: New ways to manage information. *Computer*, 31(1), pp. 57-64.

Bye, A., Hollnagel, E. and Brendeford, T. S. 1997. Task allocation: A functional Modelling approach. In: M. Modarres (Ed.), Proceedings of the Fifth International Workshop on Functional Modeling of Complex Technical Systems, pp. 195-207.

Ditlea, S. 2000. The PC goes ready-to-wear. *IEEE Spectrum*, October, pp. 34-39.

Endsley, M. R. 1995. Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), pp. 32-64.

Guerlain, S., Lee, J., Kopischke, T., Romanko, T., Reutiman, P. and Scott, N. 1999. Supporting collaborative field operations with personal information processing systems. *Mobile Networks and Applications*, 4(1), pp. 37-48.

Kaber, D. B. and Endsley, M. R. 1998. Team situation awareness for process control safety and performance. *Process Safety Progress*, 17(1), pp. 43-48.

Martin, T. 1993. Considering social effects in control system design – a summary. In: Proceedings of IFAC World Congress, 7, pp. 325-330.

Reigstad, M. and Skourup, C. 2000. Interaction with wearable computers in process plants. In: P. C. Cacciabue (Ed.), Proceedings of 19th European Annual Conference on Human Decision Making and Manual Control, pp. 167-171.

Siewiorek, D., Smailagic, A., Bass, L., Siegel, J., Martin, R. and Bennington, B. 1996. Adtranz: A mobile computing system for maintenance and collaboration. IEEE Personal Communications, 3(1), pp. 14-25.

Skourup, C. 1999. Operation Support in Process Plants: A Framework for utilising an Experience Repository. Ph.D. thesis, 1999:66, Norwegian University of Science and Technology.