NDM 2013

Hervé Chaudet
Liliane Pellegrin
Nathalie Bonnardel (Eds.)

Eleventh International Conference on Naturalistic Decision Making

Marseille, France, May 21-24, 2013, Proceedings

ARPEGE Science Publishing

Hervé Chaudet Liliane Pellegrin
Nathalie Bonnardel (Eds.)

Eleventh International Conference on Naturalistic Decision Making, NDM 2013

Marseille, France, May 21-24, 2013

Proceedings

ARPEGE
Science Publishing

ARPEGE SCIENCE PUBLISHING
Conservatoire national des arts et métiers
292 rue Saint-Martin
F-75141 Paris Cedex 03

www.arpege-recherche.org/index.php?page=editions-arpege

ISBN 979-10-92329-00-1

Organization

Co-Chairs

Hervé Chaudet, SESSTIM (Sciences Economiques et Sociales de la Santé & Traitement de l'Information Médicale), UMR 912, Aix Marseille Université, France

Liliane Pellegrin, Centre d'épidémiologie et de santé publique des armées (CESPA), Service de Santé des Armées, Marseille, France

Nathalie Bonnardel, Centre de recherche en Psychologie de la Connaissance, du Langage et de l'Emotion (PsyCLE, EA 3273), Aix Marseille Université, France

Associates Organizing Committee (Doctoral Consortium)

Françoise Anceaux, DEMoH-PERCOTEC, Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines (LAMIH, UMR 8201), Université de Valenciennes, France

Christine Chauvin, Equipe LESTIC , Centre de Recherches en Psychologie, Cognition & Communication (CRPCC, EA1285), Université de Bretagne Sud, France

Scientific Committee

Françoise Anceaux, DEMoH-PERCOTEC, Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines (LAMIH), UMR 8201 Université de Valenciennes et du Hainaut-Cambrésis (UVHC), Valenciennes, France

Nathalie Bonnardel, Research center PsyCLE (EA 3273), Aix Marseille Université, France

Hervé Chaudet, SESSTIM UMR 912, Aix Marseille Université, France

Christine Chauvin, Lab-STICC (UMR CNRS 6285), Laboratoire en sciences et techniques de l'information, de la communication et de la connaissance (EA1285), Université de Bretagne-Sud, Lorient, France

Françoise Darses, Département Adaptation aux contraintes des systèmes opérationnels, Institut de Recherche Biomédicale des Armées, Service de Santé des Armées, Brétigny-sur-Orge, France

Rhona Flin, School of Psychology, Kings College, University of Aberdeen,, Aberdeen, United Kingdom

Anne-Claire Macquet, Service Recherche, Institut National du Sport, de l'Expertise et de la Performance, Paris, France

Leena Norros, Systems Research Centre, VTT Technical Research Centre of Finland, Espoo, Finland

Liliane Pellegrin, CESPA, Service de Santé des Armées, Marseille, France

Neville Stanton, Transportation Group-Engineering and the Environment, University of Southampton, Southampton, United Kingdom

Peter Thunholm, Department of Military Studies, Swedish National Defence College, Stockholm, Sweden

Agnès Van Daele, Sciences Psychologiques et de l'éducation, Université de Mons, Mons, Belgium

William Wong, Interaction Design Centre School of Engineering and Information Sciences, Middlesex University, London, United Kingdom

Additional Rewievers

Paul Ward, University of Greenwich, Chatham, Kent, United Kingdom

Contents

DOCTORAL CONSORTIUM

SENSEMAKING

Operators' conceptions of Procedure Guidance in NPP Process Control

Leena NORROS, Marja LIINASUO, Paula SAVIOJA
VTT Technical Research Centre of Finland
tel. +358405616774
forename.surname @vtt.fi

ABSTRACT

Introduction: Resilience in NPP process control depends on balance between operators' autonomy for decisions and pre-defined guidance.. Finding balance is a strategic decision by the organisation, but finally it takes place at the sharp-end, by the operating personnel. Balancing was studied by analysing the role and usage of procedures. **Method**: Conceptions of 62 control room operators at two Finnish nuclear power plants (NPPs) were queried under three themes relevant to the autonomy-guidance balancing. Answers were classified into theory-based categories, i.e., interpretative, confirmative or reactive, which are epistemic attitudes, portraying more or less potential for resilience in process control activity. **Results and discussion:** In both NPPs, the confirmative orientation, emphasising the importance of acting according to rules prevails. Orientations reflect operator roles and their demands. Interpretative orientation that values operators' competence but does contrast autonomy and procedures should be actively supported to facilitate resilience in NPP activity.

KEYWORDS

Resilience; procedure guidance; autonomy; NPP operators.

INTRODUCTION

In safety-critical domains like nuclear power production there is a good reason to develop different kinds of organisational routines: Commitment to obeying these routines facilitates transparency and predictability in the organisation. The routines also support control of changes in practices and technologies so that the impacts of these changes on the very complex organisation may be comprehended in sufficient breath and depth. These and further reasons for organisational routines are analysed in a thorough review article by Becker (2004). One form of organisational routines in safety critical environments are operating prodecures which are important resources for the personnel in normal daily work, and are particulary significant tools in coping with demanding disturbance or accident situations.

Advances in automation technology, and especially the digitalisation of process automation including improved possibilities of presentation of process information, appear to increase interest to standardise further the human-automation interaction. As an example we may take the French computerised control room of the N4 nuclear powr plant (NPP) in which implementation of computerised emergency operating procedures led to increased procedure control of operations in distrubane and accident situations. Operating experience and human factors studies indicated, however, that procedure control probably was too tight and inoptimal causing some unanticipated difficulties to the operating crews to understand the process situation while applying procedures during the event and coordinating own actions appropriately (Filippi, 2006).

While it may be effective and sufficiently safe to rely on standardised ways of taking the plant into control in desigh-based accidents, concerns have arisen in particularly after the Fukusihima nuclear accident concerning the capability of the operating organisations and personnel to act in innovative ways in very complex and un-anticipated accidents. A too straight-forward reliance on stadardisation as a means to improve safe operations has also been one of the issues raised by the so-called resilience engineering school (Hollnagel, 2006). The claim is that a continuous balancing between innovative and standardised forms of acting is needed for maintaining the system within safe boudaries of operations. Also Papin discussed the issue of creating resilience in the safety-critical organisations (Papin, 2010). Papin's perspective is the design of complex systems. On the basis of analyses concerning optimal operator guidance the author concluded that when searching for resilience in NPP processes, the problem is to find a justified balance between the actors' autonomy and guidance.

Most studies that raise the problem of balancing between guidance and autonomy consider the issue on the level of organisational safety management. The intention of our study was to try to understand resilient behaviour on the level of the sharp-end actors, in our case main-control room operators daily practices. We assumed that the

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing.
ISBN 979-10-92329-00-1

operators have conceptions of the constraints regarding striking balance between autonomy and guidance, and that they have developed practices to do that in their daily work. The operators' conceptions of an appropriate balance can be uncovered by investigating operators' opinions of the role of procedures and of the importance of professional skills in their work. The operators' conceptions are of course influenced by the position that the organisation has taken to the very issue.

METHODS

We conducted a simulator study focusing on an analysis of operators' activity in disturbance and accident situations. As part of the study, interviews about the NPP operators' professional orientation, i.e. epistemic attitude towards work, were accomplished, including operators' conceptions of the role of procedures in work and professional expertise. Data was gathered via observing activity and interviews. This paper focuses on the interview data.

Interview Characteristics

The interview was conducted in two Finnish NPPs during the years 2008 and 2009. In the first NPP (NPP1) all operators (12 shifts, 44 operators including two trainees) were interviewed. In the other (NPP2), half of the operators (6 shifts, 18 operators) were interviewed. Interviews took place in the beginning of a simulator session and lasted about 20 minutes in maximum. A number of themes were covered as the scope of the interview was wider than just procedure usage. Here, only a subset of questions is focused on, namely the ones that is related with the dimension of autonomy vs. guidance in operator's work.

The dilemma of the balance between autonomy and guidance in operator's work was dealt in the interviews with the themes of

(1) Control of action
(2) Initiating action
(3) Concept of a good operator.

Control of action and Initiating action are directly related to the question of autonomy and guidance whereas the Concept of Good Operator sheds light to the question from a more general viewpoint revealing the operators' basic assumptions concerning the demands of their work.

Data concerning the three above-mentioned themes was acquired by posing the following interview questions:

- What is the role of procedures in process control?
- Do such situations exist to which no dedicated procedures exist?
- Do the procedures determine the course of actions totally in some situations?
- Are alarms the primary starting point for action?
- How would you characterise a good operator?

The interviews were accomplished individually in a face-to-face contact and they were audio recorded. Eventually, the recorded interviews were transcribed into written protocols and analysed.

Interview Data Analysis

In each reply, the operator's conception is classified in one of the three categories: reactive orientation, confirmative orientation and interpretative orientation. This classification draws on the ideas of the American pragmatist philosopher Charles Sanders Peirce (Peirce, 1998). In his writings he proposes that in coping with the contingencies of the environment a human being has a possibility to survive via his/her interpretative effort, i.e. via personal attempts to make sense of the environment by developing hypotheses and forming beliefs about it. A contrary basic form of facing the environment is reactivity, i.e. by being directly steered by the events of the environment. On the basis of this distinction we developed the classification of people's orientation to their work, and used the classification with regard to different aspects of work orientation (Norros, 2004). Interpretativeness is characterised by presence in the particular situation and questioning the observed phenomena, building hypotheses concerning the situation and future events. Reactiveness reflects passivity and lack of hypotheses concerning the situation; confirmativeness is an intermediate position and is characterised by mapping the situation with known possibilities and acting in a predefined ways. In the following, the analysis is described according to the factor in question. Also an example is provided for each reply type. The theory-based claim is that interpretative qualities are assumed of such practices that are capable of facilitating resilience in action (Norros, 2012).

Control of action

The theme Control of action focuses on the operators' conceptions of agency in process control and how agency is seen to be distributed among the operator himself and the procedures. The issue of control was investigated by four different sub-themes (in italics below).

Role of procedures : This sub-theme was studied by the question "What is the role of procedures in process control". The replies were categorised according to the following principles :

- Interpretative: Procedures provide framework and method according to which the operations are performed. For example: "We operate according to them. They provide such a framework that they support [our work]."
- Confirmative: Procedures support actions, procedures are needed in rare or demanding situations; procedures unify the ways of working and quarantee safety. For example: "They are central. About everything is done according to the procedures. Of course, routine tasks are such that procedures are not needed. But procedures are important. And we must work according to them."
- Reactive: Procedures support if the operator is insecure or when he has problems with memory; procedures prevent memory-dependent or other mistakes; procedures define the correct procedures and their order; put-down of procedures. For example: "Procedures support my work and provide security, you can have a look at them whenever in doubt."

Coverage of procedures : For this sub-theme, the question "Do such situations exist, to which no dedicated procedures apply" was used. The replies were categorised as follows :

- Interpretative: States the principle that it is not possible to define everything with procedures. For example: "Usually, procedures determine but there are situations for which there are no procedures, then we think what to do. You cannot write a procedure for every situation, it is impossible".
- Confirmative: Surely there can be some unexpected situations or small incidents for which there are no procedures. For example: "Yes, and there can be also such situations that we have not met yet. When we changed the ejector… (the interviewee describes a situation for which there was no procedure)."
- Reactive: Possibly such situations exist which have no procedures but such situations do not come to the interviewee's mind; procedures describe the total of work practises except for some routines or revision. For example: "Yes I guess there are, but I don't remember right now any example."

Focus of procedure control. This sub-theme was studied by the question "Do the procedures determine the course of actions totally in some situations". The replies were categorised in the following way :

- Interpretative: The focus is on the control of process and process disturbancies; the grounds for procedures are found interesting. For example: "No, I don't think so, it can never be like that [that procedures would determine the proceeding of actions totally in some situations]. Of course procedures are the framework for what we do but if it, if you feel that this is not applicable, and of course for us it is the shft supervisor who conducts this ballet, so we talk with him, that now it feels like this will not go according to procedures".
- Confirmative: The focus is on tasks that the procedure is for; what should be done in each situation. For example: "Well it [the role of procedures] is very important. Whatever we do, we have procedures for it. Especially in disturbancies we get an alarm, alarm in the system, so then we start to look from the procedures for distrubancies for what to do regarding this alarm. And then, similarly, were it start-up or shut-down or a scheduled testing in question, what ever, for everything we have prcedures and we perform according to them."
- Reactive: Procedure is a means to guide human; we operate according to procedures; procedures determine operations. For example: "Well… In principle you can say like that, because we have to operate according to procedures (…)".

Procedures and professional competence : The sub-theme was studied, again by the question "Do the procedures determine the course of actions totally in some situations". Even if the question is the same as with the previously described subtheme, its analysis is different. The replies were categorised by the following principles :

- Interpretative: The usage of procedures is understood to facilitate work and the operators' understand that the situation may deviate from the one that the instruction is based on ; thus, procedures must be trusted on and procedures must be used professionally,.
- Confirmative: You are obliged to perform according to procedures, procedures are not questioned. For example: "Everything is done according to procedures. They are important, because they are designed in advance and produced for taking care of the work".
- Reactive: Procedures and professional competence are perceived as opposites, procedures compensate for experience and vice versa. For example: "Yeah, it can be that when you haven't been here for a long time, yes and such situations that repeat very seldom: you might be forced to perform according to procedures in such situations".

Initiating action

While the control of action focused more on the operators' conception of agency in on-going activity, this theme was tuned to identify how activity is initiated in the first place. For finding out what triggers actions, a following question was asked: "Are alarms the primary starting point for action". The replies were classified as follows :

- Interpretative: Repudiates the idea presented in the question that alarms would be the main trigger for action. Operator identifies that observation and monitoring are active actions, he considers weak signals, and finds anticipation most important in acting. For example: "Not really, you have to monitor the process, it doesn't… Changes are primary. By following the process we strive for, if some process is starting to change we want to see what it originates form. An alarm may follow later. It is the monitoring of the process that is primary and alarms only come after that".
- Confirmative: Alarm or other external factor such as some deviation in the process is the starting point, hihglights context dependence. For example: "Alarms at the latest, then we start action. We may get indication also before alarms. When possible we try to react upoin them immediately if we see such things. Some temperature is climbing or a blockage (…)".
- Reactive: Alarms trigger action. For example: "Well yes they are. Depends on the disturbance, if it is a big one, these are gone already, it has no meaning. [If the disturbance is minor] then it goes according to alarms".

Concept of a good operator

The theme Concept of a good operator was studied as an indicator of what the operators within their community of practice consider as professional acting and hold meaningful and valuable (drawn from the idea of "internal good of practice" by Alasdar McIntyre (1984)). It was assumed that the issue of autonomy would probably be reflected in the conceptualisation of a good operator. This them was studied by the question "What are the qualities of a good operator". The replies were classified in the following way.

- Interpretative, personal qualities: Explains features of practice that have a positive contribution to work and its purposes. For example: "Good education and level of knowledge and skills so that you understand what happens in the plant; you have to have an overview so that you can interpret very small incidents [whether it is serious or not]. Calm, capable of analysing so that when it gets tough you can think reasonably".
- Interpretative, qualities describing action: Personal interpretation, understanding. For example: "but a good operator… knows the process, knows what affects and where it affects, can piece together these what happens next, these consequences".
- Confirmative, personal qualities: Connected to work in a superficial way. For example: "Takes care of his worries and employees, does what is needed and ordered, may not perform independently of others".
- Confirmative, qualities describing action: Relates to rules. For example: "Can read this process, knows the procedures, can use them correctly. One can perform correctly in rapid situations".
- Reactive, personal qualities: General characteristics of performance that are not connected to work. For example: "Accurate and prompt and conscientious".
- Reactive, qualities describing action: Superficial, reacting to demands from the outside. For example: "Takes care of his duties conscientiously".

RESULTS

The results of the analysis of the interview data are presented from the two plants separately and, also by role (shift supervisor, reactor operator and turbine operator) within the plant. Due to limited number of pages, the results are presented in the level of themes only.

Results for the theme Control of Action

Regarding the conceptions related to the role of procedures, the share of confirmative orientation of both was prominent in both NPPs (see Figures 1a and 1b). Operators also think that a procedure will be found for disturbance situations, except for unexpected or small disturbances. According to the majority of operators of both plants, the procedures steer accomplishment of the task and must be used without questioning it at all.

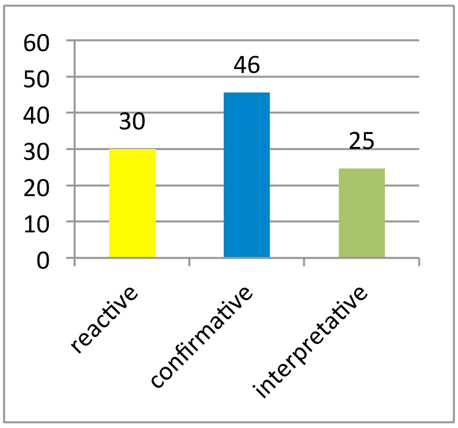

Figure 1a. Conceptions related to Control of Action in NPP1; N= 44. The share of replies as percentage in each category (reactive, confirmative and interpretative).

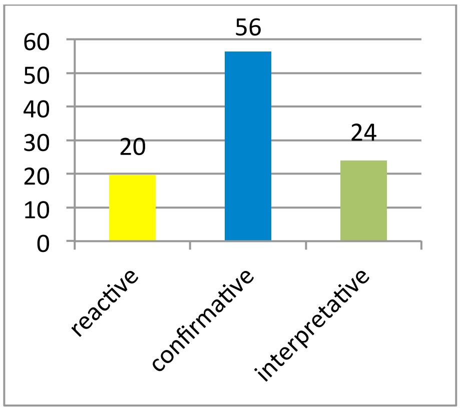

Figure 1b. Conceptions related to Control of Action in NPP2; N=18. the share of replies as percentage in each category (reactive, confirmative and interpretative)

When scrutinising the replies from the perspective of the various operator roles, some differences were found in the percentages. In both NPPs, the turbine operators expressed interpretative orientation the most and reactive the least. The contrast to reactor operators' far less interpretative orientation is clear. (see Figures 2a and 2b).

The higher share of interpretative orientation among turbine operators in the control of actions may be due to the task demands in turbine operations: The turbine systems are typically not safety classified and, hence more, freedom in operations is allowed for on-line decisions. Moreover, a great number of active operations from turbine process are typically needed in the controlled reduction of power and cooling the plant. In contrast, caution and patience in dealing with the reactor systems is evidently reflected in a stronger confirmative orientation what regards the reactor operators' control of action.

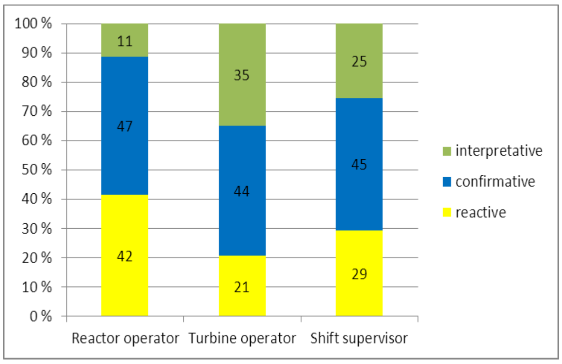

Figure 2a. Conceptions related to Control of Action classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP1; N=44. The share of each category (reactive, confirmative and interpretative) shown as percentage.

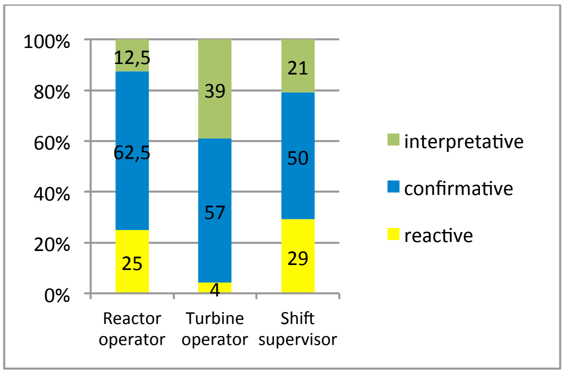

Figure 2a. Conceptions related to Control of Action classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP2; N=18. The share of each category (reactive, confirmative and interpretative) shown as percentage.

Results for the theme Initiating action

The conceptions of the operators of the two power plants about how actions are initiated are depicted in Figures 3a and 3b. The histograms indicate that in NPP1, most operators found the alarms (reactive orientation) or other external events (confirmative orientation) as triggers for actions, whereas operator-driven anticipatory identification of changes in the process were mentioned quite rearely. In NPP2, the share of conceptions were equal in all categories so that the same amount of operators found weak singals and proactive work to be the important way of working (interpretative orientation) as the amount of operators that valued alarms as a primary signal for action (reactive orientation) or found external factors in general as triggering events (confirmative orientation).

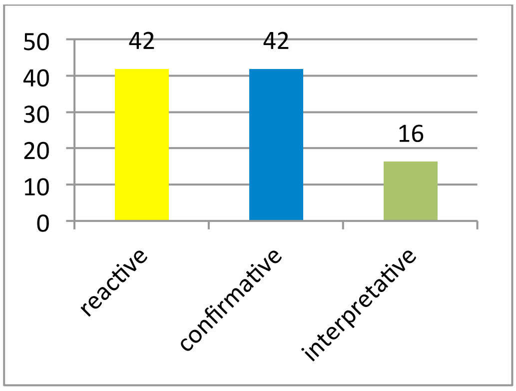

Figure 3a. Conceptions related to Trigger for Action in NPP1; N=44. The share of replies as percentage in each category (reactive, confirmative and interpretative).

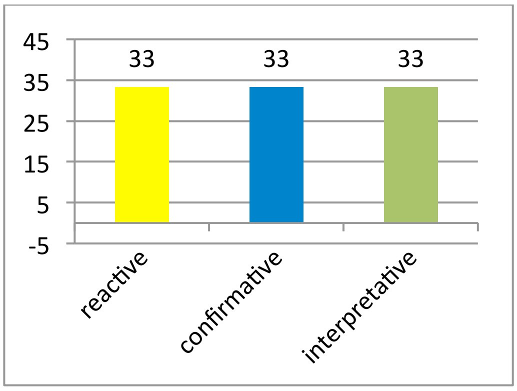

Figure 3b. Conceptions related to Trigger for Action in NPP2; N=18. The share of replies as percentage in each category (reactive, confirmative and interpretative).

The results concerning Trigger of Action were also arranged acccording to operator roles in both plants (see Figures 4a and 4b). Evident difference between the plants become visible. In plant 1 reactive orientations are typical for all roles, and the share of interpretative orientation scarce. In NPP2 reactor operators and shif supervisors of the portray high numbers of interpretative orientation with regard to initiating action, which speaks of proactiveness in work. Wea may also notice that in both plants the turbine operators show least interpretativeness and most reactiveness. If we se compare these results of Trigger of Action to those with regard to Control of action (Figures 2a and 2b) it appears that the found interpretative orientation by turbine operators in action is connected to reactiveness in starting action.
The results complete the earlier interpretation that the role and demands of work reflect readily in the orientation. Turbine opeators tend to be automous in acting when the situation demands that, whereas the shift supervisors and reactor operators express proactiveness and anticipation the process sistuation, while they are confirmative and cautios what regards action. This interpretation holds in particular with regard to Plant 2.

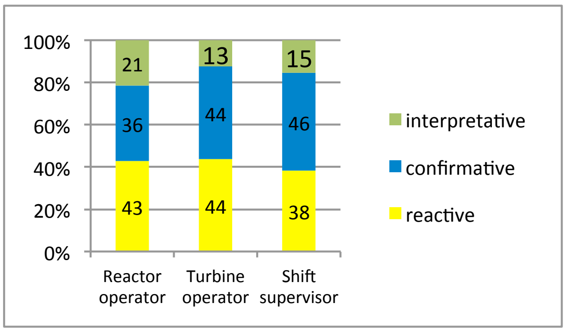

Figure 4a. Conceptions related to Initiating Action classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP1; N=44. The share of each category (reactive, confirmative and interpretative) shown as percentage.

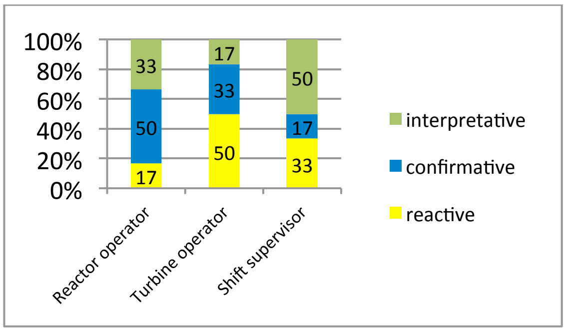

Figure 4b. Conceptions related to Initiating Action classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP2; N=18. The share of each category (reactive, confirmative and interpretative) shown as percentage.

Results for the theme Concept of Good Operator

Opinions regarding the qualities of a good operator were emphasised slightly differently in the two NPPs. In NPP1, the share of interpretative orientation was prominent, meaning that operators valued qualities justified by their effect on work practices and that they produced personal interpretations of good personal qualities (opposed to valueing the meeting of external demands) (see Figure 5a). In NPP2, the opinions of confirmative type were expressed more often (see Figure 5b). Such opinions are characterised by connecting good qualities to work in a superficial way and that the performance according to rules, "doing it correctly" is regarded as of highest importance. It is interesting that in NPP1, where orientations with regard the starting and control of action tended towards confirmative or even reactive direction, the operators still find interpretative qualifiers of practices and professional identity most appropirate. In the case of NPP2 the tendency for a dominantly confimative orientation in acting is strengthened by a dominantly confimative orientation also what regards an ideal operator.

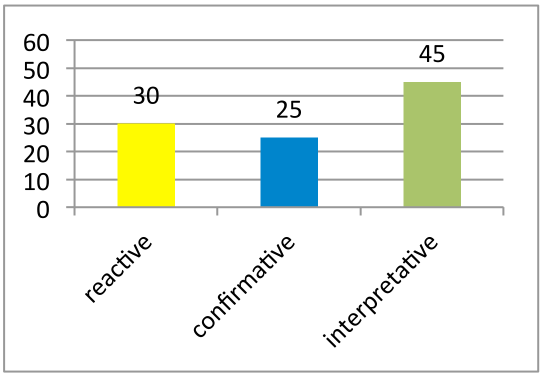

Figure 5a. Conceptions related to Concept of Good Operator in NPP1; N=44. The share of replies as percentage in each category (reactive, confirmative and interpretative).

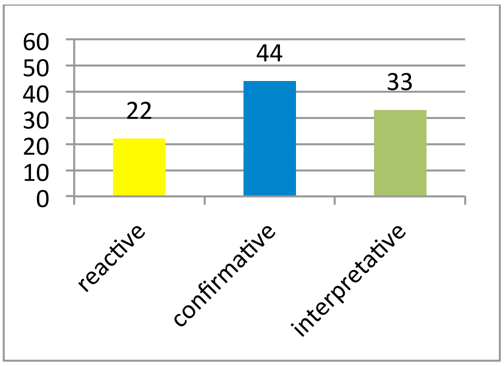

Figure 5b. Conceptions related to Concept of Good Operator in NPP2; N=18. The share of replies as percentage in each category (reactive, confirmative and interpretative.

Investigating the replies from the perspective of operator roles, it becomes clear that the stronger interpretative orientation in the NPP1 compared to NPP2 appears to be due to the clearly stronger interpretative orientation of the turbine operators in the plant 1. In both plans the shif supervisors conceptions of good operator are both strongly interpretative. (see Figures 6a and 6b).

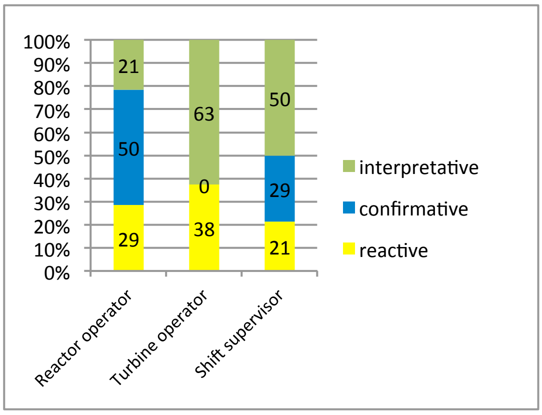

Figure 6a. Conceptions related to Concept of Good Operator classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP1; N=44. The share of each category (reactive, confirmative and interpretative) shown as percentage

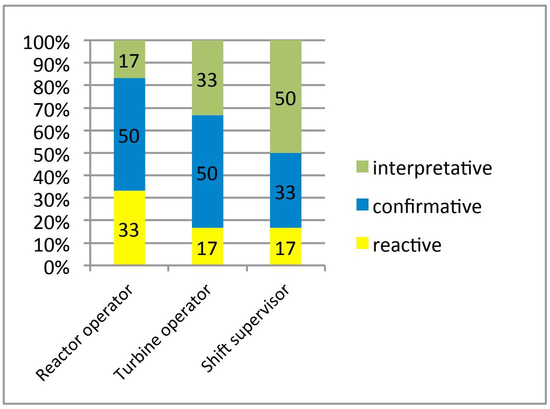

Figure 6b. Conceptions related to Concept of Good Operator classified according to the operator role (reactor operator, turbine operator, shift supervisor) in NPP2; N=18. The share of each category (reactive, confirmative and interpretative) shown as percentage

Finally, we summed up the findings with regard to all indicators of procedure orientation by both plants. The results indicate that in NPP1 the overall share of interpretative orientations was 27%, confirmative 41% and reactive 32%. In the NPP2, the same percentages are interpretative 27%, confirmative 50% and reactive 22%. , The distributions are not statistically different ($\chi 2$= 7,61, df=2, p=0.472), indicating that overall the orientations to procedures are equal in both plants. When the data from both plants is pooled and the differences in orientations between operator roles were analysed, statistically significant differences were found with regard to the Control of action, specifically in the Focus of Procedure Control (F2,58 = 4,565, p=0,014) and Procedures and Professional Competence (F2,58 = 5,333, p=0,007). The work orientations regarding Focus of Procedure Control (t(41,884)=2,865, p=.015) and regarding Procedures and Professional Competence (t(35,067)=-3,076, p=0,008) indicate higher level of interpretativeness by turbine operators compared with reactor operators.

DISCUSSION AND CONCLUSIONS

The conceptions of procedure guidance and operator autonomy express the operators' personal interpretation of how these demands, as well as goals of the production should be met. In both NPPs, the confirmative orientation, emphasising the importance of acting according to rules and procedures, prevails. This corresponds to the official principles of NPPs.

It can be argued, however, that an interpretative attitude towards initiating and control of action bring forward features which could support appropriate acting under changing demands and in possible unexpected situations. Hence interpretativeness is a qualification that would facilitate on-line resilience in the system. Our results indicate that operators themselves value interpretative features as those charactering a good operator. Against this, the overall share of interpretative conceptions being 27% in both NPPs appears unnecessary low. Some reference to this result is provided from an analysis of the conceptions of safety and the role of procedures to maintain safety among two other safety-critical domains: Pezzullo and Filippo (2009) identified conceptual differences among drivers transporting hazardous materials. They found that only about 17 % of the expressed conceptions belong to the category of "green workers" which has connections to our definition of interpretative orientation. In an analysis of expert anaesthetists orientation to their work Klemola and Norros (1998) applied a corresponding method as used here and found that 33% of the expressed conceptions portrayed an interpretative attitude.

Especially low share of interpretative orientation among shift supervisors and reactor operators, and stronger interpretative orientation among the turbine operators with regard to control of action, but a stronger emphasis on interpretative and anticipatory practices in initiating action by reactor operators and shift supervisors, demonstrate that process characteristics and the role-based task demands are reflected in orientations in a consistent way. Due to the possible role of interpretativeness for resilience the development of this epistemic attitude should be supported beyond the level of its spontaneous development in daily activities.

According to our understanding, the interpretative orientation does not posit procedure-based action and professional competence against each other. Instead, such an orientation means that procedures are used reflectively, both using professional competence and taking the situational demands and constraints into account. Thus, we claim that interpretative orientation goes beyond the dichotomy of guidance and autonomy. It enables the comprehension of procedures and their intelligent use as part of professional competence and as an expression of resilient acting. Interpretative orientation would be especially important in responding to beyond design bases situations, for which no procedures exist. From this point of view the relative low share of this orientation should be of concern by the power plant operative management.

The results of this study will in the next step be compared with results of the same operators' actual practices in proceduralised process control situations (Savioja, Norros, & Salo, submitted). The characterisation of orientations and practices with regard to the strength of interpretative attitude these portray is a novel way of comprehending resilience in the sharp-end activities, which process has so far not been sufficiently elaborated empirically in the resilience engineering literature.

REFERENCES

Becker, M. C. (2004). Organizational routines: a review of the literature. Industrial and Corporate Change, 13(4), 643-677.

Filippi, G. (2006). Studying computerized emergency operating procedures to evaluate the impact of strong procedure guidance on operators' work practices. Paper presented at the American Nuclear Society international topical meeting on nuclear plant instrumentation, controls, and human-machine interface technologies (NPIC&HMIT), Albuquerque, New Mexico.

Hollnagel, E. (2006). Resilience - the challenge of the unstable. In E. Hollnagel, D. Woods & N. Leveson (Eds.), Resilience Engineering. Concepts and precepts (pp. 9-17). Aldershot: Ashgate.

Klemola, U.-M. (1998). Analysis of the anaesthetist's activity as the basis for the development of the content of professional expertise (in Finnish). Finnanest, 331, 273-281.

MacIntyre, A. (1984). After Virtue: Study in Moral Theory (2nd ed.). Notre Dame, Indianana: University of Notre Dame Press.

Norros, L. (2004). Acting under Uncertainty. The Core-Task Analysis in Ecological Study of Work (Vol. Publications 546). Espoo: VTT, Available also URL: http//www.vtt.fi/inf/pdf/.

Norros, L. (2012). Analysis of work practices from the resilience engineering perspective. Nuclear Safety and Simulation, 3(4), on-line.

Papin, B. (2010, 14.-19.3.2010). Operator guidance in plant operation: Characterization and orientations for better team performance and socio-technical systems resilience. Paper presented at the Enlarged Halden Programme Group Meeting, Storefjell, Norway.

Peirce, C. S. (1998). The Peirce Edition Project. Introduction. In T. P. E. Project (Ed.), The essential Peirce. Selected philosophical writings (Vol. 2, pp. XVII-XXXVIII). Bloomington and Indianapolis: Indiana University Press.

Pezzullo, L., & De Filippo, R. (2009). Perceptions of industrial risk and emergency management procedures in hazmat logistics: A qualitative mental model approach. Safety Science(47), 537-541.

Savioja, P., Norros, L., & Salo, L. (submitted). Identifying resilience in proceduralises accident managment activity. Safet Sciene.

Tactile Reasoning: Hands-on vs Hands-off – what's the difference?

Stephan TAKKEN[1] and B. L. William WONG[2]
[1]Utrecht School of Art & Technology, Netherlands
[2]Interaction Design Centre, Middlesex University, UK

ABSTRACT

This study investigates tactile reasoning in relation to sense making in the context of designing a new non-traditional information environment. We define tactile reasoning as an interaction technique that supports analytical reasoning by the direct manipulation of information objects in the graphical user interface (GUI). When people directly manipulate data, for example, by moving individual pieces of information to create temporary groups or sequences, or eliminating pieces of information from a group. We hypothesise that this can enhance their sense-making and analytical reasoning ability by helping them discover new explanatory relationships created by the re-arranged pieces of information. Our study used a card sorting type of task where participants were either allowed or not allowed to touch and manipulate the cards to look for information or to construct groups of information that could provide meaningful explanations. The results showed that manipulation has a positive effect on analytical reasoning performance. In the more difficult task conditions, participants who were allowed to use their hands were observed to invoke sense-making strategies 99 times in comparison with the 50 by participants who were not allowed to use their hands.

KEYWORDS

Tactile reasoning, analytical reasoning, sense-making.

INTRODUCTION

In this paper we briefly report on our observation of sense-making strategies invoked by participants in a study that investigates how tactile reasoning affects the sense-making task. We define tactile reasoning as an interaction technique that supports sense-making by the direct manipluation of information objects in the user interface. We believe that when one is presented with a set of information that can be freely moved, manipulated, grouped or re-arranged in a visuo-spatially manner, this interaction method can help us discover meanings or relationships. Such actions are externalisations of our mental processes, and have been referred to as epistemic actions (Kirsch and Maglio, 1994). By off-loading cognitively demanding mental computations in appropriate ways, we can obtain insight by '… organising intelligence that places the full set of clues in a unique explanatory perspective' (Lonergran, 1957, page ix). In this study we used a card sorting task as where participants were either allowed or not allowed to touch and manipulate the cards to look for information or to construct groups of information that could provide meaningful explanations. We believe this is a close enough parallel of the electronic index cards that can be individually manipulated in the user interface we have developed in our research for investigating novel user interfaces. This research prototype, called INVISQUE – Interactive Visual Search and Query Environment – was originally designed to support information search and discovery in electronic library resources (Wong, et al, 2010). INVISQUE is being extended for use in a variety of complex information analysis and sense-making tasks such as intelligence analysis. Other studies found that it improved the information search and sense-making performance of low literate users (Kodagoda, et al, 2012). As part of this research, we wish to investigate how the generation of insight can be facilitated by aspects of the user interface and interaction design that create the ability to manipulate information objects in a user interface. In the study reported here, we investigate the aspect of being able to freely manipulate information.

BACKGROUND

In this study we were interested in understanding the differences in the sense-making task that may arise from the ability to manipulate, re-organise and move the cards. Sense-making has been defined as '… a motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively' (Klein et al, 2006). It is a cyclic process that does not have a clear start and end point as sometimes suggested where data => information => knowledge => understanding (Ackoff, 1989).

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing.
ISBN 979-10-92329-00-1

There are several models of sense-making (e.g. Pirolli and Card, 2005; Weick, 1995), but for reasons of brevity, we will briefly describe Klein et al.'s (2007) Data-Frame Model of Sense-making. They propose a non-linear model comprising four main cognitive activities that can happen at any time and can transition fluidly between activities: (i) Connect, when new data, say representing a situation, connects with our frame, it enables us to understand what the situation is about. We may hear explosions, and part of our current frame (understanding) is that it is New Year's Eve, so those explosions followed by the bright flashes in the sky suggests it is a fireworks display; (ii) Elaborate refers to the activity of finding out more information to improve our understanding of the situation or issue; (iii) Questioning refers to the activity of determining the accuracy or validity of one's conclusions or assumptions about the situation; and (iv) Re-framing refers to the activity of changing our understanding given new data or a new understanding of the data. Can the physicality of interaction affect this sense-making process ? We briefly review some of these studies next.

Kirsch and Maglio (1995) describe how complementary strategies, i.e. the complementary use of external elements to support cognitive process, reduce cognitive load in mental computations. For example, the use of pencils, paper or hands can reduce cognitive load to make a task easier. In the study they showed participants a set of cards depicting currency coins, i.e. quarters, dimes and nickels. The participants had to compute the total amount in dollars and cents. In one condition they were allowed to use their hands to organise the cards, while in the second condition they were not. The results showed that the complementary strategy of using hands improved performance. The notion of complementary strategies as used by Kirsh supports the view that when we off-load cognitive work into the environment, we can improve mental computation performance (Wilson, 2002).

Kirsch and Maglio (1994) also explain that the external actions that accompany these complementary strategies are known as epistemic actions, i.e. '… physical actions that make mental computation easier, faster, or more reliable, … and performed by an agent to change his or her own computational state' (p513-514). These epistemic actions have developed and are used in order to uncover information that is hidden or hard to cognitively compute. For example, sorting nuts and bolts before assembling a piece of furniture, or in more abstract tasks such as algebra and arithmatic, we simplify the mental computation by manipulating external symbols where intermediate results '… which could in theory be stored in working memory, are recorded externally to reduce cognitive loads' (p514).

Antle et al (2009) differentiates even further in a study where she analysed children's hands actions when solving a jigsaw puzzle. She described three types of manipulation: First direct placement actions, which basically means moving information in the correct position. Second she describes indirect placement actions as a complementary pragmatic action that are usually '...part of a trial and error approach to visual search' (p2); and third, exploratory actions*,* where information is being manipulated to facilitate the simplification of a task by changing the environment such as rotating information to the correct orientation. These are epistemic actions. These studies explore and describe the ways people use external elements to simplify tasks.

Maglio et al. (1995) used tiles from the game Scrabble to see if epistemic actions aid in the discovery of anagram solutions within two different sets of seven tiles. The study showed that the use of hands to manipulate the tiles provided people an edge over those who did not use hands. However, the former only applied in the conditions where the sets of tiles had a low frequency of anagram solutions. This implies that there might be a threshold on the applicability of complementary strategies in making sense of information. Manipulation might aid in situations where the information is ambiguous and patterns are difficult to discover. This threshold is, however, not the focus of our study. We are interested in the differences in how people make sense of data and how complementary strategies influences sense making.

METHODOLOGY

Participants

There were 24 participants (13 males and 11 females) with ages ranging from 20 to 58 years with a mean of 30 (*SD* = 9.29). All participants were PhD students, minimising variation in reasoning skills.

Materials and apparatus

Five sets of cards were used in the experiment, one set with 10 cards for a practice session and four sets with 22 cards each for the main experiment. The four sets were used in a counter-balanced way to minimise learning effects. The cards depicted animals, living environments and vegetation. For example, one set would consist of 16 animals, 3 kinds of vegetation and 3 living environments.

Design

The experiment was based on a 2 x 4 within subjects, related samples design. The primary dependent variable was the number of cards found; other dependent variables were the time to complete the task (in seconds) and the quality of categories. The first related measures independent variable is the problem type, with two levels (an easy 'known-item search' to a hard 'ill-formed query') with a total of four questions divided in these two levels. The second related measures independent variable is physical manipulation, either hands could be used or they could not. The conditions were counterbalanced between participants.

Tasks

There were two types of tasks, a 'known-item search' task and an 'ill-formed query' task. Due to the nature of the known-item search task, being a visual search on a small number of features, we regarded the known-item search as a low difficulty task. This task consisted of two questions where participants were first asked to find information on two and then three features. For example, "*Identify carnivore mammals.*" (E1) and "*Identify carnivore mammals in warm climates*" (E2).

The ill-formed query tasks were intended to be higher difficulty due to the large amount of possible groupings and categories. This task also consisted of two questions "*Construct one or more food chains*" (D1) and "*In how many ways can you categorize the cards*" (D2). As these tasks have a higher degree of difficulty we expected the use of hands to improve performance.

Figure 1. The experiment setup.

Procedure

Each participant performed a practice session to get familiar with the experiment setup and method. This trial was identical to the actual experiment but used a smaller card set (i.e. 10 cards) and an easier task (i.e. one feature search). Participants carried out the experiment individually, they sat at a table and were provided with a brief introduction to the experiment on a laptop. From hereon in separate subsequent screens the participant saw, **a)** which set of cards to lay out on the table which were then provided by the experimenter, **b)** whether they could use their hands or not and **c)**, the time limit for every task (120 seconds). The participant was asked to think out loud during the trials and was not allowed to take notes.

Table 1. Observed cognitive strategies hands-on, hands-off

	Sense-making strategies observed (n = 24)	**Hands-on number of observations**	**Hands-off number of observations**
D1	Exploration of and experimenting with data	20	8
Construct one or more food chains	Create, fill and finalize	8	3
	Task immersion	12	0
D2	Exploration of and experimenting with data	19	10
In how many ways can you categorize the cards?	Create, fill and finalize	19	8
	Create and illustrate	3	13
	Inverted pyramid categorisation	14	8
	Task immersion	4	0
	Total number of observations	99	50

RESULTS

After transcribing the video and audio data, we employed the Emergent Themes Analysis approach (Wong, 2003) to analyse the data. Drawing from Grounded Theory, ETA seeks to discover higher level activities, behaviours and critical cues of individuals, and to then identify common patterns and interesting themes that may emerge across the 24 participants. The results are summarised in Table 1, focusing on the more difficult tasks (D1 and D2).

Exploration of and experimenting with data

In the hands-on condition, when participants were asked to construct a food chain, participants tended to explore and experiment with the data. Participants appear to use the ability to manipulate the data to search for relationships in the dataset. For example participants seem to search for relationships in the card set by moving cards together. Visually indexing the dataset gives them a general indication of what data is available; they then can start to search for relationships in the dataset.

> "*Maybe, the bird eats the spider*" [moves the parrot and the spider to the top of the workspace]
>
> (Participant 22)

In the hands-off condition participants were observed to be less likely to explore the dataset and came up with conclusions early on in the task. Stating findings early on in the task suggests that participants are less concerned with experimenting and exploring as in the hands-on condition. Though there might have been a form of mental experimenting and exploring participants seldom verbalized this process and merely stated their findings. Participant 3 for example spend some time thinking in silence for a few moments and then came up with a few conclusions. This first part of the excerpt "Ok, I think I have some" indicates the participant to have made an initial 'rough idea' of the groups. After this initial connection is made the groups then get elaborated and finalized.

> "*Ok, I think I have some... The fly, spider, frog... frogs eat spiders, they can also eat flies so I guess you can put the frog as top of the chain being, fly, spider, frog and also frog, eh sorry, spider, frog. So that's two*"
>
> (Participant 3)

Create, fill and finalize

In D2, participants were asked to categorize the dataset any way they found meaningful. In the hands-on condition participants seem to explore the data set by creating categories. They then gradually assign cards to those categories and meanwhile create additional categories. However this is not a consistent behaviour across all participants. Some create a category and then fill it completely with all the cards they consider to be a member of that category.

> *"Mammals would be easier but that's quite a big category I think"* [moving the panda, fox, lion, monkey, tiger, mouse and rhino to the bottom of the workspace] [adds horse]
>
> (Participant 14)

Create and illustrate

The 'Create and illustrate' strategy is a variation of the 'Create, fill and finalze' strategy. In the hands-on condition, categories are created and then illustrated with a few cards that belong to that category as if to test if that category is valid. Once deemed valid, participants continue adding cards to previously created categories to update or finalize them.

> [moves shark and killer whale together] [adds fish] [adds polar bear] [3 other actions]
> [adds penguin to the group with shark, killer whale, fish and polar bear] [2 other actions]
> [adds crocodile to the 'water' group]
>
> (Participant 2)

In the hands-off condition, participants tend to create categories and illustrate them with a few examples without finalizing that category (i.e. filling it with all cards that belong to the category).

Again there seems to be a relationship with the 'explore and experiment' strategy of the D1 task. The 'create and illustrate' strategy seems to be a simplified version of the more elaborate strategies in the hands-on condition.

> *"Okay so there's sea creatures, there are birds, ehmmmm.... animals that live in North America naturally. And then I guess there's this picture that looks like the plains of something, in Africa or wherever. So some of them will live there, so you can group them by area where they live, environment where they live, whether or not they fly ehmmm.."*
>
> (Participant 22)

Inverted pyramid categorization

Participants tend to begin with categorizations that can be applied to large portions of the card set. Afterwards the categories get increasingly more specific and usually have fewer members. For example, starting the task by separating plants from animals and then breaking the animal category down into more specific categories like: mammal, reptile or birds. This behaviour can be explained by using Klein's Data-Frame model. When participants visually index the data set they seem to primarily identify a limited number of over arcing frames that fit the data. Subsequent these frames can then be subdivided into smaller and more specific frames. The card set, for example, can be divided in two large categories: animate and inanimate entities. The animate category can subsequent be divided in animals and for example, plants.

"Animals and plants"
"Mammals and reptiles, mammals and reptiles. And bird and beast"
"and also herbivore, and carnivore"
(Participant 11)

Task immersion

Participants were asked to use the Think Aloud Protocol (TAP) during the tasks. They are expected to verbalize what they think, what conclusions they draw, how they are solving a particular problem or just how they are interpreting the data. This helps us understand how the process of task completion progresses. Participants in the Hands-off condition were only allowed to point to cards in case they did not know how to identify them. Since they were not allowed to physically move the cards the participants in general talked more and used some occasional pointing. We observed that participants in the Hands-on condition tended to forget to verbalise their thoughts. They seem 'immersed' in the task. Not being able to manipulate the data (Hands-off) might put less strain on the cognitive load of participants, and who are in general more likely to verbalize their thoughts.

CONCLUSION AND DISCUSSION

Although not statistically conclusive, we have observed differences between how Hands-on and Hands-off participants work with the cards, and in particular, in the more demanding D2 task. Hands-on participants showed a greater tendency to explore, try out assumptions and play around with the data than Hands-off participants. While the exploration behaviour itself does not appear to correspond directly with the activities of the Data-Frame Model, we suggest that exploratory actions are part of the process of seeking data to match one's frame. This would be part of the process of CONNECT as defined by the Data-Frame model: attempting to see what data is relevant and how they might be organised to make sense. The shuffling around of cards is a trial and error process of testing the conceptual or meaningful fit between cards and categories or thematic organisation. Some of this activity is CONNECT, as they attempted to make a 'FRAME' in order to understand what is present in the data set in order to therefore direct one's searches for that particular type of data (e.g. this set of cards is about vegetarian animals and not carnivores). The actions can also be ELABORATE actions as they search for other possible amplifying relationships, e.g. supportive of other cards in that category such as by placing more of the same type of animals in the carnivores category. Finaly there were REFRAME actions where participants decided a category or a member of the category is wrong and then re-organised or created a new category.

In the hands-on conditions our participants had off-loaded what would have been a demanding mental computation – setting up several 'lists', identifying commonality and assigning cards with the common features to the respective mental list – to an external representation of the physical placement and grouping of the cards. This is in line with findings that epistemic and complementary strategies (Kirsh and Maglio 1994; Kirsh 1995) can minimise mental computation workload. In designing systems for aiding the sense making process we can then start to identify several guide lines. Exploration of data is an important aspect in the process of solving a problem. It therefore is vital for users to being able to manipulate data and organize it in such a way so the system can function as an external memory. Designs should facilitate the transfer of mental idea's into the world. Besides functioning as a external memory to decreasing the cognitive workload it also serves as a platform to test hypotheses, setting aside intial unfitting data to simplify the dataset and play with data to facilitate serendipitous discoveries and subsequent possible solutions to the problem being worked on. A highly adaptable environment that affords manipulation in a constraint-free workspace would, so we suspect, greatly increase the ability to solve problems and possibly decrease the chances to err.

ACKNOWLEDGMENTS

This work was partly funded through the EPSRC *Making Sense* Project Grant Number EP/H023135/1. The work also benefited from the input of Dr. Neesha Kodagoda, Dr. Chris Rooney and Dr. Jane Barnett who provided valuable assistance to the undertaking of the research. The term 'tactile reasoning' was first suggested by our colleague, Dr Tinni Choudhury as a way of describing the interaction that occurs at the INVISQUE user interface.

REFERENCES

R.L. Ackoff (1989). "From Data to Wisdom," J. Applied Systems Analysis, vol. 16, 1989, pp. 3–9.

Antle, A.N., Droumeva, M. and Ha, D. (2009). Thinking with hands: An embodied approach to the analysis of children's interaction with computational objects. In *Proceedings of Extended Abstracts of Conference on Human Factors in Computer Systems* (CHI '09), ACM Press (Boston, USA, 2009), 3961-3966

Bargh, A. J., Chen, M., Burrows, L. (1996). Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action. *Journal of Personality and Social Psychology Vol. 71, 230-244*

Fjeld, M., & Barendregt, W. (2008). Epistemic action: a measure for cognitive support in tangible user interfaces? Paper presented at the Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008).

Kei, D., Andrienko, G., Fekete, J.-D., Gorg, C., Kohlhammer, J., and Melancon, G. (2008) Visual Analytics: Definition, Process, and Challenge

Kirsh, D. & Maglio, P. (1995) Complementary strategies: Why we use our hands when we think, *Proceedings of the seventeenth annual conference of,* 1-7.

Kirsh, D. & Maglio, P. (1994) On Distinguishing Epistemic from Pragmatic Action, *Cognitive Science,* 18 (4) (October): 513-549.

Klein, G, Phillips, J. K., Rall E. L. (2007) A Data-Frame Theory of sensemaking, *Expertise out of context*.

Lonergran, B. J. F. (1957). Insight: A study of human understanding. London: Longmans, Green and Co. Ltd.

Maglio, P., Jose, S., Chernicky, B., Kirsh, D. (1995) Interactive Skill in Scrabble

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis Proceedings of the 2005 International Conference on Intelligence Analysis, 2-6 May 2005, McLean, VA (pp. https://analysis.mitre.org/.../Final.../206_Camera_Ready_Paper.pdf): Mitre Corporation.

Stelmaszewska, H., Wong, B. L. W., Barnett, J., Rooney, C. (2011) Scrabble Report 2011 (unpublished), IDC Technical Report IDC-TR-2011/01, 30 pages.

Wilson, M. (2002) Six views of embodied cognition, *Psychonomic bulletin & review,* 9 (4) (December): 625-36.

Weick, K. E. (1995). Sensemaking in Organisations. Thousand Oaks, CA: SAGE Publications.

Wong, B. L. W. (2003) Critical Decision Method data analysis, In *The handbook of task analysis for human-computer interaction*, ed. Diaper, Dan, Stanton, and Neville, 327-346. Lawrence Erlbaum Associates.

Evaluation of display devices for dismounted soldiers utilizing intel from unmanned vehicles

Tal ORON-GILAD
Ben-Gurion University of the Negev

ABSTRACT
Introduction: this evaluation is part of a larger study aiming to develop display devices for supporting dismounted soldiers in intelligence gathering missions from unmanned vehicles. The focus here was on adding attention-switching considerations between the device and the soldier's immediate environment, which could influence performance, decision-making and task-load in ways unobservable in lone interface studies. **Method:** consisted of three steps: scenario generation based on interviews with SMEs, development of scenarios in a synthetic environment, and evaluation of them using predefined macrocognitive measures and experienced infantry soldiers. **Results and discussion:** higher mental workload and temporal demands relative to prior lone interface studies were found. Variations in performance as a function of the unmanned vehicle feeds tested were found. Altogether, it has been shown that the use of macrocognitive functions improves the sensitivity of evaluation and provides new insights into the utility of unmanned vehicles' video feeds to the dismounted soldier.

KEYWORDS
Sensemaking; Military; Unmanned Vehicles; Dismounted Soldier; Urban Terrian

INTRODUCTION

This research is part of a larger study, aiming to develop display devices to support dismounted soldiers through intelligence gathered from Unmanned Vehicles (UVs). Until now, our lab has focused on testing UAV/UGV video-feed interfaces that are removed from the soldier's hostile and exigent environment. The central motivation for the current new research approach was the addition of attention-switching between the interface and the soldier's immediate environment, which could influence performance, decision-making and task-load in ways that are unobservable in lone interface studies. The goal was to create a research tool that not only discerns superiority between interfaces in terms of performance measures but additionally diagnoses strengths and weaknesses of each system through the quantification of macro-cognitive Naturalistic Decision Making (NDM) functions in synthetic and field task environments.

Unmanned Vehicles became an integral part of today's modern armed forces and their utilization for airborne and ground reconnaissance is constantly improving. In recent operations, dismounted soldiers relied on video from unmanned aerial vehicles (UAVs) to locate targets or to engage with targets. Since most of the future military operations will be held in and around urban environments, the military will have to promote the development of new doctrines, training structures and equipment specifically suited for urban operations (Collyer, 2003). Since dismounted soldiers cannot carry the relatively large displays used in stationary environments or combat vehicles, smaller interfaces are needed. Previous studies (Minkov, Perry and Oron-Gilad, 2007; Oron-Gilad, Redden and Minkov, 2011) have shown task dependent differences among display devices, video feeds and operational tasks (Ophir-Arbelle et al. 2012) with regard to soldiers' performances.

Warfighters at lower echelons typically have multiple task demands, including tactical movement, scanning the local area, firing their weapons, and communicating with team members (Mitchell, 2009). In addition to these tasks, these warfighters observing areas of interest may be required to monitor information feeds from unmanned assets such as unmanned aerial vehicles (UAVs) operated by others, unattended sensors (USs, e.g., Balloons), and unmanned ground vehicles (UGVs) that are operated by others or that they are supervising. Even before the addition of these monitoring tasks, warfighters have proven to be overwhelmed with quality information and multiple demands for attention (Mitchell et al. 2004). Poor decision-making, slower response times, and overall poor performance can result because the individual is too focused on processing information rather than performing tasks (Wickens 2008, 2002). The requirement to monitor multiple video feeds increases the probability of information overload unless system design can mitigate the burden. Redden (2002) found that the information requirements of soldiers at lower echelons were quite different from those at upper echelons and that those requirements varied according to specific tasks. Thus, it is important to understand the information

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

requirements of warfighters operating at lower echelons and to provide them only the information and the video feeds that are needed for the task at hand so they do not become overloaded.
Due to extensive training, substantial experience, quick tempo of decision making and usage of satisficing solutions, a NDM approach seemed to be most suitable for this framework of evaluation. Macro-cognition is the cognitive phenomena that affect naturalistic tasks and settings (Klein et al., 2003). Klein and his colleagues have continuously demonstrated the high utility and applicability of their macro-cognitive model to command and control military intelligence, surveillance and recon (ISR) tasks. Of particular relevance to the current effort was a report by Baxter et al. (2004) which dealt with introduction of new information technology into the combat area. In this report,within the concept of macrocognition, seven macro-cognitive functions were identified and combined into a macrocognition model. Furthermore, the authors have developed methodologies for eliciting macrocognitive task requirements from subject matter experts (SMEs). Their work included collection of data from Land Warrior specialists, development of a synthetic task environment to elicit macrocognitive requirements associated with future combat systems, and an evaluation with operational-experienced subjects. This approach seemed highly relevant to the domain of our interest, as it is the introduction of new devices and capabilities that is of concern, and also within the realm of the dismounted soldier. As such, this approach has been adopted for the current evaluation in hope to fully incorporate in fore coming experimentation.

METHOD

The experimental method consisted of three steps; 1) scenario generation via comprehensive structured interviews with SMEs, 2) development of experimental scenarios in the synthetic environment, and 3) a small scale experimental evaluation of the synthetic environment and of defined NDM measures by experienced infantry soldiers.

Participants

SMEs

SMEs in the interview phase were three experienced male reserve Israeli Defense Force (IDF) soldiers between the ages of 23-25; a Squad commander, a Platoon sergeant, and an Elite commando unit combat soldier. All three had extensive MOUT-related experience and were in active duty in the past 6 months prior to the interview.

Participants

Participants in the evaluation phase were five additional male IDF reserve soldiers between the ages of 23-25 with a mean age of 23.8 ±0.8 years. All were Ben-Gurion University undergraduate students, whom actively serve in the reserve forces of the IDF in infantry related roles. During their mandatory service they had a mean of 16 months of active operational mission duty at front line bases. All had basic experience in navigation and in hostile force apprehension operations. They were screened for average experience level (i.e., not experts) in usage of military computerized systems and/or in deciphering of UV video feeds.

Scenario Generation

To generate the experimental scenarios and define an overarching mission for following experiments, extensive structured interviews were conducted with the three SMEs. The main concern was that the outcome of the interviews should lead to the development of a simulated operational mission that would afford the use of objective measurement of macro-cognitive constructs, while minimizing the many possible confounding variables of the rich synthetic environment.
Interviews were conducted using the Cognitive Task Analysis methodology and knowledge audits (Baxter et al., 2004; Peterson et al., 2000; Phillips, et al., 2001) while grounding the analysis on Klein et al's. (2003) NDM macro-cognitive functions. CTA allowed extraction, verbalization and representation of tacit MOUT decision making processes and procedural knowledge. CTA aided not only in the understanding of behavioral aspects of ground soldiers but also revealed MOUT field best-practice procedures. The interviews entailed detailed descriptions of prototypic and specific MOUT each soldier had experienced during service. Specifically, they were asked to describe the mission, its requirements, the distribution of friendly forces, the nature of threats present, and an incremental walkthrough of the operation as it ensued.
The interview emphasized the verbalization and extraction of macro-cognitive dimensions of Identification, Sense Making, Coordination, Common Ground, Uncertainty Management, Attention Management and Orientation utilizing macro-cognitive probes as described by Baxter et al., 2004. Four archetypical missions were revealed and at the end a static team hostile-force apprehension mission was chosen as a methodologically feasible for simulation, as specified in the following section.

Experimental scenarios in the synthetic environment

Utilizing MAK tools VR-Forces and VR-Vantage, a synthetic simulated environment was chosen. Eight soldiers (in pairs, two on each side of the building, as shown in Figure 1) were located around a building where it has been reported that a suspect has been seen. The participant served as one of the soldiers in each respective

scenario. The operational task was to apprehend the suspect once he is seen. However, the course of action (COA) varied, depending on the location of the participant. If immediate contact could be made, it was the participant's pair who was responsible to take action, if the suspect has been seen via the UV feed but was not within immediate contact, the participant was to report this information to his peers. If there was a conflict, or in case of doubt, the participant was to report to a commander. The conflict could arise, as shown in Figure 2, from two pairs having immediate contact with the suspect, which could lead to fratricide, or from a situation where the suspect has been seen in one of the windows or doors of the building without making an attempt to escape.
Using VR-Vantage and Camtasia, each scenario variation was also filmed from the perspective of a UGV, UAV or Baloon (unattended static sensor) to provide the additional video feeds to support the mission (as shown in Figure1; right).The recorded UV video feeds matched the timings in the synthetic environment precisely.
Altogether 12 variations of the same scenario settings were created. In 9 of them video feed from an UV was available (three variations of feeds with UAV, UGV and Baloon, respectively). In the remaining 3 variations there was no video feed (control). Each scenario lasted for approximately 3 minutes.

Figure 1. Map layout of the simulated operational task (left). The green dashed line marks the path of the UGV, and the orange arrow marks the current point of view of the UAV, as shown in detail on right. The scene from the perspective of the UAV (right). The suspect's building is marked for emphasis.

Figure 2. Aerial view of an example of a conflict (left). Both Team O and Team D (marked as yellow ellipses) have line of sight of the suspect leaving the building (marked in a red ellipse). The map (right) emphasizes the common field of view of both teams.

Experimental Evaluation

Apparatus

The operational scene representing the participant's immediate “real” threat environment was presented on three 22" conjoined monitors, as shown in Figure 3. A hand held monocular display (HHMD) was used to display the UV feed and the map of the area (similar to Figure 2, but without the graphic emphases). The HHMD has a resolution of 1280 x 1024 pixels and two input buttons.The green response button on the HHMD was used to report identification.

NDM Measures

Of the full microgocnitive model described in Baxter et al. (2004), four macro cognitive functions were evaluated within the current operational mission because they could be more easily defined and measured in the simulated context. Problem identification: identification that there is a need to respond to a situation (e.g., the escape of the suspect from the building). Coordination: did the participant allocate the problem to the correct entity, i.e., when applicable, and participants did not have to respond directly to the presence of the suspect, whether he shared the information with the correct teammates, or with the commander. Sense-making: measured post-event by a questionnaire asking the participant why he chose to responed in this particular way. Common-ground: since it was possible that each participant will relate in another way to the same problem, and will therefore respond differently, it was important to asses gaps between the master solution and the individual responses. For example, the view of a suspect from a door or a window could have different meanings to different participants in their assessment of whom to report to.

Figure 3. Participant's immediate conflict environment (left) displayed on three adjacent screens and the HHMD Display which was used to display the UV video feeds. The HHMD Interface (right) affords video feed views and the right green button was used to report identification.

Procedure

Each participant received instructions about the operational task, the location of the forces, and of the expected COA in each situation. They were briefed to watch over a suspected house that was said to contain a hostile-force. They were briefed as to the distribution of their team of three friendly forces in a manner such that each solider was positioned facing a unique side of the house. Specific rules of engagement were developed according to the interviews and were explained to the participants prior to the onset of the evaluation. The rules of engagement served to assure expected behaviors under uncertain situations and as formal military directives. Subjects were told to leverage the UV video feed when appropriate in order to detect targets and report their whereabout to a friendly force whom had a line of sight on the target or might be in immediate danger. Following the briefing, a training scenario was conducted. Only when the participant stated that he has fully understood the mission, the use of the HHMD and the various feeds available, and the COA did the experiment begin. In a partially counterbalanced order, each participant saw 12 scenarios. He was told beforehand which UV feed was available for the scene, if at all. A subjective questionnaire was filled together with the participant at the end of each scenario in order to assess sense-making and attain the "best practice" solution for each scenario. The solution was appropriated through a crowd-sourced approach of the subjects. At the culmination of the experiment, each subject filled out a raw NASA-TLX questionnaire, a subjective evaluation of the difficulty of the mission at hand, and a usability questionnaire regarding the HHMD.

RESULTS

Due to the small number of participants, descriptive statistics and correlations were used.

Overall success rate, workload, perceived usefulness and mission difficulty

Success rate across all 12 scenarios was on average 6.5 (54%, std=2.5) per participant. Thus, variation among participants was high with the most succesful one scoring 9 out of 12 and the least succesful one scoring 3 out of 12. The raw values of the NASA TLX administered at the end of the evaluation revealed high to medium scores for mental demand (6.4 std=2.9 on a scale of 1-10), temporal demand (7.2 std=2.2) and frustration (5.0 std=3.8). A dissociation was observed between success rate and percieved workload. Thus, those who rated workload as higher did not necessarily have lower perfromance. Similar dissociation was observed with regard to perceived mission difficulty, as the ones who rated the task as most difficult did not have the worst performance and vice versa. Scores for the item "*I could effectively complete my mission using the HHMD*" were moderate to low (4.2 std=2.3 on a scale of 1-10) indicating again upon the difficulty of the task and the limited perceived usefullness of the device by some of the participants.

Macrocognitive functions

Since there is no meaning to statistical analysis on a small sample, perfromances on the three macrocognitive function were evaluated as the summation of the 12 events across all five participants, i.e., (12 times 5) 60 events altogether.

Identification and coordination

As a group, participants responded to 50 out of 60 events (83%). Response times varied between 1 to 13 seconds. Shorter problem identification times do not necessarily indicate on better performance as the correctness of the response in terms of coordination may have been hampered. Therefore, the responses were divided into three categories; succesful identification of the suspect and correct allocation of the mission to the appropriate teammates (good coordination, 43%), partial success (sub-optimal solution, 13%), and erronous responses meaning no response or a completely wrong allocation of the mission to a teammate (bad coordination, 27%), as shown in Figure 4. It can be seen that sucessful identifications generated shorter response times than erronous responses,.

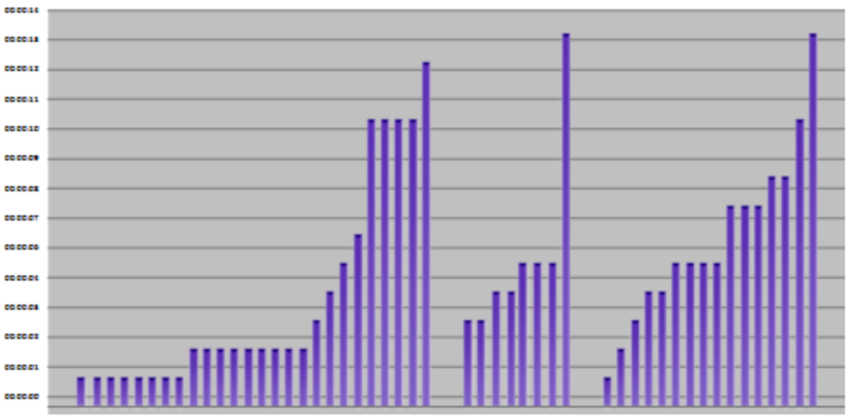

Figure 4. Response time [0-13 seconds] to an event as a function of the correctness of the coordination. The leftmost cluster of 26 events are the good coordination ones, the central cluster of 6 events are the partial success ones, and the rightmost cluster of 28 events are the erroneous responses.

Recall that for each participant there were 3 events per each type of video feed condition, UGV, Baloon and UAV and none. Altogether 15 events (3 for each by five participants) per condition. Success rate as a function of the the video feed available was 67% for UGV, 60% for none, 57% for Baloon and 53% for UAV. The UGV condition generated the highest correct response rate while the UAV generated the least. Response times as a function of the availability of the UV feeds is shown in Figure 5. Figure 5 further indicates that the Baloon video feed provided the fastest response times, while the UAV slowed responses.

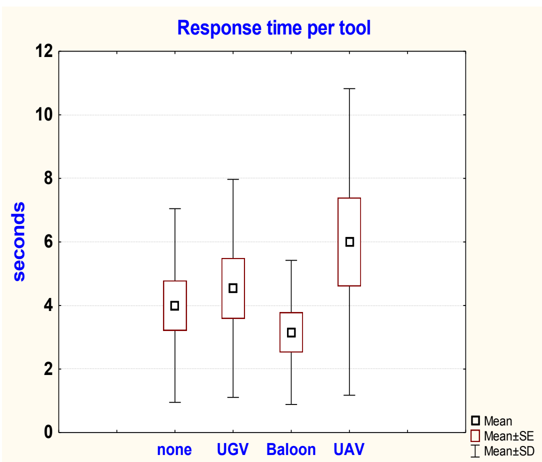

Figure 5. Mean response time a function of the UV feed available or none.

Sense-making

Sense maing was evaluated based on the macrocognitive questionnaire administered at the end of each scenario. Participants who were sucessful, always gave explanations that were similar to the ones initially given as a master solution by SMEs. It was also possible to note that partial success and erronous responses were always followed by a different explantion than the one given in the master solution. The two major sources for failure were fratricide and mis-evaluation of the danger of the situation (i.e., perceiveing an immediate dangerous situation as less dangerous and vice versa). Orientation failures were negligible (less than 6%).

Furthermore, when examining the utility of the UVs in supporting the valuation of less dangerous situations it was found that the UGV assisted participants in valuating the danger of the situation more often than the other UVs or the control condition and fewer errors of this type occurred when it was available than in any one of the other three conditions (13% for the UGV, as opposed to 29% respectively for the other three).

With regard to misevaluation of immediate danger, which obviously was best observed in the synthetic environment itself (without the need for any assistive intel) it was found that indeed, the availability of the UV feed, regardless of its kind, degraded performance and increased incidents where participants missed immediate events right in front of them. Specifically, altogether participants missed 5% of the immediate danger events with no UV feed, 11% with the UGV or UAV, and 16% with the Baloon, indicating perhaps that they over relied on the Balloon's video feed for identification of threats. On the contrary, identification of fratricide incidents was facilitated by the presence of either one of the UV feeds (11% miss each) while misses reached nearly 16% without the availability of additional video feed intel.

CONCLUSION

The goal of this effort was to explore possible macro-cognitive errors that may arise among dismounted soldiers utilizing UV video feeds. Unlike our previous studies which have examined the sole use of display devices with UV video feed interfaces (e.g., Oron-Gilad et al. 2011, Ophir-Arbelle et al. 2012), here performance was evaluated in a more holistic way, while mimicking the operational immediate threat environment that dismounted soldiers face. Although only a relatively small number of participants were used, the results provide an indication for the suitability of the experimental methodology, the sensitivity of the objective measures and some preliminary cognitive trends. First, higher mental and temporal demands was reported relative to prior lone interface studies conducted at our lab. Thus, indicating that the synthetic environment was generating additional load, closer to what dismounted soldiers face in reality. Since this manipulation was valuable, future studies will be conducted utilizing a synthetic environment projected on a dome semi-immersive environment which can further enhance the feeling of the immediate environment.

Not surprisingly, no correlation was found between participants' success and their perceived workload, perceived task difficulty or perceived usefulness of the HHMD. A correlation was found between response and sense making explanations. For erroneous or partially erroneous responses sense-making post questioning was always incomplete or wrong. While, there is nothing novel in such findings for the NDM research community, they may indicate that the experimental paradigm works and is worthy of further pursue. Participants' major difficulties were in estimating the severity of the identified event. This indeed corresponds to relationship between coordination and sense-making that the theoretical model predicts. When responding correctly, shorter response times were observed than when responding erroneously or in partial success. This can be explained by examining the relationship between identification and coordination, a relationship that has not been examined extensively in the literature in the context used here.

With regard to the utility of the video feed types, differences in performance were found, depending upon the type of feed (or none). The UGV video-feed seemed to provide better assessment of non-immediate danger. Furthermore, coordination scores, as measured by totalling the scores of each participant who used the UGV, when this feed was available, were higher than for the other tools. This finding is in line with findings in the literature that ground views improve the situation awareness of operators (e.g., Oron-Gilad et al. 2011). The UAV video-feed seemed to harm identification response times. This could have been attributed in part to the inherent rotational movement pattern of a UAV hovering around a building. However, it may have been worsened by the specific flight model used in the simulation. Future studies, should aim to improve the flight model of the UAV so that the image transmitted will be more stable.The Balloon which gave a static wide angle field of view of the scene facilitated faster identifications, but also demonstrated higher reliance of the participants on it, as their immediate environment was neglected more when this feed was available. Unmanned static sensors may seem as a very valuable asset for dismounted soldiers; however, it is less likely that such stationary cameras will be available at all times for most MOUT scenes in the real world. When no UV feed was available, participants were more aware of their immediate environment and less distracted.

Furthermore, although not intended to be so, scenario variations were not equal in difficulty. It became evident that each scenario was more than just the sum of the independent variables permutation, due to the combination of unaccounted variables in the synthetic environment itself (for example, the angles of the building, the locations of the openings, etc.) This reinforced that in future studies one would have to strive to control many of the inherent variables in the environment, e.g. as the number and complexity of objects on each flank of the suspected house.

To sum, the use of macrocognitive functions improves the sensitivity of evaluation and provides new insights into the utility of unmanned vehicles video feeds to the dismounted soldier. Future studies will further use this framework for larger scale evaluations while utilizing more measures.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Laboratory through the Micro-Analysis and Design CTA Grant DAAD19-01C0065, Task Order 113 (Michael Barnes, technical monitor). The views expressed in this work are those of the authors and do not reflect official army policy.

REFERENCES

Baxter, H. C., Phillips J., Shafer, J. Klein G. & Mosier K. (2004). MINDPRINT: Developing the Soldiers and Leaders of Objective Force and Beyond. *Research Report 1816,* US Army Research Institute for the Behavioural and Social Sciences.

Collyer, R. (2003). Human Performance Issues in Urban Military Operation (DSTO-GD-0380). *Australian Government Department of Defense*, Defense Science and Technology Organization.

Klein, G., Ross, K. G., Moon, B. M., Hoffman, R. R., Klein, D. E., & Hollnagel, E. (2003). Macrocognition, *IEEE Intelligent Systems*, 18, 3, 81-85.

Minkov, Y., Perry, S., & Oron-Gilad, T. (2007). The effect of display size on performance of operational tasks with UAVs. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting (pp. 1091–1095)*. Santa Monica, CA: Human Factors and Ergonomics Society.

Mitchell, D. (2009). Workload analysis of the crew of the Abrams V2 SEP: Phase I Baseline IMPRINT model. *Technical Report no. ARL-TR-5028*. Aberdeen Proving Ground, MD: US Army Research Laboratory.

Mitchell, D., Samms, C., Glumm, M., Krausman, A., Brelsford, M. & Garrett, L. (2004). Improved Performance Research Integration Tool (IMPRINT) Model Analyses in Support of the Situational Understanding as an Enabler for Unit of Action Maneuver Team soldiers Science and Technology Objective (STO) in support of Future Combat Systems (FCS). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Oron-Gilad, T., Redden, E. & Minkov, Y. (2011). Robotic Displays for Dismounted Warfighters: A field study, *Journal of Cognitive Ergonomics and Decision Making*, 5, 1, 29-54.

Ophir-Arbelle, R., Oron-Gilad, T., Borowsky, A. & Parmet, Y. (2012). Is more information better? How dismounted soldiers utilize video feed from unmanned vehicles - attention allocation and information extraction considerations. *Journal of Cognitive Ergonomics and Decision Making*. Published online before print May 2, 2012.

Peterson, B., Stine, J., L. & Darken, R. P. (2000). Eliciting knowledge from military ground navigators. *In Berndt Brehmer, Rannan Lipshitz and Henry Montgomery (Eds.), How Professionals Make Decisions.* Washington, D.C: American Psychological Association

Phillips, J., McCloskey M.J., McDermott, P. L., Wiggins, S. L. Battaglia, D. A. Thordsen, M. L. & Klein G. (2001). Decision-Centered MOUT Training for Small Unit Leaders.

Redden (2002) Virtual Environment Study of Mission-Based Critical Information Requirements, *Technical Report no. ARL-TR-2636*, Aberdeen Proving Ground, MD: US Army Research Laboratory.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159–177.

Wickens, C. D. (2008). Multiple Resources and Mental Workload, *Human Factors*, 50, 3, pp. 449–455.

Collaborative decision making in critical situations

The role of endogenous and exogenous uncertainty in critical incident decision making: An example from hostage negotiation

Claudia VAN DEN HEUVEL[a], Laurence ALISON[a] and Nicola POWER[a]
[a]*University of Liverpool*

ABSTRACT
Introduction: This study examined the impact of uncertainty during a simulated live hostage negotiation training event for police officers. **Method:** The nature and origin of uncertainty was assessed using observational methods of performance (at the time and by reference to video footage), decision logs and post-incident simulated recall semi-structured interviews. **Results and discussion:** Thematic analysis revealed a dichotomy of uncertainty sources. Early situation assessment phases of decision making, uncertainty was characterised by endogenous sources (uncertainty about the problem situation) such as concern for victim's welfare. During the plan formulation and implementation of actions, uncertainty was characterised by exogenous sources (uncertainty about management and team) such as low role understanding and trust. Overall, exogenous uncertainty (75%) occurred more frequently than endogenous uncertainty (25%). Naturalistic decision making studies should utilise this dichotomy to assist structured recommendations for uncertainty management research.

KEYWORDS
Uncertainty management; critical incidents; teams; role understanding; trust

INTRODUCTION

Police-related critical incidents are high risk, high stake and uncertain events that can leave a long term negative commercial, behavioural or emotional effect on a community (Alison and Crego, 2008). During critical incident decision making, decision making can be degraded when there is perceived uncertainty about the *situation*, potential *options* and projected *outcomes* (Lipshitz & Strauss, 1997). Van den Heuvel, Alison and Crego's (2012) SAFE-T model of expert critical incident response phases (Situation Assessment, Formulation and Execution of plans, Team learning) found that uncertainty could lead to decision inertia and failures to act. For example, uncertainty during the initial situation assessment phase caused repeated and redundant requests for the same information. In addition, uncertainty during plan formulation and execution phases was found to induce choice deferral. It is important to understand the sources causing uncertainty in order to provide applicable interventions to overcome it and foster effective critical incident decision making.

This paper explores a unique case study of crisis decision making. It presents a model of uncertainty that: (i) operationally categorizes sources of uncertainty; whilst (ii) offering preliminary descriptive accounts of how uncertainty may influence critical incident decision processes during different decision making stages. The model synthesises Lipshitz & Strauss' (1997) three-tier model of uncertainty into two uncertainty sources: endogenous uncertainty (relating to the problem situation) and exogenous uncertainty (relating to management and team processes). It then maps these two types of uncertainty onto the SAFE-T model (van den Heuvel et al., 2012) of decision making to pinpoint when and how they may derail strategic decision making.

DICHOTOMISING UNCERTAINTY

Endogenous (to the problematic situation) uncertainty

In line with naturalistic approaches to decision making, *endogenous* uncertainty is derived from uncertainty about the problem event that one is responding to (Klein, 1993). Such uncertainty may be a product of environmental characteristics specific to the problem incident such as ambiguous information, intense time pressure and high perceived risk (Orasanu & Connolly, 1993). Endogenous uncertainties are situational characteristics which can induce doubt in the decision maker. They can be based upon what one is dealing with in the *present* vis-à-vis situation assessment ("what are we dealing with and what are the risks?") along with what they envision to be facing in the *future* vis-à-vis prospective cognitive modelling and anticipatory thinking

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

("How will this event evolve in the future ?") (Klein, Snowden, & Pin, 2007). High endogenous uncertainty is likely to exist when situational information is sparse, overwhelming, contradictory or novel.

When sensemaking fails then decision makers experience dissonance due to the feelings associated with having a non-definitive and inadequate situational understanding (Koppenjan & Klijn, 2004). During a hostage situation, poor situational understanding can occur due to the fast-paced, unstable and emotionally salient nature of the incident (Nohrstedt, 2000). This perceived lack of control over the potential consequences of one's choice generates significant emotional distress (Nohrstedt, 2000) and increases experienced uncertainty, with an attendant detrimental effect on a decision maker's cognitive ability to assess situations and problem solve in subsequent decision phases (Klein, 1993). Situational uncertainty is extremely difficult to resolve in such dynamic incidents; the information received is often inherently complex or contradictory and additional information is unavailable (Fiore *et al*, 2010). Attempts to reduce endogenous uncertainty via information seeking are thus problematic and so strategic decision makers tend to rely on coping mechanisms (van den Heuvel, Alison & Power, in press). The present study provides a detailed account of the role of endogenous uncertainty during a simulated hostage negotiation scenario.

Exogenous (non-situation specific) uncertainty

Exogenous uncertainty stems from ambiguity surrounding management and team processes. Exogenous uncertainties can derive from confusion over the the expectations from one's own and others' performance within the decision making team which can influence the planning and execution of decisions and actions. Critical incident command teams are often interdisciplinary and cross-functional; they involve individuals from different backgrounds, with different roles and with different levels of operational experience (Nohrstedt, 2000). The ad hoc nature of such teams can produce limitations in interoperable functioning. This is due to role and responsibility confusion, both regarding one's own role and in interpreting the roles of others (Nohrstedt, 2000). Poor role understanding can reduce confidence (Shanteau, 1997) and self-efficacy (Bandura, 1997). This can cause problems for dynamic decision making, where high self efficacy and confidence are useful for effective goal setting and action planning (Olsen, Roese & Zanna, 1996). A lack of role clarity can produce confusion about individual's goals, actions and responsibilities, and is predicted to be a key source of exogenous uncertainty.

Poor understanding of a team member's role can also impede decision making by reducing interpersonal trust in that individual's perceived competence (Bonaccio & Dalal, 2010). Specifically, a lack of inter-positional knowledge (knowing how each other's roles interact) can reduce the perceived reliability of other team members' judgments, decisions and advice (Budescu & Rantilla, 2000) due to erroneous expectations for what others should be doing. Mayer et al., (1995) defined trust as a willingness to be vulnerable to the actions of another party, underpinned by three perceived antecedents relating to the trustee's: i) *ability* to perform the task; ii) *benevolence* to the best interests of the trustor; and iii) *integrity* to act consistently. In high risk situations such as hostage events, where an Incident Commander's (team leader) area of expertise is often different from their advisors, they must trust the advisors' abilities and intentions to provide accurate advice on potential courses of action (Bonaccio & Dalal, 2010). Trust is therefore predicted to be a source of exogenous uncertainty withint hese incidents.

METHOD

Participants

Police officers (n=16) from the United Kingdom participated in one of two live hostage negotiation simulation training events. Trainees were experienced police officers who might be called out to an actual hostage scenario. Participants (who ranked from inspector to chief superintendents) included first responders, hostage negotiators, negotiator coordinators (NC), tactical advisors and incident commanders (IC). In addition, 14 role players (one police officer and 13 civilians) also participated at the event playing the roles of hostage takers (n=3) and hostages (n=4) across two training nights. The main hostage taker was played by a police officer with previous experience enacting this role. The additional two hostage takers were played by civilians who were given direct instructions on their role. The remaining role players acted as hostages and were also directed.

Procedure

Data was collected at the two training exercises. The scenario involved hostages being held on a bus within a tunnel over a number of hours at night. A detailed exercise plan was designed in conjunction with police negotiation trainers. Key events and information injects were adhered to on both nights; however, to maintain fidelity and realism, the timing of injects was adjusted to reflect trainees' decisions. The same narrative arc was achieved with trainees' making similar key decisions (Table 1).

Table 1. Key decisions from both training nights

	What happened during the event?	How was the event resolved?
Event one	Hostage taker demanded a car three times A hostage had their thumb cut off A hostage was shot and killed	Release of the four hostages Main hostage taker left bus with aim of being shot but was successfully detained
Event two	Hostage taker demanded a car three times A hostage was 'seriously injured' by being thrown across the bus A hostage was shot and killed	Release of two hostages Release of two fellow hostage takers Main hostage taker committed suicide on the bust

Data Collection and Analysis

Decision Logs

Both the IC and NC completed decision logs during the scenario. Endogenous sources of uncertainty (relating to the problematic event) were measured quantitatively via a rating scale on perceived risk along with ratings for concern about the victims, other police officers, the hostage taker and themselves. Exogenous sources of uncertainty (relating to management and team processes) were measured quantitatively by ratings on perceived quality and trustworthiness of advice received from others (IC only) and by confidence in decisions. Confidence was used as an inverse measure of uncertainty so that low confidence represented high uncertainty (Kuhthau, 1991).

Videoed Strategic Command Meetings

Strategic command meetings during the scenario were recorded and transcribed. Transcriptions were divided into frames from one source (e.g. IC) to a target (e.g. NC) resulting in 735 discussion frames. Thematic analysis was utilised to code uncertainty sources (Braun & Clarke, 2006) with frames firstly coded semantically for specific sources of uncertainty (e.g. concern for victim; role understanding) and then split into two theory-driven categories (i.e. endogenous; exogenous). Discussion frames were also coded to reflect the decision phase that participants were discussing using the SAFE-T model to indicate whether discussions were about situation assessment, plan formulation or plan execution (van den Heuvel et al., 2012). Coding for decision phases (Kappa = .63; $p < .001$) and uncertainty sources (Kappa = .58; $p < .001$) had moderate yet significant inter-rater reliability between two independent coders.

Post-event Interviews

Interviews with participants (n=5) were conducted in the weeks following the training event. Interviewees were shown their videoed strategic command meetings and asked to identify and particularly uncertain decisions and why they had felt uncertain. Interviews were recorded, transcribed and thematically analysed. As above, semantic codes for uncertainty sources were initially coded semantically and then themed into theory-driven exogenous and endogenous uncertainty categories.

RESULTS

Decision-phase analysis of uncertainty

Overall, 25% (n=43) of discussion frames focussed on endogenous uncertainty and 75% (n=134) on exogenous uncertainty (Table 2). During situation assessment discussions, 37% of discussions were about endogenous uncertainty which was higher than the base rate probability of 25% across the scenario. During plan formulation and plan execution, 82% and 95% of discussions were about exogenous uncertainty which was higher than the base rate probability of 75%.

Table 2. Proportional representation of uncertainty type during each decision phase

	Situation assessment	Plan formulation	Plan execution	Total
Endogenous uncertainty (about problematic situation)	29 (.37)	14 (.18)	1 (.05)	43 (.25)
Exogenous uncertainty (about management and team)	49 (.63)	66 (.82)	19 (.95)	134 (.75)
Total	78	80	20	178

Sources of endogenous uncertainty

Both risk perception and concern for others were used as quantitative measures for endogenous uncertainty. Risk was perceived as high throughout the entire scenario (M = 8.83, SD = 1.60) with officers significantly more concerned about victim safety (M= 9.45, SD=1.50) than with the safety of other officers (M= 5.61, SD=2.87), hostage takers (M= 4.08, SD=2.75) or themselves (M= 4.76, SD= 2.87) (t(37)= 38.79, p<.001).
Qualitative analyses of interviews also found victim safety to be a key source of endogenous uncertainty. When victim safety was ambiguous, officers were found to rapidly re-evaluate their situation assessment to try and overcome uncertainty. This was evident when unprovoked violence occurred during the scenario. As two NC's noted during interview: "*I didn't feel uncertain at any point UNTIL he allegedly cut a finger off. Until then I couldn't see any rationale for an outcome that wasn't positive*" *(NC)*; and *"It was as the scenario became more difficult and there was more risk to individuals - that was when [my] decision making changed. Because it's harder because now you are playing with people's lives." (NC).*

Sources of exogenous uncertainty

Confidence and trust in advice measures were used to quantitatively assess exogenous uncertainty about management and team processes. On average, trainees were confident in their decisions (M = 7.37, SD = 2.08) but recorded low trust (M = 4.71, SD =.83). Correlations found that trust was not significantly related to confidence (r =.12, p >.05), but a positive trend in the data did suggest that low trust was associated with low confidence (i.e. uncertainty).

Poor role understanding

Qualitative analysis of interviews indicated that a lack of understanding in one's role was a prevalent source of exogenous uncertainty: *"I should be in a stage where I know exactly what I should be doing, and I didn't".* Poor role understanding was primarily due to perceived insufficient experience: "*...it's out of my comfort zone, I am not used to performing that role" (NC),* which manifest itself as either role confusion or role corruption. Role confusion involved officers not knowing what decisions and actions were their responsibility: "*The NC had never done the role before so he did not have the experience. I think he showed some confusion as to what was his or somebody else's responsibility" (IC).* Role corruption involved officers engaging in behaviour that was outside of their responsibility by: "*trying to cross over into other people's roles and responsibilities*" *(IC).* A recurrent theme involved individuals constantly having to remind themselves of their role: *"In my decision making I was constantly spinning the Conflict Management Model. But for me, in the role of Coordinator I thought 'wait there, that is not for you in your role, that is for [the IC]'... it's knowing to keep you negotiator hat on... You often get a bit of role corruption"(NC).*

Poor interpersonal trust

Qualitative analysis of interviews indicated that poor interpersonal trust was also a prevalent source of exogenous uncertainty: *"No I don't think it (uncertainty) had anything to do with the situation, it was the interpersonal side of things" (NC).* Poor trust appeared to be a product of the perceived inability of team members to perform their roles: *"I would be considering whether X was the right individual to perform that role." (IC).* If officers judged their colleagues as incompetent, they tended to distrust that individual's advice, which was potentially exacerbated by erroneous role understanding. Alternatively, perceived ability improved trust*: "I found it easier to deal with the experienced one because I could see what he was doing. With the novice I was spending time training them from scratch" (IC).*

Perceived benevolence ('the extent to which a trustee is believed to want to do good for the trustor'; Mayer et al., 1995, p.718) also reduced trust and increased exogenous uncertainty. ICs often perceived information provided by advisors as being in pursuit of individual rather than collective goals by: "*trying to assert control over the situation and run it the way they wanted it to run rather than allowing me to be the decision maker" (IC).* Trust in advice was reduced and exogenous uncertainty increased when advisors were perceived to lack benvolence and be focussing on their own personal goals by: "*making decisions to influence me instead of giving me all the information and options" (IC).*

DISCUSSION

This paper has presented a hostage negotiation simulation as a case study for assessing uncertainty in critical incident decision making. It has built upon existing theories of uncertainty (Lipshitz & Strauss, 1997) and naturalistic decision making (van den Heuvel *et al*, 2010) by synthesising sources of uncertainty into two categories: *endogenous* uncertainty (relating to the problematic situation) and *exogenous* uncertainty (relating to management and team processes). The quantitative and qualitative results illustrated that endogenous uncertainty caused decision makers to re-evaluate their situation assessment, and stemmed primarily from ambiguity over the welfare of innocent parties. Exogenous uncertainty, which was the more prevalent form of uncertainty, arose during the plan formulation and execution phases of decision making and stemmed primarily from poor role understanding and lack of trust.

Poor role understanding stemming from misunderstanding one's own and other team members' responsibilities has been associated with increased stress and frustration (Kuhlthau, 1991) and reduced satisfaction in decision making (Rizzo, House & Lirtzman, 1970). When a team member failed to meet (potentially erroneous) role expectations, their perceived competence and ability degraded (Bonaccio & Dalal, 2010). In addition, poor trust and a low perceived ability and benevolence of other team members increased uncertainty during the formulating and executing of decisions (Mayer et al., 1995). This potentially reduced team cohesion and willingness to share information (McKay, 1991), where NCs were unwilling to *share* information (Koppenjan & Klijn, 2004), and ICs were also reluctant to *seek* information (Sniezek & Van Swol, 2001). Moreover, received advice was suspected of being biased towards personal ulterior motives (Petty & Cacioppo, 1979). This focus on information source rather than message content has been identified as a frequently occurring issue which can signitifcantly derail critical incident decision making (Rake & Njå, 2009).

IMPLICATIONS

It is suggested that exogenous uncertainty may improve with role-exposure training. Interoperable training may: i) aid understanding of one's own and team members' roles (Fiore *et al*, 2010); and ii) enhance interpersonal trust and information sharing (Bonaccio & Dalal, 2010). Reducing exogenous sources of uncertainty (relating to management and team processes) may indirectly improve coping with endogenous uncertainties, where the team can respond cohesively to the problem environment. It is hoped that by dichotomising sources of uncertainty as endogenous (to the problematic situation) or exogenous (to general management and team processes) that naturalistic research into uncertainty can be more objectively structured. Future studies exploring endogenous uncertainty can highlight specific elements of problematic situations which may derail decision making, whereas studies exploring exogenous uncertainty can address the psychological processes involved in team management which may derail decision making in critical incidents.

REFERENCES

Alison, L. and Crego, J. (2008). Policing Critical Incidents: leadership and critical incident management. Devon, England: Willan Publishing.

Bandura, A. (*1997*). Self-efficacy : the exercise of control. New York: W.H. Freeman.

Bonaccio, S. and Dalal, R.S. (2010). Evaluating Advisors: A Policy-Capturing Study Under Conditions of Complete and Missing Information. Journal of Behavioral Decision Making, 23, 227–249.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.

Budescu, D. V., and Rantilla, A. K. (2000). Confidence in aggregation of expert opinions. Acta Psychologica, 104, 371–398.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. Human Factors, 52(2), 203-224.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G.A. Klein, J. Orasanu, R. Calderwood, and C. Zsambok (Eds.), Decision Making in action: models and methods (138-147). Norwoord, CT: Ablex.

Klein, G., Snowden, D., and Pin, C.L. (2007). Anticipatory Thinking. Proceedings of the Eighth International NDM Conference, Pacific Grove, CA, June 2007.

Koppenjan, J. F. M., and Klijn, E. (2004). Managing uncertainties in networks : a network approach to problem solving and decision making. Routledge, London.

Kuhlthau, C. (1991). Inside the search process: seeking meaning from the user's perspective. Journal of the American Society for Information Science, 42 (5), 361-371.

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision making analysis. Organizational Behavior and Human Decision Processes, 69, 149-163.

Mayer, R.C., Davis, J.H., & Shoorman, D.F. (1995), An integrative model of organizational trust. Academy of Management Review, 20(3), 709-734.

McKay, J. (1991). Assessing aspects of object relations associated with immune function: development of the affiliative trust-mistrust coding system. Psychological Assessment, 641-647.

Nohrstedt, D. (2000). Intuition and Avoidance in Foreign Policy Crises: How Decisionmakers Respond to Different Types of Uncertainty. Paper presented at the SUPUI seminar, The Swedish Institute for International Affairs.

Olson, J., Roese, N., and Zanna, M. (1996). Expectancies. In E. Higgins and A. Kruglanski (Eds.), Social Psychology: Handbook of basic principles (pp. 211-238). London, England: Guildford Press.

Oransanu, J. and Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Oransanu, R. Calderwood, and C. E. Zsambok (Eds.). Decision Making in Action: Models and Methods (pp. 3 –20). Norwood, CT: Ablex.

Petty, R. E., and Cacioppo, J. T. (1979). Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. Personality and Social Psychology Bulletin, 5, 173–176.

Rake, E.L. & Njå, O. (2009). Perceptions and Performances of experienced incident commanders. Journal of Risk Research, 12(5), 665-685.

Rizzo, J.R., House, R.J. and Lirtzman, S.I. (1970). Role Conflict and Ambiguity in Complex Organizations. Administrative Science Quarterly, 15(2), 150-163.

Shanteau, J. (1997). Competence in Experts: The Role of Task Characteristics. Organizational Behavior and Human Decision Processes, 53, 252- 266.

Sniezek, J.A. & Van Swol, L.M. (2001). Trust and expertise in a judge advisor system. Organizational behaviour and Human Decision Processes, 82, 288-307

van den Heuvel, C., Alison, L., & Crego, J. (2012). How uncertainty and accountability can derail strategic 'save life' decision in counter-terrorism simulations: a descriptive model of choice deferral and omission bias. Journal of Behavioral Decision Making, 25(2), 165-187.

van den Heuvel, C., Alison, L., & Power, N. (in press). Coping with uncertainty: police strategies for resilient decision-making and action implementation. Cognition, Technology & Work.

Team Decision Making: eliciting the structure of interdependences when returning to periscope depth

Neville A. STANTON[a] & Kevin BESSELL[b]
[a]Transportation Research Group, Faculty of Engineering and Environment, University of Southampton, Southampton, SO17 1BJ, UK
[b]BAE Systems, Defence Information, Training & Services, Alvington, Yeovil, Somerset, BA22 8UZ, UK

ABSTRACT

Introduction: This paper is concerned with the elicitation and representation of team decisions when returning a submarine to periscope depth. The paper is following the trend to examine the team decisions rather than those of individuals, although the contributions of individual roles are considered. **Method**: The Decision Ladder from Rasmussen's Cognitive Work Analysis is used to elicit the team decisions in the first instance. **Results and discussion:** The elicitation of team decisions using the Decision Ladder was verified by experts as a valid representation. Further analysis was conducted using Evaluation and Execution Matrices which mapped the data gathered by the Decision Ladder approach. The SOCA-DL was used to capture the contribution of different roles. The resultant analysis was proposed as a benchmark for comparing future concepts as well as an approach to develop requirements for interface, procedure, job aid, decision support and simulator design, as well as decision making training.

KEYWORDS

Decision ladder; team decisions; submarine; representations; interdependencies; periscope depth

INTRODUCTION

Naturalistic Decision Making research has shifted focus from the individual to distributed decision making (Smith and Dowell, 2000; Stanton and Wong, 2010), whereby the decision-making process is distributed across team members (Zsambok, 1992; Stanton et al, 2010). One important aspect of this research lies in understanding the nature of distributed decision making so that effective decision support can be provided (Klein and Miller, 1999; Klein, 2008). This paper presents a case study focusing on decision making in naval teams during the return to periscope depth exercise. The aim of the paper is to model the distributed decision-making process in this context and examine the insights that may be gleaned from the models. The project began with observations of the Return to Periscope Depth (RTPD) scenario in a training simulator (e.g., travelling from 60 metres depth up to 17 metres). The analysis was further developed with two Petty Officers who had recently returned from sea. Review and validation of the analysis was undertaken with members of the training staff. Finally, the analysts presented the representations to the Maritime Warfare Centre. The attendees included a number of ex-submariners who were able to verify the analysis. Part of the complexity of the problem of returning to the surface safely is the dependency on passive sonar as a means of detecting surface vessels. Additional complexity comes from working as a team comprising personnel from the sound room, control room and ships control, as well as contributions from the rest of the ship. To analyse this system required the application of a method that would assist in identifying key features of the work and be able to clarify the constraints. Cognitive Work Analysis (CWA) was developed to analyse complex socio-technical systems such as those found in nuclear power generation (Rasmussen, 1986). This development came from the realisation that an in-depth understanding of the interrelations of social systems and technical systems was required to fully appreciate how constraints act upon the working of system functions (such as the communications and activities in and between the sound room and control room on a nuclear submarine). These systems are made up of numerous interacting parts, both human and non-human, operating in dynamic, ambiguous and often safety critical domains. The complexity embodied in these systems presents significant challenges for modelling and analysis and most methods are not well designed to capture the complexity of the interrelations and analyse the layers of interconnection within socio-technical systems (Jenkins et al, 2009; Vicente, 1999; Rasmussen, 1986). The semi-structured framework presented within CWA helps to guide the analyst through considerations of the various levels of constraints acting on systems and the effects that they can have upon the way in which work can be carried out. The CWA process has been criticised for being complex and time consuming. In order to address these concerns, and in an attempt to provide some level of guidance and expedite the documentation process, the HFI DTC (Human Factors Integration Defence Technology Centre; www.hfidtc.com) has developed

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

a CWA software tool. Built to run on the Microsoft .net framework, the tool uses a familiar windows-based interface and interfaces directly with Microsoft Office applications. The tool provides templates to allow documents to be created that describe each of the five phases described in the CWA framework, providing a structure for those unfamiliar with the technique (Jenkins et al, 2009). The software tool allows data to be passed between these phases, expediting the documentation process, and facilitating updates and changes.

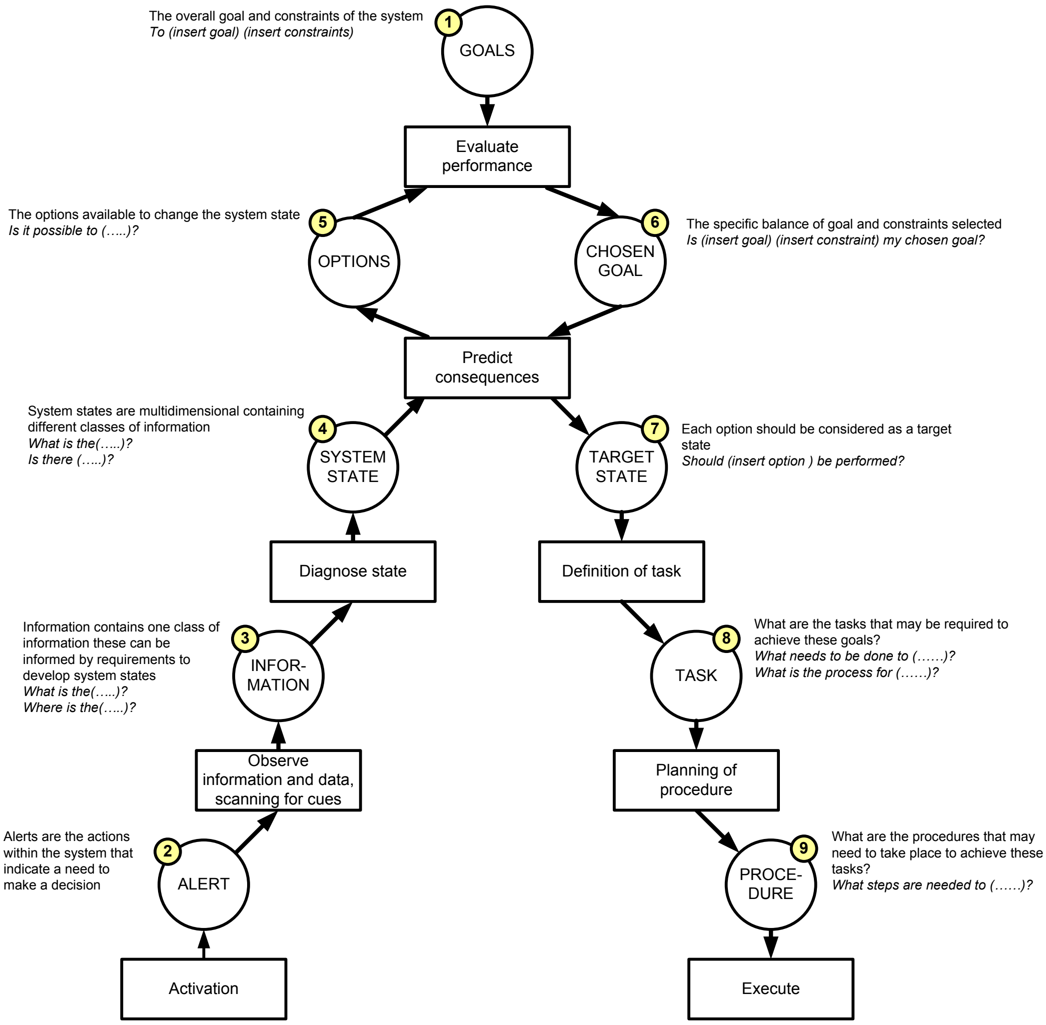

Figure 1. The decision ladder (from Rasmussen, 1986).

THE REPRESENTATION

Rasumussen's decision ladder was used as the method for eliciting and representing the deicions made by the teams (Jenkins et al, 2010). The decision ladder has two different types of node: the rectangular boxes represent information-processing activities and the circles represent states-of-knowledge that result from information-processing activities (see Figure 1). The decision ladder can be used to describe both levels of expertise and novelty of the decision processes. Novice users are expected to follow the decision ladder in a linear fashion, whereas, expert users are expected to link the two halves by short-cuts. The left side of the decision ladder represents the observation and information gathering activities to identify the system state, whereas, the right side represents the planning and execution of tasks and procedures to achieve a target system state. In between identifying the system state and target state are the options selection activities to meet the desired goal(s). In experts or proceduralised activities, observing information and diagnosing the current system state immediately signals a procedure to execute. This means that rule-based shortcuts can be shown in the centre of the ladder. On the other hand, effortful, knowledge-based goal evaluation may be required to determine the procedure to execute; this is represented in the top of the ladder. There are two types of shortcut that can be applied to the ladder; 'shunts' connect an information-processing activity to a state of knowledge (box to circle) and 'leaps' connect two states of knowledge (circle to circle); this is where one state of knowledge can be directly related to another without any further information processing. It is not possible to link straight from a box to a box as this misses out the resultant knowledge state. In an attempt to better understand decision making, the decision ladder can be used to develop prototypical models of activity. It is important to draw the distinction between typical and prototypical work situations. People tend to describe what they find to be normal, usual ways of doing things, representing an intuitive averaging across cases – typical situations. Conversely, prototypical work situations are developed from actual data from context specific cases. This then forms a set of prototypical activity elements, defined by either problem to solve or situation to solve within, which, in varying combinations can serve to characterise the activity within a work system.

This Decision Ladder analysis was undertaken with subject matter experts though interviews, as follows:

- The experts were introduced to the decision ladder model and asked to describe their overall goal in operating the system
- The experts were asked to talk the analysts through the process of making a decision about when to return to periscope depth. The experts were guided to start the description by indicating what might first draw their attention to the need to return to periscope depth (the alert).
- The experts were then asked to list the artefacts that they might use to gather information.
- The experts were asked to explain how they used this information to diagnose the current system state.
- The experts then described the options available to them.
- The experts explained how they would balance the competing constraints on their goals.
- Based upon the goal selected the experts then listed the target states available (options) and selected the target state they would take.
- This state was then broken down into a series of tasks.
- The tasks were then broken down into high level procedures.

Once recorded, the notes were read back to the expert, and at each stage of the decision ladder, the expert was asked to capture all other elements that would be available. For example, list other possible reasons why they might have to return to periscope depth.

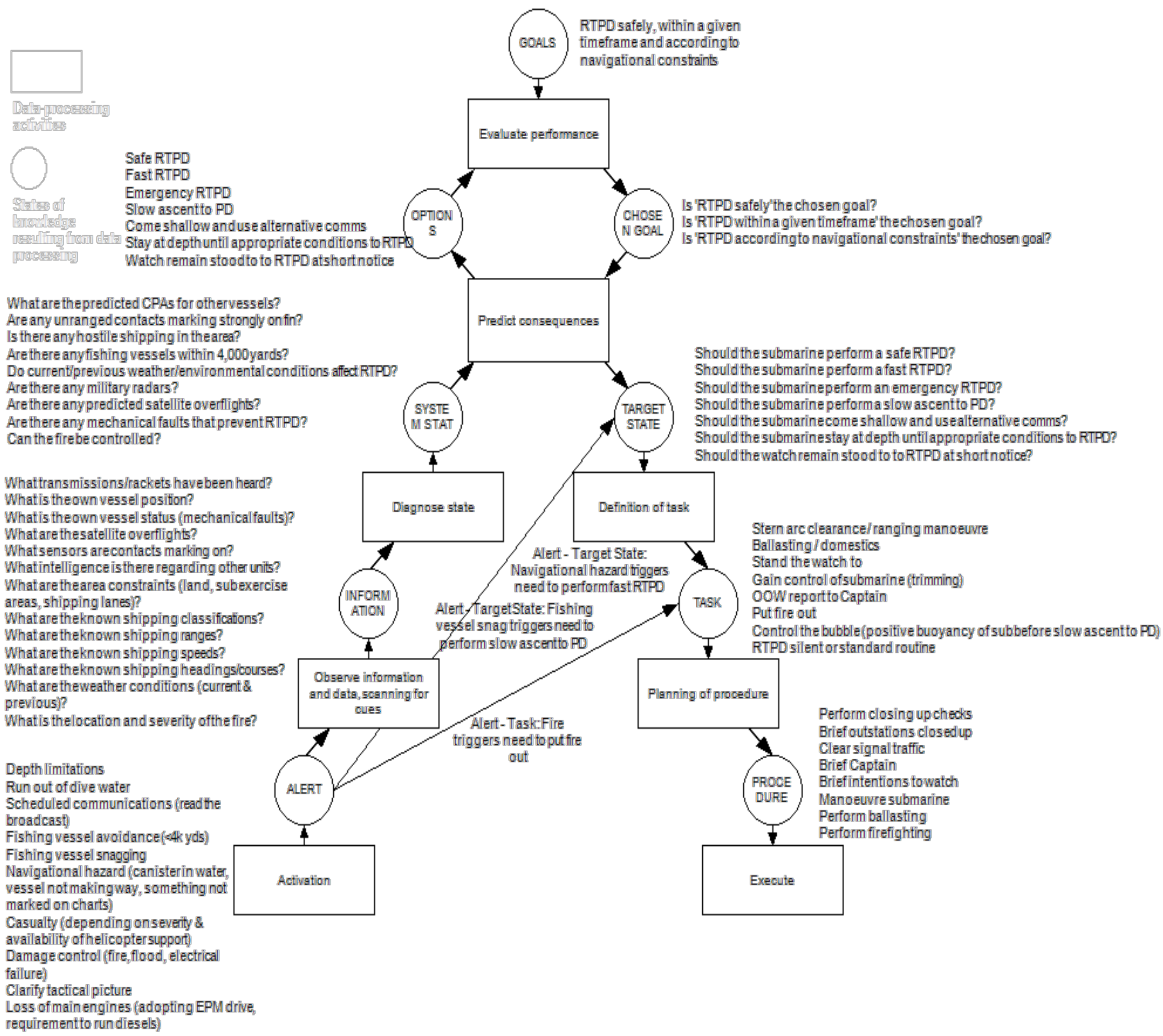

Figure 2. The decision ladder for returning to periscope depth in normal peacetime operations.

THIRTY TWO DECISIONS

The information gathered using the methods described above was used to generate the model presented in Figure 2. In order to constrain the analysis, the experts recommended that the ladder be completed for a specific scenario, with the submarine acting autonomously, conducting standard operations, under general peacetime conditions, in a home waters environment. As can be seen from Figure 2, even under these conditions a multitude of factors influenced the decisions involved in determining how to proceed for RTPD. Starting with the goals, the experts identified three constraints acting on the overall goal of returning to periscope depth; to return safely, within a given time frame and according to navigational constraints. Given that these constraints can be in conflict (the need to return safely may mean that the time frame cannot be met, for example) this then leads to three possible choices of goal.

A number of alerts, or reasons why the submarine would need to RTPD, were elicited, ranging from the need to make a scheduled communication to a casualty or fire onboard. The information that might subsequently need to be gathered included known shipping classifications, ranges, speeds and so forth, as well as the own vessel position and status in terms of any mechanical faults. The system state elements then reflect what can be ascertained from the available information. So, for example, combining information about own vessel with shipping ranges, headings, etc. enables calculation of the Closest Points of Approach (CPAs) for other vessels.

Options identified were of two types; either different ways in which a RTPD could be conducted (safe, fast, emergency or a slow ascent), or, interestingly, alternatives to returning to periscope depth, namely coming shallow to use alternative communications, staying at depth, or leaving the watch 'stood-to' ready to RTPD at short notice. These options also represented the available target states, selected on the basis of the chosen goal (RTPD safely, within a given timeframe or according to navigational constraints). Tasks listed covered the range of actions that might be necessary to achieve the possible target states, such as standing the watch to and ranging all contacts, along with the procedures for carrying out these tasks (e.g. manoeuvring the submarine). Owing to time constraints it was not possible to cover the full range of tasks and procedures, and reference to other phases of the analysis shows that there is more that could be added here. Indeed, further development and refinement of the decision ladder would be the first step in any follow-on work that aimed to take advantage of this particular output.

Finally, a number of 'leaps' were identified by the experts – shortcuts connecting two states of knowledge. One leap connecting the 'alert' and 'task' states of knowledge represented the fact that as soon as a fire is detected, it must be put out. Two leaps connected the 'alert' and 'target state' states of knowledge. Firstly, encountering a navigational hazard, for example something not marked on the charts, might trigger the need to perform a fast RTPD. Secondly, snagging a fishing vessel automatically leads to the target state of performing a slow ascent to perisciope depth, in order to minimise the risk to those onboard the snagged vessel. These leaps highlight the fact that the journey around the decision ladder does not necessarily involve travelling up one side and then down the other in a linear fashion. They also show how the decision making process can be iterative. In the case of detecting a fire, the leap to formulating the task of putting out the fire would immediately be followed by a return to the left hand side of the decision ladder in order to assess the severity of the fire and determine the correct course of subsequent action. This is also true when some of the target states relating to doing something other than RTPD are selected. If the decision is made to stay at depth, for example, then the situation will continue to be monitored until the system states are such that it becomes possible to RTPD.

Once the decision ladder had been completed, it was then possible to examine how the various elements related to each other. The left hand leg of the ladder contains information on the team's 'evaluation' of the state of the world (e.g., gathering information to understand the state of the system and identify the options available to them) whereas the right hand leg conatins information of the team's 'execution' activities (e.g., the target states and tasks they can implement when the option has been selected).

The decision ladder in Figure 2 shows the complete range of information, system states, tasks and so forth that can be involved in a RTPD, i.e. it is prototypical, as described previously. It can be useful to determine the relationships between each of these to understand which items of information contribute to which system states. This can be done for both 'legs' of the decision ladder. It should be noted that, in line with the formative nature of the CWA approach, the fact that elements are related does not mean that they 'do' influence each other, but rather that they 'could'. Firstly, the information, system states and options can be related to each other, as shown in Figure 3, with the black cells indicating a relationship.

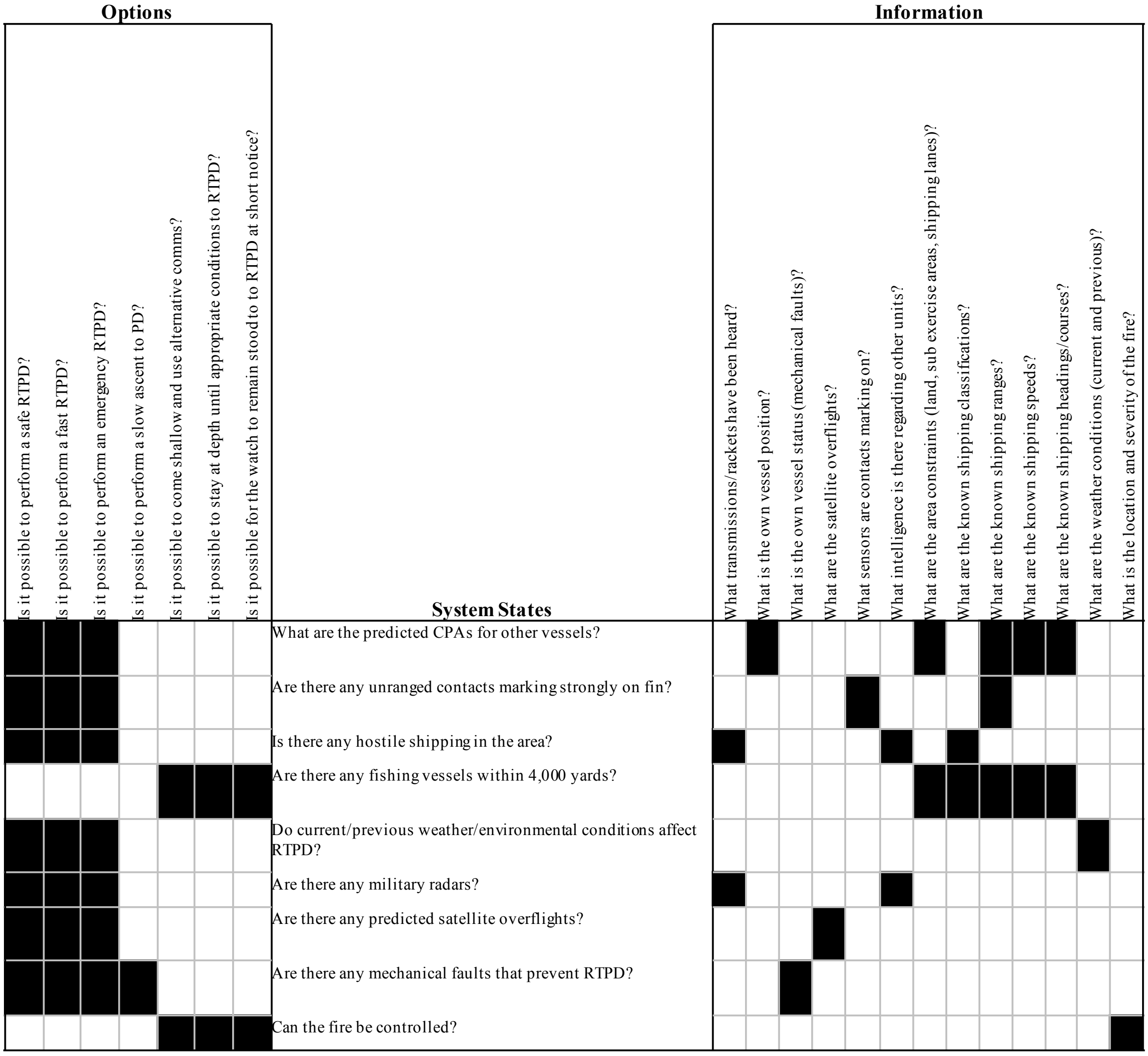

Figure 3. Matrix from the left hand leg of the decision ladder linking information to system states and options.

Relating the information and system state elements shows what information could be necessary to diagnose each system state. So, for example, determining whether there is any hostile shipping in the area requires some or all of: the transmissions or rackets that might have been heard; any intelligence regarding other units; and currently known shipping classifications. Whether there are any mechanical faults that would prevent RTPD, on the other hand, can be ascertained simply from information about the own vessel status.

The linking of the system states and options indicates how different system states might constrain the available options. It would appear that there is a basic split between states that constrain whether or not it is possible to RTPD, and states that constrain whether it is possible to choose some alternative. Where these constraints are in conflict, the chosen goal may come into play and influence the decision about the most appropriate target system state. For example, when a fishing vessel is detected within 4,000 yards, during peacetime and within home waters, a submarine is required to immediately RTPD (i.e. within a given timeframe). If, however, at the same time an unranged contact is marking strongly on fin (the contact is being detected on a particular sonar array, indicating that it may be extremely close) it could be unsafe to RTPD. In this case there is a choice to be made between returning to periscope depth and staying deep; the decision maker may choose 'RTPD safely' as the goal, leading to a decision to stay deep until the contact has been ranged. Alternatively, if 'RTPD within a given timeframe' is chosen as the goal, the submarine may RTPD whilst accepting the risk of the unranged contact.

Figure 4 relates the chosen goal, target states and tasks for the right hand leg of the decision ladder.

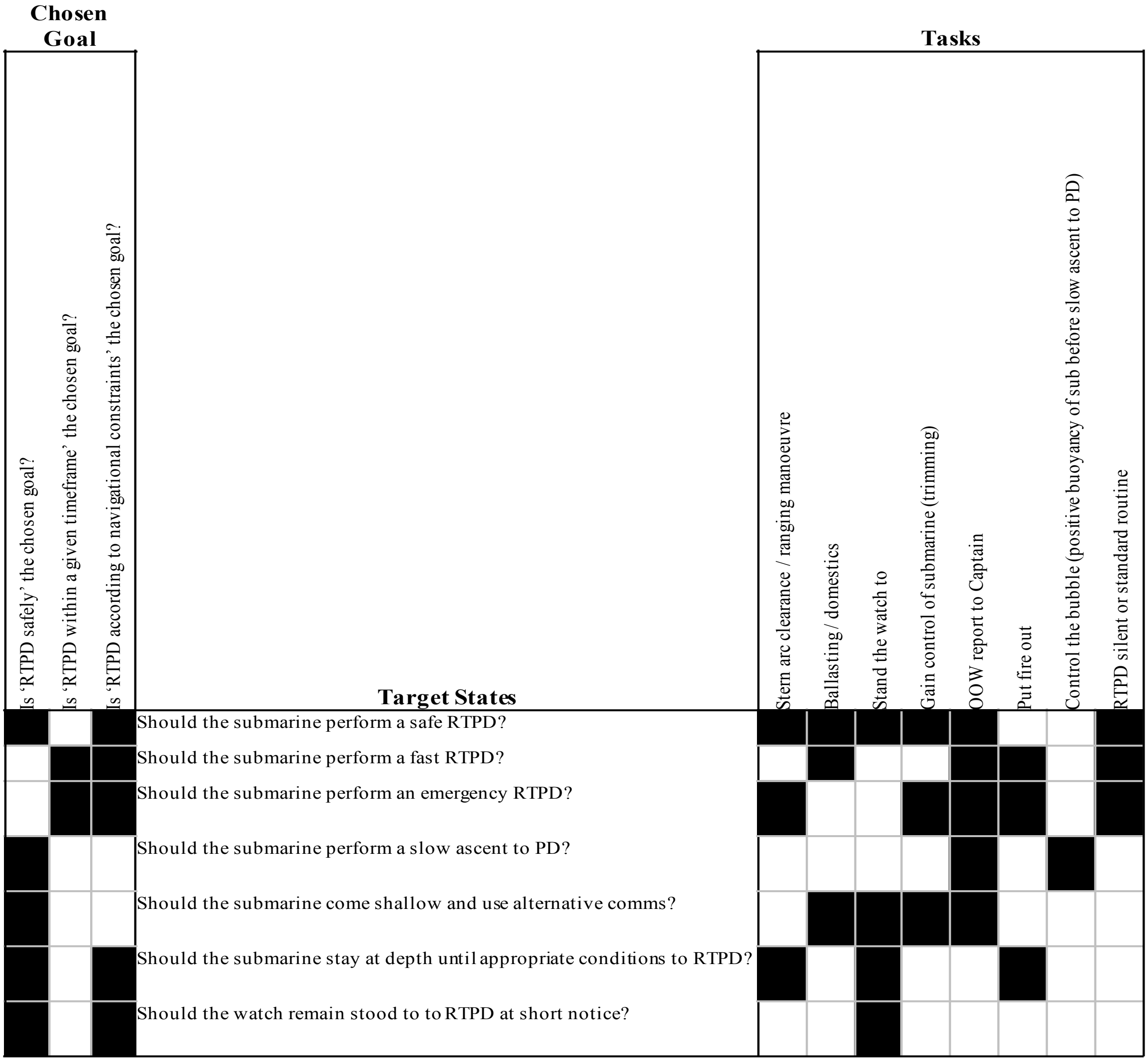

Figure 4. Matrix from the right hand leg of the decision ladder, linking chosen goal to target states and tasks.

It is apparent that, as expected, the chosen goal has an influencing factor on the way in which the submarine will attempt to return to periscope depth. For example, if safety is the overriding constraint, the RTPD will be neither fast nor emergency, meaning that the standard set of procedures will be followed. Conversely, the submarine will only ever perform a slow ascent to PD or come shallow as a result of some safety consideration (a fishing vessel snag or adverse weather conditions being the likely causes respectively).

Social Organisation and Cooperation Analysis (SOCA) addresses the constraints imposed by allocation of specific agent roles to functions in given situations. The objective is to determine how the social and technical factors in a socio-technical system can work together in a way that enhances the performance of the system as a whole. SOCA is concerned with identifying the set of possibilities for work allocation, distribution and social organisation. SOCA explicitly aims to support flexibility and adaptation in organisations developing designs that are tailored to the requirements of the various situations. In this way SOCA supports the idea of dynamic allocation of function, such that function allocation can transfer between agents as the situation changes. Flexible organisational structures are superior to rigid ones because they can adapt to local situations. Rather than defining a single or best organisational structure, SOCA is concerned with identifying the criteria that may shape or govern how work might be allocated across agents. Such criteria might include:

- Agent competencies
- Access to information or means for action
- Level of coordination
- Workload
- Safety and reliability
- Availability

The first stage of the process is to define the key agent roles in the system. The role reflects the work at any given time rather than a particular person or machine. It was observed in the control room studies that roles may change depending upon who is available, which shows that this idea has already been embraced. A list of the key roles and their related coding can be seen in Figure 5. SOCA is used to determine which roles are involved in information processing activities, and the resultant states of knowledge, at each stage of the decision making process. It can be seen in Figure 5 that the Captain or Executive Officer (XO) is solely responsible for evaluating the various options available against the overall goal, and then deciding against which constraints the goal will be achieved, leading to a desired target state. Conversely, a large number of varied roles can contribute towards alerting to the need to RTPD in the first place and then gathering the required information to diagnose the current system state. Similarly, many different roles are involved in executing the tasks and procedures resulting from the target state determined by the Captain or XO. One noticeable difference in comparing the two legs of the decision ladder is that the Planesman, Ship Control Officer and For-d Staff are only involved in responding to the target state, as might be expected given that their roles relate primarily to the control of the submarine. The Tactical Picture Supervisor, Office Of the Watch and Sound Room Controller, on the other hand, are involved at all stages of the process (with the exception of those reserved for the Captain or XO, as described previously).

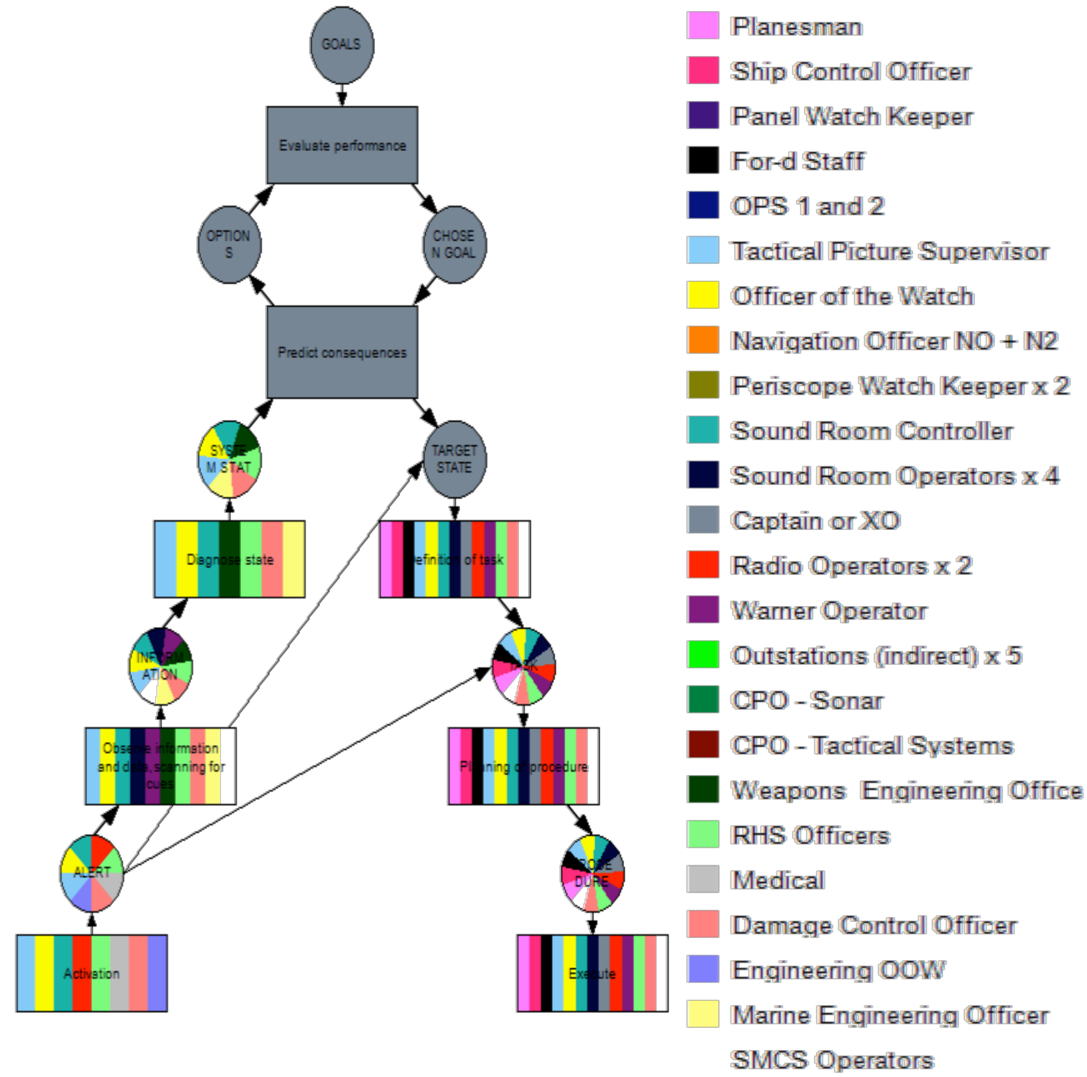

Figure 5. Social Organisation and Cooperation Analysis Decision Ladder (SOCA-DL) for returning to periscope depth in the control room.

CONCLUSIONS

RTPD is much proceduralised due to the inherent complexity and its safety critical nature. This means that many of the functions are unique to particular situations. Nevertheless, performance differences of experts were uncovered in the decision analysis, enabling the shortcut 'shunts' and 'leaps' across the ladder. Only two shortcuts were identified in this analysis (leaps across the decision ladder in emergency situations) which is due to the highly proceduralised, safety-critical, nature of the tasks. The analyses also show the activities for RTPD vary according to the situation (e.g., safe, fast emergency or slow assents).

The decsions ladder has shown utility in both the elecitation and representation of the decisions made by teams. Futher analysis is possible by using the matrices of the 'evaluation' and 'execution' legs of the ladder as well as identifying who is involved in the decisions (in the SOCA-DL phase). The mapping of information to states and options and the mapping of goals to states and tasks revelas how decisions are made in the submarine. Linking these elements in this way is useful because it provides an insight into how different pieces of information are required to determine the system states that inform option selection, and how different goals lead to various target states and their associated tasks. Once these relationships are understood they can be used for a variety of purposes such as informing interface, procedure, job aid, decision support and simulator design, as well as decision making training.

Thus the method has quite a lot of flexibility to describe quite complex team tasks. Making all of this complexity explicit is useful, as it allows the analyst to understand and question the nature of decision making and task structure. As the project was focused on next generation systems, the representations were verified with experts and will be used to support design of new system concepts. The task of returning to periscope depth currently focusses on eliciting constraints on the manoeuvre in general, and avoiding other vessels in particular. The obervational studies revealled that the team are looking for reasons not to return to the surface all the way up, only continuing with the transit in the absence of evidence to the contrary. At the moment, up to 13 roles are involved in this process, so it is quite a labour intensive activity, which requires good coordination and communication to be successful. Future concepts will investigate more efficient methods of designing the socio-technical system against this benchmark.

ACKNOWLEDGMENTS

This work from the Human Factors Integration Defence Technology Centre was partially funded by the Human Sciences Domain of the U.K. Ministry of Defence Scientific Research Programme. The authors are grateful to the submarine training team at Talisman in Devonport and to the Maritime Warfare Centre in Portsmouth for their help in this study.

REFERENCES

Jenkins, D. P., Stanton, N. A., Walker, G. H. and Salmon, P. M. (2009) Cognitive Work Analysis: Coping with Complexity. Ashgate: Aldershot.

Jenkins, D. P., Stanton, N. A., Salmon, P. M. and Walker, G. H. (2011) A formative approach to developing synthetic environment fidelity requirements for decision-making training, Applied Ergonomics, 42 (5), 757-769.

Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H. and Rafferty, L. A. (2010) Using the decision-ladder to add a formative element to naturalistic decision-making research. International Journal of Human-Computer Interaction, 26, (2), 132-146.

Klein, G. (2008). Naturalistic decision making. Human Factors, 50, 456–460.

Klein, G., & Miller, T. E. (1999). Distributed planning teams. International Journal of Cognitive Ergonomics, 3, 203–222.

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

Smith W., & Dowell, J. (2000) A case-study of coordinative decision-making in disaster management. Ergonomics, 43, 1153–1166.

Stanton, N. A., Rafferty, L. A., Salmon, P. M., Revell, K. A., McMaster, R., Caird-Daly, A. and Cooper-Chapman, C. (2010) Distributed decision making in multi-helicopter teams: case study of mission planning and execution from a non-combatant evacuation operation training scenario Journal of Cognitive Engineering and Decision Making, 4 (4), 328–353.

Stanton, N. A. and Wong, B. L. W. (2010) Explorations into naturalistic decision making with computers. International Journal of Human-Computer Interaction, 26, (2&3), 99-107.

Vicente, K. J. (1999) Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Mahwah, NJ: Lawrence Erlbaum Associates.

Zsambok, C. E., Klein, G., Kyne, M. M., & Klinger, D. W. (1992). Advanced team decision making: A developmental model. Fairborn, OH: Klein Associates.

Rehearsing for a Major Accident in a Metro Control Centre: A Naturalistic Analysis of Situation Awareness

Mikael WAHLSTRÖM[a,b], Hannu KARVONEN[a] and Leena NORROS[a]
[a]*VTT Technical Research Centre of Finland*
[b]*Helsinki Institute for Information Technology HIIT*

ABSTRACT

Introduction: Drawing on Peircean philosophy, this paper argues that shared situation awareness and decision making can be viewed as a process consisting of 'semiosis' or 'perception-action cycles': action is derived from perception of a 'sign' while each course of action implies meaningful basis for further action. **Method:** We analysed observation data of an accident rehearsal in a metro traffic control centre. The method includes interpretation of how courses of action are linked to each other and to situation awareness; it provides an overall description of how situation awareness is constructed in activity. **Results and discussion:** The process of creating situation awareness was heavily mediated; several individuals were needed for conveying knowledge from the scene of accident to the emergency response. The overall interaction pattern is explainable with the division of tasks in the control centre. The results point to practical ideas on how to streamline the flow of communication.

KEYWORDS

Situation awareness; transportation; accident management; metro traffic control; interaction analysis

INTRODUCTION

It is commonly thought that maintaining adequate situation awareness (SA) is a prerequisite for good decision making in safety-critical work. This paper proposes a theoretical background and method for studying situation awareness in work teams. We see SA as a phenomenon that progresses or builds-up hand-in-hand with work activity. This activity, in turn, is seen from a 'cultural' and 'ecological' perspective: activity is seen as intentional and attached to the 'meanings' offered by the environment of an actor. The method we are proposing consists of a meticulous step-by-step (or 'meaning/action-by-meaning/action') analysis of actions: each action potentially constructs SA and is a 'sign' for further actions. We demonstrate this method with a case study on metro control room work during a major accident rehearsal. Practical ideas on how the metro control work might be organised more efficiently are provided.

Situation awareness

The concept of SA has been discussed since the late 1980s (Endsley, 1988). At that time, the concept was used especially in the fighter aircraft domain to describe the pilot's observation of the opponent's moves and anticipation of future moves (Spick, 1988). After the success in the aviation domain, the concept diffused into various other safety-critical domains, such as transportation control centres (Golightlya, Wilsona, Loweb & Sharplesa, 2010), energy production control rooms (Burns, Jamieson, Skraaning, Lau & Kwok, 2007) and emergency medical dispatch (Blandford & Wong, 2004). Perhaps the most common definition of SA is by Endsley (1988, p. 97) according to which it is 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future'.

To apply SA in teamwork, the literature entails concepts such as 'shared situation awareness' (Endsley, Bolte & Jones, 2003), 'common ground' (Clark & Brennan, 1991), 'team cognition' (Salas et al. 2004) and 'team situation awareness' (Endsley, 1995). All of these concepts refer to shared understanding of a situation. Endsley and Jones (1997) have defined shared SA as the degree to which team members possess the same situation awareness. Team SA, in turn, as described by Endsley (1995), is the degree to which every team member possesses the SA required for his or her responsibilities. Team SA can be viewed to be constructed via collaboration, communication and co-operation between team members, and it also requires 'shared mental models' (Salmon, Stanton, Walker & Jenkins, 2009), that is, organized bodies of knowledge that enable members to anticipate each other's actions and to perform functions from a common frame of reference (Cannon-Bowers, Salas & Converse, 1993). All this seems theoretically coherent but blind spots in the literature

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

have also been suggested. Salmon et al. (2009, p. 24) write in their review of the literature that '[i]t seems logical to assume that an increased level of teamwork will lead to enhanced levels of team SA; however the specific relationship between team behaviours and attributes and team SA remains largely unexplained'. Collaborative making of shared/team SA has been addressed with varying abstract models in which interaction and collaboration between team members is a central variable (Salmon et al., 2009), but it is this interaction that demands further elaboration.

This assumption of deficiency in the literature, however, can partly stem from lack of sufficient synthesis between study lines. A number of studies exist that have drawn from ethnography, ethnomethodology and conversation analysis to explore interaction and use of tools (Heath & Luff, 2000). Studies of this type have addressed the issue of collaborative sense making of situations. For example, in a study of a London Underground line control room it was found that workers rarely provide explicit information to each other. Instead, they monitor each other 'peripherally' and make their own activities visible for others with subtle gestures and glances directed toward the tools used and by talking 'to oneself' (Heath & Luff, 2000, pp. 88–124). Similar reciprocal monitoring has been found in dispatch centres (Whalen & Zimmerman, 2005) and air traffic control (Harper & Hughes, 1993; Mackay, 1999).

The approach of our study parallels with ethnomethodological / conversation analytical workplace studies in applying a naturalistic analysis of activity. Workplace studies or 'situated action models', as studies of this type have been dubbed elsewhere (Nardi, 1992), however, are *not* concerned with 'meanings' attached to activity (Heath & Luff, 2000, p. 18). In contrast, we infer how courses of activity relate to and produce an actor's understanding of a situation. Theoretical basis for this approach is explained in the following.

Semiotic analysis of activity that creates situation awareness

According to the philosopher Charles Sanders Peirce (1998), people connect themselves to the possibilities of the environment through continuous perception-action cycles. To explain this, one may first consider the three linked elements in Peirce's analysis of signs, or, semiotics. First comes the *sign*, which can be considered the physical or 'vehicular' element needed for conveying meaning, for example, an image or a sound (e.g., smoke as a sign for fire). The *object*, in turn, is what the sign refers to; the fire that is signified by smoke. The sign/object (e.g., ink/letter) relation would not exist, however, without the third essential element, the *interpretant*: it can be thought to refer to understanding of the sign (Atkin, 2010). Let us take an example from the field of control centre work. A 'sign' could be a beeping sound and the 'object' related to this sound could be an alarm if noticed by an interpreter, such as, a control centre worker. Two points should be emphasised here. First is the relation between the interpretant and object. There are, of course, different kinds of interpretants or manners of comprehensions and therefore different kinds of objects. One operator might perceive the same sound as something alarmingly grave while another might perceive it as a simple glitch, which does not require much further action; in any case, the reaction would depend on operator's perception of the reality. The second point to be emphasised is the 'cyclic' and social aspect of Peirce's model: each action attached to an interpretant (e.g., a reaction to the alarm sound) can be seen as a sign/object if perceived by, say, another worker within the control centre. A reaction to an alarm by a worker would, indeed, influence the actions of other workers. These actions would then serve as further signs/objects and the perception-action cycle continues in the social environment of the control room. Figure 1 illustrates Peirce's model and the example above.

As implied in the example above, one may see that analysing these perception-action cycles has clear potential for inferring how SA develops within a work team. It progresses as linked actions and perceptions take place: an alarm sound first signifies one thing, and then, as displays are observed and calls are made, along with short verbal exchanges within the team, SA gradually builds up and is maintained.

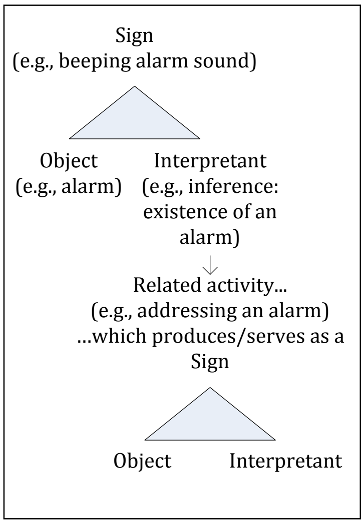

Figure 1. Illustration of Peirce's model

There is a myriad of ways to conduct semiotic analysis: the connections between signs, objects and interpretants can be examined in practically any given way. The approach adopted in this study draws from Norros's (2005; Savioja, Norros, Salo & Aaltonen, forthcoming) semiotic analysis of work. For Norros, 'action', what a worker does, is both a reflection of interpretant (worker's understandings and style of interpreting things) and a sign (especially for the other workers) and therefore a central point of analysis. Additionally, she argues that semiotic analysis of work activity should be preceded by a 'core-task analysis', that is, analysis of the essential content of the work activity: the aims, meanings and challenges of work. It is assumed that the major determinant of activity is its purpose, and therefore, it should be the major element considered by researchers of work activity. Referring to the cultural-historical theory of activity (Leont'ev, 1978), she also emphasises that these objectives should be understood in their social-historical context: people's actions reflect identities, hierarchies and organisational aims. These ideas are in line with Peirce's ideas on the analysis of signs: without understanding those who interpret (and their goals) it is impossible to understand signs/objects and consequent actions. The bottom line here is that researchers should become knowledgeable of the relevant particularities of a certain work context, such as, terminology, goals, tools, formal hierarchy and procedures if they are to do semiotic analysis on work activity.

In principle, the semiotic model has potential to complement work place studies (Heath & Luff, 2000) or situated action models (Nardi, 1992) by emphasising the element of 'meaning' in the analysis of interaction, this being an issue, which practitioners of these models de-emphasize. On the other hand, Peirce's model is perfectly compatible with the concepts of 'macro-cognition' (Schraagen, Militello, Ormerod & Lipshitz, 2008) and 'distributed cognition' (Hutchins, 1995), which are also popular among those studying real-life work contexts; these concepts emphasise the union of thinking and environment and imply the analysis of group efforts, culture and tools in examining how people accomplish cognitive tasks in naturalistic settings. Sign/object/interpretant-triad suggests unison between cognition and environment.

In the approach taken by this paper SA is central in the semiotic analysis of work activity: it is both a motive explaining activity (achieving SA is a goal) and a 'sign/object/interpretant' (that is, understanding of certain signs) needed for activity. Arguably, this approach addresses the challenge, identified by Salmon et al. (2009), of linking SA with actual team behaviour and team attributes.

THE SETTING

To understand the domain of metro operation and its traffic control work, several interviews were conducted before the actual accident rehearsal was studied. While a detailed analysis of the Helsinki Metro and its challenges have been presented elsewhere by us (Karvonen, Aaltonen, Wahlström, Salo, Savioja & Norros, 2011), in the following we depict shortly the accident rehearsal scheme and the traffic control room.

Accident rehearsal

In October 2009, a major rehearsal was organised in the Helsinki Metro. In the rehearsal scenario, due to construction work, metro trains coming from east to west have to be guided exceptionally to the south rail instead of the typical north rail; the metro system entails only a single forked line. In one of the metro stations near the city centre, called Hakaniemi, there is a long railway turning point, which can be driven with the speed of 60km/h (37.3mph). The accident train departs from Hakaniemi station and accelerates into the turning point with too much speed. This causes the train to derail and crash to the tunnel wall. Thirty of the passengers are injured. Finally, after the crash, a fire breaks out with large amounts of smoke filling the tunnel. In reality, the accident rehearsal train was simply stopped on the rails, and smoke was generated with smoke machines.

The rehearsal started at 00:30 (i.e., after the end of actual metro operation) and lasted for one and a half hours. In the rehearsal scenario story, the accident took place at 23:00, when the metro traffic is still operative. The passengers recruited for the accident train were actors, each playing their own assigned role in the scene of the accident (e.g., those without proficiency in Finnish, unconscious, disabled, in a wheel chair, under the influence of alcohol, etc.). According to the Helsinki Metro organisation, the accident rehearsal was the biggest in the history of the organisation. The goal of the rehearsal was to practice metro rescue tasks and co-operation between different actors during a major accident.

Traffic control room

During the accident rehearsal, the control room personnel consisted of a team that would have also been normally on shift at that time. The team included the following workers:

- Traffic Controller 1 (TC1), responsible for traffic control, also during the accident rehearsal;
- Traffic Controller 2 (TC2), responsible for traffic control, during the rehearsal responsible for taking care of the accident train;
- Traffic 1 (T1), otherwise a regular traffic controller, but during an accident is being called 'Traffic 1' and is responsible for taking care of the accident train's driver at the scene of the accident;
- Technical Controller (TeC), responsible for the technical control of the metro system's equipment (e.g., electricity control); during accidents it his/her responsibility to contact the emergency response centre.

Normally, traffic controllers' tasks include starting the traffic, inserting schedules, taking care that the drivers are in the right trains and securing that the rails are in order for the trains. More minor tasks include keeping a record on how long distances the trains have travelled, making passenger announcements, observing the platforms through CCTVs, instructing the train drivers and sending help to the field if necessary. Regular traffic controller work has been portrayed in previous publications (Heath and Luff, 2000; Smith et al., 2009; Theureau and Filippi, 2000). The main responsibilities of a technical controller, in turn, include electricity,

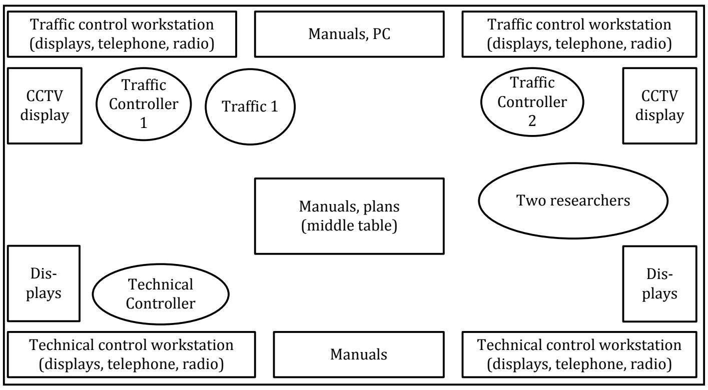

Figure 2. The control centre setting during the accident rehearsal

escalator and air conditioning management, and monitoring of devices. Figure 2 portrays the seating locations of the team during the accident rehearsal. The traffic controllers had workstations in the two corners of a large U-shaped table while TeC sat at another similar table. T1 sat next to TC1 when the accident rehearsal started. Two researchers were located next to a table between the U-shaped tables.
Traffic controllers' main tools include telephones, radio phones, platform announcement equipment, different manuals and check-up lists, CCTV camera monitors and their controllers, interlocking system's displays and controllers, and a general overview display of the traffic situation from which one can see where each train is located. TeC's main tools include electricity management displays, tunnel air flow monitoring system, telephones and radio phones.

OBSERVATION STUDY

Data collection

The metro control centre work was observed by two researchers while a third researcher observed the accident site (they were the authors of this paper). Three cameras were used for the control room observation. A stationed camera focused on TC2 during the whole rehearsal. Second camera was stationed to the middle of the room; it was vertically rotatable and focused mainly on the traffic controllers. The third camera was hand-held, focusing on where was deemed to be most activity. The recording started several minutes before the rehearsal; the whole rehearsal was video recorded. The length of the rehearsal was one hour and 15 minutes when calculated from the first communication of the accident to the end of the rehearsal.

Analysis

The aims of analysis were, first, to document and depict direct observations of the interactions and other meaningful events during the accident rehearsal. We also wanted to interpret why each course of action took place, that is, what kind of 'interpretant' or 'meaning' could be explicative of an action. Additionally, and also drawing from the idea of 'meaning/action'-cycles discussed above, we wanted to see how the interactions are interconnected: who interacts with whom and how these interactions promote situation awareness. Finally, the aim was to give some development recommendations for the transport company.
The analysis progressed in three phases. First, the video recording was transcribed: we wrote down what was said by whom and when during the accident rehearsal. Second, the transcribed data was categorised into 'events' in sequential order; an event refers here to a short reaction or period of activity by an individual (or by any agent, including non-human). For example, a question or request can be considered an event. Second, these events were listed to a spreadsheet for further annotation. In line with the aims of the analysis, this annotation included the following: 1) short paraphrasing of the content of the event (e.g., shortly the main content of said things), 2) the source of event data or the 'sign source'; this included both the device and/or the individual through which the event manifested, and 3) a short interpretation of the action was marked down; the assumed reason for the event was inferred here. After this, 4) the emerging situation awareness – produced by the event – was inferred and marked down. Finally, 5) the rightmost column of the spreadsheet was reserved for intuitive notes on anything that might make the analysis useful in practice. The video data was examined recurringly when these annotations were made.
After watching the whole video for several times, this detailed analysis was done for two crucial episodes in the making of SA. First, we wanted to study the first moments during which SA was construed within the control centre team. The question here was what was needed to be done prior calling the regional emergency centre. Second crucial episode, which immediately followed the first one, was making this call. Interactions took place while help was alarmed. We distinguish these two episodes because they entail different vantage point to the question of SA: in the first one it is examined how SA is produced within the control centre while the second was about 'transferring' SA for the regional emergency centre. Altogether, the detailed analysis was done to two minutes of the observation material.
The analysis method draws from a study by Norros, Hutton, Liinasuo, Määttä, Tukeva and Immonen (2009): the aim is to identify and name the relevant elements in chains of communication (activities, understandings and tools in particular). The method version of our study, however, provides a more refined presentation of the events during which SA is created by a work team.

Results

Table 1 portrays analysis of the first 50 seconds during which the view of the accident situation progressed within the control centre team. The first event and 'second count' starts from the moment where a radiophone alarm sound is addressed by TC2. The last event (no. 16), in turn, indicates the point in which the emergency response services are called. Overall, Table 1 portrays how situation awareness gradually builds up in interactions. Information first diffuses from the metro driver to TC2 (no. 2) who then repeats driver's utterances (no. 3). Although driver's voice is in principle audible for the whole team through a loudspeaker, it is quite possible that TC2's repetition serves as an 'outloud' further disseminating the message. They are short shouts of information that are not directed to any specific person. Previous studies report their use for construction of SA in control centres (Wahlström, Salovaara, Salo & Oulasvirta, 2011), and making an outloud can be considered more efficient than addressing a colleague because they allow that colleagues do not have to interrupt their tasks to deliver a response (Heath et al., 1993). Repeating out loud information that comes via radio arguably also serves the purpose of certainty: when the main content is repeated there is little chance for misinterpretation. This is necessary especially when understanding the message cannot be confirmed with body language, such as with affirmative nods. Both TC2 (no. 3) and the metro driver (nos. 5, 10, 15) repeat each other. In one occasion, the driver actually repeats himself, informing twice that he/she is hurt (compare nos. 2 and 10); this is perhaps a response to the lack of repetition by TC2, as this piece of information remains unrepeated by him/her (in no. 3).

Table 1. Events at traffic control room prior calling the emergency response centre (Episode 1)

No.	Time	Direct observation / *'sign'*	Sign source	Action (who involved) / *'interpretant'*	View of situation (among different actors) / *'object'*	Notes
1	0 seconds	Alarm sound rings	Alarm sound of a communication radio	Answering the radio call (TC2)	Something is wrong; someone wants to be in contact	
2	1–8s	Driver (D) (loudspeaker): train fell off the track, leaning against the wall, minor injury,	Driver / radio loudspeaker	Informing about the accident (D -> TC2)	train off the track, leaning against the wall, driver having minor injury	redundancy
3	8–11s	TC2 (shouts): train off the track	Traffic Controller 2	Outloud / repetition (TC2, D, whole control room)	[as above]	redundancy
4	12–16s	T1 (shouts): *cut the electricity between Hakaniemi and Kaisaniemi*	Traffic 1 / traffic control display	Request to Technical Controller (T1 -> TeC)	[as above] + approximate location of the derailed train + electricity will be cut	
5	12–16s	D (loudspeaker): *train derailed after Hakaniemi*	Driver / radio loudspeaker	Adding location information (D -> TC2)	[as above] + more detailed location of the derailed train	redundancy
6	15–16s	TeC: *should we call?*	Technical Controller	Asks from about calling to emergency response (TeC -> T1)	[as above]	questions 1
7	16–18s	T1: *call emergency services*	Traffic 1	Confirmation on calling ES (T1 -> TeC)	[as above] + help will be called	response 1
8	18–27s	TC2: *turning voltages down and sending help*	Traffic Controller 2	Informing the driver (TC2 -> D)	[as above] [now also for the driver]	
9	20–21s	TeC: *where is it exactly?*	Technical Controller	Asks about exact location (TeC -> T1)	[as above]	question 2
10	27–32s	D (loudspeaker): help is coming and I am hurt myself too	Driver / radio loudspeaker	Repeats what is known (D -> TC2)	[as above]	redundancy
11	27–30s	T1: *between switches A4-B4*	Traffic 1 / traffic control display	Response to question (T1->TeC)	[as above] + more exact location	response 2
12	32–34s	TeC: *it is both sides* [of the track]*?*	Technical Controller	Additional question on location (TeC-> T1)	[as above]	questions 3
13	34–40s	TC2: *can you inform the passengers on what has happened and not to leave the train?*	Traffic Controller 2	Gives guidance to the driver (TC2 -> D)	[as above]	guidance
14	34–37s	T1: both sides, at the switch	Traffic 1	Answers question (T1 -> TeC)	[as above]	response 3
15	40–45s	D (loudspeaker): *informing passengers and now smoke is coming from the train*	Driver / radio loudspeaker	Repeats and delivers more information (D -> TC2)	[as above] + smoke is coming from the train	redundancy
16	46–50s	TeC: this is a rehearsal call	Technical Controller / telephone	Calls the emergency response (TeC)	[as above]	

Alongside the interaction between the driver and TC2, a parallel exchange took place between T1 and TeC. It is TeC's responsibility to call the emergency response services but prior doing this a short series of questions and answers were made (Table 1, nos. 6 and 7; 9 and 11; 12 and 14). TeC does not have the means to gather the information required via his/her own workstation; therefore, s/he consults T1 who checks the needed information from the screens of TC1.

TeC's need for making questions to interact with the emergency response is visible also in Table 2. This table is a more reduced presentation of the events in the control centre than Table 1 since it only entails the interactions made by TeC during the emergency call; there were other interactions taking place as well, but they are not presented here for the purpose of clarity and simplification. In the beginning of the emergency call TeC explains the basic information about the situation (Table 2, no. 1); then repeats and further specifies the situation (no. 2). Questions are made for TC2 (nos. 3 and 4; 7 and 8).

Overall, the development of situation awareness and transferring it to the emergency response was a joint effort. Explicative of these interactions is the division of tasks between traffic controllers and TeC: the latter makes the emergency call while traffic controllers interact with metro drivers. Information reaches the emergency centre indirectly through exchanges involving the driver, traffic controllers and TeC.

The concepts 'sign', 'interpretant' and 'object' are marked on the tables to clarify the way in which Peirce's model is visible beneath our analysis. 'Sign' corresponds roughly with direct observations. While the 'vehicular' sign element of interaction is, in principle, directly visible, the workers' view of situation and related actions are the two interrelated elements that require interpretation during the analysis. The column 'Action', which is to describe the event and to grasp the basic reason of activity, reflects the 'interpretant' concept because assumedly people's actions and inferences are closely interrelated. On the other hand, the concept 'object' is needed here to represent the 'representational' content 'produced' in people's inferences and activity: see the column 'View of situation'. Each row in the tables corresponds to a sign/object/interpretant-triangle plus related action combination as presented on Figure 1. Actually, the example given in Figure 1 parallels with the event no. 1 on Table 1.

Table 2. Interactions of Technical Controller during an emergency call (Episode 2)

No.	Time	Direct observation / *'sign'*	Sign source	Action (who involved) / *'interpretant'*	View of situation (as shared with the emergency centre) / *'object'*	Notes
1	0m 46s – 1m 03s	TeC (speaks to phone): this is a rehearsal call, we have a train from Kaisaniemi to Hakaniemi and the driver let us know that the train has been derailed at the switch,	Technical Controller / telephone	Calls the emergency response centre (ERC) (TeC -> ERC)	train off the track, train location	
2	1m 05–35s	TeC (speaks to phone): yes towards Kaisaniemi, that is, the train is in the tunnel, the driver himself is hurt at least and there are passenger on board, and yes we have turned down electricity from both sides	Technical Controller / telephone	Repeats and explains further for emergency response centre (TeC -> ERC)	[as above] + more about train location, injured driver, possibly injured passengers, electricity has been cut from the track	specification
3	1m 39–42s	TeC (shouts): wait a second, how many train cars do they have	Technical Controller	Asks information on emergency response's behalf (TeC -> TC2)	[as above]	question 1
4	1m 42–44s	TC2: one train car pair, that is, two train cars	Traffic Controller 2	Answers the question (TC2 -> TeC)	[as above]	response 1
5	1m 44–46s	TeC (speaks to phone): two train cars, that is, one train car pair	Tech Controller / telephone	Provides requested information the for emergency response centre (TeC -> ERC)	[as above] + the amount of trains	
6	1m 46–47s	TC2: one train car pair	Traffic Controller 2	Repeats what has been said (TC2 -> TeC)	[as above]	redundancy
7	1m 53–55s	TeC: are there any signs of smoke	Technical Controller	Asks information on emergency response's behalf (TeC -> TC2)	[as above]	question 2
8	1m 55–56s	TC2: there is smoke in the train	Traffic Controller 2	Answers the question (TC2 -> TeC)	[as above]	response 2
9	1m 56–57s	TeC (speaks to phone): *yes there is smoke*	Technical Controller / telephone	Provides requested information for emergency response centre (TeC -> ERC)	[as above] + there are signs of smoke	

Practical implications

It is arguably a good practice to repeat information in safety critical communication. This redundancy (see 'notes' in Table 1 and 2) is likely to provide robust or fault-free communication. In contrast, the fact that information is mediated through several actors might cause mistakes. The information content might vary as it passes through different individuals; think of a game of Chinese whispers. Further, one may question the efficiency of mediated communication. For example, in this case, it took a minute and twenty seconds for the information on smoke to reach the emergency response (from event no. 15 in Table 1 to event no. 9 in Table 2).

For streamlining the flow of communication, one may first consider a change in responsibilities. The call to emergency response could have been done by the driver himself or by TC2. These options have their disadvantages, however. The driver might be seriously injured or not to have all the information needed. TC2 could make the call but then s/he would have two parallel tasks: 1) to give guidance for the driver (see Table 1 no. 13) and to talk with the response services. Further, there are pieces of information that can be accessed only by TeC. The emergency response might be interested on issues that are in TeC's domain, such as, tunnel wind flow directions and electricity being switched on or off in different areas and equipment. Later in the video data, a worker of rescue services, an incident commander, arrives to the control centre room and discussion takes place between him and TeC. They have to be sure, for example, that smoke is not killing the rescuers; this depends on air flow directions in the tunnel. In other words, there are good reasons as to why TeC is the connection point between the emergency response and the metro control.

A solution, however, might be a big shared screen with all the most important pieces of information, such as trains' locations and statuses. TeC would hence have much of the same information as traffic controllers. As this screen would be shared by all, it might also facilitate the discussion on the actualities at the metro track. A previous study reports the use of a screen of this kind in a rally control centre (Wahlström et al., 2011).

Even more technologically advanced possibilities can be imagined: the traffic control screen might be shared directly between the metro control centre and the emergency response; hence, the control centre workers could explain the situation with the help of a shared visual presentation. This might be beneficial since the discussion between the emergency response and TeC took quite long; TeC had to repeat and specify information (see Table 2, event no. 2). Further, it might be beneficial for TeC to have a predefined list of the issues that the emergency response needs to know; s/he could provide all the necessary information directly and not after inquiries made by the emergency response (see Table 2, no. 7). Also, given that TeC has to discuss with the traffic controllers, the physical seating arrangement is not optimal; the workers had to make 180° turns for interacting with one another. In any case, it would be better if the seating order promotes rather than hinders communication.

DISCUSSION

This study would have been a somewhat different if we would have applied 'situated action models' (Nardi, 1992) or 'workplace studies' (Heath & Luff, 2000) instead of the semiotic model of work activity. While those 'conversation analytical' methods aim to describe specific (and often very short) events in exact detail (featuring, e.g., intonations and lengths of pauses between words), the method used in this study featured somewhat less precise descriptions of 'functionally relevant episodes', that is, episodes in which certain function or task is accomplished by a work team. Arguably, the analysis method used in this study is more readily adoptable for those with practical rather than academic aims when compared to situated action models. It seems that conversation analytical workplace studies serve as an antithesis to 'individual centred' or 'cognitively orientated' human-computer interaction studies: by pointing out the collaborative nature of actual work practices one may, indeed, criticise 'the conventional wisdom of HCI which places the single user and their cognitive capabilities at the centre of the analytic domain' (Heath & Luff, 2000, p. 122). One may, however, also argue that cognitive concepts such as 'situation awareness' are intuitive (because people commonly use these concepts) and therefore they allow to present information in a readily understandable manner; this is exemplified by the two tables of this study showing how 'situation awareness' progresses with respect to time and activity. Avoiding the use of cognitive concepts can thus be problematic in practical terms. Conversation analytical findings are burdensome for the reader, which limits their practical usefulness.

Theoretically, in turn, a 'non-cognitive' (Heath & Luff, 2000, p. 18) approach to interaction is insufficient as some meanings, of course, underlie any human activity. Hence, interaction analysts should adopt a theoretical approach that takes into consideration the operative connection between individual perception, action and environment. Peirce's triadic model is applicable here; it enables to analyse the joint creation of meaning of a situation. Overall, although Salmon et al. (2009) are not entirely correct in their suggestion that relationship between team behaviours and SA remains unexplored – situated action models have addressed this issue – it is still justified to seek for new models and methods addressing the link between behaviour and situation awareness, as was done in this paper.

ACKNOWLEDGMENTS

We thank HKL, the Helsinki transport company, for cooperation.

REFERENCES

Atkin, A. (2010) Peirce's Theory of Signs, in E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy *(Winter 2010 Edition)*. Retrieved from http://plato.stanford.edu/archives/win2010/entries/peirce-semiotics

Blandford, A. & Wong, B. L. W. (2004) Situation awareness in emergency medical dispatch, *International Journal of Human-Computer Studies*, 61, 4, 421-452.

Burns, C., Jamieson, G., Skraaning, G., Lau, N. & Kwok, J. (2007) Supporting situation awareness through ecological interface design, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 51, 4, 205-209.

Cannon-Bowers, J. A., Salas, E. & Converse, S. A. (1993) Shared mental models in expert team decision making, in N. J. Castellan, Jr. (Ed.), Individual and group decision making: Current issues, pp. 221-246, LEA Hillsdale, NJ.

Clark, H. H. & Brennan, S. E. (1991) Grounding in communication, in L. B. Resnick, J. M. Levine & S. Teasley (eds.), Perspectives on socially shared cognition, pp. 127-149, APA, Washington, WA.

Endsley, M. R. (1988) Situation awareness global assessment technique (SAGAT) *Proceedings of the National Aerospace and Electronics Conference (NAECON),* New York, NY.

Endsley, M. R. (1995) Toward a theory of situation awareness in dynamic systems, *Human Factors*, 37, 1, 32-64.

Endsley, M. R., Bolte, B. & Jones, D. G. (2003) Designing for situation awareness: An approach to human-centered design, Taylor & Francis, London, UK.

Endsley, M. R. & Jones, W. M. (1997) Situation awareness, information dominance, and information warfare, Wright-Patterson AFB, United States Air Force Armstrong Laboratory.

Golightlya, D., Wilsona, J. R., Loweb, E. & Sharplesa S. (2010) The role of situation awareness for understanding signalling and control in rail operations, *Theoretical Issues in Ergonomics Science*, 11, 1-2, 84-98.

Harper, R. H. R. & J.A. Hughes (1993): "What a F-ing System! Send 'em All to the Same Place and Then Expect Us to Stop The Hitting": Making Technology Work in Air Traffic Control. In G. Button (ed.), Technology in Working Order, pp. 127-144, Routledge, London, UK.

Heath, C., Jirotka, M., Luff, P. & Hindmarsh, J. (1993) Unpacking Collaboration: the Interactional Organisation of Trading in a City Dealing Room, *Proceedings of ECSCW 1993*, Milan, Italy.

Heath, C. & Luff, P. (2000) Technology in action. Cambridge University Press, Cambridge, UK.

Hutchins, E. (1995). Cognition in the wild, MIT Press, Cambridge, UK.

Karvonen, H., Aaltonen, I., Wahlström, M., Salo, L., Savioja, P. & Norros, L. (2011) Hidden roles of the train driver: A challenge for metro automation, *Interacting with Computers,* 23, 4, 289-298.

Leont'ev, A. N. (1978) Activity, Consciousness, and Personality. Prentice Hall, Englewood Cliffs, NJ.

Mackay, W. (1999) Is paper safer? The role of paper flight strips in air traffic control. *ACM Transactions on Computer-Human Interaction,* 6, 4, 311-340.

Nardi, B. (1992) Studying context: A comparison of activity theory, situated action models and distributed cognition, *Proceedings of East-West HCI Conference*, St. Petersburg, Russia.

Norros, L. (2005) The concept of habit in an activity-theoretical analysis of situated actions, *Theoretical Issues in Ergonomics Science*, 6, 5, 385-408.

Norros, L., Hutton, R., Liinasuo, M., Määttä, K., Tukeva, P. & Immonen, A. (2009). HF-based design inputs to COPE technology - Conceptual and empirical considerations of Common Operational Picture, Deliverable No D2.2. Project report. Retrieved from http://cope.vtt.fi/publications.htm

Peirce, C. S. (1998) The Essential Peirce, Selected Philosophical Writings. Indiana University Press, Bloomington, IN.

Salas, E. & Fiore, S. M. (2004) Team cognition: Understanding the factors that drive process and performance, American Psychological Association, Washington, DC.

Salmon, P., Stanton, N., Walker, G. & Jenkins, D. (2009) Distributed situation awareness: Theory, measurement and application to teamwork. Ashgate, Surrey, UK.

Savioja, P., Norros, L., Salo, L. & Aaltonen, I. (forthcoming) Identifying Resilience in Proceduralised Accident Management Activity. Manuscript in peer-review.

Schraagen, J. M., Militello, L. G., Ormerod, T. & Lipshitz, R. (2008) Naturalistic Decision Making and Macrocognition. Ashgate, Hampshire, UK.

Smith, P., Blandford, A. & Back, J. (2009) Questioning, exploring, narrating and playing in the control room to maintain system safety. Cognition Technology & Work, 11, 4, 279-291.

Spick, M. (1988) The Ace Factor: Air Combat and the Role of Situational Awareness, Naval Institute Press, Annapolis, MD.

Theureau, J. & Filippi, G. (2000) Analysing cooperative work in an urban traffic control room for the design of a coordination support system, in Luff, P., Hindmarsh, J. & Heath, C. (Eds.), Workplace Studies-Recovering Work Practice and Informing System Design, pp. 68-91, Cambridge University Press, Cambridge, UK.

Wahlström, M., Salovaara, A., Salo, L. & Oulasvirta, A. (2011) Resolving safety-critical incidents in a rally control center, *Human-Computer Interaction,* 26, 1&2, 9-37.

Whalen, J. & Zimmerman, D. H. (2005) Working a call: Multiparty management of calls for help, in C. Baker, M. Emmison & A. Firth (Eds.), Calling for help: Language and social interaction in telephone help lines, pp. 309-345, John Benjamins, Amsterdam, the Netherlands.

Healthcare and Naturalistic Decision Making

Use of a Standardized Mnemonic, Formal Training, and Experience Improves Situation Awareness during Surgical Resident Handoffs

Grace A. NICKSA MD[a,b], Ryutaro HIROSE MD[a], Linda M. REILLY MD[a] and Lygia STEWART MD[a,b]

[a] *Department of Surgery, University of California, San Francisco, 513 Parnassus Avenue, Room S-321, San Francisco, CA 94143*

[b] *San Francisco VA Medical Center, Surgery (112), 4150 Clement Street, San Francisco, CA 94121*

ABSTRACT

Introduction: Surgical training regulations in the United States limit physician daily work hours and this creates frequent transfers of care (patient handoffs) between physicians. This observational study assessed factors associated with effective surgical resident handoffs. **Method**: A single investigator observed 141 surgical resident handoffs (junior residents, senior resident, moonlighter handoffs; evaluated on a Likert scale 1-5, 5=best). Observation occurred before and after formal handoff training. **Results:** Formal handoff training and increased resident experience (i.e. senior residents) was associated with improved situation awareness during handoffs. Training significantly improved junior resident's handoffs, moonlighter's ability to receive handoffs, and minimized the effect of distractions during handoffs. **Discussion:** The use of formal training increased the proficiency of junior resident patient handoffs to that of senior residents many years more advanced in their training. Formal education accelerates the natural course at which junior surgical residents and moonlighters acquire proficiency in conducting efficient patient handoffs.

KEYWORDS

Practical Application; Cognitive Field Research and Cognitive Task Analysis; Situation Awareness / Situation Assessment; Patient Safety; Handoffs; Surgical Resident Training

INTRODUCTION

Prior to 2003, a single surgical resident team managed the care of patients on their service (with attending surgeons). All surgical residents on the team knew the patients and alternated taking in-hospital call. This meant that there was always a member of the team in the hospital, so handoffs were unnecessary. This required surgical residents to be on call frequently (every 2-3 nights) and work long hours. In 2003, the Accreditation Council for Graduate Medical Education (ACGME) limited resident work hours to 80 hours/week, and mandated that residents go home within 6 hours of 24 hour call and receive one day off a week. This change required surgical residents to cross-cover other surgical services; and led to the use of moonlighters (physicians paid to work overnight or weekends). This increased the number of handoffs from one physician to another, and also meant that the physician managing the patient was more frequently not a member of the primary surgical team. In subsequent years, the ACGME further limited first year surgical residents work hours to <16 hours/day (Antonoff, Berdan, Kirchner, Krosch, Holley, Maddaus, D'Cunha 2013); this further increased the number of patient handoffs among residents. It has been estimated that in one month a first year surgical resident will perform >300 patient handoffs, and patients hospitalized for 3 days will have their care signed out from one provider to another >15 times (Antonoff et. al 2013).

Effective communication is essential in effectively transferring critical patient information between providers. Poor communication has been shown to be a contributing factor in malpractice claims, poor health care outcomes, sentinel events, and medical errors (Riesenberg, Leitzsch, Massucci, Jaeger, Rosenfeld, Patow, Padmore, Karpovich 2009). Effective patient handoffs require an understanding of the reason for patient hospitalization and anticipation of potential problems. Situational awareness is critical to this process defined as 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' (Endsley & Garland 2000; Flin, Yule, Paterson-Brown, Maran, Rowley, & Youngson 2007). It is not sufficient to just retrieve data; surgical residents need to be able to comprehend the information and "story" given to them, and anticipate future problems. Often, situational awareness is achieved after many years of training and learned through extensive clinical experience.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Residents have reported inconsistent training on performing patient handoffs and formal assessments have shown serious deficiencies in completeness, accuracy, and lack of situational awareness (Antonoff, et al 2007). The ACGME now requires residency programs to monitor patient handoffs to facilitate continuity of care and improve patient safety. Surgical residency training has responded to the challenge of reduced work hours with increased educational training to enhance safety and organizational efficiacy of trainee's work. We conducted an observational study on the effectiveness of patient handoffs before and after implementing a patient handoff mnemonic and conducting formal patient handoff education on efficient handoffs in our surgical residency.

METHODS

A single investigator observed 141 surgical patient handoffs, evaluating physicians giving and receiving handoffs. The observer was a surgeon (GAN), who had recently completed training. The observer was a non-participant in the handoff process. The observer did give feedback after the handoff when directly solicited for feedback by residents or moonlighters.

Handoffs consisted of a face to face discussion of each patient on the service, with the ability for the handoff recipient to ask questions. Physicians utilized a detailed handoff instrument that included the reason for hospitalization, medical problems, surgical procedures, important events during hospitalization, laboratory data, pathology, and radiology results. Physicians involved in the handoff discussed the patient's status, active medical problems, surgical procedures, the plan of care, critical medications, important events during hospitalization, things that need to be done after the handoff, and any additional information the physician coming on needed to be aware of.

Handoffs were observed for factors influencing situational awareness. Three surgical resident groups were compared: junior residents (post-graduate year 1 and 2 from medical school), senior residents (post-graduate year 3, 4, and 5 from medical school), and moonlighters (paid physicians that care for patients overnight or weekends). Most moonlighters are residents working per diem evenings and weekends while doing 1-3 years of research. Many surgical training programs require 1-2 years of research between the 2nd to 4th year of surgical residency during which time these residents will "moonlight". The observer evaluated factors associated with situation awareness during handoffs and used a survey to track factors influencing situation awareness during handoffs including: the time of day, day of the week, the handoff participants, year of training of the surgical residents and/or moonlighters, whether a situational overview was provided by the individual(s) giving signout, facilitating and distracting features that occurred during handoffs, miscommunication, effective aspects of handoffs, and whether the participants had received formal handoff training. Handoffs with prominent situation awareness for one or more patients were also noted as a separate entry.

We compared levels of situation awareness among providers who had or had not received formal didactic education on conducting a patient handoff (Table 1). Only junior residents and moonlighters received the formal didactic training; senior residents did not receive formal handoff training. Many first year surgical residents received handoff training prior to starting residency (6/21/12) and were observed for a total of 7 months. Some first year surgical residents did not receive handoff training prior to their first day of work; these residents were trained in their second year. Second year residents, third year residents, and moonlighters were trained from Sept-Nov 2012; this allowed for observation 4 months prior to training and 3 months after training (Table 1).

Table 1. Number of Residents and Moonlighters Observed Before and After Training

Giving A Handoff	Junior Resident	Moonlighter	Senior Resident
Before Training	31	11	30
After Training	57	12	
Receiving A Handoff			
Before Training	22	25	46
After Training	25	23	

The formal didactic training included use of a handoff pneumonic (IPASS). The pneumonic elements included: I=illness (is the patient stable or critically ill), P=patient summary (reason for admission, events during hospitalization), A=action list (things to do, timeline for completion), S=situation awareness (understanding what is going on with the patient, perception of important information, comprehension, and interpretation of issues, anticipation, and planning for potential problems), and S=synthesis. The training contained lectures and group discussions on handoff rational, appropriate handoff structure/content, and how to implement the handoff mnemonic.

We used a Likert scale (1-5, 5=best) to score situation awareness during handoffs (Table 2) The individual(s) giving the handoff was also graded on actively engaging the individual(s) receiving the handoff, ensuring mutual understanding of critical information, and appropriately addressing concerns. Situation awareness was evaluated on resident's and moonlighter's ability to fully understand a patient's illness and current status with comprehending and communicating concerns for the patient in the near future and what potential steps may need to be taken in caring for the patient.

Table 2. Situation Awareness Scoring

SA Level	Likert Scale	Manifestation of Situation Awareness
Level 1 Perception	1	Poor understanding of the patient's illness, unaware of the significance of critical information that may change the course of events for a patient.
	2	Understands the nuances of the patient's illness and current condition and therefore possible problems that may arise in caring for a patient but unfamiliar with treatments and appropriate actions to take if these problems were to arise
Level 2 Comprehension	3	Understands the nuances of the patient's illness and current condition, comprehension of how it may affect them, and appropriately anticipates these problems/outcomes.
Level 3 Projection	4	Understands the significance of a patient's illness and condition, grasps the possible outcomes of what could happen given their condition, anticipates these problems/outcomes, knows the appropriate treatment, and sometimes has a plan to address them.
	5	Understands the significance of a patient's illness and condition, grasps the possible outcomes of what could happen given their condition, anticipates these problems/outcomes, proficient in the treatment of these problems, and clearly communicates the plan to address them if they were to arise

Data analysis included descriptive statistics and group mean calculations of situation awareness scoring. Fisher's exact test was used to define significance between the groups. SPSS Version 12/0 (Windows) was used for all statistical analyses, $P<0.05$ was considered significant. This study was approved by the IRB (IRB # 12-09953 University of California, San Francisco and San Francisco Veterans Affairs Medical Center).

RESULTS

Formal Handoff Training improved many aspects of the signout process. Junior residents more frequently provided a situation overview during handoffs after training (6.5% vs. 37.5%, $p<0.005$) as did moonlighters after training (18.2% vs. 50%); while senior residents gave a situation overview 50% of the time (Figure 1).

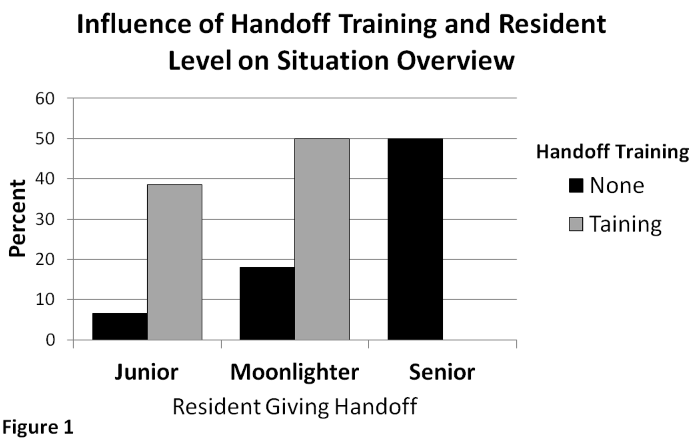

Figure 1

Training showed a profound impact on junior resident's situation awareness in giving a patient handoff (3.39 vs 3.8, $p<0.03$, Figure 2). Moonlighters also showed improvement, albeit not significant, in their situation awareness after training (3.27 vs 3.83). Senior residents demonstrated a high level of situation awareness (score 4.63 out of possible score 5) during handoffs without formal training.

Distractions during signout were lower in mornings than evenings (0.71 vs 1.83, $P=0.009$). Prior to handoff training, residents demonstrated less situation awareness when distractions (pages, calls, consults, etc.) were present (3.65 vs 3.95, $P=0.182$); after handoff training, situation awareness scores with distractions were the same as those without distractions (Figure 3). The number of omissions during handoffs decreased after training for junior residents (77.5% to 48%) and moonlighters (45% to 25%). Miscommunication was less frequent after training of junior residents (61% to 39%) and moonlighters (46% to 17%); while it occurred in 7% of handoffs by senior residents.

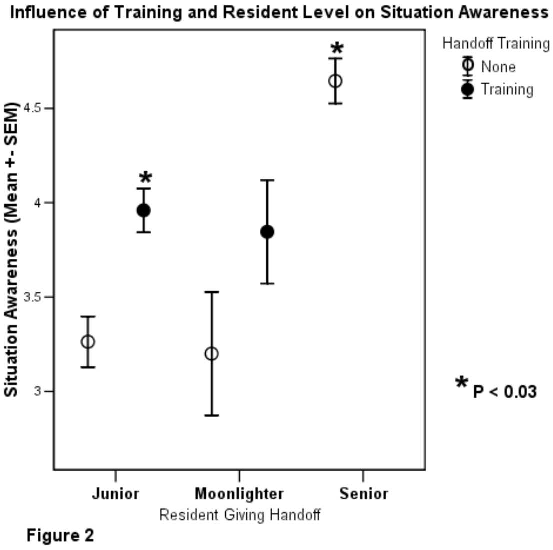

Figure 2

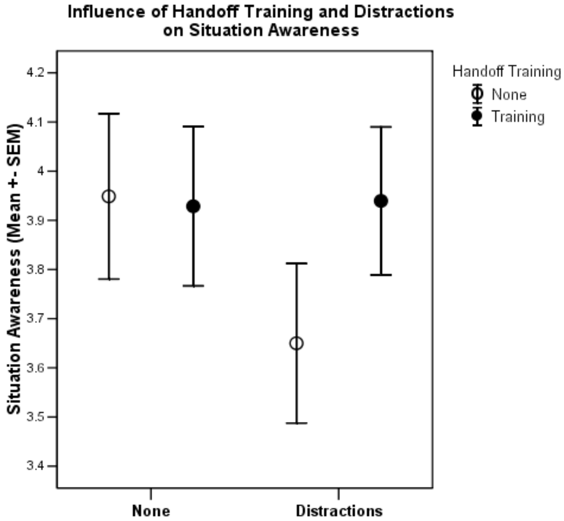

Figure 3

Situation awareness when plotted over time showed improvement for all groups(Figure 4). Junior residents had an exponential increase in situation awareness from November to January correlating with completion of handoff training and clinical experience. Moonlighters demonstrated increased situation awareness after completing handoff training in October. Senior residents started off with a moderate level of situation awareness that exponentially increased after their first month of residency, and then maintained a high level.

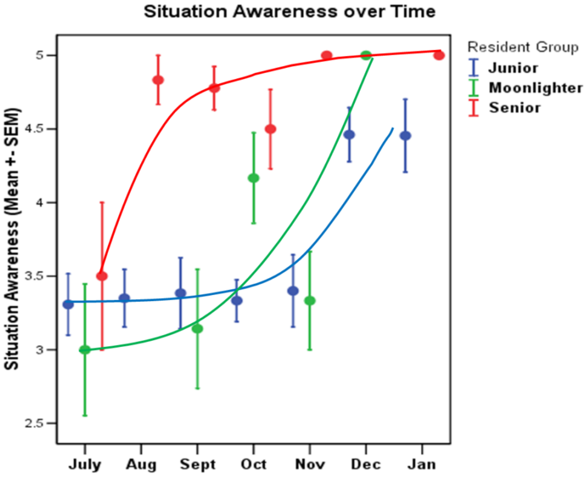

Figure 4

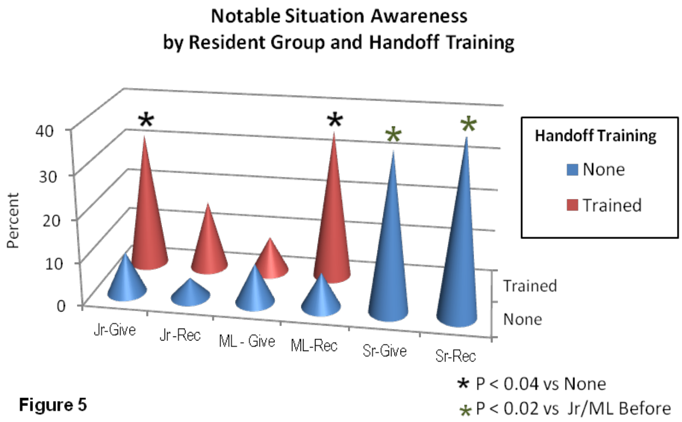

Figure 5

The observer looked for instances of particularly effective situation awareness (for one or more patients) during handoffs, and these were noted as a separate entry. These instances of excecptional situation awareness were more common among handoffs *given* by Junior residents (J-Give) after handoff training and among handoffs *received* by moonlighers (ML-Rec) after handoff training (Figure 5). Senior residents displayed a high level of situation awareness in both giving and receiving handoffs without formal training (Figure5).

CONCLUSION

In the current era of limited resident work hours, physician handoffs have increased exponentially. This makes handoffs a key part of patient safety (Borman, Jones, Shea, 2012). Studies of handoffs have reported decreased patient safety due to increased number of handoffs and problematic handoffs (Lee DY, Myers EA, Rehmani SS, Wexelman BA, Ross RE, Belsley SS, McGinty JJ, Bhora FY, 2012; Kitch, Cooper, Zapol, Marder, Karson, Hutter, Campbell, 2008; Charap 2004). In one study, 59% of residents reported that patients had been harmed due to problematic handoffs, and 12% reported this as major harm (Kitch, Cooper, Zapol, et al. 2008). Another study estimated that the most important piece of patient information was not communicated 60% of the time, and that residents overestimated their handoff effectiveness (Chang, Arora, Lev-Ari, D'Arcy, Keysar 2010). These studies highlight the need for graduate medical education to focus on improving resident's situation awareness during handoffs. The medical literature describes three key tenets of handoffs: face-to-face uniterurupted communication, communication of data that is factual and umambibuous, and the need for formal didactic and interactive training in handoffs (DeRienzo, Frush, Barfield, Gopani, Griffith, Jian, Mehta, Papvassiliou, Rialon, Stephany, Zhang, Andolsek, 2012). Others, have documented key strategies for successful handoffs (Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO, 2004); many of these strategies have been incorporated into the handoff process. The focus on safe transitions of care stress the importance of implementing innovative models in surgical curricula to enhance trainees' situation awareness. The current study showed that formal handoff training resulted in improved situation awareness. Surgical residents and moonlighters demonstrated more effective communication, greater focus on key elements, decreased omissions, and more successful handoffs following training.

Handoff training also seemed to provide resilience during distractions. Before training, distractions were associated with decreased situation awareness; but after training, residents were able to maintain the same level of situation awareness despite distractions. This is a critical skill as physicians are frequently distracted by pages, nurses, other physicians, and phone calls during handoffs. Training facilitated resident's ability to focus on critical issues. The mnemonic provided a script and mental image of critical information, allowing residents to proactively direct handoffs.

Handoffs are different in mornings vs evenings. The evening handoff is a lengthy process between a junior resident and moonlighter lasting 1.5-2 hours. This involves discussing the days events, operations, active and evolving problems, and care that needs to be completed overnight. The morning handoff takes about 30 minutes, and usually involves moonlighters signing out overnight events with a summary of things completed. This delgation of duties may explain why handoff training led to more noteable handoff situation awareness among junior residents when *giving* handoffs, while training led to more noteable situation awareness among moonlighers when *receiving* handoffs (Figure 4).

A limitation of this study is that there was only one observer of the handoffs. This may have led to some bias in the results as another observer's results were not used to validate the results obtained. There may have been a Hawthorne effect as well; the observer noted an increase uses of the mnemonic and focus in giving the signout when her observation of the handoff was known to the residents/moonlighters. This limitation was modified by observing handoffs without the participant's awareness.

Residents were keen to improve their handoffs and frequently solicited feedback from the observer. This periodic feedback reinforced resident's handoff skills. This feedback has been described as an important 'cue→strategy' association (Zsambok, Klein 1997). This meant that residents and moonlighters were getting additional "coaching" on handoffs and may explain some of the improved situation awareness over the course of the study. It could also be that over time residents and moonlighters gained more clinical experience, which improved handoff ability. But experience alone cannot explain the increase in situation awareness, because moonlighters have 1-3 years more clinical experience than junior residents, yet they started at a level of situation awareness similar to junior residents, and improved following training.

Senior residents started with lower situation awareness in July corresponding to the transition from junior to senior resident. They then quickly transitioned (within a month), without handoff training, to achieve proficiency; this likely resulted from years of clinical experience. Their years of clinical experience facilitated development of Level 3 Situation Awareness; the ability to project future events from current events (Endsley, Garland 2000; Patterson, Roth, Woods, chow, Orlando 2005). Situation awareness is an essential skill that surgical residents obtain over time, enhancing their decision making; but this study demonstrated that formal training can expedite this process.

In health care, we use critical cues and intuition to direct the care of patients. This intuition comes from pattern matching and recognition (Klein 1998). Situation awareness requires time and is gained with increasing experience. Senior surgeons might describe it as an inclination, gestalt, or impression of what might happen given the information provided; much of the knowledge is tacit rather than explicit. Resident's ability to assess clinical conditions quickly and accurately is mostly learned through clinical experience, not instruction. But, training and providing a framework for handoffs (with situation awareness as a frames) led to improved handoffs. Our data demonstrates the value of didactic handoff training and mnemonic use to improve surgical handoffs and suggests this approach can accelerate proficiency in situation awareness.

REFERENCES

Antonoff MB, Berdan EA, Kirchner VA, Krosch TC, Holley CT, Maddaus MA, D'Cunha J. (2013) Who's covering our loved ones: surprising barriers in the sign-out process. The *American Journal of Surgery, 205, 1, 77-84*.

Borman KR, Jones AT, Shea JA. (2012) Duty Hours, Quality of Care, and Patient Safety: General Surgery Resident Perceptions. *Journal of the American College of Surgeons, 215, 70-79.*

Chang VY, Arora VM, Lev-Ari S, D'Arcy M, Keysar B. (2010) Interns Overestimate the Effectiveness of Their Hand-off Communication. *Pediatrics, 125, 3, 491-496.*

Charap M. (2004) Reducing resident work hours: unproven assumptions and unforeseen outcomes. *Ann Intern Med.* 140(10): 814–815

DeRienzo CM, Frush K, Barfield ME, Gopwani PR, Briffith BC, Jiang XS, Mehta AI, Papvassiliou P, Rialon KL, Stephany AM, Zhang T, Andolsek KM. (2012) Handoffs in the Era of Duty Hours Reform : A Focused Review and Strategy to Address Changes in the Accreditation Council for Graduate Medical Education Common Program Requirements. *Academic Medicine, 87, 403-410.*

Endsley MR, Garland D.J. (2000) *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc

Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. (2007) Teaching surgeons about non-technical skills. *Surgeon, 5, 2, 86-89*.

Kitch BT, Cooper JB, Zapol WM, Marder JE, Karson A, Hutter M, Campbell EG. (2008), Handoffs causing patient harm: a survey of medical and surgical house staff. *Jt Comm J Qual Patient Saf. 34(10):563-70.*

Klein G. (1998) Sources of Power *How People Make Decisions.* Cambridge, MA: The MIT Press.

Lee DY, Myers EA, Rehmani SS, Wexelman BA, Ross RE, Belsley SS, McGinty JJ, Bhora FY. (2012) Surgical residents' perception of the 16-hour work day restriction: concern for negative impact on resident education and patient care. *J Am Coll Surg. 215(6):868-77*

Patterson ES, Roth EM, Woods DD, Chow R, Orlando J. (2004) Handoff strategies in settings with high consequences for failure: lessons for health care operations. *International Journal for Quality in Health Care, 16, 2: 125-132.*

Riesenberg LA, Leitzsch J, Massucci JL, Jaeger J, Rosenfeld JC, Patow C, Padmore JS, Karpovich KP. (2009) Residents' and Attending Physicians' Handoffs: A Systematic Review of the Literature. *Academic Medicine, 84, 2: 1775-1787.*

Zsambok CE, Klein G ed. (1997) *Naturalistic Decision Making.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Health IT-related obstacles and facilitators in coordinating care for patients with chronic illness: A longitudinal study

Pascale CARAYON[a,b], Sarah KIANFAR[a,b], Ann SCHOOFS HUNDT[a], Janet TOMCAVAGE[c], Doreen SALEK[c], Janet PETRUCCI[c] , Peter HOONAKKER[a], Bashar ALYOUSEF[a,b], & James WALKER[c]

[a]*Center for Quality and Productivity Improvement, University of Wisconsin-Madison*
[b]*Department of Industrial and Systems Engineering, University of Wisconsin-Madison*
[c] *Geisinger Health System*

ABSTRACT

Introduction: Coordinating care for patients with chronic illness is a major challenge for healthcare as these patients interact with multiple providers and healthcare organizations during their journey. Health information technology (IT) has been proposed to support care coordination, but can bring its own challenges. In this longitudinal study we examined the obstacles and facilitators experienced by care managers in using various forms of health IT. **Methods:** We interviewed 14 hospital and outpatient care managers who coordinated care for patients with chronic illnesses across multiple transitions of care. These interviews were conducted over a period of one year. **Results and discussion:** Care managers experienced both obstacles and facilitators in their use of health IT for care coordination. Despite technical and human factors problems with the technologies, care managers described several benefits of health IT.

KEYWORDS

Care coordination; health information technology; performance obstacles and facilitators; longitudinal study.

INTRODUCTION

Care of patients with chronic illness accounts for approximately three quarters of all healthcare expenditures in the US (Bodenheimer, 2002). It has been difficult to design and implement sustainable interventions to coordinate care for these patients as they transition among hospitals, home, clinics, and nursing homes. Health information technology (IT) can play a key role in supporting care coordination; however, the use of health IT to improve care for patients with chronic illness can be challenging (Alyousef et al., 2012; Bates, 2010; Carayon et al., 2012). In this study we focus on the role of care managers who coordinate care for patients with chronic illness and identify the obstacles and facilitators that they experience in using various forms of health IT to access, share and manage patient information.

BACKGROUND

Care coordination for patients with chronic illnesses such as heart failure (HF) or chronic obstructive pulmonary disease (COPD) can be very complex. These patients often journey through multiple transitions of care as they are cared by numerous providers and healthcare organizations. These transitions pose unique coordination challenges because of the large variety of healthcare professionals and organizations involved in patient care. Numerous studies have identified care quality and patient safety issues in these care transitions (Holland & Harris, 2007; Kahn & Angus, 2011; Nelson & Carrington, 2011).

Health IT has been suggested as a key component of care coordination; but numerous challenges have been identified in the use of health IT for care coordination for patients with chronic illness (Bates, 2010; Carayon et al., 2012; O'Malley, 2011). Challenges include communication and sharing of information among clinicians and healthcare organizations, and information overload (Bates, 2010). O'Malley et al. (2010) interviewed 60 physicians and staff in 26 physician practices about their experience with electronic health records (EHR), i.e. one type of health IT, and the use of EHR to support coordination activities. Both EHR-related facilitators and obstacles to care coordination were identified. For instance, the EHR can provide immediate access to patient information (facilitator), but it may be hard to find information in the EHR (obstacle). Further understanding of care coordination challenges and opportunities for improvement offered by health IT is necessary.

Many different models have been proposed to improve care coordination. Recently, the patient-centered medical home (PCMH) has been developed as a new way of organizing primary care to benefit all patients, but more

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

particularly patients with chronic illness. According to Stange et al. (2010), PCMH is “a team of people embedded in the community who seek to improve the health and healing of the people in that community. They work to optimize the fundamental attributes of primary care combined with evolving new ideas about organizing and developing practice and changing the larger health care and reimbursement systems.” (page 602) Care managers play a key role in the PCMH model; they are nurses who coordinate care and ensure that care gaps are closed and patients receive high-quality, safe care.

Care managers perform a range of activities to coordinate care, such as reviewing and reconciling medications, and ensuring that patients have a follow-up appointment with their primary care physician and that appropriate referrals (e.g., home health services) are in place (Brown, Peikes, Peterson, Schore, & Razafindrakoto, 2012; Maliski, Clerkin, & Litwin, 2004; Oliva, 2010). Care managers also educate patients about their diseases so that patients manage the unique signs and symptoms of their disease and are more actively engaged in their own care. Numerous care managers’ activities rely on health IT applications, so that care managers have access to patient information, share information with others who care for the patients, and document their work (Alyousef et al., 2012; Carayon et al., 2012).

Care coordination has been defined as: “the deliberate organization of patient care activities between two or more participants (including the patient) involved in a patient’s care to facilitate the appropriate delivery of health care services. Organizing care involves the marshalling of personnel and other resources needed to carry out all required patient care activities and is often managed by the exchange of information among participants responsible for different aspects of care.” (McDonald et al., 2010) (page 4). Care coordination manages the interdependencies between activities and people (Klein, 2001; Malone & Crowston, 1994), such as sharing patient information. Klein (2001) has identified several features of successful coordination, including preparation, planning and monitoring. Research on care coordination has not demonstrated whether health IT can support these features of successful coordination.

The purpose of this study is to assess the health IT-related obstacles and facilitators experienced by care managers in coordinating care for patients across transitions of care. In conducting this study we aimed to identify whether and how health IT can support the key features of successful coordination identified by Klein (2001).

STUDY CONTEXT

The Keystone Beacon project provides the context for this study. The Keystone Beacon project aims to improve care coordination for patients with chronic illness (Heart Failure or HF and Chronic Obstructive Pulmonary Disease or COPD) in Central Pennsylvania, US: “Keystone Beacon is demonstrating how integrated care can be coordinated community-wide through the use of patient-focused technology and specialized nurse care managers.” (https://www.keystonebeaconcommunity.org) The specialized nurse care managers are located in hospitals, clinics and a transition of care center. Inpatient or hospital-based care managers support and coordinate care for Beacon patients until their discharge. Outpatient care managers that work in participating physician practices manage HF and COPD patients as long as necessary. A transition of care (TOC) center was created to support care coordination for patients discharged from hospitals that did not have an outpatient care manager. TOC care managers call patients every week – or as needed – for 30 days post-discharge. To perform their job care managers need to access and share information and, therefore, use multiple health IT applications, including the case management software (primarily for documentation and sharing of information among the care managers), the local health IT (e.g., EHR in a hospital) and the health information exchange technology that supports exchange of patient information among participating organizations.

METHODS

Study design

In the context of the Keystone Beacon project, we used a qualitative approach using interviews of Beacon care managers in all three settings: hospitals, clinics and TOC. The interviews were conducted at five different times over a one-year period (May’2011 to May’2012) in order to capture changes in performance obstacles and facilitators from the time that the first care managers were hired (May’2011) to one year later (May’2012).

Setting/sample

A total of 14 care managers were interviewed: 5 inpatient care managers, 2 outpatient care managers, 5 TOC care managers and 2 ‘float’ care managers who worked in various settings (e.g., to help when a care manager was on vacation). At each of the five interview times, between 4 and 11 of the 14 care managers participated in data collection.

Data collection instrument

We used a semi-structured interview guide to collect information on performance obstacles and facilitators experienced by the care managers. Over one year, 33 interviews were conducted for a total of about 54 hours. Two researchers conducted the interviews. The interviews were audio-recorded (all with interviewee consent).

The interviews were then transcribed and any mention of names and healthcare organizations was removed from the transcripts.

Data analysis

We analyzed the interview data using the NVivo© qualitative data analysis software. The transcripts of the interviews were imported into NVivo© and were coded to identify health IT-related obstacles and facilitators. Researchers identified health IT-related obstacles or facilitators and described them in separate annotations. Several transcripts were analyzed by multiple researchers to ensure consistency of coding. Multiple meetings took place to resolve any discrepancy or disagreement in coding and to refine the coding scheme.

RESULTS

Analysis of the interview data produced quantitative data in the form of the number of care managers reporting health IT-related obstacles and facilitators, and the number of separate obstacles and facilitators. We also conducted a qualitative analysis to describe how the nature of health IT-related obstacles and facilitators changed over time.

Longitudinal analysis of health IT-related obstacles and facilitators – Quantitative analysis

A summary of the longitudinal analysis of health IT-related obstacles and facilitators is presented in Figure 1. The percentages of care managers who reported health IT-related obstacles (below the 0% horizontal line) and facilitators (above the 0% horizontal line) did not vary significantly over time as shown in Figure 1.

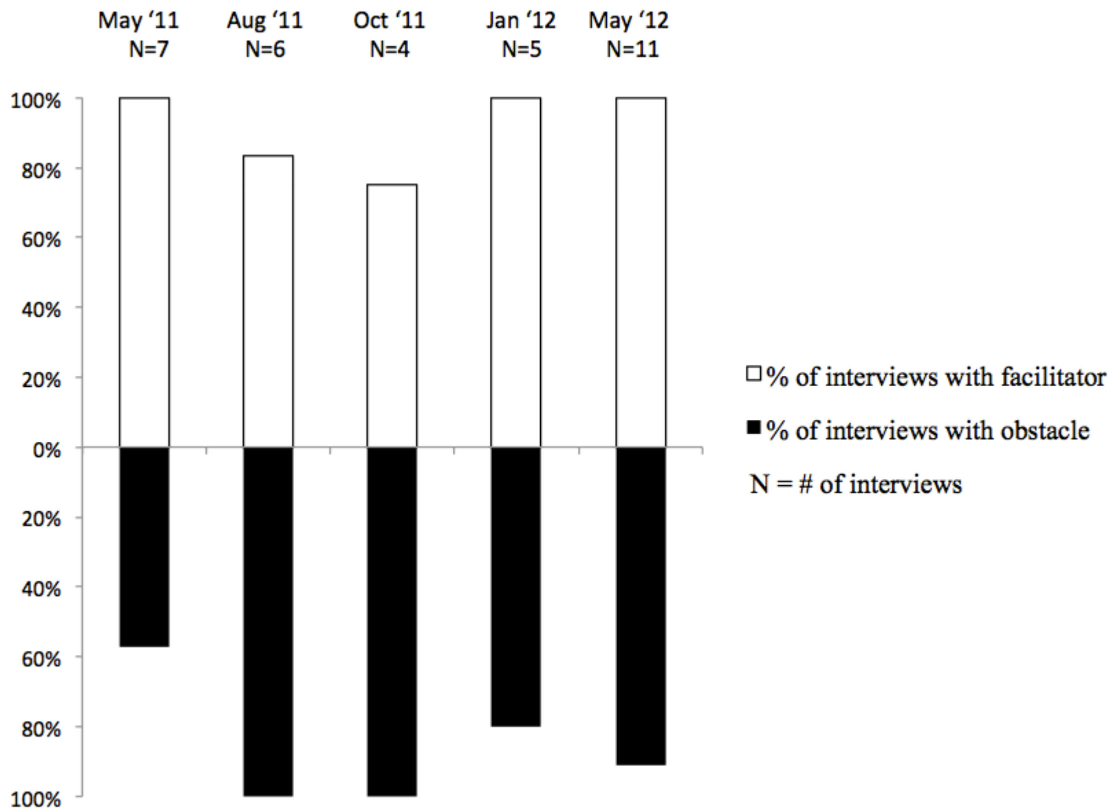

Figure 1. Health IT-related obstacles and facilitators experienced by care managers

Table 1 shows the number of separate health IT-related obstacles and facilitators experienced by the care managers at each of the five data collection times. During the first data collection (May'2011), a large number of obstacles was reported in contrast to a small number of facilitators. The number of obstacles descreased by the second round of interviews, but then did not change much over time, whereas the number of facilitators tended to increase.

Table 1. Separate health IT-related obstacles and facilitators over time

	May'2011	August'2011	October'2011	January'2011	May'2012
Number of interviews	7	6	4	5	11
Number of separate obstacles	31	11	17	12	19
Number of separate facilitators	6	21	10	20	29

Longitudinal analysis of health IT-related obstacles and facilitators – Qualitative analysis

In the early data collection care managers described several obstacles related to the difficulty of learning to use various health IT applications. During the first couple of months on the job several care managers talked about not having access to the health IT applications that they needed to perform their job. These obstacles decreased over time as care managers became more skilled at using various health IT applications and they were given access to better tools. Over time care managers talked about access to different health IT applications as a facilitator.

Some obstacles decreased over time. Initially, care managers often had problem with patient information that was not accessible, up-to-date, complete or timely. This obstacle decreased over time as, in particular, the health information exchange technology was used by an increasing number of clinicians and increasing number of patients provided authorization to share their information.

Some obstacles were consistently experienced over time. For instance, care managers often talked about technical problems, such as slow connections and having to reboot their computer. They described several usability issues with the different health IT applications. For instance, they talked about a poorly designed case management software (e.g., lots of clicks, dropdown menus with insufficient options) and challenges in locating information in some of the EHRs they used in the local hospitals or clinics.

The use of multiple health IT applications was described as both an obstacle and a facilitator. Care managers had to double document in the case management software as well as the local EHR of the hospital or clinic. This double documentation is an obstacle created by the lack of inter-operability of the various health IT applications. On the other hand, care managers described the benefit of having access to multiple health IT applications because it provides them with more complete information on their patients: they can develop a better mental model of what is happening with their patients. Specifically, access to multiple health IT applications is useful for outpatient care managers who can monitor various care activities for their patients and follow up on scheduled care activities.

As care managers developed expertise with the health IT, they talked about the benefits of the technologies for sharing information with other clinicians and providers. In particular, inpatient care managers talked about their increasing ability to share patient information with outpatient and TOC care managers. This sharing of information is critical as patients move through various care transitions such as hospital to home.

DISCUSSION AND CONCLUSION

Coordinating care for patients with chronic illness involves the management of interdependent activities performed by multiple people located in various healthcare organizations (e.g., hospitcal, clinic) and other settings (e.g., home). In the Keystone Beacon project, care managers were at the core of the care coordination network; they were the 'glue' between the disparate healthcare professionals and organizations that care for these patients. To perform their job they needed to access and share information and, therefore, used multiple health IT applications to support their functions of monitoring and communication. Care managers described how the multiple health IT applications allowed them to develop a mental model of what was happening with their patients. Over time care managers began to describe the benefits of health IT to share patient information with other care managers and clinicians. However, they also experienced a number of health IT-related obstacles, such as reliability (e.g., slow connections), usability (e.g., locating information) and access.

Coordination can be particularly challenging when information needs to flow across multiple organizations. This is a major problem in health care because of the lack of inter-operability of health IT applications. Despite this and other technical and human factors problems, care managers identified many benefits of health IT for care coordination. As the need to provide more efficient care coordination becomes more apparent, we need to continue to refine our understanding of care coordination to develop usable and useful health IT applications.

ACKNOWLEDGMENTS

Partial funding for this research was provided by the US Office of the National Coordinator through the Beacon award program [award No. 90BC001301].

REFERENCES

Alyousef, B., Carayon, P., Hoonakker, P., Hundt, A. S., Cartmill, R., Tomcavage, J., . . . Walker, J. (2012). Care managers' challenges in using multiple health IT applications. In The Human Factors and Ergonomics

Society (Ed.), *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp. 1748-1752). Santa Monica, CA: The Human Factors and Ergonomics Society.

Bates, D. (2010). Getting in step: Electronic health records and their role in care coordination. *Journal of General Internal Medicine, 25*(3), 174-176.

Bodenheimer T, W. E. H. G. K. (2002). Improving primary care for patients with chronic illness. *Journal of the American Medical Association, 288*(14), 1775-1779.

Brown, R. S., Peikes, D., Peterson, G., Schore, J., & Razafindrakoto, C. M. (2012). Six features of medicare coordinated care demonstration programs that cut hospital admissions of high-risk patients. *Health Affairs, 31*(6), 1156-1166.

Carayon, P., Alyousef, B., Hoonakker, P., Hundt, A. S., Cartmill, R., Tomcavage, J., . . . Walker, J. (2012). Challenges to care coordination posed by the use of multiple health IT applications. *Work: A Journal of Prevention, Assessment and Rehabilitation, 41*(0), 4468-4473.

Holland, D. E., & Harris, M. R. (2007). Discharge planning, transitional care, coordination of care, and continuity of care: Clarifying concepts and terms from the hospital perspective. *Home Health Care Services Quarterly, 26*(4), 3-19.

Kahn, J. M., & Angus, D. C. (2011). Going home on the right medications: Prescription errors and transitions of care. *The Journal of the American Medical Association, 306*(8), 878-879.

Klein, G. (2001). Features of team coordination. In M. McNeese, E. Salas & M. Endsley (Eds.), *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments* (pp. 68-95). Santa Monica, CA: The Human Factors and Ergonomics Society.

Maliski, S. L., Clerkin, B., & Litwin, M. S. (2004). Describing a nurse case manager intervention to empower low-income men with prostate cancer. *Oncology Nursing Forum, 31*(1), 57-64.

Malone, T. W., & Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys, 26*(1), 87-119.

McDonald, K. M., Schultz, E., Albin, L., Pineda, N., Lonhart, J., Sundaram, V., . . . Malcolm, E. (2010). Care Coordination Measures Atlas. Rockville, MD: Agency for Healthcare Research and Quality.

Nelson, J. M., & Carrington, J. M. (2011). Transitioning the older adult in the ambulatory care setting. *AORN Journal, 94*(4), 348-361.

O'Malley, A. S. (2011). Tapping the unmet potential of health information technology. *The New England journal of medicine, 364*(12), 1090-1091.

O'Malley, A. S., Grossman, J., Cohen, G., Kemper, N., & Pham, H. (2010). Are Electronic Medical Records helpful for care coordination? Experiences of physician practices. *Journal of General Internal Medicine, 25*(3), 177-185.

Oliva, N. L. (2010). A closer look at nurse case management of community-dwelling older adults: Observations from a longitudinal study of care coordination in the chronically Ill. *Professional Case Management, 15*(2), 90-100.

Stange, K. C., Nutting, P. A., Miller, W. L., Jaen, C. R., Crabtree, B. F., Flocke, S. A., & Gill, J. M. (2010). Defining and measuring the patient-centered medical home. *Journal of General Internal Medicine, 25*(6), 601-612.

Decision Support for Pararescue

Laura G. MILITELLO[a], Matthieu BRANLAT[b], Christen E. LOPEZ[a], Katherine E. WALKER[c], Victor FINOMORE[b], Jill RITTER[b], SMsgt Robert BEAN[c], and Maj David MCGRAW[d]

[a]*Applied Decision Science*
[b]*U.S. Air Force Research Laboratory*
[c]*The Ohio State University*
[d]*U.S. Air Force*

ABSTRACT

Introduction: This paper describes a project to articulate cognitive requirements for pararescue jumpers (PJs), an elite rescue unit of the United States Air Force. PJs work with other US Air Force Special Forces to conduct personnel recovery missions in hostile or denied areas. They work in a fast-paced, high-risk, dynamic environment. **Method**: We conducted a cognitive task analysis including interviews, observations, and literature review to better understand the cognitive challenges PJs must manage. We extracted cognitive requirements from existing literature, and created cognitive demands tables based on cognitive task analysis interviews and observations. We conducted card sort activities and met with subject matter experts to determine priorities for design solutions. **Results and discussion:** We created an initial concept for an *Accountability and Ex-fil Planning Support Tool*.

KEYWORDS

Practical application; cognitive task analysis; decision making; uncertainty management; situation awareness; military.

INTRODUCTION

This paper presents a project to develop decision support for pararescue jumpers (PJs). PJs work in a fast-paced, high-risk, dynamic and uncertain environment. PJs are a core component of the Guardian Angel Weapon System (GAWS), made up also of Combat Rescue Officers (CROs) and Survival-Evasion-Resistance-Escape (SERE) specialists. GAWS is a reactive force with a primary mission of personnel recovery in hostile or denied areas; they rescue people and retrieve equipment in difficult conditions. GAWS operates in all phases of service, joint, and coalition operations, in any location and any environment. They are able to act independently or in conjunction with aircraft. GAWS personnel are trained to conduct rescues in a variety of contexts, including confined spaces, collapsed structures, and swift water. They are proficient in close-quarters battle defense, mountaineering, deep dives, and deploying from aircraft. They operate under sovereign rescue option in peacetime and at war to rescue anyone, anytime, to capture and represent the actions PJs perform during rescue missions, and present one *design concept* resulting from this research program.

PJs assigned to guardian angel units usually deploy as a team, with one Team Leader, a CRO, and two to three additional PJ team members. Site management duties fall jointly to the Team Leader (personnel and patient management) and the CRO (coordination with air and/or ground assets), while other PJ team members establish security, locate personnel, and administer medical treatment. This paper is primarily focused on the tactical level of pararescue operations, high-level cognitive challenges associated with the coordination and management duties of the Team Leader and CRO.

The intent of this research program was to identify a set of requirements to drive future research and develop solutions to support the GAWS mission set. We describe *high-level cognitive requirements* to drive future research programs, *critical cognitive requirements* to inform concept generation and design, *functional diagrams*
In collaboration with the US Air Force Research Lab (AFRL), we applied cognitive task analysis (CTA) methods to identify cognitive requirements that will drive design concepts and assessment strategies. It is important to point out that AFRL has a strong emphasis on envisioning and creating solutions for the future, often exploring extensions of existing technology, tailoring off-the-shelf solutions, and inventing first-of-kind technologies. However, by leveraging new technologies, new challenges may be introduced, thus making real-world operations more difficult. For example, mobile and wearable computing devices allow for more information to be transmitted to operators in the field; however, these technologies also create new opportunities

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

for information overload, distraction, and technology failure. As depicted in Figure 1, the cognitive requirements identified in this project are intended to help balance this tension by:

- grounding design concepts in realistic mission scenarios, and
- creating scenario-based evaluation strategies that allow researchers to explore impact of proposed solutions on complex metacognitive activities.

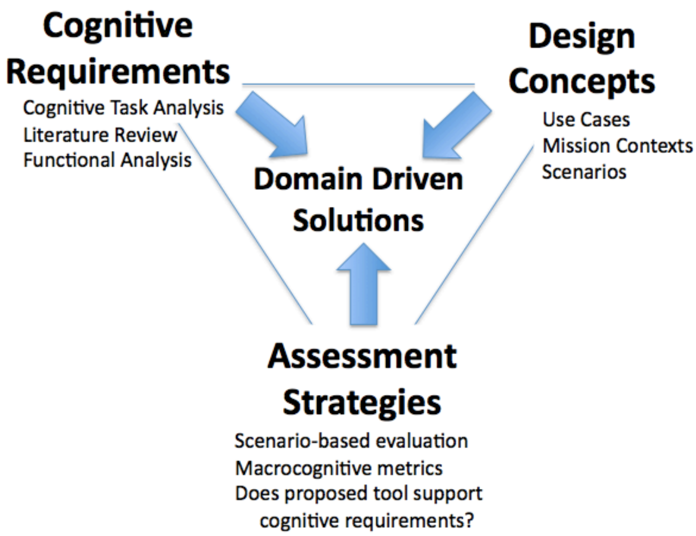

Figure 1. CTA informs cognitive requirements, design concepts, and assessment strategies to deliver domain-driven solutions

METHODS

We conducted a CTA of PJs and CROs. We began with domain familiarization activities, conducted interviews with experienced PJs and CROs, and generated scenarios as a tool to further elicit knowledge from subject matter experts (SMEs). Each of these activities is described in turn.

Domain Familiarization

To become familiar with the domain, we attended two workshops, a demonstration of a PJ rescue response, toured a PJ facility, and observed training sessions. In addition, we reviewed existing literature about the pararescue domain including existing capability gaps (Deep Dive, 2010 ; Irick, n.d.), requirements (AFSOC, 2011 ; OAR, 2008), descriptions of the PJ/CRO career fields, and training manuals for pararescue personnel. From the literature, we extracted cognitive requirements.

Interviews

For this project, we used three complementary interview techniques to conduct the CTA: the Critical Decision method (Crandall, Klein, Hoffman, 2006), Knowledge Audit (Militello & Hutton, 1998), and scenario-based interviews. At least two interviewers were present for each interview, which lasted between 60 and 90 minutes. Two small group interviews occurred opportunistically, consisting of PJs, CROs, and one Combat Controller (CCT). Ten interviews were conducted individually, allowing for more in-depth exploration of incidents. Overall, there were 16 participants: 13 PJs, 2 CROs, and 1 CCT. See Tables 1 and 2 for participant demographics. CTA interviews focused on eliciting a broad range of challenging incidents across diverse mission contexts.

Table 1. Roles and experience of individual interview participants

Participant	Role	Years of Experience
1	PJ	20
2	CRO	Not collected
3	PJ	8
4	CRO*, PJ	5 (PJ), 6 (CRO), 17 (Marines)
5	PJ	5
6	PJ*	11 (PJ), 7 (AF Maintenance)
7	Instructor*, PJ	5 (PJ), 3 (Instructor)
8	Instructor*, PJ	3 (PJ), 2 (Instructor)
9	Instructor*, PJ	28 (PJ), 7 (Instructor)
10	Instructor*, PJ	4 (PJ), 2 (Instructor)

**Indicates current position*

Table 2. Make-up of group interviews

Group Interview 1	Group Interview 2
2 PJs	3 PJs
1 CCT	1 CRO
6 researchers	1 acquisition rep

Scenario Generation for Knowledge Elicitation

In parallel with interviews, a type of scenario known as a staged-world (Woods and Hollnagel, 2006), was created to further explore one highly challenging mission context. A scenario encapsulates domain knowledge and embodies the challenges that operators face in real-world operations. The staged world is intended to capture typical characteristics of the work environment (e.g., division of roles, channels of communication), situations encountered (e.g., mission type), and challenges in conducting operations (e.g., pace of activities, uncertainty). This approach represents a collaborative effort with domain practicioners rather than relying solely on researchers' understanding and imagination. In order to adequately capture knowledge of the domain and of the complexity of operations, our process included progressive co-construction with different SMEs.

The first steps consisted of establishing a basic mission structure including the type of emergency and contextual elements. During these interviews, participants were asked to elaborate on how they would handle the mission, what could happen, and what was challenging. Researchers injected a few pacers and probes opportunistically. This included presenting new pieces of information as the mission progressed, and the inserting unanticipated events. Based on the information and feedback gathered during these sessions, a more elaborate skeleton was developed. One-on-one interviews with SMEs provided additional details. We began with discussions and white board drawings. Over time, as more details were added, a map representation was used to set the stage for later discussions. Currently, PJs and CROs mentally build and maintain a picture of the situation. The map allowed for easier discussion of the placement of personnel, casualty collection point, and aircraft landing zones.

Through these iterations, the scenario evolved into a mass casualty situation in a semi-urban environment, in which an explosive device of unknown origin resulted in the partial collapse of a public building. This type of mission was chosen because SMEs saw it as highly relevant and representative of some of the difficult situations they face. Challenges include high levels of uncertainty, the need to prioritize treatment and evacuation of many victims, and potential security concerns.

Analysis

Three levels of analysis were conducted during this effort. The first level of analysis included creating decision requirements tables based on incidents elicited during interviews. Table 3 presents an excerpt from one decision requirements table.

Table 3. Excerpt from SME-validated decision requirements table based on a mass casualty scenario-based interview

Decision	Information Considered	Information Sources	Why difficult?
1. To launch or not? Made by: - Commander, PJ-TL, CRO, pilots, Intel personnel, SERE; everyone who could be involved in the mission - Commander makes ultimate decision (go/no-go)	Mission Priority + Risk • What happened? • How many casualties? • Is the area secure, or are there enemy personnel in the area? • What was the situation like in the past 24 hours that lead to this event? • How does the indigenous population feel towards the U.S.? • Maps, photos of the area • Timeline of the mission (keeping within the "golden hour") • Intel provides: o Images from 1-2 days prior, month before, etc. o Cultural indicators o Temperature of the region (i.e., tense, peaceful, aggressive)	Intel brief	"Last thing you want to do is create more casualties" (Want to go in every time to rescue, but sometimes may be too risky to carry out the mission)

A second level of analysis included a card sort. Cognitive requirements from the literature review were combined with decision requirements from the interviews. The entire research team participated in a card sort exercise to prioritize and categorize the list of combined requirements. Researchers identified topics believed to be of potentially high impact for PJs, topics that are high priority for AFRL, topics that are not directly relevant to current goals, requirements that are already being addressed by other development efforts, and topics that could provide good opportunities for collaborating with other branches of AFRL. This prioritization activity will help focus future design efforts.

A third level of analysis focused on creating functional diagrams articulating functions, roles, and relationships in the context of the mass casualty scenario that was generated.

FINDINGS

High Level Cognitive Requirements

Findings from the card sort included 13 cognitive requirements. Cognitive requirements fall into three task categories summarized in Table 4. These high-level, broad requirements will drive future research programs.

Table 4. Cognitive Requirements by Task Category

Medical Support	General Mission Support	After-Action Support
• Medical treatment • Medical assessment • Medical equipment limitations • Health IT • Medical training for operating environments	• Emergent mission planning/re-planning • Coordination during mission planning • Scene management • Exfil prep and management • Hazard-specific assessment • Big picture situation awareness	• AARs, debriefs, reporting • Assess effectiveness

Critical Cognitive Requirements

From the larger set of cognitive requirements, we identified six critical cognitive requirements that describe important aspects of the pararescue domain that are core components of effective performance, but not necessarily task or function specific. Because they are core aspects of work, they serve as a backdrop for more easily observed and quantified activities. These high-level cognitive requirements will inform all aspects of concept generation, design and evaluation. Table 5 summarizes these critical cognitive requirements and important implications for design.

Table 5. Critical Cognitive Requirements

	Context	Implications for Design
Fluid Command	Teams of PJs quickly adapt roles as dictated by the situation and the expertise of the personnel available in a manner that is not hindered by rank or other hierarchical conventions. A PJ with the most relevant experience or the best SA may move into a leadership role even if the team started out with a different configuration	Information technology should allow for easy and quick handoff among team members. Interfaces should allow a PJ/CRO to rapidly assess the situation from a new perspective as roles shift.
Team Coordination	The PJ Team Leader and the CRO work closely to coordinate the team internally and externally. The team leader focuses on coordinating PJ activities. The CRO focuses on coordinating with aircraft, additional security forces, and others outside the PJ team. This happens under time pressure, within the constraints of radio communication and hand signals, often in the dark with a need for stealth.	Interfaces designed to support Team Leader and CRO roles should facilitate shared situation awareness, and allow for rapid, real-time exchange of information.
Emergent Mission Planning	For each mission, the GAWS team must decide whether to launch, and if so, must determine roles and equipment needs, how they will get to the site, and how they will get out. Within these larger decisions are important judgments and assessments made based on whatever information is available. Often these decisions are made in a time span as brief as 10 minutes.	Tools to support emergent mission planning should present information that is: • Easy to scan • Easy to re-orient after interruption • Up-to-date and time-stamped • Highlights most relevant information • Provides ability to search for additional information, if time allows
Distributed re-planning	Plans routinely change in the course of an individual mission. Changes are relatively easily shared within the PJ team who are all on site and experiencing the same conditions. However, changes to the plan may also impact aircraft and other vehicles involved in the rescue, other PJ teams preparing to respond, and additional security forces.	Tools to support planning and re-planning will need to be better than current strategies involving simple, direct communication.
Scene Management	Managing the scene is a complex activity that includes maintaining security, locating, rescuing, assessing, and treating isolated personnel, making sure all personnel are accounted for, and arranging transport for patients and the team. Scene management is primarily the responsibility of the Team Leader, but requires careful coordination with all team members.	Much of scene management is currently accomplished using direct radio communications and sharpies. Proposed tools will have to exceed the capabilities of these simple tools, and be as simple and reliable.
Big picture situation awareness	A key part of PJ training is learning to avoid the narrowing of focus associated with stress. The tension between task focus and big picture situation awareness is apparent in many PJ tasks including treating patients, designing and building pulley systems to move patients and equipment, and jumping from airplanes.	Proposed tools will need to facilitate ease of transition from task focus to big picture awareness. It will be important to avoid developing software tools that engage and distract users.

Functional Diagrams

Drawing from data collected using the staged world exercise, we created functional diagrams for each of the major phases of a rescue mission: planning, en-route/infiltration, on-scene, exfiltration, and reconstitution. These diagrams capture goals, sub-goals and resources present in the domain, and how these elements are related to

each other. Knowledge of interdependencies allows for better anticipation of side-effects (sometimes undesired) that the introduction of new technology could have in this domain.
Figure 2 presents the basic elements of the representation. Functions, Information and Roles are the three types of elements used in the diagram. Elements are related to each other using colored arrows to show relationships and the flow of information. Different types of functions are depicted using five types of border/shading combinations representing different phases of the action cycle as depicted in Figure 2.

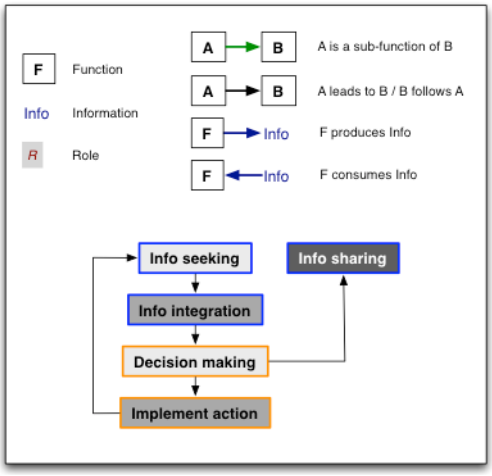

Figure 2. Basic elements of the functional representation

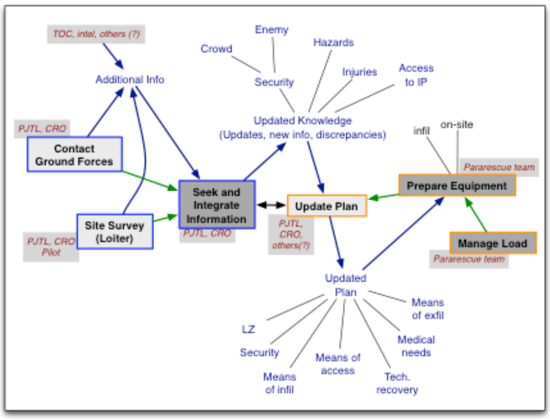

Figure 3. Functional description of the initial "En-route" phase

Figure 3 is the functional representation of the en-route phase of the mission. The team has departed from the air base and is traveling to the site where they will perform the rescue. Essential activities during this phase include gathering revised or additional information about the situation in order to update the plan accordingly.
During the en-route phase, radio communications allow for information exchanges with roles at the base (e.g., intelligence units), and also with ground forces to obtain updated information. When it is possible, a site survey is conducted from the aircraft to gather visual information about the scene and update the plan with first-hand information about topography, building layout, security situation, etc.

Design Concept

CTA findings resulted in several design concepts to support tactical level decision making for PJs and CROs. A tool to support Accountability and Ex-Fil Planning (Figure 4) is one example. We developed a design concept to aid in tracking personnel on-scene (accountability) and ensuring that no one is left behind. One reason the research team targeted this particular function is because of the cognitive challenges associated with trade-off management at this phase of the mission. Personnel must prioritize patients in terms of urgency, and plan evacuation of patients and troops relative to the resources available. Another reason for targetting these functions is that the use of a wearable computing device to support Accountability and Ex-Fil Planning is more feasible than for many other PJ functions. These functions generally take place from the vantage of the collection point, an area selected because it is somewhat protected. In addition, these functions are conducted by personnel tasked with maintaining a big picture view of personnel and activities in the area. The proposed tool might actually free cognitive resources as it may reduce the time and effort required to mentally construct a map of the area including the location of personnel and equipment, and then capture key components on paper.
Note that rather than providing automated answers to the complex problem of ex-fil planning and implementation, the purpose of the application is to support the operators' expertise through capturing and representing critical information. Operators typically use some form of hand-drawn diagram for this purpose (the specifics of it varying between operators). The computer-based application addresses important limitations of paper-based artifacts and offers additional capabilities: better support for dynamic information, links to data feeds for automatic entry, and the ability to share the big picture representation between users across the system. Such capabilities relate directly to the critical cognitive requirements such as team coordination, distributed re-planning, scene management, and big picture situation awareness.
Based on the cognitive requirements and our understanding of the domain, we developed a set of four parameters to guide our design concept. The first was to design a simple interface that is easy to scan and requires limited input with the intent of supporting big picture situational awareness (see Table 5). The environment is often noisy, chaotic, and dangerous. It is important that the interface does not present a distraction (thus limiting situational awareness), and can be easily read and manipulated. The second was to explore strategies for providing geolocation data of key players on the scene including the location of injured personnel, critical equipment, and team members. The third was that the interface must support dynamic accountability tallies. Accountability requires that the PJ team leader/CRO is able to track the number of injured

personnel expected compared to the number of injured personnel recovered, as well as the numberof personnel who have been evacuated from the scene. The fourth was to support dynamic ex-fil planning including matching resources available to ex-fil needs. The last three parameters were articulated to support scene management, team coordination, and distributed re-planning (Table 5).

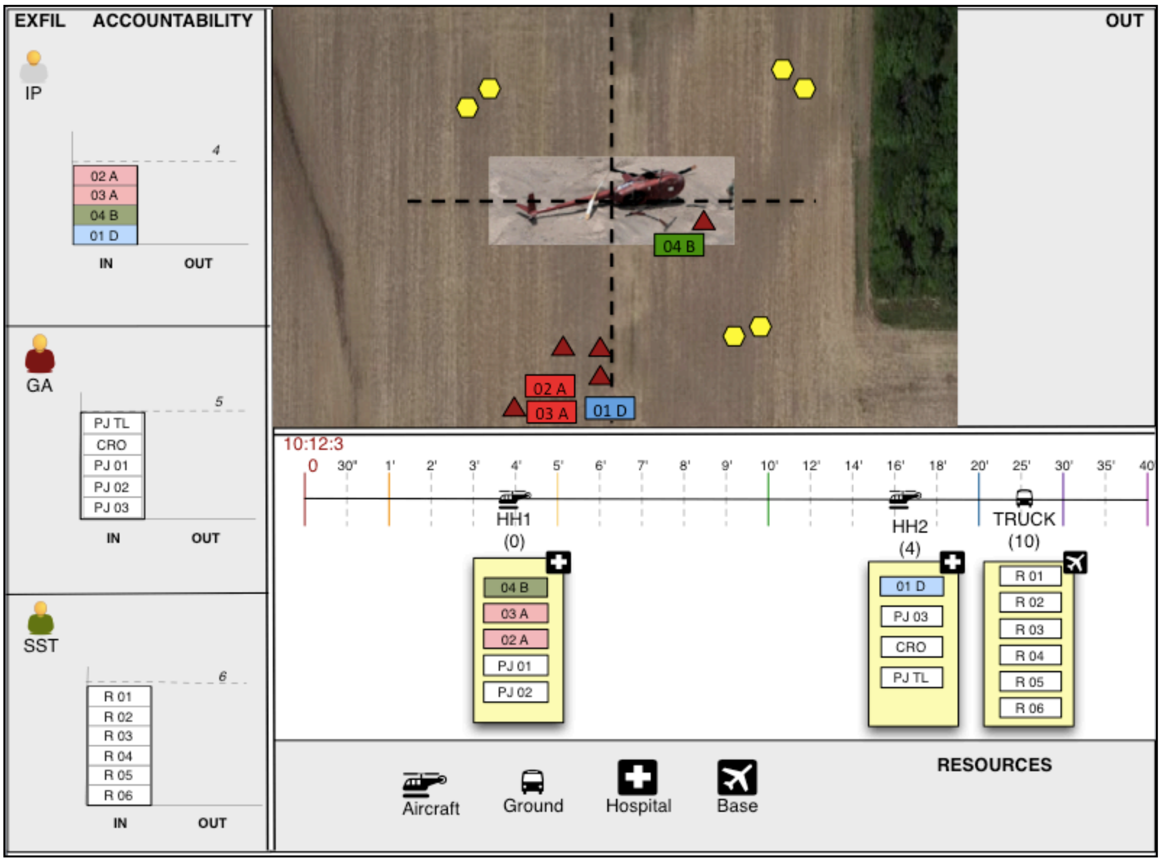

Figure 4. Accountability and Ex-Fill Planning design concept

The design concept shown in Figure 4 is designed to support Accountability and Ex-Fil Planning. The MAP panel aids in tracking where each injured person is found, and where team members are currently located. We leveraged the grid pattern commonly used by PJ Team Leaders to track the location of personnel in relation to the crash site or other agreed upon focus. By placing the crash site at the center of the grid, operators can easily create a grid using distance markers and/or use a clock metaphor to locate people and objects. As an injured person is recovered, s/he will be tagged with an RFID tag and a number and triage code will be assigned (A, B, C, D). When the injured personnel icon appears on the MAP panel (via RFID), a corresponding icon will also appear in the EXFIL ACCOUNTABILITY panel.

The EXFIL ACCOUNTABILITY panel on the left of the screen shows a tally of personnel on scene (the IN column) and those who have been ex-filled (the OUT column), categorized by function (i.e., injured personnel, Guardian Angel team, security team). A dotted line is used to indicate the number of personnel expected to be on-scene in each category.

The SYNCHRONIZATION panel in the lower portion of the screen is designed to support ex-fil planning. As the CRO becomes aware of a vehicle en-route to the incident site, it is placed on the timeline with an estimated time of arrival. The CRO can then drag personnel icons from the EXFIL ACCOUNTABILITY panel to each vehicle to create an ex-fil plan.

As each vehicle leaves the scene, the OUT panel on the upper right of the screen will be populated, indicating which personnel have been ex-filled in each vehicle. Figure 5 shows a scenario in which all personnel have been ex-filled.

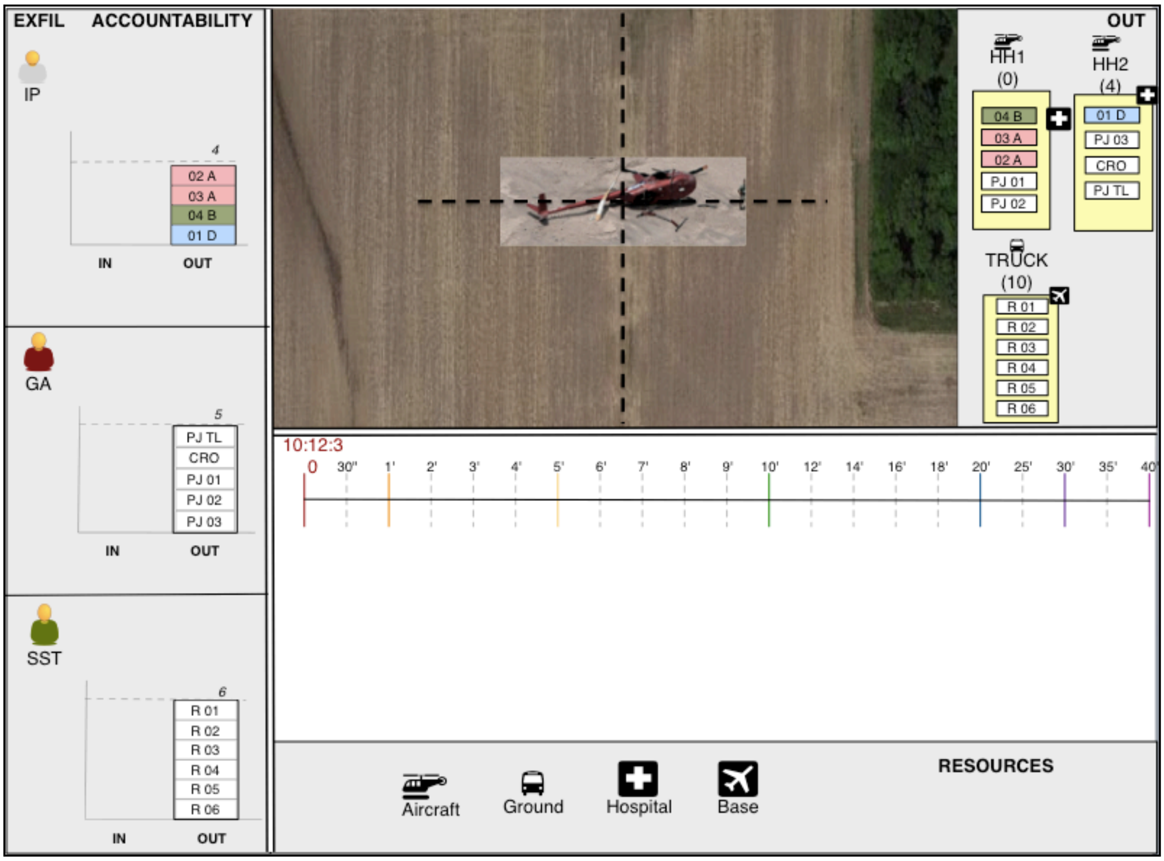

Figure 5. Accountability and Ex-fil Planning Support Tool depicting a scenario in which everyone has been evacuated

CONCLUSIONS AND NEXT STEPS

This paper described the use of CTA to identify cognitive requirements for the GAWS community to drive design activities and create scenario-based evaluations. CTA data will aid researchers in exploring the impact of proposed solutions on complex metacognitive activities such as team coordination, fluid command, emergent mission planning, distributed re-planning, scene management, and maintaining big picture situational awareness.

The CTA findings suggest that it will be important to be selective in introducing technological solutions into this domain, even though there have been many advances to create more robust, reliable, lightweight, wearable and portable technology. The fast-paced, high-risk environment in which GAWS operates requires an active awareness of the environment at all times, which even well-designed and well-intentioned technologies can hinder.

Next steps include using the requirements and scenarios articulated here to develop prototypes and conduct evaluations. The team intends to leverage AFRL's simulation capabilities, which include an immersive virtual environment. The mass-casualty scenario and others will provide an important foundation for the evaluation of design concepts. Efforts are currently targeted at transforming the detailed mass-casualty story into a script that will be programmed into the virtual environment. In addition to being a vehicle of testing prototypes, this type of simulation capability will also allow researchers to further investigate decision making in the pararescue domain and identify additional requirements. This is particularly useful in this domain because direct access to real world operations is unlikely.

ACKNOWLEDGMENTS

We would like to thank all the PJs and CROs who participated in this research. This project was sponsored by the Air Force Research Laboratory – Battlespace Acoustics Branch (AFRL/RHCB) under contract no. FA8650-08-D-6801.

REFERENCES

Air Force Medical Service Agency Air Combat Command Surgeon General (2011). Deep Dive, V1 [MS Excel workbook].

Air Force Special Operations Command (2011). AFSOC FY13 capability gap list: Derived from FY12-37 prioritized deficiency list (PDL). Department of Defense.

Crandall, B., Klein, G., Hoffman, R.R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Irick, Capt. ACC/A8RG. Guardian Angel (GA) F-Study Out Brief. Presented to Headquarters, Air Combat Command.

Militello, L.G. & Hutton, R.J.B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics Special Issue: Task Analysis*, *41*(11), 1618 -1641.

Task Analysis (2008). PJ Occupational Analysis Report (OAR).

Woods, D. D., & Hollnagel, E. (2006). *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. Boca Raton, FL: Taylor & Francis/CRC Press.

Emotion & Decision-Making

Naturalistic decision making and emotion in refereeing: affect at the heart of judgment

Géraldine RIX-LIEVRE, Michel RECOPE, Simon BOYER,
and Mathis GRIMONPREZ
Clermont Université, Université Blaise Pascal, ACTé, EA 4281, Clermont-Ferrand, France

ABSTRACT

Introduction: This study examines emotions in decision making during sport refereeing. We study how emotions contribute to referees' judgment acts. We seek to determine whether 1) a negative affect arises from the referees' relationship to the unfolding game; 2) judgment acts are able to end this unsatisfactory situation. **Method**: We have investigated four games in a professional rugby championship. The game was filmed from the stands, the referee was equipped with a head-mounted camera, and a self-confrontation interview was realized. We analyzed the situations in which the referees reported negative affect. **Results/discussion:** In some cases, affective dissatisfaction was related to an abnormal situation. The referee showed what was abnormal and overcame his dissatisfaction through a judgment act. In other cases, affective dissatisfaction was related to a paradoxical situation. The return to a satisfactory state occurred when a norm became predominant to determine the normality/abnormality of the situation within its unfolding.

KEYWORDS

Decision Making; Sport Sciences; Appraisal; Judgment act; Self-confrontation interview; Rugby referee

INTRODUCTION

Referee decision making is an emblematic example of complex cognition occurring in real-world context; this context has all the characteristics of the environment in which the NDM movement is interested in (Mascarenhas, Collins, Mortimer, & Morris, 2005). Our contribution considers especially how emotions influence the process of decision making during refereeing performance.

According to the NDM community (Klein, 2008), referees' decision processes are studied in naturalistically settings to consider their complexity and effectiveness (Rix-Lièvre, Boyer, Récopé, 2011). Referees are not only the guarantors of the rules, but they also manage the game (Mascarenhas, et al., 2005). Referees' many goals must all be taken into account (Mascarenhas, O'Hare, & Plessner, 2006). Their judgments result from an active process of decision making that handles the entire context of the match (Mascarenhas, et al., 2009; Mascarenhas, et al., 2005; Rix-Lièvre, et al., 2011). Our study was conducted from this perspective: studying referees' naturalistic decision making.

Although many studies have focused on referee decision making, few have taken their emotions into account (Philippe, et al., 2009). Studies on referees' emotional state focused on the sources of stress and their coping strategies (Dorsch & Paskevich, 2007; Voight, 2008) and didn't support a strong relation between decision making and emotions. Few studies focused on the consequences of errors upon affective and cognitive processes of refereeing (Philippe, et al., 2009). Nevertheless, the emotions that arise during a game are rarely taken into account. When emotions are studied, they are generally assessed through questionnaires or scales filled out before or after the game. These studies are thus not able to document referees' affective and cognitive functioning during a game: they only provide information on how the referees usually relate to it. We address this issue.

Our study focused on referees' decisions considering them as judgment acts (Rix-Lièvre, et al., 2011). Referees' judgment acts show and impose on the players what is possible by making reference to both what is ideal and what is legal; in so doing, they establish and qualify events (Ricoeur, 2000). Every judgment act describes a situation in a very particular way; this description is performative in the sense that it brings forth what it describes the way it is described (Rix-Lièvre, et al., 2011). In other words, a foul is called not just in relation to the reality of a fact, but it also depends on the referee's relation to the players' actions. This does not mean that referees make the game or invent the goals, but it does underline that even though players' activities pre-exist judgment acts, it's the way referees show them that makes them game events —goals, tries, fouls, etc. To understand how referees construct what they show and impose on players, the relationship between referees and

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

players' actions must be examined. However referees do not deal with an objective world of objects and events: individuals compose their own situations. The environment to which actors react is thus originally centred on and by them. According to Canguilhem (1989), a situation is always experienced as an ordeal in the affective meaning of the term; that is, an individuated relation to what is normal or abnormal for the person, depending on his or her own norms. These are the expression of a preference and an instrument of the intention to substitute a satisfactory state for a disappointing one. Every perception is already an evaluation of an affective nature (Lazarus, 2001); every situation is in this way more or less satisfactory/unsatisfactory according to the actors' own norms. This perspective suggests the interest of studying referees' decisions or judgment acts by focusing directly on affective functioning. Although emotion is often considered as a bias, a disturbance, that alters cognitive functioning, we examine what affects referees —what is satisfactory and unsatisfactory for them or the valence that they spontaneously confer on players' activity— and how it contributes to what they show and impose to the players. We sought particularly to study: (1) how referees' negative affects arise especially whether these arise from their relationship to the unfolding game, (2) whether a judgment act helps them –or not– to put an end to an unsatisfactory situation.

METHOD

This study used a long-term ethnographic approach and other more specific methods. We worked with four experienced rugby referees. These investigations took place during official games in the top level French professional rugby championship.

For each investigation, (1) the game was filmed from the stands and (2) the referee was equipped with a head-mounted camera to record a perspective close to his own subjective view of the situation. After the game, a *subjective re situ* interview (Rix & Biache, 2004) was conducted. The referee was asked to describe his activity and tell the researcher what had been important for him during his actions. The head-mounted video fosters "an experiential immersion" (Omodei, Wearing, & McLennan, 1997). This interview allowed the referee to share his experience and let the researcher understand what made sense for the referee at the moment he was refereeing, what compelled him, what affected him, and what mattered to him.

The referee's verbalizations during the interview were first transcribed. Next, we removed from the corpus all the justifications and general comments. We then identified and selected those passages in which the referee made note of a state of dissatisfaction, a negative affect. At last, we combine the interview excerpt with the description of the corresponding moment in the game. For each excerpt, we then proposed an interpretation as close as possible to what the referee had made explicit during the interview, emphasizing what was unsatisfactory for him and what his acts of judgment imposed.

RESULTS

The analysis of the interview excerpts in which the referees reported a dissatisfaction, negative affect, revealed two distinct cases. In the first case, dissatisfaction was related to an abnormal situation. The referee showed what was not normal for him and put an end to the situation: the dissatisfaction was resolved by a judgment act. In the second case, dissatisfaction was not only related to an abnormal situation, but also to a paradoxical relationship to the environment.

Dissatisfaction regarding an abnormal situation

The first category concerned all cases where dissatisfaction was related to an abnormal situation. In each case, the interpretation indicated what was affecting the referee, the source of his dissatisfaction —the nature of a contact, modes of play for grounding the ball, scrum positioning— and what he did in order to make the situation satisfactory again. In such cases, the negative valence of the situation depended on the game that was produced.

This affect was based on a value that defined what is normal and abnormal: norm depreciates everything that the reference to it may not hold for normal (Canguilhem, 1989). The norm and its strength were manifested strong conviction during verbalization, expressed by abrupt sentences (Récopé, Fache, & Rix, 2008): *"we have to, have to, have ...", "stop, stop, stop ..."*. In such cases, it should be noted that the abnormality of the situation was related to a single norm. In these moments, the players' activity was characterized with reference to a norm that predominated and imposed its requirements. The referee ended the abnormal situation by blowing his whistle. The judgment act resolved his dissatisfaction and brought the game back within a normal scope: it substituted a satisfactory state for a disappointing one for him. This leads to the conclusion that what affects the referee provides the basis for his judgment acts and what he shows and imposes on the players. The referee, in and by his judgment acts, points what is abnormal and indicates this to the players to better circumscribe what is possible within the dynamics of the game.

Discomfort regarding a paradoxical relationship to the environment

In the type of case described above, the negative valence of the situation is relative to the abnormality of the game as it is unfolding and the judgment act substitutes for the referee a satisfactory state for an unsatisfactory state. But other cases were observed. In these cases, dissatisfaction arose from a paradoxical relationship to the environment and a tension experienced by the referee regarding what he needed to do in order to act properly *hinc et nunc*. The referee experienced tension between two poorly compatible, even contradictory, perspectives that he was simultaneously aware of. This situation was quite paradoxical: normal by one value and abnormal by another. The referee thus experienced ambivalence arising from the tension between different norms. The norms may have concerned what is acceptable in a game; what a referee has to do to perform properly; what is important to progress in refereeing levels; what players, spectators, supervisors are expecting; and so on. We did not identify clearly competing norms, but rather frequent tensions between different norms. The tension between norms produced a negative affect that persisted until the resolution of the norms debate (Schwartz, 2011). The return to a state of satisfaction required that one quality became more salient in the dynamics of the situation. Thus, the normality of the situation is established. If it became clearly abnormal, the referee ended it by blowing the whistle. Compared with the previous result, the judgment act was being constructed within a larger time frame. What the referee considered as abnormal was not captured in the instant, but had been developing within the dynamics of the situation. If the situation became normal, the referee allowed the game to go on. The negative affect arising from the tension between norms thus did not necessarily lead to an intervention by the referee. The discomfort persisted as long as the situation aroused ambivalence and remained unclear.

DISCUSSION

The results of this exploratory study are discussed in three parts: (a) the origin and place of affect in the judgment acts of referees, (b) the connections between perception, decision, and emotion in refereeing, (c) the role of norms and values in the judgment activity of referees.

Our second result leads us to reconsider how referees' negative affects arise. If the referees' negative affects are sometimes related to their relationship to the unfolding game, it could also be embedded in their evaluation of their paradoxical relationship to the environment in his entirety. In other words, the origin of referees' negative affects can be understood from a meta-level. Similarly, complexity must be considered to conceptualize the place of affect in referees' judgment acts. As we expected, some judgment acts instantaneously clear up referees' dissatisfaction and get the play back into a normal framework. But every negative affect did not give place to judgment acts. The case of negative affect related to norms debate shows this: the negative feeling fades when the debate is resolved and the situation becomes clear. If an abnormal situation is confirmed or is becoming amplified and more prevalent, a judgment act ends the persisting negative affect. But if normalcy is becoming more predominant, a judgment act is no longer needed to lift a referee's negative affect. However negative affect appeared to be at the heart of the referees' judgment acts, several processes are distinguished. We established that the situations showed by the judgment acts are characterized as abnormal spontaneously in the instant or within a longer time frame that was related to the dynamics of the referee's situation.

About the connections between perception, decision, and emotion in refereeing, Dosseville et al. (2011, p58) suggested that "the perception of [players'] action provokes an emotion that will guide the decision". Our results support partially this proposal. Negative affect was the basis for the referees' judgment acts: what is perceived as abnormal in the moment or in the dynamics of the situation leads the referee to blow the whistle. But it is not the perception of action that causes an emotion; affect is at the very heart of referees' perceptions of player activity. The notion of "appraisal" helps to clarify our position: every perception is an assessment of what is happening in terms of personal well-being (Lazarus, 2001). This appraisal is not the result of reflection, calculation, or deliberation; it is instantaneous and results from a mostly unconscious and embodied process. These proposals might lead some authors to denounce bias in refereeing judgments and to find an affective explanation of biases in perception and/or categorization. Instead we emphasize the "implacable logic of emotion" and its rationality (Ibid, p59). According to Lazarus, it is in fact this rationality that is needed to investigate. As this rationality is based on tacit and embodied knowledge, more researches are necessary to better understand this rationality-in-action in the subjective experience of referees. More broadly, this leads to consider the interest of a cross-fertilization of naturalistic decision making movement and appraisal approach. Indeed, appraisal could provide an understanding of situation awareness' process and emotion should be incorporated as knowledge in decision making. It seems to be an alternative to an algorithmic rationality (Kahneman, Klein, 2009).

The referees' norms and values were established early on as references for what was affecting them (Canguilhem, 1989). Our results show that a referee's negative affect qualify/denote the unfolding of the play; abnormal play is specified relatively to a norm, with this norm providing core of the referee's relationship to his environment centred on the game unfolding. But our results also suggest that negative affect may be related to the ambivalence that referees experience: their situation may be normal according to a certain norm –the game has to manage itself- and abnormal for another –the game have to be cleared. Negative affect therefore originates in norms and in norms debates. Although we did not identify any particular tensions between norms, Schwartz (2011) noted the importance of examining the nature of the norms debate. He distinguished between prescribed/exogenous norms and endogenous norms, which are those specific to an actor and constructed by experience. He assumed that activity refers to a polemical relationship between exogenous and endogenous norms. This orientation suggests perspectives for future work on the activity of referees. Refereeing is based on many exogenous norms: game rules, federation guidelines, the criteria used to assess refereeing performance, and so on. Although Schwartz (2011) acknowledged that humans cannot possibly be completely directed by exogenous norms, it is equally unthinkable to referee without these norms being more or less operative. Therefore, future researches have to undertake the norms debates at the heart of refereeing activity, to identify their nature, and to study how they could be resolved. The study of how referees' judgment acts are constructed in connection with both what is legal and the referees' personal norms and values should help to provide empirical support for the proposals of Ricoeur (2000).

ACKNOWLEDGMENTS

This study was carried out within the PERF Arbitrage activities and with the support of Rugby French Federation.

REFERENCES

Canguilhem, G. (1989) The normal and the pathological, Zone books, New-York.

Dorsch, K. D. & Paskevich, D. M. (2007) Stressful experiences among six certification levels of ice hockey officials, *Psychology of sport and exercise*, 8, 585-593.

Dosseville, F., Laborde, S., Traclet, A., & Koffi, P. E. (2011) Refereeing decisions in soccer and physical performances in ecological situation, *STAPS, 93*(3), 51-60.

Kahneman, D. & Klein, G. (2009) Conditions for intuitive expertise, *American Psychologist*, 64(6), 515-526.

Klein, G. (2008) Naturalistic decision making, *Human Factors*, 50(3), 456-460.

Lazarus, R. S. (2001) Relational meaning and discret emotions, in K. Scherer, A. Schorr & T. Johnstone (Eds.), Appraisal processes in emotion: theory, methods, research, Oxford University Press, New-York.

Mascarenhas, R. D., Button, C., O'Hare, D. & Dicks, M. (2009) Physical performance and decision making in association football referees: A naturalistic study, *Open Sports Sciences*, 2(9), 1-9.

Mascarenhas, R. D., Collins, D., Mortimer, P. & Morris, R. L. (2005) A Naturalistic Approach to Training Accurate and Coherent Decision Making in Rugby Union Referees, *The Sport Psychologist*, 19, 131-147.

Mascarenhas, R. D., O'Hare, D. & Plessner, H. (2006) The Psychological and Performance Demands of Association Football Refereeing, *International Journal of Sport Psychology*, 37, 99-120.

Omodei, M., Wearing, A. & McLennan, J. (1997) Head-mounted video recording: a methodology for studying naturalistic decision making, in R. Flin, et al. (Eds.), Decision making under stress: emerging themes and applications, 137-146, Ashgate, Aldershot.

Philippe, F. L., Vallerand, R. J., Andrianarisoa, J. & Brunel, P. (2009) Passion in referees: Examining their affective and cognitive experiences in sport situations, *Journal of Sport & Exercise Psychology,* 31, 77-96.

Récopé, M., Fache, H. & Rix, G. (2008) Proper norm and bodily exercise: a case-study of volleyball player, *Corps*, 4, 105-110.

Rix, G. & Biache, M-J. (2004) Recording in *situated subjective* perspective and the *subjective re situ* interview: a method for the constitution of experience, *Intellectica*, 38, 363-396.

Rix-Lièvre, G., Boyer, S. & Récopé, M. (2011) Referee cognition as it occurs : different kind of judgment acts, Proceedings of the 10th International Conference on Naturalistic Decision Making, Orlando, Florida.

Schwartz, Y. (2011) Body-self: why this concept? Body-self, activity, experience, *Travail et Apprentissage*, 7, 146-177.

Voight, M. (2008) Sources of stress and coping strategies of US soccer officials, *Stress and Health*, 25, 91-101.

Effect of Emotions on Cognitive Processes: Are Proficient Decision-Makers Inoculated Against the Influence of Affective States?

Lorenzo P. LUINI, Francesco S. MARUCCI
"Sapienza" University of Rome - Department of Psychology

ABSTRACT
Introduction: Eventhough several studies has shed light on the effect of emotions on cognitive processes such as decision-making, research on naturalistic decision making has so far been virtually silent on the issue of emotions. Thus, the aim of the present research was to investigate how affect states influence the performance of a group of proficient decision-makers, police officers, in the ability to detect armed targets. **Method:** Two groups of subjects, civilians and police officers, where presented with a First Person Shooting Task. Pictorial stimuli were presented before the task to elicit specific emotions and to investigate the effect on particpants' performance confronting their results. **Results and discussion:** Whereas police officers resoulted able to better control the effect of arousing stimuli and negative affect states than civilians, their latencies where longer, a strategy that could be dangerous and resulting in becoming the victim in an hypothetical shooting incident.

KEYWORDS
Shooting Bias; weapon Bias; emotion; cognitive control; decision-making; stressful situations.

INTRODUCTION

Several categories of workers (i.e., police officers, personal security experts, etc.) with different personal characteristics and training backgrounds usually need to carry a gun or weapon. Weapon holders are, therefore, aware that one day the decision to pull the trigger may be necessary.

Contrary to common belief, the sudden decision to shoot or not usually happens unexpectedly, and generally routine activities, i.e., traffic stops or security checks, change immediately into a more complex situation.

In fact, police officers are routinely engaged in activities at different levels of risk but the majority of incidents involving police officers' use of firearms occurs while responding to a routine call for service in what are called spontaneous incidents (Burrows, 2007).

Therefore, in order to decrease the number of potential incidents, law enforcement personnel and security experts are trained in standard operating procedures or instructed in rules of engagement (Rahman, 2007).

Fortunately, not all the weapon holders have to experience the physiological and psychological arousal associated with a real life shooting incident but they do perform, as part of their training, tactical shooting training on the firing range or with virtual weapons training systems.

In fact, as part of police officers' training, they perform tactical exercises, including different scenarios, in order to be ready to properly respond when using their weapons (Burrows, 2007).

Nevertheless, as noted by Nieuwenhuys & Oudejans (2011), regular police training focuses predominantly on the technical, tactical, and physical aspects of performance and mostly neglects the role of psychological factors such as stress, anxiety and emotions.

Decisions such as those to shoot and take a life are called split-second decisions since mostly they have to be taken very quickly (a few milliseconds), despite the tremendous implications they might have on the lives of people involved.

Moreover, the improper use of lethal force, especially by police officers, has received wide coverage in the media in the past decade and raised concerns by the public because it involves innocent people being killed.

As a result, many researchers started investigating this phenomenon, especially in the USA, where much of the research has been extensively conducted (Barton, Vrij, & Bull, 1998).

Debate has focused on personal qualities in the use of lethal force and the factors that might affect the processes involved in the decision to shoot.

The decision to shoot (or not to shoot) depends on a subjective process that is influenced by many variables, by the perception of characteristics inherent to the decision task itself, and the decision environment (Beach, Mitchell, Deaton, & Prothero, 1978).

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Thus, researchers tried to explain DM in shooting by investigating personal variables such as ethnicity, gender, education and experience, and personality characteristics such as attitude and perceptions with controversial results.
Researchers from the Social Psychology domain such as Correll, Park, Judd, & Wittenbrink (2002) and Correll et al. (2007), focused on the influence of ethnicity of the target using a First Person Shooting Task (FPST) which involved the detection of hostile targets, providing robust evidence of a racial bias both in the psychological criterion participants adopted in deciding to shoot or not to shoot, and in the speed with which participants made the decisions.
Payne (2001) suggests that weapon misidentification may hinge on the distinction between controlled processing and automatic accessibility of the stereotype while Correll et al. (2007) argue that targets are quickly categorized according to race, and the relative stereotypes including the concept of danger are promptly activated.
In actual fact, fatal incidents are difficult to explain only with the Weapon Bias (Payne, Shimizu, & Jacoby, 2005; Payne, Burkley, & Stokes, 2008) or the Shooter Bias (Correll, Park, Judd, & Wittenbrink, 2002; Correll et al., 2007) hypotheses, and the study of other situational factors should be taken into account.
Nevertheless, to our knowledge, only limited research have been conducted concerning the effect of emotions on the decision to shoot (see Nieuwenhuys & Oudejans, 2011; Oudejans, Nieuwenhuys, & Raab, 2009).
Emotion such as fear includes many components, for instance, changes in the psychological state, cognitive processing, and subjective distress (Barlow, 2001, 2002).
Since the nature of police work often involves decision-making (DM) in the midst of a threatening situation and the extent to which an officer perceives the situation as being a threatening influence on his or her shooting decision, and threatening situational encounters endemic to police work elicit negative emotions such as fear and anger (Pole et al., 2001) that are associated with heightened psychological arousal, the aim of this research is to test whether the components of the emotions are sufficient to alter the performance in a First Person Shooting Task (FPST) (Correll, Park, Judd, & Wittenbrink, 2002).
Thus, one of the aims of this research is to investigate whether arousal is sufficient to alter the perception in general, and specifically of threats (i.e., armed target).
Recent studies suggest that heightened negative affect usurps cognitive resources that facilitate controlled decision-making (Kleider & Parrott, 2009).
Attempts to regulate negative emotions impair performance suggesting that heightened negative affect and arousal act as a load on cognitive resources and impair performance on capacity-demanding tasks such as DM.
In this study, the affective valence, another component of emotions such as fear, was tested because there is evidence (Schimmack & Oishi, 2005) to suggest that, even if elicited with emotional pictures, they could produce interference in a primary cognitive task.
Furthermore, both physiological responses and activation of cortex vary for different specific contents within the same valence category (Bradley, Codispoti, Cuthbert, & Lang, 2001; Bradley et al., 2003).
According to Bradley, Codispoti, Cuthbert, & Lang (2001), psycho-physiological differences exist between categories of a negatively valenced picture.
Moreover, the evolutionary threat hypothesis (Pratto & John, 1991) assumes that stimuli which threaten survival have more adaptive value than the detection of other stimuli and that this adaptive pressure have resulted in a specific detection mechanism for stimuli that presented a threat to survival (Schimmack & Oishi, 2005).
In this research threat stimuli, such as assaults, were compared to other affective stimuli to investigate if the they have a different effect on performance in a FPST.

METHOD

Subjects and Materials

95 participants: 42 white EU citizens (32F, 10M; Age: M = 26.4; SD = 8.9); 53 White police officers (45M, 8F; Age: M = 39.2; SD = 5.55) participated in the study voluntarily.
184 digitalized photographs of human figures varying in Ethnicity (Black/White/Neutral) were created with Photoshop® in order to build a set of bodies with different postures without any racial cue. To every figure a face was added to the top in order to create an harmonic human figure varying in ethnicity. The Neutral set of people was created by replacing the face with a grey baseball cap.
Prime stimuli included thirty-six human figures (four different relaxed postures), with twelve White faces, twelve Black faces and four Neutral; target stimuli (172) were represented by the same human figures but holding either four weapons or four tools (Left/Right hand) (See Figure 1 and 1a).
Two groups of 6 (G1 and G2) and three groups of 4 (G3, G4 and G5) digitalized 20 x 16 cm International Affective Picture Sistems (IAPS)' pictures were created. Stimuli were selected according to IAPS standard ratings and the groups, homogeneous for Valence and Arousal, named as follows: Group 1 (L_P): Arousal = Low and Affective Valence = Positive; Group 2 (H_P): Arousal = High and Affective Valence = Positive; Group 3 (H_NG): Arousal = High and Affective Valence = Negative; Type = Generic; Group 4 (H_NA): Arousal = High and Affective Valence = Negative; Type = Assault; Group 5 (H_NAA): Arousal = High and Affective Valence = Negative; Type = Armed Assault.

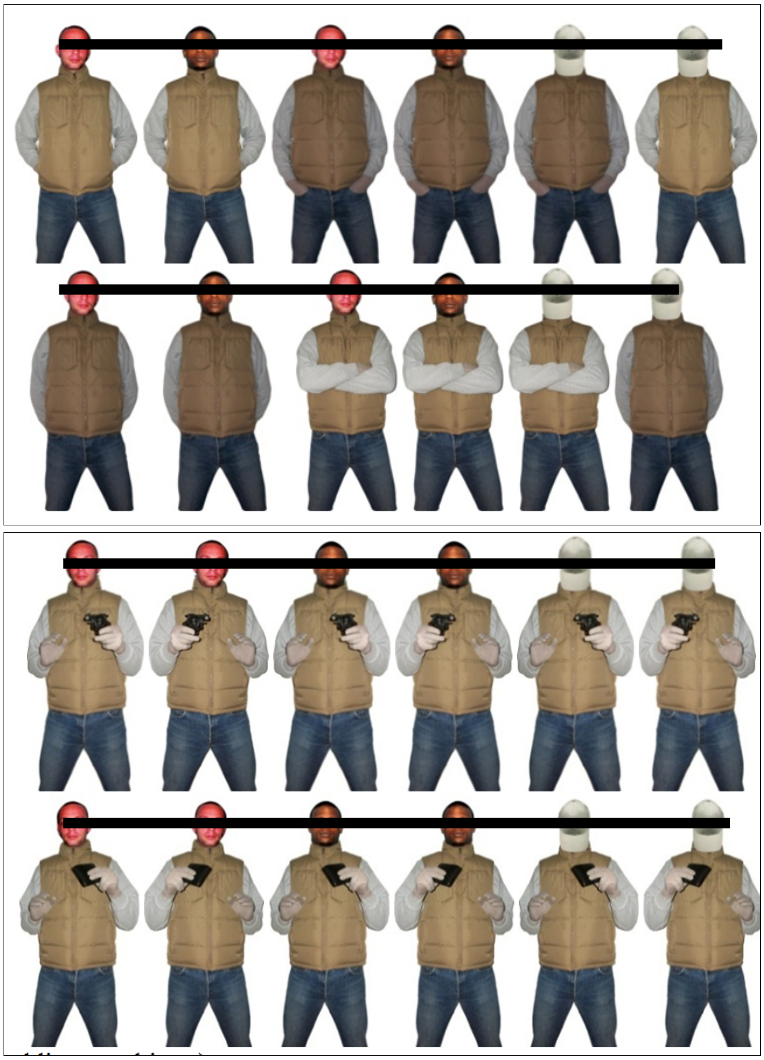

Figure 1 and 1a Sample stimuli: relaxed postures and holding up objects

Procedure

After the completion of a training phase of 20 trials to familiarise themselves with the task, all participants were warned they were beginning the experimental phase and informed about the lasting and the different phases of the test.

Participants then entered three experimental blocks in which they were presented with test trials. In the main task (Task 1) a prime varying in Ethnicity (Black/White/Neutral) was used and after 200 ms this was replaced by the target (Armed/Not Armed).

Participants had to respond fast pressing a response key to shoot (space bar) and not to shoot (C or M). After 200 ms the object was masked and after 500 ms a "Faster" sign appeared if any response key were pressed. The "Faster" banner was placed in order to hide the object leaving the face visible. Fixation cross and feedbacks were placed at face high.

Participants received a red "Wrong Armed/Not armed target" as negative feedback or a black "OK Armed/Not armed target" as positive feedback that remained on the screen till the next trial was presented, thus the effective deadline was 900 ms after the mask was superimposed on the target.

After 700 ms another trial would start (see Figure 2).

Three blocks of test trials were presented to participants: Task 1 (90 trials) as described above; for Task 2 participants were warned that next they were going to perform a similar task to Task 1 but that it would be anticipated by a picture randomly selected from group G1 or G2; for Task 3 as for Task 2 but pictures were randomly selected from group G3, G4 or G5. Participants were warned to ignore the first picture (IAPS pictures) and to respond to the target.

Altogether, participants had to perform 542 trials: 20 training trials (using only Neutral human figures); 90 trials for Task 1; 216 trials for Task 2; and 216 trials for Task 3.

Reaction time was recorded in milliseconds and, considering Signal Detection Theory (SDT), perceptual sensitivity (d') and response bias (c) were computed using Hit and False Alarm (FA) rates to measure the effect of Object, Ethnicity, Arousal, and IAPS pictures Type on participants' performance.

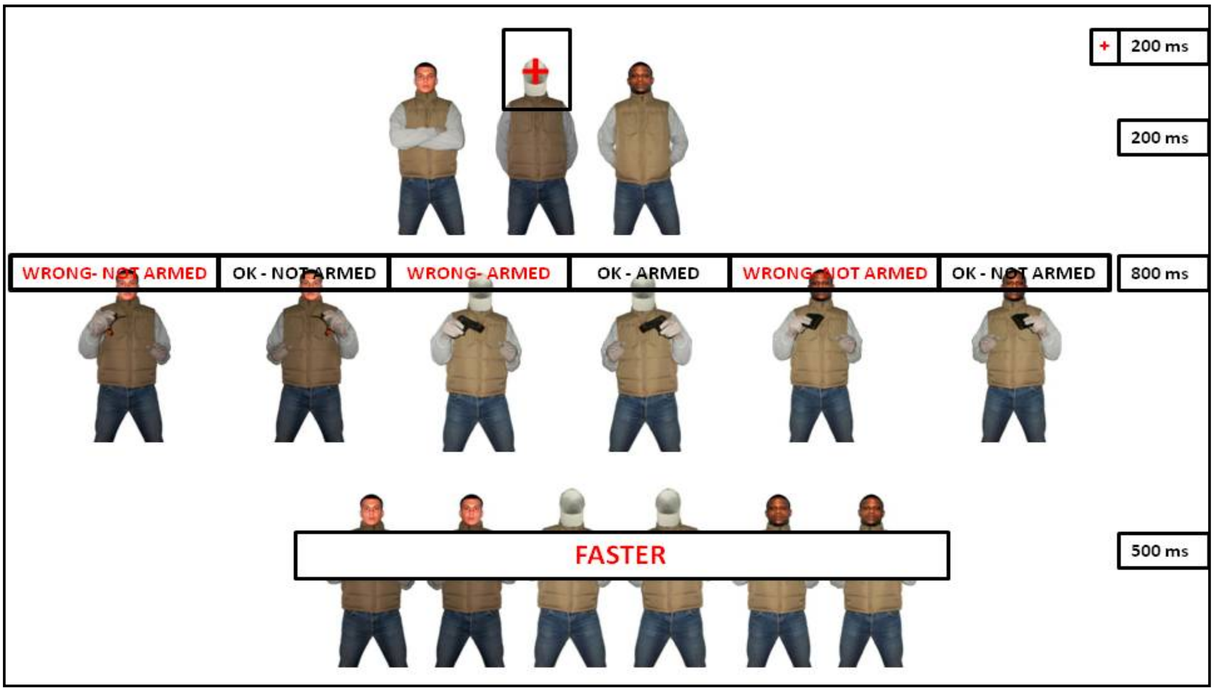

Figure 2 Experiment procedure. Stimuli presentation and timing

RESULTS

Civilians and the FPST

Data were collected on 42 white EU citizens (32F, 10M; Age: M = 26.4; SD = 8.9) in order to verify if the same pattern of researchers' results in this domain (Racial/Weapon Bias) could be generalised for ethnical groups where immigration histories and minority groups are different.

Results for the non-IAPS picture control condition partially replicated findings from the literature of reference (Correll, Park, Judd, & Wittenbrink, 2002; Correll et al., 2007).

In fact, in the control condition, although participants displayed no racial bias, they consistently had shifted their attention toward armed targets (see Table n. 1).

Table 1. RT Means and S.E. for correct responses and type of target (Hit and Correct Rejections [CR])

TARGET	MEAN (ms)	S.E.
Armed	423** r = .69	9.455
Unarmed	475	10.768

**P < .01

Moreover, pictorial stimuli presented before the main task had a dramatic effect on racial bias and target categorization and made ethnicity more salient whereas no racial bias was displayed. That is, a difference was found either for Blacks or Whites as compared to Neutrals (see Table n. 2).

Results of the study demonstrate how arousal and valence interact differently with ethnicity, supporting the hypothesis that arousing cues could be transferred from one source to another, not taking into account the valence.

In fact, from one point of view, low arousing positive valenced IAPS pictures made participants "un-armed" target-oriented (visually tuned in), and extended this tendency to misperceptions of weapons as tools without any ethnical orientation; conversely, high arousing stimuli, even if positive, visually tuned participants to being more aware of (thus more ready for) armed Black targets.

Table 2. RT Means and S.E. for ethnicity. Participants responded faster when a human face were displayed in comparison with neutral stimuli.

ETHNICITY	T2	MEAN (ms)	T3
Black	436* r = .31		439* r = .35
White	435* r = .31		438* r = .31
Neutral	445		448

*P < .05

Another variable that affected the tendency to shoot and modulated associations is the content of the stimuli.

In fact, armed assault pictures lowered the level of FA rate values for Blacks as compared to Neutral targets, assault enhanced Hit rates for Blacks, whereas generic type of IAPS pictures enhanced Hit rates for White targets. Moreover, in general, Armed Assault IAPS pictures favoured the tendency to shoot, whereas Assault type had little effect on shooting behaviour and Generic IAPS pictures only slightly affected the tendency to shoot (see Table n. 3).

Table 3. IAPS picture type presented before the FPST: Hit Rate Means and S.E. for the FPST

IAPS PICTURE TYPE	MEAN	S.E.
Armed Assault	.890	.013
Assault	.863	.017
Generic	.883	.012

We argue that when a weapon was present in the stimuli (environment) the eventuality of a shooting reaction become more salient, thus, attention shifting toward weapons was enhanced, whereas when a weapon was not present, a shooting reaction was suppressed. When information of weapons was not clearly present and its absence was not definitively excluded, a higher level of awareness was kept, in order to detect the eventual presence of a weapon and it resulted in a higher level of vigilance as compared to a non-weapon condition.
Finally, it is interesting to note that racial bias was sensitive to the content of the stimuli present in the environment (IAPS pictures) supporting the hypothesis that associations between social categories, concepts, and affective states influence visual processing mechanism.
Real-world objects perception and face attendance are fundamental aspects of vision and information extracted by the environment affect personal mood state and, thus, visual tuning and information processing. Mistakes, or misperceptions, and the way perceivers structure visual stimuli are directed by associations (often automatic) that are influenced by personal affective state.
Another issue was addressed was whether the same pattern of response bias would be displayed by Italian police officers.

Police officers in action

The same task was performed on 53 White Italian police officers (45M, 8F; Age: M = 39.2; SD = 5.55) to investigate if their specific training or the general exposure to weapons and crime-related stereotypes would have been comparable to civilians' effects on their performance in the FPST.
It is interesting to note that for the control condition (non-IAPS picture) a difference was found in RTs within responses where objects were misperceived. In fact, participants showed they were faster in misperceiving a tool as a weapon compared to the opposite condition (see Table n. 4).

Table 4. RT Means and S.E. for control trials showing participants were faster in wrongly misperceiving as armed a taget that was holding up an innoquous.

RESPONSE TYPE	MEAN (ms)	S.E.
Hit	439** r = .77	9.466
Correct Rejections	498** r = .89	10.389
Miss	495	10.557
False Alarm	422	13.114

**P < .01

Even though it is difficult to determine if participants responded according to what they "saw or believed they saw", their anticipation in seeing guns, because armed targets were the most relevant category in the task, was extended to responses, probably faster probability-based, with the effect of predicting more weapons (armed targets) than tools (un-armed targets) resulting faster but less efficient (shorter latencies for FA responses).
Police officers responded faster in shooting armed targets than un-armed ones either in the non-IAPS picture control condition or in the two conditions where IAPS pictures were presented before the main task. In fact, police officers were never affected in their reactions to shooting armed targets faster (see Table n. 5 and n. 6).

Table 5. RT Means and S.E. for correct responses and type of target (Hit and Correct Rejections [CR])

OBJECT	MEAN (ms)	S.E.
Armed	445** r = .81	7.828
Unarmed	497	7.641

**P < .01

Table 6. RT Means and S.E. for correct responses and type of target (Hit and CR, Task2 & T3)

OBJECT	T2 MEAN (ms)	T3
Armed	455* r = .66	454** r = .58
Unarmed	479	479

*P < .05; **P < .01

Moreover, police officers demonstrated they were able to use stimuli to enhance their performance. In fact, they were more accurate (higher Hit rates) and less prone to misperceive objects displaying a decrease in errors, specifically, deadly ones (FA rates).
Results for police officers support the hypothesis that being more motivated in detecting armed targets may alter the threshold at which weapons would be detected as well as predicted in fact, armed assault pictures negatively affected police officers' FA rates demonstrating they were more sensitive to armed assault IAPS pictures than

the other two categories (generic and assault) and more prone to misperceive the threat of armed targets in that condition.

DISCUSSION AND CONCLUSIONS

It is comforting that real weapon holders showed they have no response bias in the final decision to shoot or not toward particular ethnical groups and, thus, demonstrated that they cognitively discriminate better between necessary information for the successful result of the primary task (i.e., shooting) and information that could be sources of lethal errors (i.e., racial cues).

The problem is that, whereas police officers showed to have no response bias in the final decision, they seem to control the effect of arousing stimuli, negative affect states, and stereotypic associations at the expense of latencies and it could cause them injuries, or being killed when confronted with armed suspects.

Our results are in line with research in the split-second decision domain, that investigate decisions such as shoot/don't shoot decision, that is consistent with the fact that police officers show a robust racial bias only in their latencies when called to a shoot/don't shoot decision.

In fact, Correll, Park, Judd, Wittenbrink & Sadler (2007) report that whereas targets' ethnicity didn't affect police officers' ultimate decision, it affects their reaction times showing that police officers overcome stereotypes, even if they process racial information and access relevant stereotypes too, avoiding their interference in final decisions.

On the other hand, evidence for racial bias in SDT parameters in police officers are contrasting. In fact, our results and those of Correll, Wittenbrink, Park, Judd, & Sadler (2007) are not consistent with other research showing police officers display the racial bias in the SDT analysis (Plant, Peruche, & Butz, 2005).

Sim & Correll (2009) argue that training is what helps in overcoming the influence of prevalent cultural stereotypes on the decision to shoot because it may improve object-based judgements or may promote cognitive control, reducing the influence of stereotypes, thanks to more controlled responses.

Interestingly, results showed how arousal and valence interact differently with ethnicity, supporting the hypothesis that arousing cues could be selectively transferred from a source to another, "overriding" the valence.

In fact, whereas positive affect state eliminated the response bias toward Blacks and weapons, positive arousing stimuli affected participants' RTs in shooting faster armed Black targets as compared to Neutrals and Whites.

Moreover, it was demonstrated that the content of the stimuli affected the tendency to shoot and modulated associations.

In general, results support the hypothesis that people are more motivated to detect weapons than innocuous objects because threat has a central role in surviving; thus, humans have developed an efficient mechanism that has to detect survival-threatening stimuli in the environment.

When humans are called to detect a potential source of threat, i.e. a gun, they perhaps enhance their ability to detect it by priming their mind, imagining the object constructing it mentally. Then, people proceed by comparing real life objects with mental image objects using a faster matching versus a slower mismatching process, and latencies of responses result accordingly.

As proposed by Neisser (1967), Marr (1976; 1982), and Marr & Nishihara (1978), there are at least two levels of visual information processing. The first level works more automatically, pre-attentive, in a parallel-processing way managing information simultaneously and will-independent. The second level is sequential and depends on the will of the observer. That is: the visual system starts to code, disaggregate in a certain number of simple characteristics that are evident and automatically detectable, then attention starts to identify and recognise the object (Marucci, 2009).

In this, affect states influence the orientation toward threatening or non-threatening stimuli facilitating, i.e., the detection of weapons or tools. It has, in fact, been shown that relevance is a predictor of attention allocation (Dalgleish, 1995; Marucci, 2009).

Furthermore, arousal makes social categories more salient (Wilder, 1981) and facilitates associations, especially stereotypic ones, and intensifies ongoing behaviour even when residual (Zillmann, 1978). Furthermore, arousal, causing incoherent cognitions and biased information-processing or attitude-changing, has important effects on attention (Easterbrook, 1959), narrowing the focus and limiting the amount of information or cues an individual could process.

When aroused, even if positively, people display bias and the response bias is exacerbated when arousal states are the consequence of negative arousing activation.

Thus, positive affect states facilitate tools detection, whereas negative states facilitate weapons detection, and arousal facilitates stereotypical associations.

In addition, this research found evidence that dissociation between arousal ratings and behavioural responses in shooting could be caused by different types of stimuli. Like Scherer, Smelser, & Baltes (2001) proposed, stimuli are sequentially evaluated not starting automatically from the evaluation of the valence or threat but rather, determining how relevant the event is for the person. The nature of the unfolding emotion process is influenced, by the initial relevance check and the consequential regulation of attention to detect the stimuli's characteristics (Scherer, Smelser, & Baltes, 2001).

As compared to civilians, police officers were slower in responding, particularly when presented with arousing stimuli, demonstrating that they had overcome any eventual stereotypic association at the expense of time. Moreover, police officers responded faster in shooting armed targets than un-armed ones either when cognitively activated or not and showing they were able to use stimuli to enhance their performance.

Whereas police officers were efficient and not biased demonstrating that they were able to discriminate better between necessary information and information that could be sources of lethal errors, to control the effect of arousing stimuli, negative affect states, and stereotypic associations, they used time (longer latencies), that was a dangerous strategy that could result in their becoming the victim.

On the other hand, our results support Correll, Wittenbrink, Park, Judd, & Sadler's (2007) hypothesis. In fact, whereas targets' ethnicity does not affect police officers' ultimate decision, it affects their reaction times showing that police officers overcome stereotypes, even if they process ethnical information and access relevant stereotypes too, avoiding their interference in final decisions.

It remains to be determined if police officers are not affected by ethnicity in their final decision because training: 1) enhances their ability inhibiting stereotype-related information; or 2) makes them more skilled at object judgements; or 3) gives them more cognitive control over the influence of activated stereotypes; or 4) makes them more able to regulate themselves enhancing their ability to extract relevant cues from the task or in controlling the selection of responses (Sim & Correll, 2009).

From this research it appears that police officers are adequately trained to disregard information that is not essential for the successful execution of a shooting task even under negative affective states, when aroused or in specific contexts.

Nonetheless, programmes, either basic or advanced, predominantly focus more on the training of action than on the training of the ability to better make decisions or properly judge the situation.

Therefore, police officers' shooting reactivity does not suffer a delay either in the action, or in the judgement, but in the decision.

However, training programmes for police officers are at present the most advanced, and result in an adequate method of training for the profession. Nonetheless, future advances in training programmes will surely have to focus on the decision (to shoot) phase, that is to say the ignition for the behaviour (shooting) in order reduce reaction times, and the eventuality the police officer would become the victim.

ACKNOWLEDGMENTS

We thank the two anonymous reviewers for their valuable and constructive comments.

REFERENCES

Barlow, D. H. (2001). Clinical handbook of psychological disorders: A step by-step treatment manual. (3th ed.). New York: The Guilford Press.

Barlow, D. H. (2002). Anxiety and its disorders: The nature and treatment of anxiety and panic. (2nd ed.). New York: The Guilford Press.

Barton, J., Vrij, A., & Bull, R. (1998). New paradigms in researching police use of firearms. Journal of Police and Criminal Psychology, 13(2), 36-41.

Beach, L. R., Mitchell, T. R., Deaton, M. D., & Prothero, J. (1978). Information relevance, content and source credibility in the revision of opinions. Organizational Behavior and Human Performance, 21(1), 1-16.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and Motivation: Defensive and Appetitive Reactions in Picture Processing. Emotion, 1(3), 276-298.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the Visual Cortex in Motivated Attention. Behavioral Neuroscience, 117(2), 369-380.

Burrows, C. (2007). Critical Decision Making by Police Firearms Officers: A Review of Officer Perception, Response, and Reaction. Policing, 3(1), 273-283.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals. Journal of Personality and Social Psychology, 83(6), 1314-1329.

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot. Journal of Personality and Social Psychology, 92(6), 1006-1023.

Dalgleish, T. (1995). Performance on the emotional Stroop task in groups of anxious, expert, and control subjects: A comparison of computer and card presentation formats. Cognition & Emotion, 9(341 - 362).

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and organization of behaviour. Psychological Review, 66, 183 - 201.

Kleider, H. M., & Parrott, D. J. (2009). Aggressive shooting behavior: How working memory and threat influence shoot decisions. Journal of Research in Personality, 43(3), 494-497.

Marr, D. (1976). Early processing of visual information. Phil. Trans. R. Soc. Lond. B, 275:483-524. .

Marr, D. (1982). Vision. San Francisco: CA: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three dimensional structure. Paper presented at the Proceedings of the Royal Society of London, London.

Marucci, F. (2009). Le immagini mentali: Carocci.

Mosier, K. L., & Fischer, U. M. (2009). Does Affect Matter in Naturalistic Decision Making? Paper presented at the Ninth International Conference on Naturalistic Decision Making (NDM9), London, UK,.

Neisser, U. (1967). Cognitive Psychology New York: Appleton-Century-Crofts.

Nieuwenhuys, A., & Oudejans, R. R. D. (2011). Training with anxiety: short- and long-term effects on police officers' shooting behavior under pressure. Cognitive Processing, 12, 277-288.

Oudejans, R. R. D., Nieuwenhuys, A., & Markus Raab, J. G. J. a. H. R. H. (2009). Perceiving and moving in sports and other high-pressure contexts. In Progress in Brain Research (Vol. Volume 174, pp. 35-48): Elsevier.

Payne, B. K. (2001). Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon. Journal of Personality and Social Psychology, 81(2), 181-192.

Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why Do Implicit and Explicit Attitude Tests Diverge The Role of Structural Fit. Journal of Personality and Social Psychology, 94(1), 16-31.

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. Journal of Experimental Social Psychology, 41(1), 36-47.

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. Journal of Experimental Social Psychology, 41(2), 141-156.

Pole, N., Best, S., Weiss, D., Metzler, T., Liberman, A., Fagan, J., et al. (2001). Effects of gender and ethnicity on duty-related posttraumatic stress symptoms among urban police officers. Journal of Nervous & Mental Disease, 189, 442-448.

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. Journal of Personality and Social Psychology, 61(3), 380-391.

Rahman, M. (2007). High Velocity Human Factors: Human Factors of Mission Critical Domains in Nonequilibrium. Human Factors and Ergonomics Society Annual Meeting Proceedings, 51(4), 273-277.

Scherer, F. M., Smelser, N. J., & Baltes, P. B. (2001). Firm Behavior. In International Encyclopedia of the Social Behavioral Sciences (pp. 5676-5681). Oxford: Pergamon.

Schimmack, U., & Oishi, S. (2005). The Influence of Chronically and Temporarily Accessible Information on Life Satisfaction Judgments. Journal of Personality and Social Psychology, 89(3), 395-406.

Sim, J. J., & Correll, J. (2009). Training moderates racial bias in the decision to shoot. Paper presented at the Paper presented at the Annual Meeting Midwestern Psychological Association Chicago: IL.

Wilder, J. (1981). The "Law of Initial Values", a neglected biological law and its significance for research and practice. In S. W. Porges, and Coles M.G.H. (Ed.), Psychophysiology (pp. 47 - 55). Strodsberg, PA: Dowden, Hutchinson, and Ross.

Zillmann, D. (1978). Attribution and the misattribution of excitatory reaction. In J. H. I. Harvey, W.J., and Kidd, R.F. (Ed.), New directions in attribution research (Vol. 2). Hillsdale, NJ: Erlbaum.

Coordination in Decision Making

A Model of Metacognition for Bushfire Fighters

Lisa M. FRYE & Alexander J. WEARING
School of Psychological Sciences
The University of Melbourne

ABSTRACT

Introduction: A large-scale bushfire is an example of a *macrocognitive* work system (Klein, et al.) in which many people (such as incident management teams, bushfire fighting crews, and residents) are required to respond to uncertain and changing conditions. In these situations people need to manage multiple (often competing) cognitive demands, and they use *metacognitive* skills to regulate their thinking. **Method:** We explored the metacognitive skills of career and volunteer bushfire fighters using human factors interviews on the fireground and visual-cued recall debriefs during command post simulation experiments (Frye & Wearing, 2011). **Results and Discussion:** These think-aloud techniques revealed an extensive use of metacognitive knowledge based on previous experiences, and previous experience was described as both a source of expertise and a source of human error (consistent with Kahneman & Klein, 2009). In this paper we use a model of metacognition to describe how expert fireground commanders regulate their thinking, and thus avoid errors associated with cognitive overload, during large-scale bushfires.

KEYWORDS

Learning and Training, Bushfire, Expertise, Decision Making, Metacognition, Cognitive Control

INTRODUCTION

Bushfires occur in Australia every year, and the worst large-scale bushfires have historically caused a significant loss of life and property. For example, the 1939 Black Friday bushfires resulted in 71 fatalities, 1300 destroyed homes, and 2 million hectares of burnt landscape (Stretton, 1939). Similarly, the 1983 Ash Wednesday bushfires resulted in 75 fatalities and over 2000 destroyed homes, while the 2009 Black Saturday bushfires resulted in the largest number of bushfire fatalities in Australia's recorded history, when 173 residents perished (Teague, McLeod, & Pascoe, 2009). The damage during these large-scale bushfires was due in part to extensive fire spotting (which occurs when embers travel ahead of the main fire front to start new fires), and loss of life was frequently associated with a sudden wind change (which rapidly turned the flanks of a bushfire into a new, and much larger, fire front; see Teague et al., p. 42). These conditions mean that the fireground is a unpredictable environment for making decisions and that all bushfires involve some level of risk.

Macrocognitive Work

It also means that large-scale bushfires are an example of *macrocognitive* work, in which: decisions are complex and often involve data overload; decisions involve risk, high stakes, and are made under extreme time pressures; goals are ill-defined and multiple goals often conflict; and decisions occur in conditions where few things can be controlled or manipulated (Klein, Ross, Moon, Klein, Hoffman & Hollnagel, 2003, p. 81). Bushfire fighters call these *'damned if you do, damned if you don't'* situations, and in these conditions people can experience 'stress, fear, panic and a collapse of clear thinking' (Putman, 1995), or cognitive overload (McLennan, Pavlou, & Omodei, 2005). Judicial investigations (such as coronial inquiries and royal commissions) have therefore recommended that practitioners need experience and mentoring (as well as technical training) to develop leadership and decision-making skills for bushfire response roles in Australia (see Teague, McLeod, & Pascoe, 2010; Johnson, 2002). These recommendations are also consistent with the advice of human factors and crisis management researchers (e.g., Ericsson, 2006; Fadde & Klein, 2011; Leonard, 2010; t'Hart, 2010).

Bushfire Fighters

However, while large-scale bushfires are high consequence events, they occur relatively infrequently, and opportunities to gain appropriate experience are therefore limited. Also, bushfires in Australia are fought by a largely part-time workforce, comprised of volunteers of the Country Fire Authority (or equivalent rural fire service in each state), and public land managers from the Department of Sustainability and Environment (or equivalent government departments in each state). This means that there are relatively few full-time bushfire fighters in rural Australia, and instead, people are deployed from their substantive roles to fight bushfires when they occur (see Frye, 2012). This is different to other macrocognitive work environments (such as aviation,

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

medicine, or urban fire fighting) where practitioners frequently perform their role as a primary occupation, and therefore develop their decision-making skills in that context. Consequently, developing and maintaining expertise for large-scale bushfires is a challenge in Australia, and training programs need to support accelerated learning on the job.

Human Factors Research

At the same time, there has been a substantial amount of research about how people make decisions (and errors) in dynamic, uncertain, and time-pressured situations (like bushfires). For example, Klein, Calderwood and Clinton-Cirocco (1988) found that expert fireground commanders use previous experiences to recognise cues in a new situation and to deploy strategies that have been successful in the past (called Recognition Primed Decision Making; Klein, 1999). Several researchers (e.g., Gigerenzer & Goldstein, 1996; Connolly, 1999) also propose that heuristics, or rules of thumb, enable experienced practitioners to make fast decisions that are good enough to get the job done (called satisficing; Simon, 1956), while Cohen, Freeman & Wolf (1996) found that *metacognition* enables experienced practitioners to recognise, critique and correct their decisions (and errors) while they are implementing a course of action (see also Anderson, Oats, Chong & Perlis, 2006). Metacognitive strategies therefore enable practitioners to operate within the constraints (or boundary conditions) that typically exist in macrocognitive work settings, and to make decision tradeoffs, such as tradeoffs between efficiency and thoroughness (bounded cognizance; Hoffman & Woods, 2011; Valot, 2002).

Bushfire Training

However, until recently this empirical human factors research has had little impact on the way that end users are trained for large-scale bushfires in Australia (cf. Frye, 2012; Owen et al.; Slijepcevic et al.; Stack & Owen, 2012). Instead, lessons learned from judicial proceedings have historically resulted in lengthy policies and standard operating procedures (SOPs), which were then integrated into standardised technical skills training. These initiatives may improve decision-making (and reduce human errors) in routine situations, but are inadequate for dealing with unique or rapidly deteriorating situations (for similar issues in aviation training, see Adams & Ericsson, 2000). This is because one of the underlying causes of human error during large-scale bushfires is *cognitive overload* (see Teague, McLeod, & Pascoe, 2009, p. 235; Johnson, 2002, p. 602).

A COGNITIVE ENGINEERING APPROACH

A cognitive engineering approach would therefore ensure that fire agencies are aware of the cognitive demands that people face during large-scale bushfires, and in turn, design effective 'technology, training, and processes to manage cognitive complexity' (Militello, Dominquez, Lintern & Klein, 2009, p. 263). To achieve this, human factors research needs to be domain-specific (for large-scale bushfires), and the findings presented in a format that can be easily understood by end-users and integrated into fire training progams (i.e., by training managers, incident management teams, career and volunteer bushfire fighters).

Method

Metacognition refers to cognition about other cognitions (Cohen, Freeman & Wolf, 1996), and Pressley (2002) recommended the use of think-aloud protocols to obtain qualitative data (such as descriptions of metacognitive knowledge), and grounded theories (such as methods of constant comparison) to identify patterns or regularities in behaviours and thoughts. To this end, we used human factors interviews on the fireground, and visual-cued recall debriefs during command post simulation experiments, to explore the metacognitive skills of career and volunteer bushfire fighters. The Central Mountain Fire studies are described in more detail at Frye & Wearing (2011).

The samples for these studies were small (n = 4 on the fireground, n = 2 in a repeated measures simulation experiment, and n = 4 in a high cognitive load simulation rated by expert observers). Nonetheless, they had a high degree of ecological validity (Frye & Wearing, 2011, p. 40), and provided a rich source of qualitative data about how experienced domain practitioners (>10 years experience) make decisions during large-scale bushfires. Two participants were also rated as experts, which means that we were able to compare the metacognitive skills of expert fireground commanders (who also reported feeling cognitively in control), with their peers (who described feeling frequently overloaded in high cognitive load conditions).

Cognitive Errors

In the high cognitive load experiments (Frye & Wearing, 2011, p. 36), experienced fireground commanders were susceptible to the same types of cognitive errors that are observed with trained novices, such as:

1. Focussing on what is happening in front of them, but losing sight of the bigger picture (or vice versa).
2. Focussing on what is happening right now, but losing sight of what might happen next (or vice versa).
3. Focussing on situational awareness, but leaving it too late to make decisions (or rapidly making decisions with inadequate situational awareness).
4. Persisting with a goal, but failing to change plans when the situation changes (or failing to establish goals and priorities).

5. Accepting responsibility, but micromanaging or failing to escalate issues to others (trained novices who lack confidence may also defer too many responsibilities to others).
6. Focusing on safety, but focussing so much on one safety issue that they lose sight of another (whereas trained novices may also overlook safety, while pursuing higher order goals).

These errors represent different types of *tunnel vision*, in that focussing (or fixating) on one aspect of the bushfire results in a blind spot to other aspects of the situation. Additional training about policies and procedures is unlikely to help, because instead, the situation requires cognitive regulation skills.

Cognitive Overload

For example, in one of our simulation studies an experienced fireground commander forgot to deploy a fire tanker to protect a bulldozer. He was not oblivious to safety issues (or the standard operating procedure), in fact he was so pre-occupied with the safety of one fire crew (facing a burnover situation) that he lost sight of the big picture, and therefore the safety of another fire crew (also facing a burnover situation). In this respect, he became cognitively overloaded and susceptible to errors.

> *Quote 1*
>
> *'I got so focussed on one (tanker) that I lost focus on the other, and I end up having to go through the same scenario again with them. Can you get out? Are you safe?...this is when I realised, Bugger!...I'd missed one of our cardinal rules…I didn't have a fire unit with the dozer…and I need to be stepping back again, to what's coming at them around the corner, but my big (problem) was the speed of the way things were happening at the incident…we should have faith in the crew leaders to (manage safety), but at the end of the day, in this position, you're responsible for that safety.'*

These types of conflicting cognitive demands are common during large-scale bushfires, and while standard operating procedures (SOPs) can help, they do not improve safety (or productivity) unless practitioners can successfully implement them in real-world situations. In this case (Quote 1) the cognitive demands of implementing multiple procedures exceeded the cognitive capacity of the operator, and he became overloaded.

Expertise and Cognitive Control

In contrast, expert fireground commanders described feeling in control during high cognitive load conditions, and they articulated a high degree of self awareness (metacognitive awareness):

> *Quote 2*
>
> *'It's a recognition of your capability, it's your own self…inner recognition of…this is what I'm capable of. It's just years of experience, you've just got to know…this is what I can do, this is what I'm capable of, but it's the…the trigger to say, flick the switch, no, this is going to go beyond me...'*

In this case, the expert fireground commander gave the (confederate) incident controller an early *'heads up'* that things were going to get beyond his span of control. Then, once the Incident Controller acknowledged the message and assigned that area of the fire to someone else, the participant '*switched off, and didn't worry about that anymore, and got back to concentrating, and prioritizing what was in my own span of control'*. In this way, he maintained cognitive control while other participants described: '*fighting a losing battle*', persisting with strategies that they described as *futile*, and experiencing cognitive overload.

Recognition Primed Decisions

Experienced fireground commanders also described making decisions automatically based on previous experiences, which is consistent with a Recognition Primed Decision model (RPD; Klein, 1999). However, while this was an effective strategy for reducing cognitive load, experts also cautioned that pattern matching could sometimes lead to errors:

> *Quote 3*
>
> *'...automatically making decisions, it just becomes automated, you just do it. I suppose I don't even think about it. Too automated sometimes (because) you do it naturally and (if) you don't question yourself you can fall into a trap…you can miss something...'*

In this respect, the participants' descriptions are consistent with previous research (see Cohen, Freeman & Wolf, 1996; and Kahneman & Klein, 2009), and highlight the need for judgement and expertise in fireground command, as well as the successful implementation of practiced rule-based procedures.

A MODEL OF METACOGNITION FOR BUSHFIRE FIGHTERS

These examples (Quotes 1, 2 & 3) show that fireground commanders need to regulate their cognitive activity during large-scale bushfires to maintain situational awareness (and notice anomalies), make decisions, and manage dynamic activity, as shown in Figure 1.

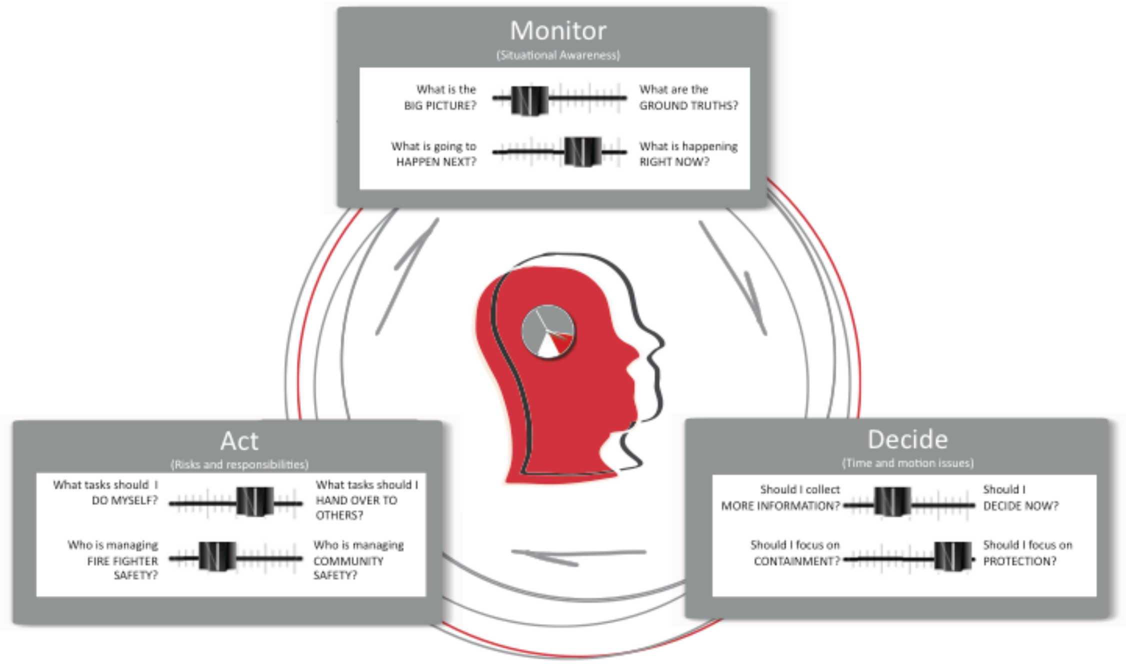

Figure 1. A model of metacognition for bushfire fighters

Fireground commanders also described using feedback from previous decisions and activity to build situational awareness and inform subsequent decisions, consistent with a metacognitive loop (Figure 1, see also Anderson, Oats, Chong & Perlis, 2006). Their descriptions were consistent and repetitive, and thus provide useful examples for training bushfire fighters about cognitive regulations skills. In particular, experienced fireground commanders described switching the focus of their attention to regulate their cognitive performance within particular boundary conditions (or tolerance limits; Hoffman & Woods, 2011; Valot, 2002).

Monitor (Situational Awareness)

For example, experienced fireground commanders gained (and maintained) situational awareness by monitoring different perspectives (big picture and ground truths), and different temporal distances (what will happen next versus what is happening right now). They adjusted the focus of their attention regularly, and seemed to be aware of potential blind spots. They communicated with other people in the chain of command to address gaps in situational awareness, and to update their own and other peoples' picture of the situation. In this respect, their descriptions are consistent with previous research (such as Cohen, Freeman & Wolf, 1996; Endsley, 1995; Hoffman & Woods, 2011; Klein, 1999).

Decide (Time and motion issues)

Similarly, expert fireground commanders regulated the tempo of their decision-making (balancing speed with accuracy), and knew when to change goals. For example, they described clear *trigger points* for shifting from an offensive goal (such as containing the bushfire) to a defensive goal (such as protecting life and property), and they quickly resorted to prepared fallback options when initial plans failed. In contrast, non-experts persisted with goals that they described as futile (often describing overconfidence and optimism biases), found the situation getting ahead of them, and implemented poor strategies.

Act (Risks & Responsibilities)

Finally, expert fireground commanders knew the limits of their own cognitive ability (see Quote 2), and described personal performance boundaries that signalled a need for change. Consequently, they did not attempt to manage everything themselves, but managed the distribution of cognitive load between themselves and other people in the chain of command (including both subordinate and superior positions; consistent with Valot, 2002). In this way, experts delegated (or escalated) risks and responsibilities to levels where they can be appropriately managed, and maintained cognitive control (consistent with McLennan, Pavlou & Omodei, 2005).

CONCLUSIONS

The current research shows that training and experience are necessary, but not sufficient conditions for developing expertise in fireground command (consistent with McLennan, Pavlou & Omodei, 2005). Indeed, many experienced fireground commanders are susceptible to the same types of cognitive errors (and blind spots) as trained novices, particularly under high cognitive load conditions. However, expert fireground commanders use metacognitive knowledge based on previous experience to monitor, decide, and act (a metacognitive loop), which enables them to keep their thinking on track and adapt to changing conditions (consistent with Cohen, Freeman & Wolf, 1996; and Anderson, Oats, Chong & Perlis, 2006). In these studies, expert fireground commanders also described higher levels of self-awareness (metacognitive awareness) than their poorer performing peers, consistent with earlier research (McLennan et al.).

Recommendations

For this reason, we recommend that fire agencies encourage deliberate practice for large-scale bushfires, so that practitioners develop the necessary metacognitive knowledge (and cognitive regulation skills) to:

- Change perspectives (e.g., between the big picture and ground truths).
- Employ anticipatory thinking (e.g., what might happen next, worst case scenarios).
- Regulate operational tempo (e.g., tradeoff between mistakes, and missed opportunities).
- Recognise trigger points for changing goals (e.g., offensive versus defensive goals).
- Recognise personal performance limits, and manage distribution of cognitive load.
- Manage risks and responsibilities appropriately within a chain of command.
- Apply rule-based procedures effectively, including watchouts and safety drills.
- Identify potential blind spots and errors in situational awareness, and develop mitigation techniques.

Furthermore, because expertise involves mental simulations (e.g., predicting likely and worst-case scenarios; Klein, 1999), we recommend development activities that are ecologically valid for the bushfire context, such as: case studies, scenarios, simulations, staff-rides, exercises, and coaching on the job (e.g., Fadde & Klein, 2011; Frye, 2012; Slijepcevic et al.). These types of activities would encourage the application of technical knowledge (including rule-based and standard operating procedures), whilst also developing the metacognitive knowledge (and cognitive regulation skills) required for work in macrocognitive environments. A significant goal should be to reduce cognitive overload, and therefore human errors, during large-scale bushfires in Australia.

Future Research

The Australian Bushfire Cooperative Research Centre has collected a large number of fireground interviews in the last 10 years. For example, the Human Factors Interview Protocol (HFIP; Omodei, McLennan & Reynolds, 2005) was used to conduct 120 interviews with bushfire fighters during the 2003-2006 bushfire seasons. A Bushfire Research Taskforce also conducted 600 interviews with survivors of the 2009 Black Saturday bushfires (Bushfire CRC, 2010) and similar taskforces interviewed residents affected by the Perth Hills bushfires (2011), and the recent Tasmanian and New South Wales bushfires (2013). These interviews could be used to further test the model described in the current research (Figure 1), and to validate, modify, or refute it's application for training about large-scale bushfires in Australia. Further analysis with larger samples would also highlight examples of fireground decision-making that can be used in fire training (e.g., as case studies, scenarios, staff rides, or exercises). With this in mind, the authors used the model (Figure 1) to analyse how survivors (bushfire fighters and residents) of a small rural community made decisions during the 2009 Victorian Black Saturday bushfires (Frye & Wearing, *in preparation*).

ACKNOWLEDGEMENTS

Richard Jones (of the firm One Eighty Seven & a Half) produced the graphic design of this model for our use in bushfire training.

REFERENCES

Adams, R., & Ericsson, A. (2000). Introduction to cognitive processes of expert pilots. *Human Performance in Extreme Environments*, 5 (1), 44-62.

Anderson, M., Oates, T., Chong, W., & Perlis, D. (2006). The metacognitive loop I: Enhancing reinforcement learning with metacognitive monitoring and control for improved perturbation tolerance. *Journal of Experimental and Theoretical Artificial Intelligence*, 18 (3), 387 – 411.

Bushfire CRC. (2010). *Research Task Force - 2009 Black Saturday Bushfires*. Melbourne: Australian Bushfire Cooperative Research Centre. Retrieved 12/11/2012, http://www.bushfirecrc.com.au.

Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Metarecognition in time-stressed decision making: recognizing, critiquing, and correcting. *Human Factors*, 38(2), 206-214.

Connolly, T. (1999). Action as a Fast and Frugal Heuristic. *Minds and Machines 9*, 479–496.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), The Cambridge Handbook of Expertise and Expert Performance* (p. 39-68). Cambridge: Cambridge University Press.

Fadde, P. J., & Klein, G. A. (2010). Deliberate performance: Accelerating expertise in natural settings. *Performance Improvement*, 49(9), 5-14.

Frye, L. (2012). A multi-agency Incident Leadership Framework. *Proceedings of the 3rd Human Dimensions in Wildland Fire Conference*, Seattle, 17-22 April 2012.

Frye, L., & Wearing, A. (2011). The Central Mountain Fire Project: Achieving Cognitive Control during bushfire response. *Journal of Cognitive Technology*, 16 (2), 33-44.

Frye, L., & Wearing, A. (*in preparation*). A model of metacognition for community bushfire response. *In Hoffrage, U. & Markewski, J. (Eds), Special issue of Modelling and Aiding Intuitions in Organisational Decision Making, to be published by Journal of Applied Research in Memory and Cognition*, San Diego, CA.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103 (4), 650-669.

Hoffman, R. R., & Woods, D. D. (2011). Beyond Simon's Slice: Five Fundamental Trade-Offs that Bound the Performance of Macrocognitive Work Systems. *IEEE Intelligent Systems*, 26 (6), 67-71.

Johnstone, G. (2002). Report of the Investigation and Inquests Into a Wildfire and the Deaths of Five Firefighters at Linton on 2 December 1998: State Coroner's Office, Victoria.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64 (6), 515-526.

Klein, G. (1999). The recognition primed decision model. In *Sources of Power*: How People Make Decisions (pp. 15-30). Cambridge, MA: The MIT Press.

Klein, G., Calderwood, R. & Clinton-Cirocco, A. (1988). Rapid Decision Making on the Fire Ground. Technical Report 796, US Army Research Institute for the Behavioural and Social Sciences.

Klein, G., Ross, K., Moon, B., Klein, D., Hoffman, R., & Hollnagel, E. (2003). Macrocognition. IEEE Intelligent Systems, 18 (3), 81-85.

Leonard, H. B. (2010). Organizing Response to Extreme Emergencies, Submission to the 2009 Victorian Bushfires Royal Commission. Doc. ID: EXP.3031.001.0018.

McLennan, J., Pavlou, O., & Omodie, M. (2005). Cognitive Control Processes Discriminate Between Better versus Poorer Performance by Fire Ground Commanders. In H. L. Montgomery, R., Brehmer, B. (Ed.), *How professionals make decisions* (pp. 209-221). Mahwah, NJ: Lawrence Erlbaum Associates.

Militello, L., Dominquez, C., Lintern, G. & Klein, G. (2010). The Role of Cognitive Systems Engineering in the Systems Engineering Design Process. *Systems Engineering*, 13 (3), 261-273.

Omodei, M., McLennan, J., & Reynolds, C. (2005). Identifying the causes of unsafe firefighting decisions: A human factors interview protocol (Bushfire CRC Project D2.3 Safety in Decision Making and Behaviour, Tech. Rep. No 1). Melbourne: Australian Bushfire Cooperative Research Centre. Retrieved 13/11/2012, http://www.bushfirecrc.com.au.

Owen, C., Stack, S., Whight, S., Pyrke, A., Duggan, P., & Staier, E. (2010). *Designing the staff ride: A vehicle for learning from wildfire and prescribed burning operations in Australia.* Report commissioned by the Tasmanian Parks and Wildlife Service. Melbourne, VIC: Bushfire Cooperative Research Centre. http://www.bushfirecrc.com/sites/default/files/managed/resource/appendix_12_d_staff_ride_manual.pdf

Pressley, M. (2000). Development of grounded theories of complex cognitive processing: Exhaustive within- and between-study analysis of think-aloud data. *In G. Schraw & J. Impara (Ed.), Issues in the measurement of metacognition* (pp. 261-296). Lincoln, NE: Buros Institute of Mental Measurements.

Putman, T. (1995). The collapse of decision making and organisational structure on storm king mountain. *Wildfire,* 4 (2), 40-45.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63, 129-138.

Slijepcevic, A., Haynes, J., Buckley, A., Swiney, M., Salter, L., Frye, L., & McHugh, P. (2012). Improving learning and development for joint agency incident management teams in Victoria. *Proceedings of the Bushfire Cooperative Research Centre & Australasian Fire and Emergency Services Authority Council Annual Conference*, Perth, 29-30 August 2012.

Stack, S., & Owen, C. (2012). *Evaluation of the 2012 Cobaw Staff Ride*, Report commissioned by the Department of Sustainability and Environment, Melbourne, VIC: Bushfire Cooperative Research Centre. http://www.bushfirecrc.com/sites/default/files/managed/resource/cobaw_staff_ride_final_report.pdf

Stretton, E. B. (1939). Transcript of evidence and Report of the Royal Commission into the causes of and measures taken to prevent the bushfires of January, 1939, and to protect life and property in the event of future bush fires. (Report No. 000002967). Melbourne: Parliament of Victoria.

t' Hart, P. (2010). Organizing for effective emergency management. Submission to the 2009 Victorian Bushfires Royal Commission. Doc. ID: EXP.3031.001.0001.

Teague, B., McLeod, R., & Pascoe, S. (2009). *2009 Victorian Bushfires Royal Commission - Interim Report* (No. 225 – Session 2006–09). Melbourne: Parliament of Victoria.

Teague, B., McLeod, R., & Pascoe, S. (2010). *2009 Victorian Bushfires Royal Commission* - Final Report Summary. Melbourne: Parliament of Victoria.

Valot (2003). An Ecological Approach to Metacognitive Regulation in the Adult. In P. Chambers, M. Izaute, P. Marescaux (Ed.), *Metacognition: Process, Function and Use* (pp135-152). Boston, MA: Kluwer Academic Publisher.

Safeguarding moral perception and responsibility via a partnership approach

Tjerk DE GREEF [a], Alex LEVERINGHAUS [b]
[a] *Interactive Intelligence Group, Delft University of Technology*
[b] *Oxford Institute for Ethics Law and Armed Conflict & Oxford Martin School, University of Oxford*

ABSTRACT

Introduction: New Combat Technologies (NCTs) create unprecedented capabilities to control the delivery of military force. But dependence on them generates difficult moral challenges impacting the decision-making process, which are only beginning to be addressed. **Method**: Some welcome NCTs explicitly whereas others plea to abandon such technologies. In order to arrive at an informed opinion about the impact of NCTs on decision-making, we need to know more about what NCTs are and how they operate. We provide a short overview of the different types of NCTs and discuss the challenges posed to moral perception. **Results and discussion:** Before NCTs can be deployed, we need to rest assured that their usage enhances, rather than undermines, human decision-making capacities. We propose a design recommendation based on a partnership approach, ensuring that operators use appropriate information in the decision-making process. There are important choices to be made, and sound design is 'design for responsibility'.

KEYWORDS

Sensemaking; military; drones; unmanned systems; moral perception; responsibility; partnership.

INTRODUCTION

One issue within the discourse over military interventions (MI) concerns the impact of new combat technologies (NCTs) on the human decision-making process. NCTs include military robots, such as unmanned aerial vehicles (UAV), computer-based targeting systems and missile defence systems. Scholars from different disciplines expressed concerns regarding implication on the decision-making process due to the distance from the battlefield and increased autonomy (Sharkey, 2010; Sparrow, 2007), potentially leading to the abdication of responsibility (Matthias, 2004). Dependence on NCTs thus generates moral and legal challenges impacting the quality of the decision-making process, which are only beginning to be addressed.

NCTs have already been used during MI. Commenting on NATO's intervention in Kosovo in 1999, Ignatieff speaks of 'virtual war', where NATO forces, supported by NCTs and more traditional air power, did the fighting, but only 'Serbs and Kosovars did the dying' (Ignatieff, 2000). NATO's service personnel, Ignatieff shows, were removed from the actual combat zones, but with the help of technology could carry out military missions. More recently, warships shelled targets in Libya to assist rebel fighters in the overthrow of the regime. Like Kosovo, intervening forces relied heavily on NCTs. Modern military technologies thus render 'boots on the ground' unnecessary.

Some scholars have welcomed explicitly the use of NCTs during MI (Altman & Wellman, 2009; Strawser, 2013). Firstly, for reasons of proportionality, they argue, the use of NCTs is desirable. Using UAVs leads to less damage and destruction than a large-scale military operation with boots on the ground. Secondly, western states are under pressure to minimize casualties amongst their own service personnel. Indeed, one of the morally and politically attractive features of NCTs is their ability to protect the lives of service personnel (Strawser, 2010).

There have also been critical voices. Ignatieff seems sceptical about the prospects of 'virtual war', while Sparrow argues that NCTs must not be deployed because they undermine a commitment to responsibility in the armed forces (Sparrow, 2007). If Sparrow's criticism is true, NCTs must not be deployed in *any* conflict.

However, in order to arrive at an informed opinion about the impact of NCTs on decision-making, we need to know more about what NCTs are and how they operate. Research in this area is in its infancy, but is likely to become more prominent. In this paper, we discuss how NCTs impact on the moral perception of those who operate them. The issue of moral perception, in fact, is crucial for a commitment to responsibility. We begin with some brief comments on moral perception. Subsequently, we give a short overview of the different types of NCTs and discuss some of the challenges they pose to moral perception. We continue by making design recommendations for NCTs. There are important choices to be made, and sound design is always 'design for responsibility' – or so we shall argue.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

MORAL PERCEPTION

The concept of moral perception is central to hold individuals responsible for their actions. Since the Nuremberg Trials, it also plays a major role in international law. Roughly, the concept of moral perception refers to the knowledge of the relevant facts in a particular situation. In holding individuals responsible for their actions, we assume that they have acted with knowledge of morally relevant facts. Conversely, in order to be exculpated from wrongdoing, an individual has to prove that he could not have acquired knowledge of the relevant facts. Since Nuremberg, combatants need to at least meet the moral perception criterion in order to be exculpated from wrongdoing (May, 2005).

However, in light of the rise of NCTs, the way in which soldiers acquire knowledge of relevant moral facts is transformed. As Ignatieff's idea of 'virtual war' implies, NCTs introduce (amongst other things) an element of distance. By removing soldiers from the combat zone, NCTs may well affect their moral perception of a particular situation, making it more difficult to, say, discriminate combatants from non-combatants. Some UAVs, for instance, transmit images they record with their sensors to a video screen where operators view them. UAVs remove their pilots from the actual battlefield physically, but in doing so it may restrict their moral perception of the situation. Depending on the context, it might be more difficult to hold operators of UAVs responsible for, say, applying force to a target. One potential reason for this is that, in case force is applied to the wrong target, operators could argue that, due to the restrictions imposed by the technology, they did not have full situational awareness and should therefore be exculpated from any wrongdoing.

This possibility leads to two immediate requirements. Firstly and from a more technologically oriented perspective, engineers designing military equipment must be sensitive to how different types of technology impact on moral perception. That is to say, they must take into account how psychological factors impact information processing and shape the perception of morally relevant facts. Secondly and from a more legally and normatively oriented perspective, NCTs must be designed in order to minimize any distortions or unnecessary restrictions of their operators' moral perception.

Overall, sound design must always be design that enhances, rather than undermines, the preconditions for individual responsibility. Ensuring this is, in our view, one of the central moral obligations of engineers and designers. Before we provide an indication what design for responsibility may look like, we provide an overview of the different types of NCTs.

NEW COMBAT TECHNOLOGIES

Within the broad category of NCTs, it is important to distinguish between two main classes:

1. Remote-controlled systems are being deployed by militaries already. Originally designed for reconnaissance and surveillance missions, UAVs are today capable of carrying a payload, which is why they are used for the morally and legally controversial practice of targeted killings. UAVs are tele-operated (they are controlled by using information provided by a video-link) from a large distance while the vehicle itself is airborne in an area of interest in or near a conflict zone.
2. The most controversial dynamic within the development of NCTs relates to what we call operationally autonomous weapons delivery systems. We define operational autonomy as the capacity of a machine to carry out specific tasks without assistance from an operator. In the future, UAVs, for instance, are expected to become increasingly operationally autonomous. This reflects a wider trend within the military, namely to reduce interaction between humans and machines.

Both NCTs have in common that they increase the distance to the battlefield. This has benefits as well as downsides. First, as we already noted, the possibility of death or serious injury amongst service personnel decreases greatly. Secondly, given that they do not face an immediate threat to their safety, the stress soldiers experience in combat is diminished. Arguably, stress affects moral perception because it influences how humans interpret their environment and frame certain issues. To illustrate the point, consider the infamous My Lai massacre that occurred during the Vietnam War. Fearing that the inhabitants of the hamlet of My Lai were Vietcong guerrillas posing as civilians, American soldiers experienced high levels of stress and, as a result, failed to apply the discrimination criterion accurately. This led to one of the worst massacres in history. A decrease in stress, then, might lead to greater awareness as well as more accurate interpretations of morally relevant facts. If this is true, the impact of NCTs on moral perception seems positive, rather than negative. That said, the reduction of stress could have negative effects. While it is correct that too much stress diminishes decision-making capacities, low levels of stress can have positive effects on an operator's alertness. It has, for example, been demonstrated that boring work conditions impact negatively on decision-making and performance (Endsley & Kiris, 1995).

There is another worry about the impact of distance on the moral perception. The quality of information that operators receive from machines may be lower than a human 'first hand' account. Operators, for instance, may receive too little information. Machines might not transmit certain information that a human would have picked up. This problem is particularly acute in operationally autonomous machines where the interaction between operator and machines is decreased. Depending on their task, operationally autonomous machines may process and filter large amounts of information themselves before passing on information to an operator. On the other hand, NCTs might supply operators with too much information, which may be more acute in tele-operated than operationally autonomous NCTs. Tele-operated UAVs have many sensors and remain airborne for long periods. Processing the amount of information may be difficult for a single operator. In both cases – the undersupply and oversupply of information – it becomes difficult for operators to filter out morally relevant facts. As a result, the deployment of NCTs may undermine a commitment to individual responsibility.

If these observations are accurate, the introduction of NCTs could go either way. It could increase or decrease the ability of individuals to acquire morally relevant facts. This is why different design options for NCTs are important.

PARTNERSHIP

How can we ensure that individuals perceive the relevant facts in a give situation? Contemporary literature shows two approaches in the development of NCTs. The first, defended by roboticist Ronald Arkin (2010), contends that NCTs should be given full operational autonomy in order to prevent war crimes. According to Arkin, automated systems lack emotional components that prompt humans to commit war crimes. Although a fair argument, taking the human completely out of the loop leads to underutilization of human experiences and capabilities. It remains a fact that specific humans capabilities are, up to date, hard to replicate by artificial intelligence. While, for example, computers are really strong in executing many calculations on large datasets, they lack creativity or a capability to recognize patterns that humans easily recognize. To do so, humans use a variety of psychological tools such as the knowledge-based reasoning mechanism (Rasmussen, 1986), which allows the human to cope with novel and unexpected situations by using fundamental knowledge (e.g. principles, physical laws) that governs the specific domain. Today's artificial intelligence technologies fail to model precisely the set of tools that allows being creative and recognizing patterns, leading to a fundamental asymmetry between human and artificial agents. It is, for instance, difficult to see how a machine could interpret human behaviour to distinguish a combatant from non-combatant whereas military operators recognize patterns-of-life to make such discriminations. Arkin is thus right to point out that humans are bad at decision-making and interpreting complex information under stressful conditions but machines presently lack the reasoning capacities that allow interpreting complex situations.

Faced with this problem, it is a common sense response to 'team-up' humans and artificial agents by integrating them into a *joint cognitive system* (Hollnagel & Woods, 2005). The partnership approach represents the second approach in the development of NTCs, which allows utilizing the strengths of one to compensate for weaknesses of the other. This move potentially enhances situational understanding and subsequent decision-making, as well as moral perception in general. Just as human team members develop different perspectives on a situation, machines and humans may develop different perspectives on a situation. Operators can use the perspective provided by their machine to check if they are missing morally relevant facts. The machine may flag up aspects of a situation that the operator might have otherwise overlooked.

The partnership approach is not entirely new (cf. de Greef, Arciszewski, & Neerincx, 2010; Johnson et al., 2011). The latest generation of partnerships is facilitated by *working agreements* (Arciszewski, de Greef, & van Delft, 2009) leading to a fine-grained division of labour between the human and machine with regard to specific tasks; this allows the human to stay in firm control of those tasks and objects that are regarded important. A human may, for example, delegate identifying unambiguous airplanes to the artificial agent while remaining in control of the more cognitive demanding ambiguous objects that potentially are legitimate targets.

The concept of working agreements were validated recently (de Greef et al., 2010). The eight navy officers who participated highly appreciated the division of labour between human and artificial agents introduced by the working agreement, especially when decisions had to be made under pressure. In this experiment, a partnership prototype was compared to a more static version resembling today's combat management workstations aboard navy frigates. The officers liked the working agreements and were relieved that they could focus on the more demanding bits while having the machine do the regular 'low risk' easy bits. The performance effects revealed a clear preference towards the partnership approach in that the identification times (a measured variable) increased in general 60%.

In light of these findings, the effect of the partnership approach on moral perception is potentially positive. Firstly, as the study shows, the partnership increases efficiency via a better division of labour, lowering the stress while keeping the human in the decision-making process. Secondly, working agreements allow for better information management. If there's a danger that the operator will not be able to filter out morally relevant facts because he either receives too much or too little information, it needs to be ensured that he gets the right amount of information. As part of the working agreement, the operator determines which information is provided and how it is managed in the decision-making process.

The partnership approach protects a commitment to responsibility within the armed forces. First, operators will be responsible for their working agreements. This raises issues about foresight, negligence and so on that we cannot tackle here. For now, it suffices to note that the operator remains firmly control of his machine – even if there's a physical distance between them. Secondly, working agreements ensure that operators receive the morally relevant facts needed to make decisions that comply with international humanitarian law, as well as key moral principles.

CONCLUSION

We started by noting that some commentators have argued in favour of the deployment of NCTs during MI. Indeed, there are some benefits associated with NCTs. But the difficulties posed by these systems must not be neglected. Before NCTs can be deployed, we need to rest assured that their usage is safe and that they enhance, rather than undermine, human decision-making capacities. During MI, the operational requirements upon NCTs and those who operate them are therefore high. Ordinary operators and their superiors need reliable information about the complex environment they operate in, especially when they are not directly present. We propose a design recommendations based on the partnership analogy, ensuring that operators use appropriate information in the decision-making process. There are important choices to be made, and sound design is 'design for responsibility'. Admittedly, the partnership approach is no magic formula and we doubt that there is such a formula. However, partnerships are a promising way forward, especially when compared to proposals for fully operationally autonomous machines.

ACKNOWLEDGEMENTS

This work is part of the research project 313-99-260 'Military Human Enhancement: design for responsibility in combat systems' that is financed by the Netherlands Organization for Scientific Research (NWO) in the research program line Societal Responsible Innovation. Parts of this publication have been also been published in the Whitehall Report "Hitting the Target?" published by the Royal United Services Institute (RUSI).

REFERENCES

Altman, A., & Wellman, C. H. (2009). *A Liberal Theory of International Justice*. Oxford University Press.

Arciszewski, H. F. R., de Greef, T. E., & van Delft, J. H. (2009). Adaptive Automation in a Naval Combat Management System. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, *39*(6), 1188–1199.

Arkin, R. (2010). The Case for Ethical Autonomy in Unmanned Systems. *Journal of Military Ethics*, *9*(4), 332–341.

De Greef, T. E., Arciszewski, H. F. R., & Neerincx, M. A. (2010). Adaptive Automation Based on an Object-Oriented Task Model: Implementation and Evaluation in a Realistic C2 Environment. *Journal of Cognitive Engineering and Decision Making*, *31*, 152–182.

Endsley, M., & Kiris, E. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors*, 381–394.

Hollnagel, E., & Woods, D. D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. Boca Raton: Taylor and Francis.

Ignatieff, M. (2000). *Virtual Wars: Kosovo and beyond*. London: Vintage.

Johnson, M., Bradshaw, J. M., Feltovich, P., Jonker.C., Van Riemsdijk, B., & Sierhuis, M. (2011). The Fundamental Principle of Coactive Design: Interdependence Must Shape Autonomy. In M. De Vos, N. Fornara, J. Pitt, & G. Vouros (Eds.), *Coordination, Organizations, Institutions, and Norms in Agent Systems VI* (Vol. VI, pp. 172–191). Berlin / Heidelberg: Springer.

Matthias, A. (2004). The responsibility gap. Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, (6), 175–183.

May, L. (2005). *Crimes Against Humanity: A Normative Account*. *European Journal of International Law* (Vol. 18, p. 310). Cambridge University Press.

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

Sharkey, N. (2010). Saying "No!" to Lethal Autonomous Targeting. *Journal of Military Ethics*, *9*(4), 369–383.

Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy*, *24*(1), 62–77.

Strawser, B. J. (2010). Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles. *Journal of Military Ethics*, *9*(4), 342–368.

Strawser, B. J. (2013). *Killing by Remote: Ethics of an unmanned military*. Oxford: Oxford University Press.

Investigate Naturalistic Decision-Making of Football Players in Virtual Environment: Influence of Viewpoints in Recognition

Yohann CARDIN[a,c], Cyril BOSSARD[a,c], Cédric BUCHE[b,c] and Gilles KERMARREC[a,c]

[a]*Centre de Recherche sur l'Éducation et la Didactique (EA 3875), Université Européenne de Bretagne, U.F.R Sport et EP de Brest*

[b]*Laboratoire en Sciences et Techniques de l'Information, de la Communication et de la Connaissance (CNRS-UMR 3192), École Nationale d'Ingénieurs de Brest*

[c]*Centre Européen de Réalité Virtuelle*

ABSTRACT

Introduction: The study of underlying processes of decision-making in dynamic situation, whether in work or in sport, is essential to the development of training tools. Virtual simulations are both key tools to study these processes and training. **Method**: Our work consisted in analysing the players' naturalistic decision-making in the virtual simulator CoPeFoot and the influence that changes of viewpoint can have on it. Behavioural data were recorded from six players in two different views (immersive and external), supplemented by verbal data collected during self-confrontation interviews. **Results:** A content analysis of the data revealed that in situations with strict time constraints, the players, to make decision, used twenty four schemata which facilitated the recognition of game situations. **Discussion:** These results points to the dynamic aspect of decision-making activity in the simulator and the consistency with the findings of studies in natural situations and the homogeneity for immersive and external views.

KEYWORDS

Decision making; dynamic situation; viewpoint; virtual reality; football.

INTRODUCTION

New technologies have always been an essential asset for researchers in sport science and have recently led them to question the value of simulation for research and training in high-level sport (Bossard, Kermarrec, Bénard, De Loor & Tisseau, 2009a). In this perspective, the virtual reality is now a scientific and technical field which exploits computing and behavioural interfaces to simulate a virtual world behaviour of entities in 3D real-time interaction among themselves and with one or several users in immersion. One of the main innovations allowed is in the development of participatory simulations that highlight the coupling between a user and the computer system. This coupling between individual and environment is also one of the theoretical principle of Naturalistic Decision Making paradigm (Klein, 2008), derived from the ergonomic psychology, and which allows the study of intuitive decision making in dynamic situation, i.e. in a unpredictable environment and where the agents are under strong time pressure. The naturalistic or intuitive decision making is defined generally as a complex cognitive process for the selection of an action or series of actions. In the context of high performance sport, especially team sports, the protagonists are often subjected to many pressures, including time that have a strong influence on their behaviour in the game. Quality of decisions is thus seen as the ability of an athlete to act at any moment of the game quickly and efficiently.

To understand this phenomenon, many studies have been developed in recent years (for a review see Bossard & Kermarrec, 2011) highlighting two different and complementary epistemological perspectives : cognitive and naturalists approaches. However, the cognitive approach, both the oldest and most predominantly adopted by the researchers does not take into account the actual context of the activity. This lack of consideration for the environment and the resulting discrepancy between what is experienced in an experimental situation and what is actually "really" lived in a natural situation has led some researchers to understand the context in which the actors are immersed during their activity. Thus the naturalistic approach, whose main characteristic is to study the complexity of human activity in the individual-environment coupling, appears to be best suited to the dynamic situations that concern us in this case.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

NATURALISTIC DECISION MAKING

The NDM approach has set the aim to improve support systems for decision-making in the military field but also in the nuclear industry and the field of civil aviation. It examines how experts, working alone or in groups in uncertain dynamic environments, identify and evaluate situations, make decisions and perform actions whose consequences are meaningful to them and their environment (Lipshitz, Klein, Orasanu & Salas, 2001). In this context, a dynamic situation is characterized by evolution, uncertainty and time pressure (Hoc, 2001) that imposes on a group of agents that interact to achieve a identified common goal (Gutwin & Greenberg, 2004). Sport situations being also dynamic situations *par excellence*, the parallel with social, professional or training situations that meets these criteria is no longer needed (Fiore & Salas, 2006). Indeed, in a team, the roles are often progressive and the majority of game situations requires the player to make decisions under time pressure.
Within the NDM approach, the Recognition-Primed Decision (RPD) model developed by Klein (1997, 2008) is particularly suited to analyse the player-situation coupling and intuitive decision-making processes.

Recognition-Primed Decision model

Klein and Brezovic (1986) refute the idea that individuals, confronted with dynamic situations, base their choice on a rational calculation or an exhaustive utility analysis (game theory and formal theory of decision). More specifically, the evaluation of the situation in course of action is based on the recognition of significant spatio-temporal configurations from their own experience. In a dynamic situation, the recognition process is based on the mobilization of a "cognitive package" that combines four types of secondary by-products: expectancies, cues, actions from experience and goals (Ross, Schafer & Klein, 2006). The "glance" of the expert in course of action, is an implicit matching between relevant cues perceived and functional structures available in memory.
Additionally, the RPD model proposed by Klein (1997, 2008) distinguishes three recognition processes used by experts faced with a dynamic situation. The "Simple match" that can be considered as a reactive process whereby the expert recognizes a situation already encountered and directly associates an adequate course of action. The second one, "Diagnose the situation" which is a process of diagnosis by comparing relevant cues from the situation encountered with several similar previously experienced situations to choose among known answers and implement an appropriate course of action. And finally, the third one "Evaluate a course of action" when the expert develops a new solution in the course of action and evaluates by a process of mental simulation in order to imagine how his actions could be integrated into the current situation.
However, the RPD model is subject to several criticisms especially with certain limits on unusual and complex events or by not explaining clearly how the "cognitive package" activated by a subject update over the situational dynamics. Thus, Ward, Williams and Ericsson (2012) showed by three experiments realised in the field of soccer that about prediction about the other players the best performances were supported by a situation model-type mechanism as proposed by long-term working memory theory rather than RPD model for example.
Despite of this, the RPD model has been validated in the field of sport in experimental conditions in handball (Johnson & Raab, 2003) and in real conditions of training or of competitive settings respectively in football, ice hockey and volleyball (Bossard, Kermarrec, De Keukelaere, Pasco & Tisseau, 2011; Bossard, De Keukelaere, Cormier, Pasco & Kermarrec, 2010; Macquet, 2009).
Thus, Macquet (2009) used RPD model to analyse the expert players' decision-making in a real volleyball game. The results show that experts in a volleyball competitive situation invest primarily the first modality of the RPD model ("simple match") to make decisions. Bossard et al. (2010, 2011) in the context of football and ice hockey, have subsequently confirmed that the time pressure exerted on the players in counter-attack situation forced them to mobilize primarily this first process. Additionally, these recent investigations under natural conditions showed that the constitutive categories of "cognitive package" of expert football or hockey players covered the four initial by-products RPD model and could be supplemented by a fifth called "knowledge".
To recognize a situation as typical and respond to it quickly, experts would have structures to maintain meaningful and effective action potentials. This hypothesis refers to the idea that people store and organize information from past experiences in abstract forms, i.e. schemata.

The schema theory

The concept of schema has been proposed in cognitive psychology ergonomic to study jointly the role of cognitive structures involved in an adaptive process and the role of contexts that affect their implementation (Anderson, Matessa, & Lebiere, 1997). Studies on naturalistic decision making suggest that the experience of the practitioner is an important factor (Klein, 1997), particularly in determining the decision-making schemata that an expert use in a situation. Researchers generally conclude that real-world decision-making is strongly schema-driven (Lipshitz & Shaul, 1997). They are reused to make quick decisions in new, similar or identical situations (Rumelhart, 1980). Schemata allow experts to interpret a situation as a whole and thus to make decisions by categorizing it efficiently as a whole pattern (Federico, 1995).
In the recent application of RPD model in sport (Macquet, 2009), the results clearly illustrate the assessment of the situation. The volleyball experts players mobilized a decision process that relied on both the evaluation of the situation and the choice of action.

The schema theory applied by Bossard et al. (2010, 2011) in recent studies on the naturalistic decision-making, respectively of expert football and hockey players during fast-breaks, highlights the role and adaptation of these background structures. The authors show the activation and permanent adaptation of schemata depending on the context. To choose the right action to perform, the main prerequisite is then the matching between the contextual invariants from the situation and the background invariants which enable to act: schemata. Additionally, when members of a group (or sports team) share experiences, they build similar schemata, which leads them to respond similarly in prototypical situations within the reference field (Piergorsch, Watkins, Piegorsch, Reininger, Corwin & Valois, 2006). Thus, schemata are reference structures, characteristics of an area of expertise.

APPLICATION OF THE VIRTUAL SIMULATIONS

In the trend of new technologies, the field of virtual reality can enable an individual to immerse himself in a completely virtual world. As part of these participatory simulations, are taken into account not only the influence that the environment has on the individual (simple simulation) but also the influence that the individual has on the virtual environment. And it is precisely this mutual influence individual-environment that is at the heart of the naturalistic approach. This is called the “co-evolution principle” which can then be connected to the notion of implicit learning (Bossard, Kermarrec, Bénard, De Loor & Tisseau, 2009b).

Several simulations have emerged and have thus been used in various studies involving physical activity and sports. In this context, three types of simulators have been created: for studying or practicing a technical gesture, for analysing strategies in sports situations and for immersing the user in sport environments. The main inconvenience in all these sports simulations lies in the fact that all of them, whatever their type, fail to really involve the three criteria which are autonomy, scalability and interaction. This is the question which was at the basis of the CoPeFoot simulator design in order to make it a participatory simulator allowing to reproduce credibly the decision-making in collaborative and dynamic situations (football counter-attacks) and combining the three indicators listed above (Bossard et al., 2009a). It is on this football simulator, whose the design model of virtual agents is the result of an analytical work on the activities of real football players during a practice (Bossard et al., 2011), that our study rested.

CHANGES OF SCENE PERSPECTIVE AND DECISION-MAKING

A number of studies have looked into the relationship in the field of sport between the viewpoint adopted by expert players and decision-making (Petit & Ripoll, 2008; Williams, Ward, Ward & Smeeton, 2008; Farrow, 2007). The majority of them was to ask participants about the decisions they would take by watching some pictures or films showing different game situations under standardized conditions. Thus, Petit and Ripoll (2008) for example, made a study based on the presentation of two video sequences in external and immersive views for two football players groups (experts and novices). The players had to make the choice to pass the ball or not to answer to the game situations and the results showed faster decisions for the experts and overall faster and more relevant decisions in immersive view than in external view. In the contrary, Farrow (2007) found water-polo players had superior decision-making accuracy with an aerial perspective relative to the player-view perspective. He explained that by the wider view of essential spatial information provided by the aerial view.

Beyond these constrasting results, this kind of study shows a large gap between what is perceived, experienced in an experimental situation and what is perceived, lived in a natural situation thus obscuring the real context (or natural) of the decision and the individual-environment coupling upon which the NDM approach rests. In this direction, the contribution of new technologies has enabled researchers to improve ecological conditions of experimental methods to get closer to natural situations. This is particularly the case of a virtual reality simulator as CoPeFoot which is a very interesting tool to conduct a qualitative study on the decision-making of football players based on the viewpoint they adopt on the game situation.

For this study, our work is based on three presuppositions: 1) decision-making is a recognition process which results of the association of the background elements and relevant cues identified in the context, 2) decision-making in dynamic situation can be described, explained and commented in continuous way by the actor exposed to the records of his own activity, 3) each experienced situation reflects the activation of a "schema", used as a benchmark to perceive and decide in course of action.

The purpose of this study is to highlight in the virtual simulator: a) the diversity of significant elements taken into account by the football players to decide quickly b) the types of recognition processes used under high time pressure c) the homogeneity of schemata activated by a group of experienced players, and d) the influence of viewpoint on football players decision-making in a virtual environment.

METHOD

Participants and procedure

The study was conducted in collaboration with six volunteers specialized in football, students from the University of Sport and Physical Education of Brest and playing all at regional level. They have been solicited

for their experience and knowledge in this sport. The average age of the players during the experiment was 21 years old and their average experience of practice was 13.8 years.
The six players were separated randomly into two teams which competed in a football match on the virtual simulator CoPeFoot. It is also important to note that players from the same team were together in the same room making it possible for verbal dialogue between them. The situation then set up on the virtual simulator consisted of a typical football training: a game of three versus three on a small field (30x40m) with a goal on both sides (1.50 m).
The experiment was conducted in four phases. First, the players participated in a phase of training session that lasted about 20 minutes during which players were able to familiarize themselves with the software commands. Then, the study strictly speaking consisted of three phases of 5 minutes during which the players competed on the network with different viewpoints. The first sequence was played in immersive view (view called "first person") during which participants adopted the viewpoint of the avatar they led (Image 1). The second sequence took place in external view, then all subjects with the same raised viewpoint allowing them to see the field as a whole (Image 2). Finally, during the third sequence, each player freely adopted one of these two viewpoints to play.

Image 1. Immersive view

Image 2. External view

Data collection

During the experiment, two types of data were collected. First, observational data corresponding to the virtual recording by the simulator of all the game sequences between the six players, allowing to replay at will all the actions carried out during the match.
Then, verbalization data collected during individual self-confrontation interviews (Theureau, 2010) conducted during about one hour at the end or the day following the experience on CoPeFoot. During these interviews, conducted by the same person, the investigator and the player watched together virtual recordings of the three game sequences following the own viewpoint of the player. This confrontation with virtual records of his activity aims to promote the recall of elements actually mobilized by the player during the game sequences studied. The researcher attempts to place the player in a posture and a mental state favourable for the explanation of his decisions through reminders on the sensations, perceptions, focalisations, concerns, emotions and thoughts that accompany each decision. During these interviews, the subject was then asked to describe and comment on his activity. The reminders focused on the actions that were significant (and therefore described and explained by the players) in the game sequences and on the events for which the researcher wanted to obtain additional information. Sharing a common sports culture between the researcher and the players has facilitated the understanding of the comments of the protagonists and has avoided the reminders leading to an explanatory style. This kind of interview is based on a true moral contract of cooperation between the players and the researcher.

Data analysis

Finally, the analysis of the data was carried out in five steps. At first, the data transcription that is to prepare them by linking the behaviours observed in the game recordings with the data obtained during self-confrontation interviews. The second step was to select and identify meaningful units, that is to say observed behaviours and passages or sentences pronounced by the player who gave information on his decisions during phases of play. To encode the meaningful units selected, the system of categories defined by the RPD model (Klein, 1997; 2008) and recently completed by Bossard et al. (2010, 2011) was used to classify the elements of players speech in five categories: plausible goals, expectancies, course of action, knowledge mobilized and relevant cues collected. Then, the task was to cut the stream of player activity to identify and distinguish the successive situations experienced by (and with the point of view of) each player. The fourth step was to make clusters of situations experienced in the same way by the same individual or several of them. Finally, this analysis ended by a validation process to ensure the validity of the data and which consisted of a "triangulation" process between two researchers familiar with the object of the study.

RESULTS

Coding and identification of meaningful units

The work of identifying and coding meaningful units (MU), corresponding to the game sequences, enabled us to count 1606 MU. According to the RPD model and its components, data analysis by theoretical categorization shows that these units are divided into the five theoretical categories expected: goal, action, relevant cue, expectation and knowledge. All these results are collected in Table 1.

Table 1. Distribution of Meaningful Units

Meaningful Units	Immersive view	External view	Total
Goals	161	109	270
Knowledge	71	46	117
Relevant cues	332	191	523
Actions	372	215	587
Expectations	71	38	109
Total	1007	599	1606

Identification of experienced situations and categorization of patterns

The cutting of each player's activity progress, taking into account the indications on the form and the meaning of the speech, we identified 424 experienced situations for all players interviewed during the study. Through a process of empirical categorization, these situations have been gathered together in 24 "schemata" that can be classified into three distinct groups: 6 in "Offensive phases with the ball" (the subject was in possession of the ball), 6 in "Offensive phases without the ball "(a team-mate was in possession of the ball) and finally 12 in " Defensive phases "(the opposing team had possession of the ball). Some schemata may be regarded as typical of expertise in football because they are activated by several players in various situations. Table 2 shows the number of typical schemata used in each type of game phase by all players with the two viewpoints adopted.

Table 2. Distribution of typical schemata classified into types of game phases

Types of game phase	Immersive view	External view	Total
Offensive phases with the ball	48	29	77
Offensive phases without the ball	86	48	134
Defensive phases	124	80	204
Total	258	157	415

Analysis of experienced situations and RPD model

The analysis of experienced situations, referring to the RPD model (Klein, 1997; 2008) showed that participants mobilized three levels of recognition process to decide in dynamic situation. Among the 424 experienced situations identified, 415 could be classified into one of the three modalities. Table 3 summarizes the distribution of recognition processes per viewpoint adopted. Table 4 summarizes the number of recognition processes per type of game phase.

Table 3. Distribution of recognition processes

Recognition processes	Immersive view	External view	Total
Mod 1 : Simple match	166	97	263
Mod 2 : Diagnose the situation	49	29	78
Mod 3 : Evaluate a course of action	43	31	74
Total	258	157	415

Table 4. Distribution of recognition processes per type of game phase

Types of game phase	Immersive view			External view			Total / Mod			Total
	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3	
Offensive phase with the ball	38	9	1	23	5	1	61	14	2	77
Offensive phase without the ball	48	13	25	21	14	13	69	27	38	134
Defensive phase	80	27	17	53	10	17	133	37	34	204

DESCRIPTION AND DISCUSSION OF RESULTS

Elements taken into account by the players to decide in action in CoPeFoot

The coding of the players' decision-making according to five by-products (expectations, goals, knowledge, actions and relevant cues) enables us to note that all these elements are involved in the decision-making of each individual. This study enables also to go into the content of these elements in depth. Goals correspond to the intentions of the player in action. Knowledge expressed by the players refers to their own activity, the strengths and weaknesses of the partners and opponents, the characteristics of the simulator (how to steal the ball into the opponent's feet, running speed with or without the ball) or general principle of the football game. Expectations mainly correspond to assumptions made by the players on the evolution of the current situation, verified or not later. The actions performed by the actors refer to their own movements but also those of the opponents and partners. Finally, the relevant cues detected from the context by players mainly concern the placement and movement of the partners and opponents. Thus, these results enable to establish the elements considered significant by subjects to decide during the simulation CoPeFoot.

In light of our results, we observe some similarities with the content of the categories proposed by other qualitative studies to describe decision-making in team sports (Bossard et al., 2010; Lenzen, Theunissen & Cloes, 2009; Macquet, 2009). Although the interpretation of the number of occurrences of the elements involved in the decision-making must be conducted carefully in a qualitative study, we observe a large proportion of data for the categories "perception" or "relevant cues" in these qualitative studies that are conducted either under natural conditions or in a virtual simulator. Expert players in team sports do not report all information collected by them to make a decision but they only call critical or significant signs in the situation (Kevin: "Player B is along the line, normally it's easier for a double team block at that moment"). These results support the idea that in dynamic situation, experts (athletes or not) devote more time to recognize the situation than compare various options for making a decision (Johnson & Raab, 2003; Klein, 2008). That seems to be the case in simulation too.

The results relating to the contents of the elements mobilized in and for action are in the continuity of those obtained in previous studies. They allow to enrich the establishment of decision-making model because they capture the characteristics of expertise in dynamic situation whether in natural or virtual environment.

Relationships between elements mobilized in action and the constraints imposed by CoPeFoot

In a general way, the results in Table 3 show that the decisions adopted by the players on the simulator primarily involve a process of “simple match” (263 of 415 cases, i.e. 63.37%) rather than processes of “diagnose the situation” (78 cases, 18.8%) or “evaluate the course of action” (74 cases, 17.83%). During game sequences, players have worked primarily on a reactive mode meaning that each decision was an implicit reaction to some signs (mainly the positioning and the movements of partners and opponents) perceived as significant in the same situation. This holistic assessment of the course of action (Lipshitz et al., 2001) is consistent with the results of studies conducted under natural conditions of work (Klein & Brezovic, 1986) or sports (Macquet, 2009).

The results in Table 4 highlight more specifically the relationship within the CoPeFoot simulator between the type of game phases played and the processes mobilized by the players. Thus, in offensive phase with ball, players have overwhelmingly used the “simple match” level (61 cases of 77, i.e. 79.22%). This is explained by the fact that being with the ball contributes to feel a strong time pressure from the opponents (who want to steal the ball quickly) but also from the team-mates and the experienced situation itself. Indeed, some situations may require to fast-forward, toward the target areas to take advantage of clear spaces and take opponents by surprise. Under this strong time pressure, the recognition process is thus more related to "course of actions" and "relevant cues" than the mobilization of general knowledge about the game or other players and assumptions on the game situation.

We can also notice that the process of assessment of the situation by mental simulation has been used very few times during the offensive phases with the ball (2 of 77 cases, 2.60%), and much more in defensive phases (34 of 204 cases, 16.67%) and especially in offensive phases without the ball (38 of 134 cases, 28.36%). This is explained by the fact that not being with the ball, the players do not feel as much time pressure and therefore consider that they have more time to make their decision. The verbalization of a greater amount of knowledge and above all expectations by the players during these phases "without ball" can also be explained by the uncertainty of the situations experienced during these game sequences that does not always favour a simple and quick recognition process. Indeed, when the ball is with another player (opponent or partner), the feeling of control over the result of the current action seems much lower than with the ball and therefore contributes, for players, to evaluate its possible evolution by mental simulation because less relevant cues are then recognized.

Time pressure, or more precisely the perceived urgency of the situation by the player, as well as the uncertainty concerning the evolution of the situation seems to be the main factors influencing the process of recognition in the CoPeFoot simulator. These results tend to confirm those noticed in previous qualitative studies (Bossard et al., 2010, 2011; Macquet, 2009) conducted in natural situation.

Learning, individual dynamic faced with the simulation

The particularity of this study of not being conducted in a natural situation highlights an important aspect of this experience which is the learning faced with the CoPeFoot simulator. Indeed during the study, the protagonists didn’t have more than 20 minutes of grip on the simulator, which does not exclude the fact that they continued this "learning" phase during the three steps of the experiment that followed.

Thus, our results show that among the five categories of elements identified, the "knowledge" expressed by the players is the only category steadily increasing over the three stages of the experiment (5.15% for the first step, 7.68% for the second and 8.24% for the third). This evolution can be explained by several aspects including especially the learning of the simulator features (e.g. Julien: "as the defenders were faster than the player who had the ball ..."). Indeed, while no acquired knowledge about the virtual simulator features was issued during the

first step conducted in immersive view, there were 7 during both phases of the experiment that followed (external view and view chosen by each player). This increase of using knowledge by players for their decision-making within the simulation can therefore be explained in part by the learning of the football simulator features. The players having, by practice, an amount of knowledge more important would be likely to use it most frequently to make decisions, influencing in this way the recognition process of the experienced situation. The verbalization of a greater amount of knowledge by the participants in these situations can be interpreted as the players' commitment to a process of seeking solutions among those available and acquired during the previous actions. The players beginning to know more typical situations then used their experience to decide quickly.
These results on the recognition process and meaningful units, even few in number, show us an important component to consider in order to interpret objectively the possible influence of the viewpoint adopted within the simulation on the players' decision-making.

Influence of viewpoint adopted on the decision-making

First of all, Table 1 shows a perfect homogeneity between the two viewpoints adopted for each of the five types of by-products identified during the interviews (chi-square test of homogeneity giving a p-value=0.977). Then, contrary to what might be expected, the external view, presenting yet more cues owing to a wider field of vision, does not cause a greater share of significant cues quoted by players to make their decisions. However, this confirms that whatever the viewpoint adopted, experienced players recognize situations quickly through the collection of only a few elements. These relevant cues picked up from external view then may not be the same as those identified in immersive view in the same type of action.
Then Table 2, displaying the distribution of types of typical schemata recorded shows once again a certain homogeneity observed between the two viewpoints (but not significant, p-value=0.813). A slight difference can just be notified about the proportion of typical schemata of “offensive phases without ball” which is lower in external view (30.57% to 33.33% for immersive view). And conversely, for the typical schemata of “defensive phases”, of which the proportion is greater for the external view than for the immersive view (50.96% to 48.06%). This slight difference could be explained by the fact that the external view allowed the players to raise more cues on the opposing team (positions and movements of the three opponents) then facilitating perhaps the activation of some typical schemata when they were in defence.
Finally, as previously with the typical schemata, concerning the recognition processes of typical situations during this experience, their proportion remains overall constant, but not significant (p-value=0.74), from one step to the next (Table 3). It just can still be noted once again, a slight difference about the recognition process “simple match” which is somewhat lower in external view than in immersive view (61.78% to 64.34%). The results confirm those obtained by Petit and Ripoll (2008). Indeed, for immersive view, experienced players seem to activate more frequently a recognition process “simple match” allowing a faster decision. This difference can also be directly connected to what has been previously analysed for the highest proportion of typical schemata of “defensive phases” recognized in external view.
However, if these results are relatively consistent between the two viewpoints, Table 4, detailing the recognition processes of game phases, shows a big difference. Indeed, the proportion of recognition processes during "offensive phases without the ball" seems to be much more variable from one view to another. Concerning this game phases, the players have much more used the diagnostic process in external view to recognize typical situations (29.2% to 15.11%) and in return, less recognition processes “simple match” (43.8% to 55.81%). This highlights the fact that in immersive view, when team-mate has the ball, the player would be more likely to quickly recognize a typical situation by a simple analysis of relevant cues. In external view, in the same situation, that player would be more likely to scan the various possibilities to make his decision. This can be explained by the greater number of elements that can be picked up by the player who, detecting more cues and having probably more possibilities, would tend to use the diagnosis process of the experienced situation.
Other work opportunities could complete these first results by another analysis more based on the concept of team cognition enabling to highlight the collective aspect involved in the two viewpoints available in CoPeFoot. In this way, the coordination models governing the collaborative activity within each team as well as the dynamic of contents shared by teammates could have been put forward (Bourbousson, Poizat, Saury & Sève, 2012). Nevertheless, despite of this lack, the sight difference between the viewpoints adopted, linked with our other results shared by other studies in natural situations (Bossard et al., 2011; Lenzen et al., 2009; Macquet, 2009) presents an interesting perspective concerning decision training on the CoPeFoot simulator. Thus, the two viewpoints could be interesting: the external view to train players to identify several alternatives by allowing them to have hindsight about the game situations and the immersive view for the training of the quick and simple recognition of game situations in order to decide quickly. A combination of both would be also possible with an experience in immersive view, then an analysing in external view and finally a return to the immersive view.

REFERENCES

Anderson, J. R., Matessa, M. & Lebiere, C. (1997) ACT-R: A theory of higher level cognition and its relation to visual attention, *Human Computer Interaction,* 12, 439-462.

Bossard, C., Kermarrec, G., Bénard, R., De Loor, P., & Tisseau, J. (2009a) Virtual Reality for Research and Training in Sport: an Illustration with CoPeFoot (Chap. 16), In C.H. Chang (Ed), *Handbook of Sports Psychology* (pp. 391-402), New York: Nova Science Publishers, Inc. ISBN: 978-1-60741-256-4.

Bossard, C., Kermarrec, G., Bénard, R., De Loor, P. & Tisseau, J. (2009b) An exploratory evaluation of virtual football player's believability, *In proceedings of 11th Virtual Reality International Conference – VRIC'09* (pp.171-172), Laval, France.

Bossard, C., De Keukelaere, C., Cormier, J., Pasco, D. & Kermarrec, G. (2010) L'activité décisionnelle en phase de contre-attaque en Hockey sur glace, *Activités*, 7(2), 41-61. http://www.activites.org/v7n1/v7n1.pdf

Bossard, C. & Kermarrec, G. (2011) La prise de décision des joueurs de sports collectifs. Une revue de question en psychologie du sport, *Science & Motricité,* 73, 3-22.

Bossard, C., Kermarrec, G., De Keukelaere, C., Pasco, D. & Tisseau, J. (2011), Analyser l'activité décisionnelle de joueurs de football en situation d'entraînement pour développer un modèle de joueur virtuel, *E-JRIEPS*, 23, 124-151.

Bourbousson, J., Poizat, G., Saury, J. & Sève, C. (2012) Temporal aspects of team cognition: a case study on concerns sharing within basketball, *Journal of Applied Sport Psychology*, 24 (2), 224-241.

Farrow, D.T. (2007) The effect of viewing perspective on decision-making performance in water-polo, *Journal of Sport & Exercise Psychology,* 29, 70-71.

Federico, P. A. (1995) Expert and novice recognition of similar situations, *Human Factors ,* 37 (1), 105-122.

Fiore, S. M. & Salas, E. (2006) Team cognition and expert teams: Developing insights from analysis of exceptional teams. *International Journal of Sport and Exercise Psychology*, 4(4), 369-375.

Gutwin, C. & Greenberg, S. (2004) The importance of awareness for team cognition in distributed collaboration. In E. Salas & S. M. Fiore (Eds.), *Team Cognition: Understanding the factors that Drive process and Performance* (pp. 177-202), Washington DC: APA Press.

Hoc, J. M. (2001) Toward ecological validity of research in cognitive ergonomics. *Theorical Issues in Ergonomics Science,* 2, 278-288.

Johnson, J. G., & Raab, M. (2003) Take the first: Option generation and resulting choices. *Organizational Behavior and Human Decision Processes,* 91, 215–229.

Klein, G. & Brezovic, C. (1986) Design engineers and the design process: Decision strategies and human factors literature. *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting*, 2, pp.771-775.

Klein, G. (1997) The Recognition-Primed Decision (RPD) model: looking back, looking forward, In C.E Zsambok & G.A. Klein (Eds.), *Naturalistic Decision Making* (pp.285-292), New Jersey: Lawrence Erlbaum Associates.

Klein, G. (2008) Naturalistic Decision Making, *Human Factors, 50* (3), 456-460.

Lenzen, B., Theunissen, C. & Cloes, M. (2009) Situated Analysis of Team Handball Players' Decisions: An Exploratory Study, *Journal of Teaching in Physical Education,* 28, 54-74.

Lipshitz, R. & Shaul, O. B. (1997) Schemata and mental models in recognition-primed decision making, In C.E Zsambok & G.A. Klein (Eds.), *Naturalistic Decision Making* (pp.293–304), New Jersey: Lawrence Erlbaum Associates.

Lipshitz, R., Klein, G., Orasanu, J. & Salas, E. (2001) Focus article: Taking stock of naturalistic decision making, *Journal of Behavioral Decision Making,* 14, 331-352.

Macquet, A. C. (2009) Recognition within the decision-making process: A case study of expert volleyball players, *Journal of Applied Sport Psychology,* 21, 64-79.

Petit, J.-P. & Ripoll, H. (2008) Scene perception and decision making in sport simulation : A masked priming investigation, *International Journal of Sport Psychology*, 39, 1-19.

Piegorsch, K. M., Watkins, K. W., Piegorsch, W. W., Reininger, B., Corwin, S. J., & Valois, R. F. (2006) Ergonomic decision-making: A conceptual framework for experienced practitioners from backgrounds in industrial engineering and physical therapy, *Applied Ergonomics,* 37, 587-598.

Ross, K. G., Schafer, J. L. & Klein, G. (2006) Professional judgments and "naturalistic decision making", In Ericsson, K. A., Charness, N., Hoffman, R. R. & Feltovich, P. J. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 403-419), Cambridge: Cambridge University Press.

Rumelhart, D. E. (1980) Schemata: The building blocks of cognition, In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading and comprehension: Perspectives from cognitive psychology, linguistic, artificial intelligence, and education* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Theureau J. (2010) Les entretiens d'autoconfrontation et de remise en situation par les traces matérielles et le programme de recherche 'cours d'action', *Revue d'Anthropologie des Connaissances, Volume 4, n° 2, 2010/2*, 287-322. http://www.cairn.info/revue-anthropologie-des-connaissances-2010-2.html

Ward, P. Ericsson, K. & Williams, A. (2012) Complex perceptual-cognitive expertise in a simulated task environment, *Journal of Cognitive Engineering and Decision Making*, First published on October 3, 2012 as doi:10.1177/1555343412461254.

Williams, A. M., Ward, P., Ward, J. & Smeeton, N.J. (2008) Task specificity, role, and anticipation skill in soccer, *Research Quaterly for Exercise and Sport, 79*, 428-433.

Expertise, Situation awareness and assessment

Thinking Inside the Box: The ShadowBox Method for Cognitive Skill Development

Gary KLEIN[a], Neil HINTZE[b], and David SAAB[b]
[a]*MacroCognition LLC*
[b]*Battalion Chief, FDNY*
[b]*MacroCognition LLC*

ABSTRACT

One way to help trainees develop expertise is to let them see the world through the eyes of experts. However, the tasks of gaining access to the expert's cognition and then of making experts available for training are daunting and impractical. Recently, however, Hintze (2008) developed a technique to allow trainees to shadow the thinking of experts. The trainees work through scenario-based materials, entering their information and decision priorities in a series of one-inch-square boxes. At the end of each scenario the trainees calibrate their interpretations with the conclusion of a panel of experts. Thus, the method does not require the a subject-matter expert to facilitate the training. Hintze used this ShadowBox method to increase expertise of firefighters, and it is currently being applied to a DARPA project for developing social interaction skills.

KEYWORDS

Training; scenario-based training; expertise.

INTRODUCTION

One way to help trainees develop expertise is to let them see the world through the eyes of experts. This "expert view" would let trainees discover what experts think is important in a situation, how they focus their attention, and also what they ignore. It would help trainees broaden their viewpoints and appreciate how subtle events might have important implications.

Bloom and Broder (1950) provided this type of perspective in a project designed to help under-performing college students do better on multiple choice tests. Bloom and Broder collected think-aloud protocols from students who did very well on these tests. Then they had under-performing students also generate think-aloud protocols as they struggled with the same items. Next, Bloom and Broder showed each under-performing student his/her own transcript along with a transcript from a successful student, and asked, "What was the difference?" For example, some under-performing students noticed that when they didn't know the answer they gave up, whereas the successful student shifted from a recall/recognition mode into a problem solving mode, trying to figure out what the answer might be, or at least trying to eliminate a few of the options. Bloom and Broder did not offer any advice. They let the under-performing students make their own discoveries about how they were falling short and what they needed to do on future tests. Their method generated significant improvements in performance. This study illustrates the impact of letting trainees see how experts (or people more skilled) view a task. The under-performing students defined deficiencies that were meaningful to them, which helped them translate training requirements into personal action.

Hintze (2008) provided a similar opportunity for firefighters. He developed scenario-based exercises coupled with cognitive task analysis materials to allow newly promoted officers to see the scenarios through the eyes of experienced officers. Hintze developed four challenging scenarios about unusual emergency situations. The scenarios were presented in a booklet and given to 14 experienced New York City Fire Department Officers who described how they would handle the decisions and wrote down the rationale for their decisions. These data were synthesized to illustrate the "expert mindset" to the experimental group.

These scenario materials were presented to an experimental group of 14 recently promoted New York State Fire Department Officers. The matched control group consisted of 15 New York City Fire Department Officers. The training materials exposed novices to situations they would not experience during routine operations. The training increased decision making performance by 18% ($p < .001$) for the experimental group over the control group. Hintze created a scoring key with 100 points as the maximum. The average score for the experimental group was 86.9, compared to 73.6 for the controls.

The Hintze study demonstrates a means of providing trainees with access to the thinking process of experts without having the experts present. It provides a means of enabling trainees to compare their responses to those

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

of the expert panel and determine for themselves what the differences are, and to speculate about the limits of their decision making.

Subsequent to this study, this methodology has been used to train additional groups of firefighters, and it has evolved into the ShadowBox method described below.

PRACTICE INNOVATION

The ShadowBox method is scenario-based. The trainees receive a booklet presenting a challenging scenario. They also have a second booklet for recording their answers at predetermined decision points. They record their answer to questions posed at each decision point in a separate small box, usually one-inch-square (but sometimes two-inches square). The trainees are essentially trying to shadow the expert panel by seeing the match between the responses they enter in the answer boxes to the answers provided by the expert panel.

Thus, for a given scenario, the initial page might describe the immediate situation including a map or photograph. At the bottom of the page is a prompt to enter any information they want to remember in the box for decision point #1, along with their rationale for what they included in the box. Once the trainees finish (they are given about 2.5 minutes for this task), they can never turn back in the booklet. All they will have to go on is what they wrote down:

(1) The scenario passage is read aloud to the group, they follow along.
(2) When the facilitator finishes reading, the trainees record what they want to remember in the ShadowBox.
(3) The trainees then answer the decision point questions.
(4) The facilitor then compares their Shadowbox entries and decision point answers to that of the experts, explaining the rationale for experts' answers.
(5) The trainees compare their responses to those of the panel, and compare the rationale. They are asked to describe the differences in the contents of the two boxes.
(6) The facilitator moves on to the next scenario update.

There are several forms of Shadow box. First, with the Attention box, trainees select what to put in the Shadow box to remember. Second, at some of the decision points the trainees have an Action Priority box; the booklet lists a small set of potential actions and the trainees prioritize these in order of importance and enter the top three into this second type of box, along with their rationale in the margin of the page. Then they learn the priorities listed by the experts, along with their rationale. Again, they have an opportunity to compare their responses and rationale to those of the experts.

A third type of box is the Information box. For some of the decision points the trainees are instructed to enter one query into the Information box—one type of information they would like to have at this point in the scenario or a question they would like to pose to the expert panel, along with their rationale for asking this question. Then they learn what the experts wrote, and go through the same comparison.

Other types of boxes are possible. There can be an Anticipation box (what is likely to happen in the next 15 minutes), an Assessment box (asking about different possible explanations of what is happening, often in form of yes/no questions), a Monitoring box (which cues should be watched most carefully) and so forth.

One of the most labor-intensive activities for implementing the ShadowBox method is to obtain responses from an expert panel. In his Master's thesis, Hintze (2008) used 14 FDNY officers as Subject Matter Experts (SME). He interviewed all of them face-to-face. Because of scheduling limitations, Hintze interviewed some of the experts 1:1, and others in small groups of 5-8 experts. These small groups were more efficient to run, and also provided valuable dialog that helped to improve the responses. The disagreements prodded the small groups to define their rationale more clearly.

Based on these interviews with experts, an answer key is prepared. The answer key consists of the consensus response for what should go into each box, along with a summary of the rationale responses. Hintze found a strong consensus for many of the boxes, but never achieved 100% convergence. In some cases the experts did not reach a strong consensus, and Hintze let the trainees know about any strong minority position that had emerged. He made it clear that there was no ground truth for any of the answers.

PILOT TEST

We applied the ShadowBox method to the task of training social skills in police officers. This pilot was part of a DARPA program, "Strategic Social Interaction Modules (SSIM)." The purpose of the SSIM program was to understand why some police and military personnel are more skilled at interacting with civilians to gain voluntary compliance, and to turn the findings into training to build interpersonal skills in police and military. The nickname for the program is the "Good Strangers" program because the intent is to develop skills of positive interactions with civilians rather than relying on intimidation to gain compliance.

Three scenarios were developed, using incidents that had been probed through the use of a Critical Decision method interview. The scenarios revolved around a domestic violence incident, an incident of managing a gang during the funeral of one of their members, and a case of persuading a suicidal man to drop his weapons. We prepared ShadowBox forms for each incident, using a mix of different types of queries. We collected calibration data using a panel of seven Subject Matter Experts interviewed separately by telephone. We conducted an initial pilot test with seven participants to modify the materials.

We collected pilot data on 16 experienced police officers working in the Spokane Washington area. Their experience level ranged from 7 to 29 years. We ran four groups of four officers. It took approximately four hours to run each group. We had two facilitators (neither of them with any police experience) guiding the groups. As the week continued, we added a wrinkle of having the officers fill out the priorities from the perspective of a "Bad Stranger," or from the perspective of a rookie. We broadened the Attention box to include inferences and questions, not just information to remember.

We collected evaluation data from 15 of the sixteen police officers. We asked them to provide ratings using a 5-point scale where 5 was high and 1 was low. When asked if the scenarios were realistic, the police evaluated them as a 4.3. When asked if the scenarios were interesting and engaging, the rating was 4.4. The question "I learned a lot from this exercise" was rated 3.8. For the question of whether the training should be delivered by an experienced police officer, seven said yes and seven said no (there were two non-responders). When asked if they would do anything differently having gone through the scenarios, the responses ranged from 50% to 87.5% who said that they would do something different, for the three scenarios.

To illustrate, one of the ShadowBox scenarios we used for the Good Strangers project involved a case of domestic violence. The actual incident was a bit unusual in that several police squad cars arrived on the scene, and when the senior officer entered the home he had three other police officers behind him. In the scenario (as in the actual incident), the senior officer gained entry to the home after knocking and getting no response. Upon entry the officer saw an agitated and frightened eight-year-old boy opened the door. The boy was the one who called 911. Sounds of a fight were clear – emanating from a bedroom on the second floor. The scenario stopped at this point to inject a decision point: what are the priorities?

The choices were: (1) Ignore the boy and focus on the screams; (2) Inform the boy why the police are there and calm him down before moving on to the threat; (3) Direct the other officers to handle the threat upstairs (4) Remove the boy to stairs outside the apartment. The scenario continued to the arrest of the abusive boyfriend, the entry of the bloodied mother who becomes outraged to learn that her son had summoned the police, and so forth.

The later decision points addressed alternative strategies for handling the mother's anger at her son and abuse of the police, and the criteria for deciding when it was safe to leave – without endangering the son who might be physicially punished by his out-of-control mother. For the purpose of this paper, we will only address the first decision point. Before reading further, you might consider how you would prioritize these options.

The panel of SMEs reviewed the four options and by a narrow margin selected option (3) as the top priority. The rationale including the following: the boy was a potential witness; the boy was a victim-by-proxy; this was an opportunity to demonstrate to the boy that police officers can be trusted to provide security; the boy is at risk in the future if the boyfriend returns; the boy's needs should be respected; the large contingent of police officers made it unnecessary for all four to go charging up the stairs. The panel gave the lowest priority score to option (1) Ignore the boy and focus on the screams.

Of the 16 police officers tested on this Decision Point, zero selected (2) as the top priority. Five of them made it their lowest priority. The ShadowBox exercise was thus very effective for confronting these police officers with the mismatch between their own tendencies and the perspective of a Good Stranger.

DISCUSSION

The method permits trainees to "shadow" experts by seeing how their own responses compared to the responses from the expert panel, and to speculate, using the rationale materials, about how the experts were thinking. For his original research, Hintze estimated that trainees needed 2 to 2.5 hours to complete the four scenarios, about 30-40 minutes per scenario. The trainees had about 2.5 minutes to fill in the box for each decision point. Hintze was the facilitator for the training sessions for his Master's thesis. He read the scenario aloud as the trainees followed along with their booklets, pausing at each decision point to let the trainees fill in the boxes for the questions posed. We needed 4 hours for our three scenarios.

The ShadowBox method is similar to techniques such as Tactical Decision Games (TDGs, Schmitt, 1996) and Decision-Making Exercises (DMXs, Klein, 2005) that present scenarios and invite trainees to respond. The ShadowBox method seems to have some advantages over these other approaches in that it provides an answer key generated by an expert panel (although, like the TDGs and DMXs it does not claim that there is any "right" answer), it provides the rationale for the experts' answers, and then it provides a convenient format for trainees to explicitly provide their responses. The contrast between trainee answers/rationale with those of the expert panel helps trainees appreciate the limits of their mental models and points the way for them to develop their cognitive skills further.

When we applied the ShadowBox method to the DARPA Good Strangers project, one skill we addressed was to take someone else's perspective, particularly when the other person comes from an unfamiliar culture. Here, the ShadowBox method was similar to "cultural assimilators," a method that has been validated in the field of intercultural communication. The method of cultural assimilators relies on a collection of real-life scenarios describing puzzling cross-cultural interactions and explanations for avoiding the emerging misunderstandings.

This is one of the most researched and accepted methods of cross-cultural training (Bhawuk & Brislin, 2000; Landis & Bhagat, 1996, Albert, 1983).

The ShadowBox method differs from techniques such as the U.S Army's "Think Like a Commander" program (Lussier & Shadrick, 2002; Ross & Lussier, 1999) which tries to develop cognitive skills via tactical scenarios. The Think Like a Commander program has defined a core set of cognitive skills for the military officer, such as keeping a focus on the mission and the commander's intent, modeling a thinking enemy, considering effects of terrain, using all assets available, considering timing, seeing the big picture, visualizing the battlefield, and remaining flexible. The training emphasizes these principles. In contrast, the ShadowBox method does not pre-define any principles. The training is all contextualized, and the trainee learns by peering into the minds of the experts—seeing their priorities and responses and the rationale they offer. In addition, the ShadowBox method can be used with non-kinetic missions such as counterinsurgency, which are different than the kinetic missions and core skills the Think Like a Commander program was designed for.

One of the important potential extensions of the ShadowBox method is to adapt the materials so that they do not require expert facilitation. In theory, the ShadowBox method could be in an on-line form without any facilitator. Participants could drag and drop items from the scenario into the answer box, speeding up the process and reducing a source of ambiguity.

Challenges for the ShadowBox method are to determine whether the scenarios used are representative of the domain to be handled, and to develop efficient means for knowledge capture with domain experts. Other extensions would be to use video representations of scenarios rather than written/text representations, and to use systematic scoring methods to enable computer-based scoring for immediate feedback. The ShadowBox format is very flexible; in addition to a training application it can be adapted for knowledge capture for a cognitive task analysis effort.

ACKNOWLEDGMENTS

This work was supported by The Defense Advanced Research Projects Agency (government contract/grant number 06-1825383). The views, opinions, and/or findings contained in this (article, poster, paper etc.) are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense. We gratefully acknowledge the assistance provided by Lois James, Washington State University.

REFERENCES

Albert, R. D. (1983). The intercultural sensitizer or culture assimilator. A cognitive approach. In D. Landis & R. W. Brislin (Eds.), (Vol. 2, pp. 186-217). New York: Pergamon.

Bhawuk, D. P. A., & Brislin, R. W. (2000). Cross-cultural training: A review. *Applied Psychology: An International Review, 49*(1), 162-191.

Bloom, B. S., & Broder, L. J. (1950). *Problem-solving processes of college students: An exploratory investigation*. Chicago: University of Chicago Press.

Hintze, N. R. (2008). *First responder problem solving and decision making in today's asymmetrical environment.* Unpublished Master's thesis, Naval Postgraduate School, Monterey, CA.

Klein, G. (2005). *The power of intuition*. New York: A Currency Book/Doubleday.

Landis, D., & Bhagat, R. (Eds.) (1996). *Handbook of intercultural training.* Newbury Park, CA: Sage Publications.

Lussier, J. W., & Shadrick, S. B. (2002). *Think Like A Commander: Captain's Edition - Prototype 1.0*: ARI Research Product 2003-01 available from U.S. Army Research Institute for the Behavioral and Social Sciences, 2423 Morande Street, Fort Knox, KY 40121.

Ross, K. G., & Lussier, J. W. (1999). A training solution for adaptive battlefield performance *Proceedings of the 1999 Interservice/Industry Training, Simulation, and Education Conference*. Orlando, FL.

Schmitt, J. F. (1996). Designing good TDGs. *Marine Corps Gazette, 80*(*5*), 96-97.

Exploring sensemaking through an Intelligence Analysis exercise

Chris BABER[a], Simon ATTFIELD[b], William WONG[b] and Chris ROONEY[b]
[a]*University of Birmingham*
[b]*University of Middlesex*

ABSTRACT
Introduction: A laboratory-based approach to studying sensemaking was developed. **Method**: Five teams reviewed a drug-smuggling scenario in order to determine which of five suspects they would arrest and the five pieces of evidence most important to their decision. **Results and discussion:** Teams emphasised different aspects of the scenario and, while there were some overlaps in suspects or pieces of evidence selected, there were interesting differences in the explanations the teams offered and the conclusions drawn. The study provided an informative approach to the study of 'framing' in sensemaking.

KEYWORDS
Sensemaking; Intelligence Analysis; Laboratory-study.

INTRODUCTION

Sensemaking is a process through which data meet background knowledge in the construction of an explanatory model of a situation. The model is the 'sense' that is inferred from the data. Klein, Moon and Hoffman (2006a,b) and Klein, Phillips, Rall and Peluso (2007) use the term 'frame' to stand for this model. They explain how frames and data are reciprocally selected and evaluated as part of a process of comprehension. A frame can be used to account for and evaluate the saliency of these data to a particular interpretation. If the data are consistent with the interpretation, then the frame is used to guide search for further data which, if consistent, will elaborate the frame. If the data seems inconsistent, then an alternative frame will need to be selected or the data somehow explained away. The process can be seen *as "…a motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively."* [Klein et al., 2006a, p.71]. What is not obvious, in this approach, is where frames originate. In order to address this question, we had teams of people performing a simulated Intelligence Analysis task in a controlled setting.

Intelligence Analysis

At its most rudimentary, Intelligence Analysis is the business of compiling material from multiple sources in order to determine the likelihood of an undesirable event (Heuer, 1999). Intelligence Analysis is complicated not simply by the need to employ multiple sources but also by the fact that the material is incomplete, inconclusive or ambiguous (Tecuci, Boicu, Schum and Marcu, 2010).

In their investigation into military intelligence analysis, Roth, Pfautz, Mahoney, Powell, Carlson, Guarino, Fichtl and Potter (2010) demonstrated that *"…the problem formulation phase, in which information requests are framed and contextualized, is one of the most cognitively challenging aspects of the intelligence analysis process."* (p.221). Roth et al. also demonstrate the importance of 'metainformation' to help frame intelligence requests through the use of assumptions based on experience. The additional knowledge and assumptions covered not only the intelligence itself but also likely Courses of Action and credibility of sources from which the intelligence was gathered. The description of this process can be considered very similar to the data-frame concept of sensemaking (see above).

Data, Frames and Narrative

Pirolli and Russell (2011) note that *"Frames can be expressed in a variety of forms including stories, maps, organizational diagrams or scripts."* [p.5]. The inclusion of 'stories' in this list echoes Weick's (1995) earlier assertion that, *"In short, what is necessary is sensemaking is a good story".* [p.61]. In previous work, we have explored the relationship between narrative and intelligence analysis (??) and in this paper we propose that we form of 'frame' that might be used in the analysis could be defined as a 'crime frame', in which the analysts predict likely courses of action that could be taken in the activity under investigation on the basis of their prior knowledge of criminal activity.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

INTELLIGENCE ANALYSIS EXERCISE

An exercise was designed to provide teams of participants with a task that was sufficiently open-ended (in terms of incomplete, inconclusive and ambiguous data) to be challenging, while, at the same time, have sufficient boundaries (in terms of named 'suspects') to be tractable within a 2½ hour session. Having different teams address the same problem provided us with a point of comparison, on the assumption that if the teams have the same materials, then any variation in performance will arise from the manner in which the teams interpret and make use of these materials.

Participants

The United Kingdom Visual Analytics Consortium (UKVAC) Summer School took place at Middlesex University's Hendon campus in September 2012. There were around 40 delegates and the majority of these participated in the study. Delegates came from a variety of backgrounds, ranging from Computer Science to military intelligence. For the exercise, delegates were divided into teams of between 5 and 8 members. As far as practicable, teams were balanced in terms of gender and experience, i.e., each team had a least one member with experience of intelligence analysis and at least one student delegate.

Overview of the Exercise Scenario

The Scenario is set in the United Kingdom and involves two locations. A small harbour on the South coast is used as the point at which drugs are smuggled into the country. A city in the North of England has a drug distribution ring, run by taxi drivers. The taxi drivers work for a company owned by a woman who also owns a haulage firm in the city (and who has been under Police surveillance for several years). This woman recently brokered the purchase of the harbour by an American company. The taxi drivers collect their drugs from a warehouse and meet in a café to share their money; the warehouse, the café and the harbour are all owned or managed by people associated with the accountant of the woman. The dates of the van rentals and yacht arrivals coincide sufficiently to suggest that the yacht arrives and has its cargo transported in a hired van to the city in the North of England. Key to the solution of this scenario is the recognition that the hired van containing the drugs is always driven by the passenger on the yacht.

Procedure

All teams were given an initial briefing and then taken to their own incident rooms to complete their investigation. These rooms were equipped with whiteboards, large notepads, computers, pens, post-it notes and paper. The initial briefing included some background information about each of the suspects. Table 1 summarises this information.

Each team was provided with a pack of 49 slides. The pack included 9 suspect cards (with picture of the suspect and their address), together with a combination of telephone logs, harbourmaster logs, maps, business accounts, witness and arrest statements, newspaper articles etc. Figure 1 provides an illustration of the types of material supplied.

Following the briefing, teams were given the following Direction: "Using the evidence provided, decide who (from Table 1) should be arrested and where the best place might be to make such arrests. Following your investigation, each team will give a presentation on their findings. The presentation from each team had to include:

1. Name of individual, or individuals, to target as Suspects.
2. The FIVE pieces of evidence that best support your proposal to 1.
3. Location of the arrest or arrests."

A Correct Solution

Grosby is a businesswoman who lives in Leeds and runs a road haulage and mini-cab firm. While she has no criminal convictions, local police have long been suspicious of her acquaintances and believe that she has links with criminal activity, particularly relating to drug smuggling and trafficking. A newspaper article states that 8 of the taxi drivers in the firm she owns were arrested for selling drugs. Calabrese (who was the chauffeur for Grosby's accountant) was sentenced to 9 years for drug smuggling. Grosby also helped set up the purchase of Exmouth Marina and intelligence suggests that there is likely to be a shipment of class A drugs being delivered to this harbour. There is evidence to suggest that the gang have shipped drugs into this harbour in the past and taken the drugs, in a hired van to a warehouse in Leeds for distribution to the tax drivers. The yacht carrying the drugs is owned by a client of Grosby's accountant. In the previous shipments, the yacht's passenger had driven the van (which had been hired by the marina manager's girlfriend). The main solution to the challenge would be to arrest Pierre Pasquidini as he drove a van load of smuggled drugs from Exmouth to Leeds. There is less evidence to arrest anyone else (although it might be possibly to arrest Martina Sarti, who hired the van or Vanessa Munoz, sub-manager of the marina for their involvement).

Table 1. Suspects

Suspects's Name	Background Information
David Pico	Denise Ajachinsky's ex-boyfriend who is still jealously and possessively monitoring her; a petty criminal with several convictions
Jake Ajachinsky	Denise Ajachinsky's brother and Pico's best friend; recently arrested with Pico
Denise Ajachinsky	See above. Lives in a wealthy part of Leeds with her father who has a large income of unspecified means.
Kenny Chiappe	Taxi driver in Leeds who had been arrested for drug dealing
Pierre Pasquidini	Travels regularly between France and UK
Vanessa Munoz	Sub-manager of the Exmouth Marina who was appointed shortly after its take-over by Atlanta Holdings.
Muriel Grosby	Owns haulage firm and taxi firm in Leeds. Known to Police for suspected involvement in a range of criminal activities but has never been successfully prosecuted.
Martina Sarti	Girlfriend of the owner of Exmouth Harbour
Jennifer Garlica	Muriel Grosby's sister-in-law. Widowed by a gangland execution of her husband.

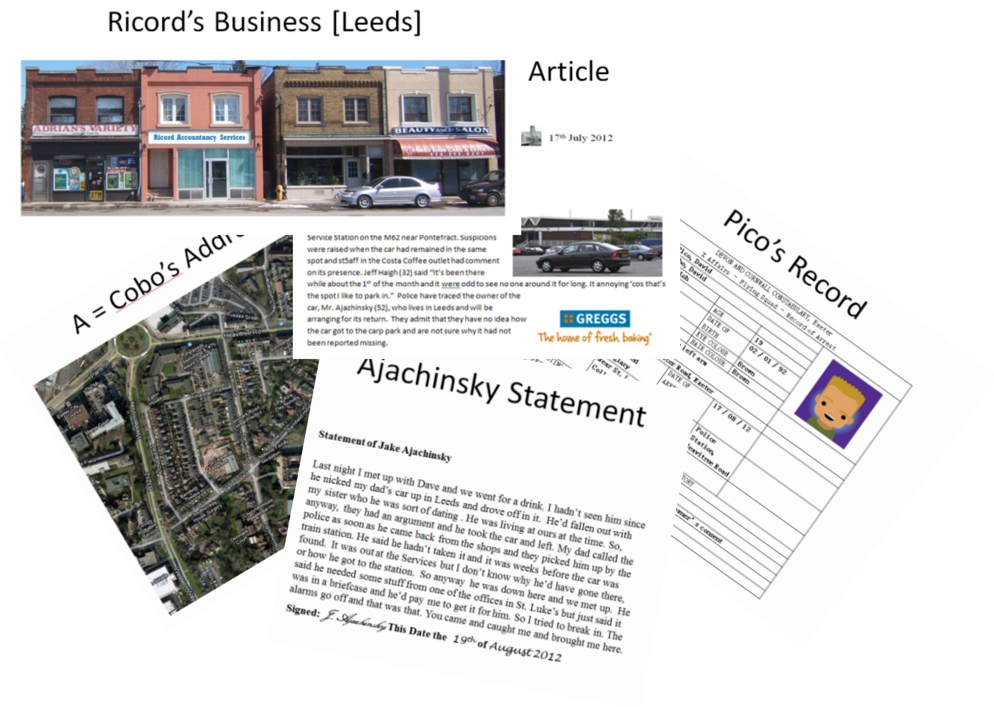

Figure 1. Sample of materials used in the exercise

Results

A video recording was made of each team presenting their conclusions, and photographs were taken at approximately 15 minute intervals of the teams during their activity. In terms of the value of the evidence provided, attitudes ranged from Team 1 who felt that they had 'no solid evidence to arrest' suspects, to Team 2 who felt that 'there seems to be enough interconnecting evidence if you can catch one of them red-handed'. The presentations were transcribed and a summary of the main answers to the three questions is shown in Table 2.
For brevity, we focus on two instances (it is worth repeating that the exercise was designed with the assumption that Pasquidini was driving the van full of drugs who should be arrested in Leeds).

Why arrest David Pico / Cobo?

Team 1 and Team 4 both noted that the accounts for Ricord, Grosby's accountant, show that David Pico is as paid as a chauffeur for Ricord. Ricord's previous chauffeur is currently in prison, having been arrested driving a van carrying drugs in the South-West of England. Team 2 and Team 5 proposed that David Pico was 'most likely' or 'obviously' driving the van (on the basis of his criminal record and the fact that he lived near Sarti).

Why arrest Muriel Grosby?

Team 1, Team 2 and Team 3 all note that Grosby's accounts look suspicious and suggest that her involvement in brokering Atlanta holdings purchase of Exmouth Marina suggests a plan for a drug smuggling operation. However, Team 5 note that the accounts might not be that suspicious because they indicate a company director receiving dividends and a UK transport company branching out from road haulage and taxis to maritime activity, and brokering a deal for a US partner.

Table 2. Evidence and Suspects in each Groups report

	Group 1	Group 2	Group 3	Group 4	Group 5
Suspect					
Pico / Cobo	x	x		x	x
Sarti		x		x	x
Angeletti				x	x
Grosby	x			x	x
Munoz		x	x		
Garlica					x
Pasquidini				x	
Perrin				x	
Condiere				x	
Jake Ajachinsky					x
Chiappe					x
Ricord				x	
Mrs Ricord				x	
F Davis		x			
Key Evidence					
Bocagnani statement	x		x	x	
Grosby accounts	x				x
Pasquidini phone logs	x	x		x	
Harbourmaster logs	x	x	x		x
Van rental logs	x	x	x	x	x
Grosby newspaper		x			
Calabrese statement				x	
Ricord accounts				x	
Taxi drivers newspaper			x		x
Other Phone logs		x	x		x
Location of Arrests					
Leeds	x		x	x	x
Exmouth	x	x			

Reasoning Activity

One approach to reasoning in this exercise would involve the team developing an explanatory model and then seek evidence to support this. One form of such an explanatory model would be in the form of a *narrative* in which certain expectations or procedures or modus operandi, can be combined. Such a narrative could involve acts such as 'smuggle drugs', 'store drugs', 'distribute drugs' etc. Thus, one of the starting points for the 'frame origination' is the task goal itself, with participants initially asking "what is happening (or has happened)?" This helps create a meta-frame which coordinates the knowledge and expectations derived from socio-cultural backgrounds, education, training, values, experiences, expertise etc. of team members. If this was the case then one might anticipate initial activity in the teams to involve the development and sharing of these narratives to help guide search of the evidence. In their presentation, for instance, teams 4 and 5 both emphasised the need to 'follow the money' rather than focus exclusively on the movement of drugs from one place to another, while other teams focused more on the transportation itself (with team 2, for instance, considering how best to organise tracking of the hire van that might be carrying the drugs). While teams developed explanatory models or narratives, we noticed that they tended to begin their investigation by manipulating the items of evidence and laying these out in different ways. This implied an alternative approach, in which teams used features of the evidence to provide a structuring mechanism that could allow specific items to be related to each other. One form of structuring mechanism involves the relationships between individuals. Such relationships, for this exercise, were defined by telephone records, business accounts or attendance at formal dinners. A structuring mechanism based on these relationships, as shown by Figure 2, shows the different types of person and different types of relationship between them.

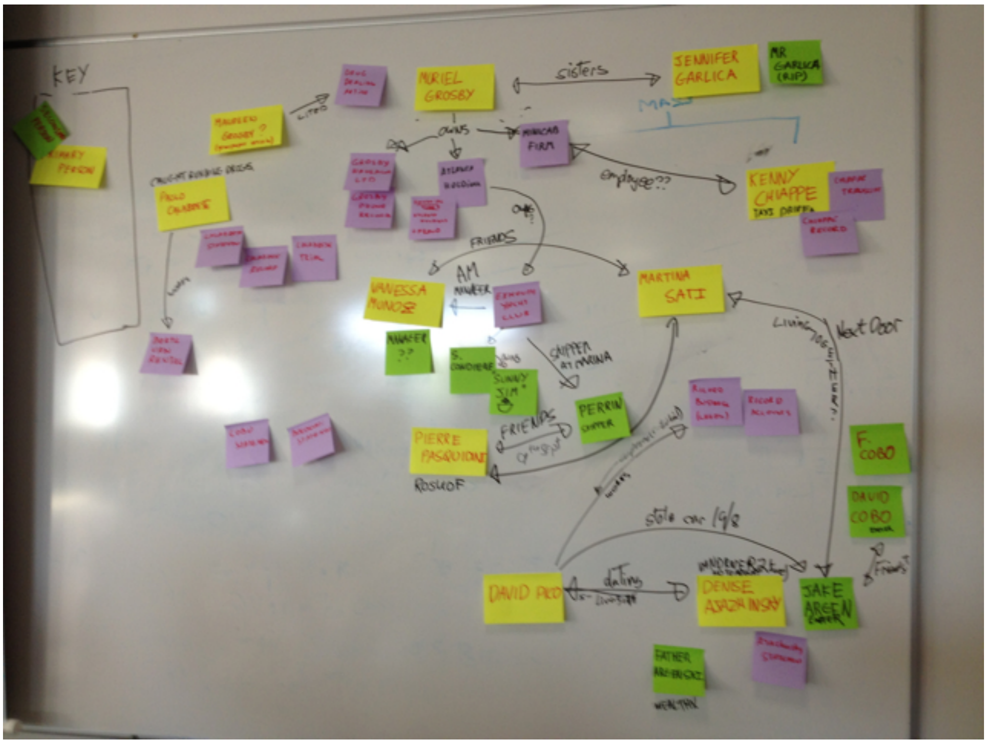

Figure 2. Colour-coded Social Network Graph.
Light: Suspect (from Table 1); Dark with arrows: Secondary person; Dark but not linked: Additional information.

Another common approach was to concentrate on the dates and times that several items of evidence contained to produce various forms of event sequence (Figure 3).

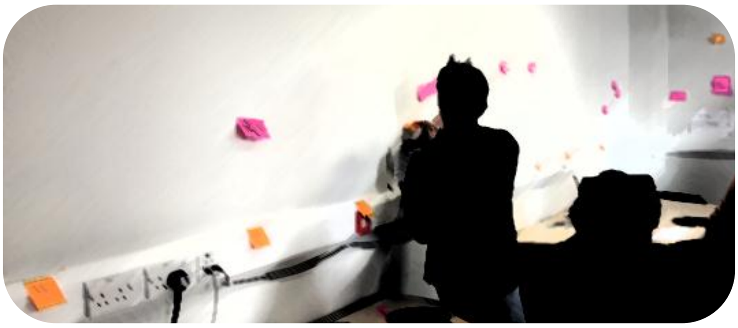

Figure 3. Using an entire wall to create a timeline
People anonymized. Post-it notes on the cable conduit are month & year markers; post-it notes on the walls are markers for summarized events. The relevant items of evidence were then blue-tacked to the wall to minimize re-copying of details.

The groups who used low-dimensionality representations seemed to get stuck at some point in the investigation, and found it hard to progress unless they shifted to a different representation. Thus, we noted that the groups who began with the construction of social networks all moved to timelines to arrange the evidence. One group who was able to integrate the multiple-dimensions, were able to eliminate and narrow their suspects list (Figure 4). Having said this, this team did not identify Pasquidini or Angeletti as the main suspects.

Figure 4. Multi-dimensional Network
Geographical information provides context. The red links represent ***communications networks*** (e.g. there were phone communications between Perrin and Martine, who is girlfriend to David). The team also used this chart to show or talk about the ***flow of goods***.

DISCUSSION

Before discussing the conclusions drawn from this work, it is worth considering how well the exercise relates to real intelligence activity. The exercise was intended to introduce people to the challenges of Intelligence Analysis and allowed exploration of sensemaking in an laboratory setting. Clearly the use of ad hoc teams with a mixture of expertise could be considered to render the exercise different to real situations. We would point out that the aim of the exercise was primarily for training purposes and so would be reluctant to argue that the conclusions reflect real behaviour of skilled analysts. The second issue relates to the material used. Here we have more confidence that the general point made is reflected in real analysis, i.e., that materials arrive in a variety of formats, that there is missing information, that there might be conflicting information etc., and that this potential confusion of disparate information sources presented the primary sensemaking challenge for Intelligence Analysis.

The main question for paper was where do the ‘frames’, in the Data Frame description of sensemaking, originate? This exercise suggests that (initially at least) ‘frames’ are constructed on the basis of data, ie., the evidence being considered. Teams began by selecting similar forms of evidence, e.g., those which contained dates of events or those which contained financial transactions or those which contained ‘phone calls. Depending on which type of evidence the team selected, the next step was to produce a visualisation of the relations between the selected pieces of evidence. Thus, for the phone logs the visualisation was a social network diagram, while for the van rentals and harbour logs it was a timeline. This suggests that rather than beginning with a frame based on a narrative of the type of crime being investigated, most of the teams employed some form of structuring mechanism based on features of the evidence. In other words, the ‘frame’ was the most appropriate visualisation for the type of data being explored. The visualisations (Figures 2, 3 and 4) that teams constructed in the exercise were used in a highly dynamic fashion. Teams would scribble over links, move post-it notes, add comments (either directly on the artefact or through the use of post-it notes stuck on top of the artefact), or insert question marks, in their process of testing and questioning the frame that the artefact was implying. From this we propose that the artefacts provide an initial ‘draft’ frame which is visualized and then amended and elaborated as the teams consider the implications of what the visualization represents or the relationship between the frame and the evidence under consideration. While this provided a perspective on the evidence, it obscured some key relationships.

Many of the teams were influenced by the case study of David Cobo (aka Pico) and did not reach the solution that the boat's passenger was always the driver. One explanation was that main focus of these teams was on the relationship between the boat arriving in port and the hiring of vans to take the drugs to Leeds. Once they had constructed a frame using the dates of events, they seemed less inclined to take the next step and consider *who* might be most likely to drive the van. This suggests that using a frame which could focus on temporal relationships between events might not have been sophisticated enough to also deal with motive and opportunity for individual action. This also suggests that the test of plausibility that teams were applying was largely driven by the material provided (or at least the most common factor in that material, which was the dates of events and transactions).

Following the draft framing, the analysis was elaborated by asking specific questions, such as "what is individual Y doing at time X?" or "what were the outcomes of event Z?" The process led to a (loose) story, as an initial hypothesis to which teams tested the 'goodness of fit' of pieces of evidence. As data or evidence to support the story become available and combined into the story, the story becomes more robust and more useful for explaining the situation. The role of the artefacts then shift from being the motivator for the story to become the externalised model of the constructed story. We propose that when this initial story is developed and strengthened, it develops in its formality and maturity to become a narrative. As this narrative is developed further, by verifying and validating the data and the relationships between data, and the rules that govern those relationships, teams arrived at a version that can withstand interrogation and became the basis for their presentations.

REFERENCES

Andrews, D., Baber, C., Efremov, S. and Komarov, S. (2012) Creating and using interactive narratives: reading and writing branching comics, *CHI'12*, New York: ACM, 1703-1712.

Baber, C., Andrews, D., Duffy, T. and McMaster, R. (2011) Sensemaking as narrative: visualization for collaboration, *Proc VAW2011,* University College of London.

Heuer, R. (1999) *Psychology of Intelligence Analysis,* Washington, DC: Center for the Study of Intelligence, CIA

Klein, G. A., Moon, B., & Hoffman, R. R. (2006a) Making sense of sensemaking: 1. Alternative perspectives. *IEEE Intelligent Systems, 21*, 70–73.

Klein, G. A., Moon, B., & Hoffman, R. R. (2006b) Making sense of sensemaking: 2. A macrocognitive model. *IEEE Intelligent Systems, 21*, 88–92.

Klein, G. A., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007) A data-frame theory of sensemaking, In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*. New York: Erlbaum, 113-158.

Pirolli, P. and Russell, D.M. (2011) Introduction to this special issue on sensemaking, *Human-Computer Interaction, 26,* 1-8

Roth, E.M., Pfautz, J.D., Mahoney, S.M., Powell, G.M., Carlson, E.C., Guarino, S.L., Fichtl, T.C. & Potter, S.S. (2010) Framing and contextualizing information requests: problem formulation as part of the intelligence analysis process, *Journal of Cognitive Engineering and Decision Making, 4,* 210-239.

Tecuci, G., Boicu, M. Schum, D. and Marcu, D. (2010) Coping with the Complexity of Intelligence Analysis: cognitive assistants for evidence-based reasoning, *Research Report 87, Learning Agents Center, George Mason University*

Weick, K.A. (1995) *Sensemaking in Organizations,* London: Sage

Investigating Constraints on Decision Making Strategies

Patrick K. BELLING[a], Joel SUSS[a], and Paul WARD[a,b]
[a]*Michigan Technological University, USA*
[b]*University of Greenwich, UK*

ABSTRACT

Recently, researchers have focused on how individuals generate different courses of action (i.e., options) on-the-fly, and predict the options to be taken by others. When generating predictive options, previous research supports the use of cognitive mechanisms described by Long Term Working Memory (LTWM) theory (Ericsson & Kintsch, 1995; Ward, Ericsson, & Williams, 2012). However, when generating response options, previous research supports the use of the Take-The-First (TTF) heuristic (Johnson & Raab, 2003). The current research investigates the effect of time constraint on option generation behavior. Our results provide further support for the use of LTWM mechanisms during prediction, but support was observed for TTF only during response without time constraint. When participants responded under time pressure, they shifted towards a LTWM-type strategy. Modifications to the cognitive-process level description of decision making during response are proposed and implications for training during both prediction and response are discussed.

KEYWORDS

Decision Making; Sport; Long Term-Working Memory; Take-The-First heuristic.

INTRODUCTION

Decision making research has historically focused on strategies for choosing between a fixed set of options presented to a participant in a task, meanwhile ignoring the process by which naturalistic decision makers actually come up with alternatives from which to choose (see Johnson & Raab, 2003; Zsambok & Klein, 1997). The latter process is henceforth referred to as option generation and is considered a critical component of decision making in many real-world domains (Klein, 1993).

In a study of handball, Johnson and Raab (2003) proposed that the Take-The-First (TTF) heuristic supports successful decision making. Participants were shown short video clips of handball play that were frozen unexpectedly at a critical decision point, immediately prior to the player with the ball taking a specific course of action (e.g., shoot, pass). Viewing from the perspective of the player with the ball, participants highlighted as quickly as possible the first option they would take that came to mind, and then any additional options they could conceive. These authors found a negative correlation between the total number of options generated and the quality of the final decision. They concluded that an initial option is generated based on association with the environmental structure. As activation spreads, other less relevant options are generated. As a consequence, generating many options decreases the likelihood that one will select the first and best option. In a subsequent study, Raab and Johnson (2007) also demonstrated that as skill level increases, fewer total options are generated. This 'less-is-more' phenomenon has been supported by other models of decision making in similar environments, such as the Recognition-Primed Decision (RPD; Klein, 1993) model.

On the other hand, Long Term-Working Memory (LTWM) theory suggests that experts build a situational model that integrates stored knowledge with new environmental information on-the-fly, producing an updated situational representation (Ericsson & Kintsch, 1995). This allows experts to accurately anticipate the situational consequences and maintain direct access to task-relevant decision alternatives. In three experiments conducted by Ward, Ericsson, and Williams (2012), participants watched video simulations of soccer play similar to those presented by Johnson and Raab (2003) from the perspective of a defender on the opposing team to the team with the ball. Participants anticipated what the player with the ball would do next and generated the options that they were considering when the video was either frozen on screen or occluded from view. Consistent with most models of expertise, players generated few options and better options first. Furthermore, skilled players anticipated situations more accurately than less-skilled players. However, unlike TTF, the number of *task-relevant* options—those considered as high quality options by a domain-specific expert panel—was positively correlated with anticipation accuracy. This relationship held even when clips were occluded rather than frozen on screen. Ward et al. (2012) suggested that the ability to build and maintain access to a well-developed situational model afforded skilled participants the opportunity to generate more, as opposed to less, task-relevant options.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Current models of expertise (i.e., TTF, LTWM) make opposing predictions regarding strategy use, especially concerning the relationship between option generation and decision quality. These may be influenced by methodological differences between studies—including phase of decision making/decision perspective (Ward, Suss, Eccles, Williams, & Harris, 2011). For instance, when deciding upon a course of action for oneself (e.g., during intervention/response), research supports the use of TTF (Johnson & Raab, 2003; cf. Ward et al., 2011). Alternatively, when assessing the course(s) of action that others in the environment might take and/or predicting the outcome of a situation (e.g., during assessment/prediction), research supports the use of LTWM mechanisms (Ward, et al., 2012; cf. Klein & Peio, 1988). Preliminary data supports this distinction (Belling & Ward, 2012) but this may also be dependent on domain (see Ward, et al., 2011).

Time constraint may also affect strategy use. For instance, Calderwood, Klein, and Crandall (1988) found that moves selected by Master players in time-pressured games did not degrade relative to those selected in standard tournament games. In contrast, less skilled chess players showed a significant decrement. Furthermore, Gobet and Simon (1996) compared Gary Kasparov's performance during standard and simultaneous chess games, which reduced time available per move. They concluded that recognition skill was primarily responsible for Kasparov's superior situational assessment and move selection because performance did not decrease under time pressure (cf. Chabris & Hearst, 2003). In the experiments conducted by Johnson and Raab, participants were permitted different amounts of time (i.e., from 45 to 6s) to view the last frame of action frozen on screen, during which participants generated options. However, time to respond with the final decision was unlimited. While time exposure to stimiuli were not experimentally manipulated in their research, no obvious skill-based differences were observed in option generation behavior across time exposures.

The current research aims to investigate the effects of time constraint on option generation behavior during both situational assessment/prediction (henceforth prediction) and intervention/response (henceforth response) phases of decision making (see Ward et al., 2011) to investigate *when* the respective strategies are employed and constrained by environmental factors. Consistent with LTWM theory, we hypothesize that the number of task-relevant options generated will be positively related to performance during prediction trials. Following TTF, we hypothesize that the total number of options generated will be negatively related to performance during response trials. Based on research that has examined the effect of time pressure on move selection, we speculate that participants will favor a TTF-type strategy when under additional time pressure, irrespective of decision phase.

METHODS

Participants and Materials. Twenty-one (17 male) recreational-level players with an average age of 19.8 years (*SD* = 1.94) participated in this study. Twenty-four video clips were created using footage of a live soccer match. The video was filmed from a slightly elevated angle above and behind the goal. Each clip lasted approximately 10 seconds and was edited to end (i.e., occluded) at a critical decision point—immediately prior to the player with the ball pursuing a course of action (e.g., shoot, pass, dribble). A blank white screen with black lines representing the field lines and ball (i.e., the occlusion image) appeared at the point of occlusion—all other perceptual information was removed (see Ward, et al., 2012). Response sheets for each trial were created that matched the image on screen.

Procedure. Participants completed 12 prediction and 12 response trials. During prediction trials, participants were asked to envision themselves as a defensive player and generate options that they were concerned the offensive player with the ball might do next. During response trials, participants were asked to envision themselves as the player with the ball and generate the options that they would consider pursuing next. Participants responded by drawing options—which consisted of any combination of players, their actions, field positions, movements, and ball position and movement—onto their response sheet using a simple notation scheme. Next, for prediction trials, they rated how likely it was that the opposing player with the ball would choose each generated option as their next move and how concerned they felt that option was to their defense. For response trials, they rated how likely it was that *they* (i.e., the participant) would pursue each generated option as *their* next move and how good they felt that option was. Because all trials were occluded at the critical decision point, no additional time was given to inspect the final image (cf. Johnson & Raab, 2003). During half of all prediction and response trials, participants were given only 10 seconds to *generate* and *mark* options on the response sheet. Time to *rate* each option was not restricted. During the remaining half of trials, no time constraint was implemented.

RESULTS

Measures. For each prediction trial, we recorded the total, task-relevant, and task-irrelevant number of options generated. To assess performance, we recorded the ability to accurately anticipate the outcome of each simulation (i.e., assign the highest likelihood rating to the actual outcome). We also also recorded other measures of performance that were potentially interesting, including the ability to generate the most threatening/concerning criterion option as determined by SMEs (irrespective of serial position/rating), rate the most threatening criterion option as being of most concern, and anticipate (albeit incorrectly) that the most threatening criterion option would happen next.

For each response trial, we recorded the total, task-relevant, and task-irrelevant number of options generated. To assess performance, we recorded the ability to select the criterion best option as the intended course of action (i.e., assign the highest likelihood rating to the best option). We also recorded other potentially interesteing measures: the ability to generate the criterion best option (irrespective of serial position/rating), and the ability to accurately rate the criterion best option as best (i.e., assign the highest quality rating to the best option).

Time constraint, performance and option generation. Paired samples t-tests were used to assess the effect of time constraint on performance. During prediction trials, no significant differences were observed between time constraint conditions in the abillity to anticipate the actual outcome (or to anticipate as the outcome the most threatening criterion option). However, participants generated significantly fewer total options under time constraint (M = 1.70, SD = 0.53) than under no time constraint (M = 1.93, SD = 0.60); $t(20) = -2.259$, $p = 0.035$. The number of task-relevant or task-irrelevant options generated across time constraint conditions did not significantly differ.

During response trials, no significant differences were observed between time constraint conditions in the ability to select the best option. As in prediction, participants generated signifcantly fewer total options under time constraint (M = 1.72, SD = 0.55) compared to no time constraint (M = 2.0, SD = 0.51); $t(20) = -2.764$, $p = 0.012$. However, unlike prediction trials, during response trials participants generated significantly reduced the number of task-irrelevant options under time constraint (M = 0.90, SD = 0.41) compared to when not under time constraint (M = 1.21, SD = 0.46), $t(20) = -2.879$, $p = 0.009$. Time constraint did not significantly affect the number of task-relevant options generated during response trials.

Correlations during prediction. During prediction trials with time constraint, there was a trend towards both the total number of options generated and the number of task-relevant options generated being positively correlated with anticipation accuracy ($r = 0.35$, $p = 0.12$; $r = 0.40$, $p = 0.07$, respectively). The correlation between number of task-relevant options generated and anticipating the most threatening criterion option as the outcome also approached significance, ($r = 0.36$, $p = 0.10$). The number of task-irrelevant options generated was not related to any measures of performance.

During prediction trials with no time constraint, the total number of options generated was not related to anticipation accuracy. There was a trend towards the number of task-relevant options generated being positively correlated with anticipation accuracy when not under time constraint ($r = 0.30$, $p = 0.19$). However, the number of task-relevant options generated was significantly and positively correlated with the ability to generate the most threatening criterion option ($r = 0.60$, $p < 0.01$), accurately rate that option as most concerning ($r = 0.59$, $p < 0.01$), and to anticipate the most threatening criterion option as the outcome ($r = 0.45$, $p < 0.05$). The generation of task-irrelevant options was significantly and negatively correlated with the ability to generate the most threatening option ($r = -0.48$, $p < 0.05$), accurately rate the most threatening option ($r = -0.49$, $p < 0.05$), and anticipate the most threatening option as the outcome ($r = -0.43$, $p < 0.05$).

Correlations during response. During response trials with time constraint, the number of task-relevant options generated was positively correlated with the ability to select the criterion best option as their intended course of action ($r = 0.63$, $p < 0.01$). Both the total number of options generated and the number of task-relevant options generated were also positively correlated with the ability to *generate* the best option ($r = 0.50$, $p < 0.05$; $r = 0.72$, $p < 0.001$, respectively). The generation of task-irrelevant options was not related to any measures of performance.

During response trials without time constraint, the number of *task-relevant* options generated was not correlated with selection of the criterion best option as their intended course. However, there was a trend towards this variable being postively correlated with the ability to *generate* the criterion best option ($r = 0.36$, $p = 0.10$). In addition, there was a non-significant trend toward both the total number of options generated and the total number of irrelevant options generated being negatively related to the ability to *rate* the criterion best option as best approached significance ($r = -0.36$, $p = 0.11$; $r = -0.37$, $p = 0.10$, respectively).

DISCUSSION

As time pressure increased, participants generated significantly fewer options during both prediction and response trials, as would be expected by current models of skilled option generation. During prediction trials, the data suggests that both types of information—task-relevant and -irrelevant options—were reduced equally. However, the correlation data provide some additional insight in this regard. In accordance with Ward et al. (2012), when under time constraint the generation of task-relevant options approached being significantly and positively related to anticipatory performance (i.e., anticipation of actual outcome and most threatening option as the outcome)—rather than negatively correlated with the total number of options as predicted by TTF. When the time constraints were removed, the strength of these correlations increased, reaching statistical significance. These data support the assertions of LTWM theory. Moreover, the provision of additional time to access the encoded representation of the situation permitted a more accurate retrieval of relevant information. These data have implications for training prediction skills: Given more time to access the encoded representation, trainees are more likely to retrieve task-relevant information more reliably. Ultimately, the goal would be to develop exercises that shift the trainee toward reliable generation of task-relevant information under increasing amounts of time pressure.

During response trials, the data suggest that the reduction in time available to process information when under time constraint resulted in the selective reduction of the amount of task-irrelevant information being generated. While this information reduction strategy appears highly consistent with the TTF heuristic—i.e., by stopping the generation of additional lower quality options—the correlation data suggest that a TTF strategy was employed only during decision making without time constraint. The observed trend was toward a negative relationship between the *total* number of options generated and generating the criterion best option (but not accurately rating or selecting the best). Counter to our expectations and the TTF prediction, when under time constraint, the number of task-relevant options generated was significantly and positively correlated with all measures of decision quality (i.e., generating the criterion best option, accurately rating the criterion best option as best, and selecting the criterion best option as their intended course of action). These data also have implications for training decision skills: If permitted more time to deliberate over their encoding of the situation, participants are less likely to generate and select the best option. Accordingly, the training goal would be to develop exercises that improve more rapid and intuitive decision making without necessarily reducing that intuition to a single option—because the number of task-relevant options generated was positively (rather than negatively) correlated with decision quality when under time constraint.

In addition to providing novel results to the field by exploring the effects of time constraint on decision making during both prediction and response, our results offer an explanation of the seemingly contradictory findings in previous research. When under no time constraint, the LTWM and TTF hypotheses provide potential mechanisms that would explain prediction and response behaviors, respectively. However, when put under time constraint, contrary to our expectations, participants shifted to a LTWM strategy during response, rather than vice versa.

REFERENCES

Belling P. K., & Ward, P. (2012). Evaluating the Take-The-First heuristic in assessing situations and decision making using an option generation paradigm in soccer. *Journal of Sport & Exercise Psychology*, *34*, S210.

Calderwood, R., Klein, G. A., & Crandall, B. W. (1988). Time pressure, skill, and move quality in chess. *The American Journal of Psychology, 101*(4), 481-493.

Chabris, C. F., & Hearst, E. S. (2003). Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science*, *27*(4) 637-648.

Ericsson, K.A., & Kintsch, W. (1995). Long Term Working Memory. *Psychological Review 102*(2), 211-245.

Gobet, F., & Simon, H. A. (1996). The roles of recognition processes and look-ahead search in time constrained expert problem solving: Evidence from grand-master-level chess. *Psychological Science 7*(1), 52-55.

Johnson, J.G., & Raab, M. (2003). Take The First: Option-generation and Resulting Choices. *Organizational Behavior and Human Decision Processes 91*, 215-29.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G.A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making inaction: Models and methods* (pp.138–147). Norwood, NJ: Ablex.

Raab, M., & Johnson, J. G. (2007). Expertise-based differences in search and option-generation strategies. *Journal of Experimental Psychology: Applied 13*(3), 158-170.

Ward, P., Ericsson, K.A., & Williams, A.M. (2012) Complex Perceptual-Cognitive Expertise in a Simulated

Task Environment. *Journal of Cognitive Engineering and Decision Making 6(4).* doi: 10.1177/1555343412461254

Ward, P., Suss, J., Eccles, D. W., Williams, A. M., & Harris, K. R. (2011). Skill-based differences in option

generation in a complex task: A verbal protocol analysis. *Cognitive Processing: International Quarterly of Cognitive Science, 12*, 289-300.

Zsambok, C. E., & Klein, G. (1997). *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Trust and Uncertainty management

Cultivating Good Will Amidst Hostility

Helen KLEIN[a], Gary KLEIN[a]
[a]*MacroCognition LLC*

ABSTRACT

Introduction: The purpose of this research was to identify the skills of Good Strangers: military personnel who are successful in building trust in civilian populations. **Method:** Twenty-four soldiers and Marines were interviewed using a Cognitive Task Analysis method that relied on critical incidents. These interviews yielded 48 critical incidents. **Results and discussion:** Analysis of these incidents showed that soldiers and Marines with a Good Stranger mindset are alert to opportunities to build trust. They also depend on the skills of perspective taking, gaining voluntary compliance, and de-escalating situations.

KEYWORDS

Trust; Cognitive Task Analysis; Culture; Mindset; Unconventional warfare.

INTRODUCTION

The purpose of this project was to understand the skills required by military personnel to gain voluntary compliance from civilians, rather than relying so heavily on threats and intimidation. Historically, U.S. military personnel were trained to be warfighters engaged against traditional national forces. They were prepared in the strategies and tactics of the battlefield that involved well-defined adversaries and clear rules of engagement. Soldiers were to adopt a 'warrior' mindset. However, United States troops are now more likely to engage paramilitary forces and nationless adversaries in unconventional actions that are less well defined and predictable. Adversaries can melt away when they are losing and re-emerge when they regain strength. In the context of unconventional warfare, a warrior mindset is still necessary but no longer sufficient.

During unconventional warfare, military personnel must take on the additional tasks of 'nation-building'. They are often charged with fostering the stabilization of local and national governance. They help rebuild infrastructure and human services. To succeed in these tasks they have to cultivate and retain good relations with citizens. These also make it more difficult for insurgents to gain a foothold, sabotage nation-building, and mount attacks against U. S. forces. Simultaneously managing two roles – warfighter and nation-builder – presents difficult challenges.

This research was designed to describe how the nation-building role plays out during the uncertain, complex, and potentially hostile setting of unconventional warfare. We asked military personnel how they managed conflicts, gained compliance, and de-escalated conflicts.

To approach these questions, we first conducted interviews with skilled police officers. Like military personnel during unconventional operations, police officers must fight crime and keep citizens safe but must also maintain safety at public events, provide emergency services, and support community development. Many police departments teach strategies for gaining voluntary compliance and de-escalating conflicts. We described police officers that were highly effective in these areas as having a "Good Stranger" mindset. They looked for opportunities to foster trusting relationships with civilians. They didn't take provocation personally but rather worked to de-escalate conflict and secure voluntary rather than coercive compliance.

Not all police officers are able to adopt the Good Stranger mindset. Some may have a personal need to dominate while others lack the social skills needed to engage strangers. These officers elicit anger from citizens and endanger both citizens and fellow officers. Peers didn't like to go out on patrol with them; when possible they are transferred to jobs away from civilians.

METHODS

Approach

The goal was to learn about the social skills and cognitive tools that best support the decisions and actions military personnel face in hostile environments. We interviewed military personnel about their field experience in potentially dangerous, hostile and uncertain settings. We listened for the cognitive and social challenges of unconventional operations as they engaged with citizens. The interviews provided a rich array of narratives reporting incidents encountered during deployment.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Sample

The sample consisted of 24 military participants with experience in Iraq and Afghanistan, and in some cases Japan, Bosnia, or other sites. They were drawn from the U.S. Marine Corps (n = 8) and from the U.S. Army, including the Army National Guard (n = 16). The participants ranged in rank from E-5 (several years of military service) to a retired Colonel.

Procedures

Using Cognitive Task Analysis (CTA) methods (Crandall, Klein, and Hoffman, 2006), we collected critical incidents the interviewee had experienced with civilians and with local military personnel during overseas deployments. The interviews lasted 39-123 minutes. The interviews were recorded and yielded 595 pages of transcripts. We asked about specific incidents that they had found to be challenging. We first asked the interviewee to provide an overview of the entire incident. We then guided each interviewee to identify key components, to describe the role of the key players, and to anchor these events in a time line.

Most interviews explored one or two incidents, a few had as many as three or four. A few lacked any incident with sufficient detail to allow scoring. We excluded one story that came prior to a participant's military experience as well as six undeveloped stories. A total of 48 incidents were used in this analysis.

We scored and recorded the specific features of each incident along with the initiating event or the assignment of the incident and the interviewee's early sense of the situation. This included goals, threats, and key participants. The incidents involved a variety of missions, such as transporting supplies, providing bank protection during money deliveries, or conducting check stops. The material that follows includes examples from our interviews. The details of the examples have been altered to guard the anonymity of participants while retaining the dynamics of the interaction.

RESULTS AND DISCUSSION

The first and most powerful impression from listening to the interviews and from reviewing the transcripts was the differences among interviewees. Many of the military interviewees could articulate the circumstances that moved them to a Good Stranger Mindset. We found examples of effective strategies for taking the perspective of others and accommodating differences. Finally we identified strategies that are successful in maintain positive and trusting relationship even amidst hostility.

The Good Stranger Mindset

Origins

The interviewees showed great variety. Different experiences contributed to Good Stranger mindsets. Early family experience was important for some people. One interviewee reported: "It's probably from my family. At dinner, you could have an argument but you always listened to what the others said. That way everyone come out getting along. Now when I walk into a meeting I'm prepared for arguments but I try to listen to others needs and come up with plans that make my actions more acceptable to people." Another interviewee said: "For this mission, I went back to how I was raised: to greet people, to smile at people, to shake hands, because that's what I knew and what I'm comfortable doing."

Some interviewees adopted a Good Stranger mindset as a rejection of negative experiences: "When I just started there was a lot of issues with hazing. Things happened that I didn't particularly care for. I decided I was going to do the opposite of what some of the people I encountered did."

Prior work experience also shaped interaction with citizens. Interviewees with experience as police, social workers, or missionaries would mention these past experiences. One officer described a less usual experience: "One summer, I was a security guard at a bar. People were there to have fun. But people can get angry if somebody pushes them and it's crowded and hot. One of the things I learned was how to defuse the tensions before you have a situation. I used this 'defuse the crowd concept' overseas. I'm good with groups where not everybody gets along as long as they have a common goal."

Finally, some interviewees were driven to be more effective. They adopted the Good Stranger mindset because it helped them accomplish their missions. One reported: "To get anywhere with Afghanistan or Iraqi nationals the nicer you are to them, the nicer they are to you." Similarly, another said, "I try to develop personal relationships and trust. It's the only way we are going to win the war with the right people." He continued, "As you worked with the Iraqis and the Afghanis you learn what works and what doesn't work. You just have to figure out who is around you and what motivates them. How do you get them to accomplish or do what you want them to do?"

Perspective Taking

The Good Strangers try to get into another person's head, to feel what another might feel and to sense what the person would see as a satisfying resolution. If a Good Stranger can imagine how it would feel if people were to push their way into his own home, he is more likely to seek a less offensive way to carry out his mission. A warfighter, who can feel the loneliness of an Iraqi man working with U. S. soldiers, is more likely to learn some local language. For example, one soldier explained how he managed interactions respectfully, "When I worked a check stop, I try to keep in mind how it would feel to have a stranger from a foreign country stopping me and demanding my identification." Another reported, "I want to treat these people like I would want my family treated at home."

Culture imposes limits on perspective taking. Interviewees often arrive in country lacking knowledge of cultural variation. They might understand the universal imperative of demonstrating respect but not the culture-specific ways of doing this. Assuming Western signs of respect to be universal can convey the wrong message. Direct eye contact may indicate respect one place while in another one should divert one's gaze to convey respect. While a U.S. officer may describe a threat in a calm controlled manor, in other cultures, a calm, unemotional style would suggest that the information was no importance or urgency. Knowledge of cultural variation fosters effective social interaction with divergent populations.

One interview reported, "I knew this one guy. Before we left stateside, he found a local restaurant run by people from the nation we were being deployed to. He'd gone and ate the food; he'd actually talk to these people. And he'd ask them, 'What's it like going over there? Do you still have family?' And they were like, 'Yeah, this is going on.' And when we'd talk to him, he'd be able to tell us; 'This is what these people are saying; it's like this over there right now.' It made me realize how important it was to think about the people we were going to help. When you can get into the heads of the people you'll encounter it's easier to work respectfully and effectively."

Preemptive Actions

Interviewees regularly reported that they tried to find time to walk around and greet people when passing through towns. Their missions in local community took on the additional goal of 'friendly presence' as they chatted over coffee at a local stall. Another interviewee reported that his unit always carried bottled water, prized by local people, to hand out. One unit carried soccer balls and initiate games with kids when their assigned missions allowed time. Finally, we were told how a routine check stop mission could turn into an information-gathering mission. After a friendly greeting, a driver going through his checkpoint used this brief and private contact to offer information about a suspected terrorist who has moved into an abandoned house in his neighborhood.

De-escalation

Some of the incidents reported show great mastery of de-escalation in difficult situations. During one tense period, rumors started that a shrine in a neighboring district had been desecrated and worshippers mistreated. An angry crowd moved towards the base. The situation was very explosive. The officer in charge was able to separate the leaders of the group and to ask them to sit down and talk. He explained that he did not believe these rumors to be true but that he would have them investigated and report back to them. As their discussions went on a long time, the crowd dispersed and the leaders went home. He quickly followed up by investigating the accusations and confirming their falsehood with photographs. In the following week, he brought the pictures to all of the leaders involved in the initial confrontation and discussion. He asked them to bring any future concern to his unit's commander. The interviewee indicated that there had been no further trouble in that region during the time he remained.

Voluntary Compliance

The military is sometimes charged with enforcing laws that are not popular. Soldiers and Marines are prepared to use their physical force to achieve compliance but this creates animosity. One interviewee described a successful incident he observed where force was avoided and voluntary compliance achieved through the use of verbal strategies. Here's how he described the incident: "We were working at a checkpoint. We stopped a black BMW and found the driver had a pistol but not the needed documentation. The man started out angry and really scared. The soldier in charge kept a defensive posture to protect himself and the others present. At the same time, we talked to the man respectfully. We told him the rules related to the various weapons and explained that since he didn't have the required documentation, he would have to leave the gun it until documentation could be produced. The man was not happy but the discussion stayed civil. He would interrupt a little bit and there would be an explanation going back and forth. In the end, we took the weapon, end of story. Instead of turning it into a knock–down-drag-out there on the side of the road, the checkpoint continued to flow."

Repair

Not every problem can be avoided. Military personnel also need skills to repair mistakes. Soon after his arrival, a military officer visited a local community to inspect progress on a construction project. While he was walking around the site with village representatives, his troops waited near their trucks. A large number of people gathered to watch the event. As he returned to the trucks with the village workers, his men stood up preparing to leave. Suddenly the sound of gunfire filled the air and screams were heard from the assembled villagers. At first, the officer thought it was an ambush and he would need to have his troops rush through the crowd to prevent further damage. Before he could issue the order, he heard, "Sir, I fired by mistake." The officer immediately assumed responsibility and ordered his medics to look at two workers who had been shot. The medics reported that one had very minor wounds and the second would require medical help. The officer immediately ordered that a vehicle immediately take the seriously wounded man to medical care and that another one take his family members to the hospital as well. His medic attended to the minor wounds. The next day he returned alone and without a weapon to repeat his regrets. He found the community to be very appreciative because of his immediate acceptance of responsibility and medical helps. The village remained friendly throughout his stay.

The Power of Trust

Trust appears to be critical for establishing and maintaining the good will of civilians and even adversaries. Mayer and his colleagues (Mayer, Davis, & Schoorman, 1995; Schoorman, Mayer, & Davis 2007) working primarily with Westerners in organizational setting, identified Benevolence, Integrity, and Ability as dimensions critical for establishing trust during encounters. Their research suggests that trust can reduce hostility; increase information flow, and garner co-operation. It eases negotiation, and increases operational effectiveness during complex and dynamic interactions. We define a Good Stranger as a warfighter who seeks to increase trust from the local populace during all kinds of encounters.

The dimensions used to assess trustworthiness vary significantly by national group and situation (Klein et al., 2013). Dimensions beyond Mayer's three, include cognition (Markus & Kitayama, 1991); social values and priorities (Hofestede, 1980) as well as Affect, Non-Verbal Communization, Dialectical reasoning, Interdependence, and Status. These cultural variations make it more difficult to interpret information and correctly anticipate responses and decisions. Cultural differences are particularly important in hostile environments where misunderstandings are common and consequences potentially fatal.

The Power of the Good Stranger Mindset

Interviewees who were vigilant for opportunities to demonstrate their trustworthiness were particularly successful during social interactions. They wanted to believe that conflict resolution was attainable. While not letting down their Warrior mindset, they looked for resolutions that would increase trust among citizens. This positive stance or mindset appears to be a prerequisite for skilled social interaction performance. Warfighters who failed to cultivate a Good Stranger mindset did not strive for empathic and respectful interactions. The Good Stranger mindset is about wanting to increase trust from beginning to end of an encounter, and boosting the trust given to the organization, not just the individual.

ACKNOWLEDGEMENTS

This work was supported by the Defense Advanced Research Projects Agency (government contract 06-1825383). The views, opinions, and/or findings contained in this paper are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense. We appreciate the support provided by Joseph D. Borders and James C. Whitacre.

REFERENCES

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to Cognitive Task Analysis.* Cambridge, MA: The MIT Press.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related value*s. Newbury Park, CA: Sage Publications.

Klein, H. A., Lin, M., Miller, N. L., Militello, L. G., Lyons, J., Stokes, C. K., Finkeldey, J., Klein, G (in review). Mapping the terrain of multinational trust.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*, 709–734.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of organizational trust: Past, present, and future. *Academy of Management Review, 32,* 344–354.

SPURS: A framework towards Scenario Planning for Unexpected events, Response, and Startle using research, horror films, and video games

Jane BARNETT[a], B.L. William WONG[a], Rick ADDERLEY[b] and Michelle SMITH[b]
[a]*Middlesex University, London, UK*
[b]*A-E Solutions Ltd, Warwickshire, UK*

ABSTRACT

Introduction: Decision makers face major challenges during crisis events as unexpected events may occur at any time. The startle reaction that may generate from these events can have a negative effect on subsequent decision-making. Therefore, it is important for personnel to train for the ability to quickly recognise and react quickly to unexpected cues, in order to regain control of the situation. **Method**: Using research, and learning lessons from horror films and video games, the SPURS framework provides design guidelines that can assist with the building and configuration of simulated training scenarios containing unexpected events and cues. **Results and discussion:** SPURS factors are chosen in order to illustrate how a simulated training scenario with unexpected events may work. As a consequence, successful training can be acquired where the trainee sees what needs to be done, quickly assesses resources, and implements timely decision-making in order to regain control of the crisis.

KEYWORDS

Learning and training; uncertainty management; planning and prediction; engineering.

INTRODUCTION

Crisis events are not pre-scripted and unexpected events can occur at any time. Therefore, decision makers face major challenges when faced with uncertainty, and time pressures (Tversky & Shafir, 2003). A potentially dangerous consequence of an unexpected event is the startle reaction, where emergency personnel may momentarily pause, unsure of what to do next, which can cause delay or failure of subsequent decision-making (Barnett, Wong, Westley, Adderley, Smith, 2011; 2012). Therefore it is important for personnel to train the ability to recognise and react quickly to cues generated by unexpected events, in order to regain control of the crisis. Simulations provide realistic training scenarios because they allow expected and unexpected cues to be manipulated, which 'force' the trainee to consider long-term projections under time pressure (Lafond, DuCharme, Gagnon & Tremblay, 2012). This paper introduces our Scenario Planning for Unexpected Events, Response and Startle (SPURS) framework, which can be used to inform the design of unexpected events arising as a consequence of information misperception (Level 1 Situation Awareness (SA)). As a consequence, it can be used to train decision making under pressure and uncertainty. As a basis, we draw from research investigating decision making, procedures vs. complexity, and information misperception (Level 1 SA), in addition to exploring how directors design unexpected events in horror films and video games.

Factors Affecting Decision Making

Unexpected events may startle personnel but also interrupt primary complex task performance because the additional cues generated by the event may lead to information cognitive overload (e.g. Altmann & Trafton, 2004). Other factors that negatively affect the accuracy and quality of decision making are time pressure (e.g. Payne, Bettman, Johnson & Luce, 1995), and confidence about decisions (Smith, Mitchell & Beach, 1982). Cognitive readiness is "the mental preparation an individual must establish and sustain to perform effectively in the complex and unpredictable environment …" (Fautua & Schatz, 2012, p. 277), which requires a high level of intelligence, affect control, and social skills in order to be able to adapt to the changing environment, recognise patterns (ibid.), and develop intuitive decision making skills (Yancy, 2006). Once acquired, emergency personnel can successfully perceive Level 1 SA factors, understand their meaning (comprehension - Level 2 SA), and predict future events and consequences (projection – Level 3 SA) (Endsley, 1988, 1995).

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Procedures, Complexity, and Cognitive Processing

Procedures require little to no decision-making (Klein, 2009) as only a low level of attention is needed to complete the task (Shiffrin & Schneider, 1977). Similarly, the SRK (Skill, Rule, Knowledge, Rasmussen, 1983) model suggests three types of decision making – 'Stick and Rudder', where skill governs decisions and behaviour are automatically and subconsciously processed and require little to no effort. Rule-based decisions are made using IF/THEN statements, which may require some analysis beforehand, whereas knowledge-based decisions are more likely to be implemented in unusual situations where mental models are formed in order to ascertain the best course of action, and under time pressure, thereby increasing the likelihood of making errors.

The training of complex tasks involves separating procedures into smaller events to ease the learning process. However, even simple procedures have unexpected changes and personnel must quickly adapt in order to regain control by switching to controlled cognitive processing (e.g. creating mental models) based on what is already known. However, experiencing these events may momentarily pause controlled cognitive processing, and cause a delay or failure to make decisions (see Barnett et al., 2011, 2012; Haus et al., 2012), particularly when important information was obscured, missing, or unavailable.

Information Misperception (Level 1 SA)

Emergency personnel may scan their environment in order to maintain a high level of SA, and to react quickly to unexpected cues (Sneddon, Mearns & Flin, 2006). However, crisis events are stressful, which may reduce the effectiveness of information distribution systems (Hermann, 1963) by omitting information, delaying decision making (Miller, 1960), or incorrectly fill information gaps (e.g. Suedfeld, 1971). Failure to understand the environment caused a series of serious incidents on an offshore oil-rig, with 66.7% of them being related to Level 1 SA information misperception (Sneddon et al., 2006). Similarly, but with regards to unexpected events, recent research (Barnett et al., 2011, 2012) found that personnel experienced startle reactions caused by information misperception, where information conflicted with existing cues, changed meaning over a short period of time, or important information was missing.

INTRODUCTION TO SCENARIO DESIGN FOR UNEXPECTED EVENTS

The SPURS framework was developed in order to create unexpected events that could be injected randomly within a larger main scenario (e.g. a train crash). However, we first investigated startle research, and horror film and video game design in order to understand how to create and maintain uncertainty, and how to design scenarios where important information changes meaning.

Table 3. The SPURS design factors.

Design Factor		A startle reaction can be created by including one or more of the following:
2	**Real-world research**	• Prioritize information – information should conflict with existing cues • Formulate handling strategy – Information should change meaning over a period of time • Communicate information – Important information should be missing
3	**Environmental design**	• Content must provide a degree of anticipation • Engagement must be prolonged so the trainee remains attentive • Problem solving should not be too difficult to avoid trainee frustration • Provide alternative possibilities to problem solving that can be evaluated • Both desirable and undesirable outcomes should be achieved
4	**Create vulnerability**	• Invaluable information must be temporarily or permanently irretrievable (e.g. burned in a fire). • Scripting variations • Stage a series of false alarms
5	**Camera angles**	Unexpected events should occur in • Wide-open spaces • Narrow confined spaces
6	**Proximity**	• People or objects should intrude into the trainee's personal space • Unexpected events should occur outside of the trainee's peripheral vision • Incorporate unexpected events that are both distant, and close to the trainee's location
7	**Obscurity and everyday objects**	• Should change their meaning at some point • Trainee tools and everyday objects can behave in unexpected ways • Partially or fully block the view of items and/or locations
8	**Sound effects**	• Implement sudden noises (bangs, shouts etc.) related to the present task, or just random
9	**Synchronization**	• The above effects can be synchronized with each other to increase complexity

Development of the SPURS Framework

The process consisted of three steps. First, using the CUUES interview methodology (Barnett et al., 2012) with emergency personnel provided us with training aims and initial design direction (Table 1, Section 1). Second, interview results revealed three factors relating to Level 1 SA (Table 1, Section 2). Third, we conducted a literature review that investigated the ways in which video game and horror film directors incorporate design factors relating to uncertainty, that as a consequence, produce audience startle reactions. The review provided us with a further seven design factors – environmental design (Section 3), creating vulnerability (4), camera angles (5), proximity (6), obscurity and everyday objects (7), sound effects (8), and synchronisation of these factors (9). The main points are summarised here.

Research suggests that the startle reaction varies as a function of level of anxiety. If an individual is highly anxious, their startle reaction will be more intense (Grillon, 2008). Therefore, it is important to create an e*nvironmental design* that will generate and maintain anxiety by designing scenes that produce fear, false hope, uncertainty, and helplessness (e.g. Frome & Smuts, 2004). Recent research (Windels et al., 2012) explored gamers' startle responses while playing horror video games. They suggested four recommendations for generating and maintaining uncertainty - *anticipation*, *engagement*, *scripting*, and *false alarms*, which in turn, which should *create vulnerability*. Information misperception can be created in a number of ways. For example, by varying scripts, where either too much (overload) or too little (filling gaps) information is provided, by incorporating false alarms (e.g. a loud crash) before the real unexpected event begins, or by limiting the use of the in-game camera to look around. Restrictive *camera angles* have successfully induced stress in horror games such as *Resident Evil* (Konami, 1996; 1998; 1999), where gamers were forced to make a series of quick decisions when choosing escape routes through zombie infested alleyways (Figure 1).

Figure 2. Game play in Resident Evil 3: Nemesis.
Source: http://uk.gamespot.com/resident-evil-3-nemesis/images/220763/.

Once uncertainty is created, unexpected events can be incorporated in various ways. For example, the *proximity* of an unexpected event may cause startle depending where it occurs in relation to the trainee's location. Early research (Landis and Hunt, 1939) suggested that startle intensity varied as a function of proximity. If it occurred near or on the edge of peripheral vision then the likelihood of startle would be higher, compared to further away where the trainee could place the event into context. *Tom Clancy's Splinter Cell* series (Ubisoft, 2002-2012) require stealth mode to progress but uses *sound effects* to add to the stress. Enemies may still spot the gamer and unexpectedly appear behind them firing bullets and shouting loudly to alert nearby colleagues. Research (Windels et al., 2012) using the horror game *Dead Space* (EA/Visceral Games, 2008-11), suggested that sudden noises produced startle in gamers (e.g. a roof vent falls to the floor), particularly when the location of the noise was unknown. Finally, the *obscuring of everyday objects* increases uncertainty and can be incorporated as information misperception. The horror game, *Silent Hill* (Konami, 1999), is a town covered in an omnipresent fog making gamer visibility limited, and increasing vulnerability. Berg, Norin Persson & Ögren (2006) found that they spent 31.5% of the time watching the edges of the fog at one-second intervals in order to determine the meaning and location of sudden noises. The fog limits the information the gamer can perceive (clues and threats), which forces them to use their imagination to replace what they cannot see (Girard, 2011).

To increase complexity these factors can be *synchronized* and/or randomized to an extent, in order to provide training for different levels of expertise. The horror film, *Alien* (Scott, 1979), contained a startle scene where the crew are searching for the alien but when opening a suspect locker, instead find their trapped angry cat (Figure 2). By synchronizing three effects - close-up visuals (proximity), sudden shouts, and the screech of the cat (sudden noises), and the low-angle camera shots (restricted view), Scott produced intense startle reactions from audiences.

When designing unexpected events using SPURS, it is not necessary to incorporate every single design factor at the same time, but instead, choose factors appropriate for specific training requirements. The following section provides an example of this.

EXAMPLE SCENARIO DESIGN USING SPURS

Assuming a main scenario is in place, unexpected events can be injected beforehand on a timed basis, or as training progresses. Initial training may consist of simple procedures, which require automatic processing. The instructor may then increase complexity by injecting a SPURS-based unexpected event, which requires the trainee to switch from automatic to controlled cognitive processing. The following example illustrates a SPURS-based unexpected event, which was injected into a main scenario designed to train decision making in the event of a train crash with a civilian vehicle. The factors used to plan the scenario are in brackets.

Training scenario

The trainee arrives on the scene and is required to perform a familiar procedural task and collect evidence surrounding the crash (procedural, automatic processing). Part of the train suddenly shifts and crashes to the ground (unexpected event 1; sudden noise; false alarm; uncertainty), and startles the trainee, who, after recovery, nears the burning vehicle where two fire fighters are extinguishing the fire. Suddenly, the vehicle explodes because the vehicle contained two hidden gas canisters (unexpected event 2; sudden noise; high anxiety; obscurity). The fire fighters are thrown backwards as a result of the explosion and the trainee relays the events to the on-scene commander (prioritise communication; communicate information; resist pressure). A panicked colleague suddenly starts shouting from immediately behind the trainee (unexpected event 3; proximity; sudden noise).

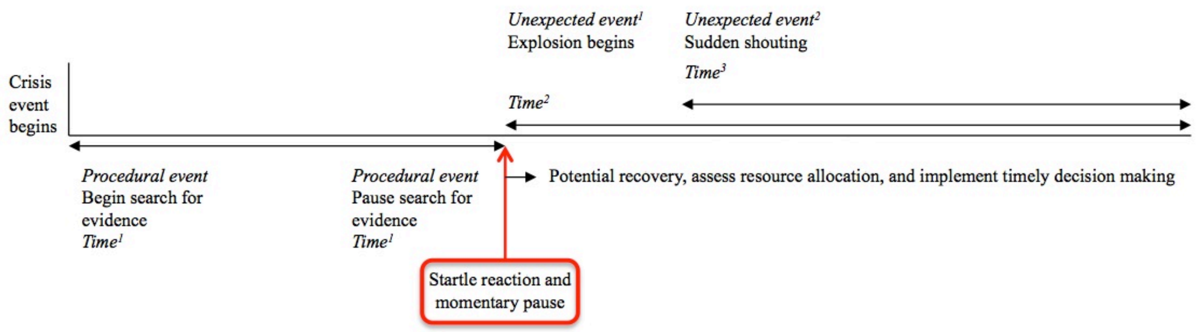

Figure 3. A timeline of the procedural and unexpected events

Training Aim: Take control of the situation by,

1. Resisting pressure from the first unexpected event in order to take control of the second and third.
2. Implementing decision making and communicating the events to the commander in a timely manner.

Outcomes:
Desirable: Completing the training aims.
Undesirable: The trainee is startled by the events, and is unable to implement timely decision making.

Synchronization
The event factors can be synchronized and randomized to a point (Table 2). For example, the explosion occurs in a wide-open area but could also be confined to a restrictive space.

Table 4. Examples of synchronizing and randomizing the design factors.

Synchronize all or some factors	Difficulty level	
	Novice	Expert
Trainee proximity	Far away from explosion	Near to explosion
Environment	Open area	Narrow area
Trainee proximity + environment	Far away + open area	Near + narrow area Near + open area
Sudden noise (explosion) Sudden noise (shouting located behind trainee)	Far Far away shouting	Near Immediately behind trainee
Proximity + Environment + Explosion + Shouting	Far	Near

DISCUSSION

This paper introduced a framework, SPURS, which can be used to create uncertainty and design unexpected events in simulated training scenarios. The factors are not exhaustive and can be adapted for further scenarios that train for task interruption (e.g. Gillie & Broadbent, 1989), inattentional and change blindness, and sensory information overload. Research suggests humans experience a wide visual array of stimuli, and are often unable to detect fully visible objects, or obvious unexpected changes to the environment (Simons & Chabris, 1999). Additionally, when two or more 'noises' compete, attention permits selective listening of relevant events whilst disregarding non-relevant information, because it cannot process all sensory input simultaneously. Therefore, incorporating dichotic listening tasks may be important for training the understanding and recall of important information.

Simulations permit trainees to learn about the nature of, and actively engage with unexpected events. Therefore, simulations should replicate real-life in order to develop the level of intuition required for effective decision making (Rousseau, 2003). Increasing event complexity permits training towards the acquisition of higher levels of expertise and complex cognitive skills such as developing mental models. Successful training can be acquired where the trainee sees what needs to be done, quickly assesses resources, and implements timely decision-making in order to regain control of the crisis.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° FP7-242474.

REFERENCES

Altmann, E.M. & Trafton, J.G. (2004). Task interruption: Resumption lag and the role of cues. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*: Erlbaum.

Baird, R. (2000). The startle effect: Implications for spectator cognition and media theory. *Film Quarterly*, *53*(3), 12-24.

Barnett, J., Wong, B.L.W., Westley, D., Adderley, R. & Smith, M. (2011). Startle points: First impressions of a preliminary framework exploring decision-making during uncertainty. Proceedings of the 10th International Conference on Naturalistic Decision Making (Florida, USA).

Barnett, J., Wong, B.L.W., Westley, D., Adderley, R. & Smith, M. (2012). Startle reaction: Capturing experiential cues to provide guidelines towards the design of realistic training scenarios. Paper presented at the 56th meeting of the Human Factors & Ergonomics Society, Boston, MA (October 2012).

Berg, Norin Persson & Ögren (2006). Looking for the invisible: Eyetracking within limited visual spaces. *Unpublished student paper*. Gotland University, Sweden.

Endsley, M.R. (1988). Situation awareness global assessment technique (SAGAT). In *Proceedings of the National Aerospace and Electronics Conference* (pp. 789-795). New York, NY: IEEE.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32-64.

Fautua, D.T. & Schatz, S. (2012). Cognitive readiness and the challenge of institutionalising the "New" versus "News". *Journal of Cognitive Engineering & Decision Making, 6*(3), 276-298.

Frome, J. & Smuts, A. (2004). Helpless spectators: Generating suspense in videogames and film. *TEXT Technology, 1*, 13-34.

Gillie, T. & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research, 50*, 243-250.

Gilpin, D.R. & Murphy, P.J. (2008). *Crisis management in a complex world*. New York, NY: Oxford University Press Inc.

Girard, P. (2011). *The fear system: Triggering tension in survival horror games*. Bachelorarbeit.

Haus, M., Rooney, C., Barnett, J., Westley, D. & Wong, W. (2011). Evaluating the effect of startling and surprising events in immersive training systems for emergency response. Paper presented at the 56th meeting of the Human Factors & Ergonomics Society, Boston, MA (October 2012).

Hermann, C.F. (1963). Some consequences of crisis which limit the viability of organisations. *Administrative Science Quarterly, 8*, 61-82.

Klein, G. (2009). *Streetlights and shadows: Searching for the keys to adaptive decision-making*. MIT Press Books.

Lafond, D., DuCharme, M.B., Gagnon, J-F, & Tremblay, S. (2012). Support requirements for cognitive readiness in complex operations. *Journal of Cognitive Engineering & Decision Making, 6*(4), 393-426.

Landis, C. & Hunt, W.A. (1939). *Startle pattern*. Farrar & Rinehart Inc.: New York.

Miller, J.G. (1960). Information input overload and psychopathology. *American Journal of Psychiatry, 16*, 695-704.

Payne, J.W., Bettman, J.R., Johnson, E.J. & Luce, M.F. (1995). An information processing perspective on choice. In J. Busemeyer, R. Hastie & D.L. Medin (Eds.), *Decision Making from a Cognitive Perspective*, New York: Academic Press.

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, 13*(3), 257-266.

Rousseau, C. (2003). Complexity and the limits of modern battlespace visualisation. *Canadian Military Journal, 4*(2), 35-44.

Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190.

Simons, D.J. & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*, 1059-1074.

Smith, J.F., Mitchell, T.R. & Beach, L.R. (1982). A cost benefit mechanism for selecting problem solving strategies: Some extensions and empirical tests, *Organisation Behaviour & Human Performance, 29*, 370-396.

Sneddon, A., Mearns, K. & Flin, R. (2006). Situation awareness and safety in offshore drill crews. *Cognition, Technology & Work, 8*, 255-267.

Suedfeld, P. (1971). Information processing as a personality model. In: M.H. Schroder & P. Suedfeld (Eds.), *Personality Theory & Information Processing*, 3-14. New York: Ronald Press.

Tversky, A. & Shafir, E. (2003). *Preference, belief, and similarity: Selected writings*. Cambridge, MA: MIT Press.

Windels, J., McAllister, G., Smith, A., White, G. & Mirza-Babaei, P. (2012). *Scary game findings: A study of horror games and their players*. Retrieved from:

http://www.gamasutra.com/view/feature/6480/scary_game_findings_a_study_of_.php?print=1

Yancy, S.D. (2006). Adapting professional military education to the current threat (AU/ACSC/5694/2006-04). Maxwell Air Force Base, AL: Air Command & Staff College, Air University.

Naturalistic Decision-Making in Sports

What Makes Basketball Players Continue with the Planned Play or Change It? A Case Study of the Relationships between Sense-making and Decision-making

Anne-Claire MACQUET[a] and Koffi KRAGBA[b]
[a] French Institute of Sports (INSEP), Paris, France
[b] University of Paris-Est, Créteil, France

ABSTRACT

Introduction: This study aimed to identify relationships between the sense players made of situations and their decisions to continue with the planned play or change it. **Method:** Seven female elite basketball players from the same team participated in the study. A match was video-recorded. Post-performance interviews were conducted separately; each player was shown the video and asked to describe her activity in relation to the events observed. The interview data were used to identify how players made sense of situations and the relationships between sense-making and decision-making. **Results and discussion:** Results showed two steps in sense-making: assessment of the current situation, and anticipation of possible situation developments to manage risk. Results also showed that risk assessment led the players to continue or change the play. Results suggest players used situation recognition and mental simulation to make sense of situations. They changed the play to cope with risks, suggesting team resilience.

KEYWORDS

Rigor/Resilience; risk management; expertise; team functioning, elite sports.

INTRODUCTION

In recent years, there has been renewed interest in team functioning and team coordination in sport (e.g., Pederson & Cooke, 2006). Nevertheless, little is known about how team members undertake the tasks they face during a game, nor how they coordinate with other team members. To suggest some answers to this, Eccles and his colleagues conceptualized team coordination in sports by taking into account social and cognitive processes (e.g., Eccles, 2010; Eccles & Tran, 2012; Eccles & Johnson, 2009). Team coordination is the process of organising team members' actions in order to achieve the most positive outcome (Eccles & Tran, 2012). Team members' actions are arranged according to three dimensions, namely, type, timing and location. The type of action depends on the situation and team members' roles and competencies. Each type of action is carried out at a particular time and locus to enable the action to be successful.

Different studies showed that coordination among team members was facilitated by: (a) a shared mental model (i.e., common knowledge held by teammates; e.g., Pederson & Cooke, 2006); (b) "basic compact" (i.e., level of commitment of each individual to support coordination; Klein, Feltovitch & Woods, 2005); (c) action interpredictability (i.e., making his/her own actions predictable for team members; Klein et al., 2005); (d) communication (i.e., intentional and unintentional exchange of information among group members; e.g., Eccles & Tran, 2012); and (e) division of labour (i.e., different types of action undertaken by specific team members according to team-level planning; Eccles, 2010). Eccles (2010) suggested that studying coordination among players might improve knowledge of team functioning. He also stressed the lack of theory and research on team functioning and team coordination and called for studies focusing on when, how and why play develops.

In team sports, team-level planning seems to play an important role in coordination (Eccles & Johnson, 2009). This refers to the playbook defining plays (i.e., structured patterns of players' coordination and teammates' actions) for offensive and defensive options, according to: (a) players' roles within the team (i.e., guard, forward, and centre in basketball); (b) players' competencies; and (c) potential game development in relation to opponents (e.g., Eccles, Ward, & Woodman, 2009). Plays are planned before being undertaken by players. Plays are flexible and adaptive to allow players to adjust when the situation develops in a different way to that anticipated (e.g., timing of an attack). Little is known about what makes players continue with the play or change it during the game. In sports training and other domains characterized by uncertainty, high time pressure and high stakes, it is important to understand when and why team members decide to continue or change the plan.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

It is widely recognized that decision-making depends on sense-making (e.g., Klein, 2009). To better understand players' decisions within the game, it seems important to study what sense players make of situations. Sense-making has been theorised in the domain of Naturalistic Decision-Making (Klein, 2009; Klein, Moon, Hoffman, 2006a, b). Sense-making is the process of analyzing events retrospectively, explaining apparent anomalies, anticipating the future, and directing exploration of information. It consists of connecting the dots comprising events, knowledge, etc., to build a frame. It determines what is considered as a dot, in relation to our goal, competences, expectations, etc. Due to Simon's boundary rationality, we cannot connect all the dots. Under time pressure, getting more information could be counterproductive and it may be better to prioritise information and jump to conclusions. The frame can be added to or changed in the light of new information. Sense-making pertains to a double cycle comprising two steps: building the frame and maintaining it, with reference to Piaget's (1954) concept of assimilation, and reconsidering the frame and enriching it by new information, with reference to Piaget's (1954) concept of accommodation. This theory of sense-making seems useful in studying what makes teammates continue or change the play during the game. The present study aimed to investigate how players made sense of situations and the relationships between sense-making and the decision to continue or change the play.

METHOD

Participants

Seven elite female basketball players from the same team volunteered to participate. They ranged in age from 16 to 18 years (M= 17.3 years, SD= 0.5 year) and had been playing basketball for 4.5 to 10.5 years (M= 9 years, SD= 2 years). They had been playing for the same team for 1.5 to 2.5 years (M=2,3 years, SD= 0.5 year). Since the present study was conducted, five of the participants have represented France during the European Championships. Players were given pseudonyms to provide some degree of confidentiality: numbers 1 to 7. The study was approved by a local ethics committee.

Data Collection

A senior French Championship match was recorded by a camera placed in the stand near the centre of the court (lengthwise).

Self-confrontation interviews were conducted with each player separately the day after the match. Interviews were conducted regarding the fourth quarter of the match (i.e., final ten minutes). This was the only period when the two teams successively led the score. Earlier, the team being studied led the score. Many coaches stress that when the score is tight, players are more involved in the game in order to prevent the opponents from scoring. It could be argued that, due to the tight score, this fourth quarter had greater uncertainty regarding situation development and match outcome, and consequently more changes to planned plays. Interviews were recorded (M= 48.2 min, SD= 2 min) and transcribed.

Data Processing

Data processing was done using the constant comparative method (Corbin & Strauss, 1990). Two researchers processed the data. Firstly, each researcher divided the transcripts into meaningful units according to the decision to continue or change the play. Secondly, researchers identified what information players connected and how they connected it to make sense of situations. Thirdly, they identified the relationships between sense-making and the decision to continue or change the play. After each data processing step, data were constantly compared until saturation was reached, which occurred when no further meaningful unit and category were identified from the data. The two researchers compared their results and discussed any initial disagreement until consensus was reached. Interview transcripts were divided into 195 meaningful units.

RESULTS

Results are presented in two parts. The first deals with the sense players made of situations. The second refers to the relationships between sense-making and decision-making.

Sense Players made of Situations

Results showed that in order to decide whether to continue or change the play, players first made sense of a situation. They connected information about players' placement and movement on the court, players' competencies, tendencies and roles, and ball trajectory. Two steps were identified: (a) assessment of current situation development, and (b) anticipation of potential situation development. Players used solely the first step, or both steps, depending on the situation.

Assessment of current situation development consisted of building up a frame of the way the situation had developed. It was based on comparison between information pertaining to the expected situation development and actual situation development. Players checked the situation was developing as expected according to the playbook. For example, Player 2 said:

"I can see them and I know where they are and what they're going to do in this situation because I know them. I check what they're doing. I'm looking at [Player 7] and seeing she's coming into this zone as she is required."

Similarities between expected and actual situation development indicated that players did not perceive any risk about situation development.

Anticipation of potential situation development consisted of reconsidering the frame created and enriching it with new information. New information related to predicting a player's possible actions and the consequences on the game. Results showed two kinds of anticipation: (a) an opponent's or teammate's risky action which could jeopardize the success of the play; and (b) a teammate's action which was not required by the play, however, which could achieve a positive outcome and avoid risk-taking in comparison to the action required by the play. In the first kind, anticipation led the player to investigate the consequences of risky action and assess whether the risk was difficult to manage or manageable. Risk was considered difficult to manage when the player thought it would have been very difficult to make the play work. It was considered manageable when the player believed she could cope and make the play work. For example, Player 4 said:

"Her opponent might catch the ball after the free-shot. I'm supporting my teammate by guarding her opponent to make the play work".

In the second kind of anticipation, the player anticipated a different situation development to that expected according to the play. Anticipation led the player to change the play in order to achieve a positive outcome without risk-taking. The player anticipated such a development as a teammate was either better placed than her or another teammate with whom she was expected to coordinate and her placement was not risky, or she was more competent in this situation than required. For example, Player 3 said:

"I was supposed to pass to [Player 1] who was guarded and I saw [Player 5] close to me. She tends to play for herself and I felt the current situation was good for her: it was an offensive dual. I showed her I was ready to help her. I know she's able to play an offensive dual easily. I know that as she's going to receive the ball, she's going to want to play by herself. She's 1,93m tall [6'4"] and nobody can stop her."

Relationship between Sense-making and Decision-making

According to the playbook, players were required to undertake planned plays so that their teammates knew what would happen and could achieve positive outcomes. They were also required to change the play when situations developed in a different way to that anticipated. Results showed that players assessed risk which could jeopardize the success of the play, in order to decide whether to continue or change the play. Results showed four levels of risk: (a) no risk; (b) manageable risk; (c) risk that was difficult to manage; and (d) risk avoidance.

Checking that actual situation development was as expected from the playbook and was not risky led players to continue with the play (41% of total decisions). For example, while checking her teammate was well placed, Player 2 said: "As I see her, I guard my opponent, like I'm supposed to do."

Assessing risk as manageable led players to continue with the play. To cope with risk and achieve a positive outcome, players decided to be more involved while facing the opponent's action (22% of total decisions). For example, Player 1 said:

"I'm not focusing on my teammates because I'm on the ball. The opponent I'm required to guard might make a screen on my playmaker. I tell her about the screen and stay focused on my defence on the ball. I must watch this opponent's every move."

Assessing risk as difficult to manage led players to change the play. Players avoided taking the risk that the play might fail (18% of total decisions). For example, Player 7 said:

"My teammate wants to set me the ball. I'm supposed to shake off my defender but she makes an "over-play", so I make a "back-door". She's defending close to me so I can't get the ball. I get away. My teammate has to change wing to pass the ball."

Risk avoidance led the player to change the play because a teammate was better placed or more competent than the player required by the play (19% of total decisions). For example, Player 7 said:

"my teammates were making screens to allow me to go ahead with the ball. I saw [Player 2] was ahead and alone because her defender was in late. So, I set her the ball.

DISCUSSION

These results are discussed in two parts: (a) consistency of results to the sense-making theory; and (b) planned decision versus emergent decision.

Consistency of Results to the Sense-making Theory

As the theory of sense-making predicts (Klein, 2009; Klein, et al., 2006a, b), players assessed the situation by connecting dots pertaining to available information on situation development, in order to build a frame. In some situations, they enriched the frame by anticipating situation development from players' current actions, placements and competencies. Information on situation development was compared to that memorised from the playbook and known about teammates' competencies and tendencies. These results suggest that sense-making was governed by recognition of situation and mental simulation, consistent with the Recognition Primed-Decision Model (Klein, Calderwood & Clinton-Cirocco, 1986).

Results showed that assessment and anticipation of situation development were governed by the players' placements and actions, and teammates' competencies and tendencies. These results are consistent with Macquet's (2009) study on volleyball players' decision-making.
Results also showed that players anticipated specific scenarios and assessed risk-taking in order to consider whether risk was manageable. If they considered risk was manageable, they continued with the play; if not, they changed. Results suggest that risk assessment refers to the anticipate-adapt perspective developed by Shapira and described by Klein (2009). This perspective allows players to cope with complex, ambiguous and unpredictable situations and to manage risk.

Planned Decision Versus Adapted Decision

Results showed that players continued with the play as long as they considered they could make it successful. As situations became too risky, or to avoid risk-taking, players changed the play in light of the opponent's and teammate's actions. Continuing with the play suggests rigor in team functioning; it allows action interpredictability (e.g., Klein et al., 2005). Adapting the play to the evolving situation suggests resilience in team functioning. The results are consistent with existing theory about resilience engineering (Hollnagel, Woods, & Leveson, 2006; Klein, 2009). Hollnagel et al. (2006) described resilience engineering as a way of designing systems and organisations to be flexible in order to cope with unpredictable risks. Instead of using safeguards against previous threats, resilience engineering aims to improve the system capacity to reconfigure in order to cope with unexpected risks. One function of resilience engineering is reliance on system adjustment capacities by preparing team members to expect to face unpleasant surprises rather than trying to predict and prevent risks. In sports, few researchers have provided evidence of experts' engagement in such a process (Horton, Baker, & Deakin, 2005; Klein, 2009).
At the applied level, it seems important to develop team resilience as "a tactic for protecting ourselves against risk" (Klein, 2009; p. 249). Beyond developing plays for offensive and defensive options, coaches might use drills involving changing conditions (e.g., timing of an attack, number of players involved in a specific play) to force players to adjust coordination to changing conditions. This would contribute to developing the flexibility of using tactical solutions in changing game conditions, and the originality of tactical solutions chosen by teammates. Coaches might also ensure that players are realistic in interpreting situations. Optimistic bias seems to play a role in risk-seeking: individuals misread the risks, leading them to be overconfident, and to give greater weighting to successes than failures (Kahneman, 2011). Overconfidence might prevent team resilience.
This study presents a limitation. It did not feature other teams and matches for comparison.
In conclusion, the data tend to support the view that sense-making and more specifically risk assessment play a key role in coordination among players, leading players to use team-level planning, or change it to adapt to evolving game conditions. The continued study of coordination will improve our understanding of team functioning and team performance.

ACKNOWLEDGMENTS

This study was supported by a grant from the French Institute of Sports (INSEP). The authors are grateful to the players and their coaches for their participation.

REFERENCES

Corbin, J., & Strauss, A. (1990). Basics of qualitative research: Grounded theory. Procedures and techniques. Newbury

Park, CA: Sage.

Eccles, D. W. (2010). The coordination of labour in sports teams. International Review of Sport and Exercise Psychology, 3, 154-170.

Eccles, D. W., & Tran, K. B. (2012). Getting them on the same page : Stategies for enhancing coordination and communication in sports teams. Journal of Sport Psychology in Action, 3, 30-40.

Eccles, D. W., & Johnson, M., B. (2009). Letting the social and cognitive merge. New concepts for an understanding of group functioning in sport. In. S. D. Mellalieu, & S. Hanton (Eds.) Advances in applied sport psychology (pp. 281-316). London: Taylor & Francis Group.

Eccles, D. W., Ward, P., & Woodman, T. (2009). The role of competition-specific preparation in expert sport performance. Psychology of Sport and Exercise, 10, 96-107.

Hollnagel, E., Woods, D. D., & Leveson, N. (2006). Resilience engineering. Aldershot: Ashgate Publishing.

Horton, S., Baker, J., & Deakin, J. (2005). Experts in action: A systematic observation of 5 national team coaches. International Journal of Sport Psychology, 36, 299-319.

Kahneman, D. (2011). Thinking fast and slow. NewYork: Farrar, Strauss and Giroux.

Klein, G. A. (2009). Streetlights and shadows: Searching for the keys to adaptative decision-making. Cambridge, Ma: Massachusetts Institute of Technology.

Klein, G. A., Calderwood, R., Clinton-Cirocco, A. (1986). Rapid decision-making on the fireground. Proceedings of the Human Factors and Ergonomic Society 30th Annual Meeting, 1, 576-580.

Klein, G., Feltovitch, P. J., & Woods, D. D. (2005). Common ground and coordination in joint activity. In W. B. Rouse, & K. R. Boff (Eds.), Organizational simulation (pp. 139-184) NewYork: John Wiley & Sons Inc.

Klein, G., Moon, B., & Hoffman, R. (2006a). Making sense of sense-making 1: Alternative perspectives. IEE Intelligent Systems, 21, 70-73.

Klein, G., Moon, B., & Hoffman, R. (2006b). Making sense of sense-making 2: A macrocognitive model. IEE Intelligent Systems, 21, 88-92.

Macquet, A.-C. (2009). Recognition within the decision-making process: A case study of expert volleyball players. Journal of Applied Sport Psychology, 21, 64-79.

Pedersen, H. K., & Cooke, N. J. (2006). From battle plans to football plays: Extending military team cognition to football. International Journal of Sport and Exercise Psychology, 4, 422-446.

Piaget, J. (1954). Construction of reality in child. London: International Library of Psychology.

Cognitive Control in the Activity of a Handball Coach

Thierry DEBANNE[a], Christine CHAUVIN[b]
[a]University Paris-Sud Orsay, CIAMS , France
[b]University of South Brittany, Lab-STICC, France

ABSTRACT

Introduction: This paper focuses on the dynamic adjustment of the cognitive control mode used by a coach during handball matches. Two main dimensions characterize cognitive control modes: one is the level of abstraction (symbolic/subsymbolic) and the other the origin (internal-anticipative, external-reactive) of the data used for control. **Method:** Verbal communications between a coach and his team were recorded during four matches and coded using the MacSHAPA software. **Results and discussion:** Analysis shows that the coach adjusts his cognitive control according to the team performance. He favors a more abstract level of control as well as more internal data when the score is no longer in the team's favor. An identical analysis will be carried out with other professional coaches. The goal of such studies is to identify the cognitive control modes that are the most satisfying according to the main features of a match.

KEYWORDS

Practical applications; Cognitive field research; Team and organizational factors in complex cognitive work; Cognitive control; Coaching activity; Sport.

INTRODUCTION

The coaching activity is dynamic and chaotic (Bowes & Jones, 2006) and generated by on-going events. This is particularly true for team sport competitive situations, which are dynamic and complex environments (Fiore & Salas, 2006) and make up a specific subclass inside the generic class of dynamic situations.

This paper focuses on coaching activities for handball teams. It describes the first results obtained in a study that aimed at modeling the activity of a handball coach during the defensive part of a handball match.

Handball is a team sport in which two teams of seven players each (six outfield players and a goalkeeper on each team) pass a ball to throw it into the goal of the other team. A standard match consists of two periods of 30 minutes, and the team that scores the most goals wins. Substitutes may enter the court, at any time and repeatedly. Team handball, like other team sports, is an invasion game, but it is also a collision contact sport, as is ice hockey, where contact is necessary and integral to playing (Silva, 1983). In this type of game, the main objective of defensive players is to stop offensive players from reaching the goal. In order to do so, defensive players are organized in different systems to attempt to push opponents away from the score area and to be numerous between the ball and the goal.

This paper is divided into four main parts. The main cognitive and cooperative features of the activity of a coach are pointed out in a first section. The method used to analyze this activity is then explained. For this paper, the activity of one coach was analyzed on the basis of four different matches. Results are presented in a third part and discussed in a fourth and last section.

THEORETICAL FRAMEWORK

The activity of handball coaching has three main features: it happens in a highly dynamic situation, it applies to a team, and it has to take opponents' behaviors and intentions into account. This activity may be described in keeping with the theoretical framework designed by Hoc and Amalberti (1995), Hoc (2001), Hoc and Amalberti (2007).

Hoc and Amalberti (1995) proposed a cognitive architecture of Dynamic Situation Management (DSM) to account for individual cognitive activities (cf. Fig. 1). Largely inspired by Rasmussen's (1986) step ladder model of diagnosis and decision-making, this architecture distinguishes three feedback loops in terms of the abstraction level and temporal span of the decision.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

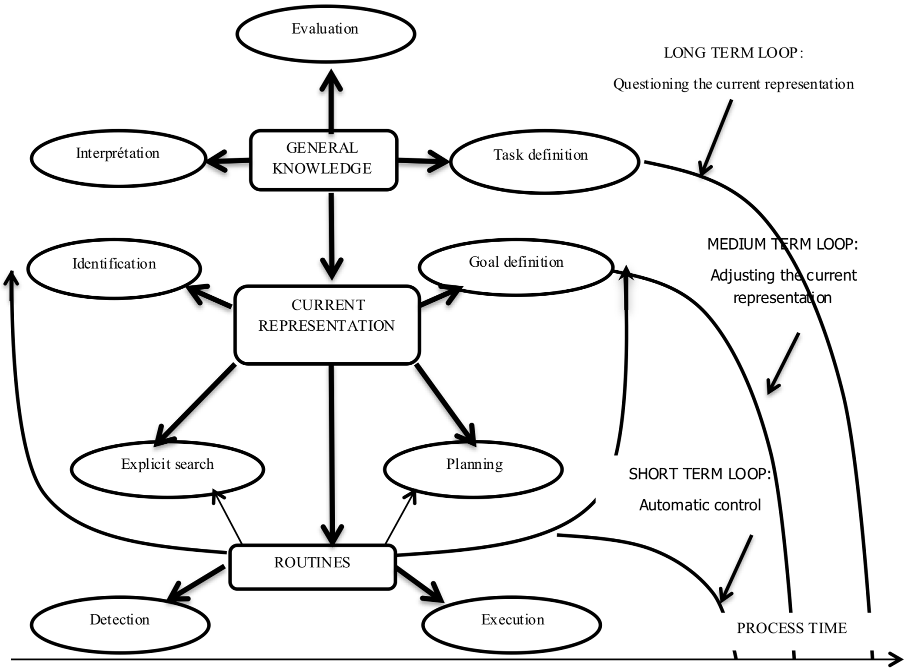

Figure 1. Dynamic Situation Management model (Hoc & Amalberti, 1995)

This model may be extended to team cooperation. Hoc (2001) thus proposed to analyze cooperation at three different levels:

Action level. At this level, cooperative activities consist in locally creating, detecting, anticipating, and resolving interference. Interference creation can be positive when it consists in mutual control or cross-checking. In this case, one agent can inform another of his or her disagreement or can give advice.

Planning level. At this level, a Common Frame of Reference (COFOR) is explicitly generated and maintained as a common representation of the situation. The situation includes the environment (currently referred to as Situation Awareness — SA) as well as the agents' activities. COFOR also includes common goals, common plans, and function allocation.

Metacooperation level. Just as the Planning level activity can facilitate the Action level activity, the Metacooperation level activity facilitates the Planning level activity. It includes high-level knowledge that is useful to other levels, such as mental models of the other agents and of oneself.

These models show that a specific system may be controled at different abstraction and temporal levels. In a more recent paper, Hoc and Amaberti (2007) distinguished the level of abstraction (symbolic/subsymbolic) and the origin (internal-anticipative, external-reactive) of the data used for control. Metacognition is considered as a means to distribute cognitive control within these dimensions in order to ensure mastery of the situation.

The consideration for the data origin remains the notion of "supervision span". This notion concerns information as well as action. From an informative point of view, the operator can have access to the process variables through a more or less limited window that can be defined in temporal, causal, or spatial terms, or at times, isomorphic terms. A restricted information span can result in difficulty to anticipate and to make decisions at the right time (Hoc, 2006).

In this paper, we propose to examine the mode of cognitive control used by a handball coach, that is, to identify its adjustments according to the situational features. Cognitive control will be analysed in terms of the abstraction level as well as in terms of the data used.

METHOD

Participant and Procedure

This paper focuses on the activity of one coach, observed during four matches of the top male professional French championship (coded M1, M2, M3, M4). He is 38 years old, has 4 years of experience as a professional coach and was selected 23 times in the French national team.

Three kinds of data were collected: (i) the main situational features (main events of each match); (ii) the coach's verbal communication with defensive players; (iii) self-confrontation interviews.

Each match was video-recorded from the first to the final whistle of the game. Verbal communication between the coach and the players was collected using the digital voice recorder connected to the microphone. Self-confrontation interviews with the coach were conducted during the week following the match.

Data Collection

The audiotapes from matches and self-confrontations were transcribed.
For the self-confrontation interviews, we used video recordings in natural settings and interview techniques of stimulated recall (Lyle, 2003).

Data Analysis

All data (situational features, coach communication, and self-confrontation data) were synchronised.
Verbal protocols (coach communication) were encoded using a general cognitive method introduced by Amalberti and Hoc (1998). It consists of inferring elementary cognitive activities from the overt behavior, the context, a general cognitive architecture, and knowledge of the application domain. Each verbal protocol was decomposed into elementary units. These units were coded using a predicate–arguments format, with the MacSHAPA software. MacSHAPA is a software tool built to help human factors investigators to carry out protocol analysis. It allows them to develop sophisticated coding schemes – relying on predicate calculus - and to use statistical routines to analyze the data once fully encoded (Sanderson et al., 1994).
The predicates were classified into several classes, according to the coding scheme defined by Hoc (2001) (see Table 1). Each predicate has several arguments, and each argument has several sub-arguments. For example, the argument "structure" has three sub-arguments: density (defenders' density), depth (movements towards opposite players), and defensive system (defensive team organization).

Table 1. Extracts of verbal protocols

Predicate	Main argument	Examples
CR-ITF (creation of an interference with a player's behavior [Action level])	Structure	*Don't remain flattened! Get out on the players!*
	Involvement	*Run! Come back!*
	Technique	*If you go back, you must fall with him.*
PLA-ACT-PLAN (plan maintenance or elaboration [Planning level])	Structure	*Change defense system, align, 0-6..*
	Involvement	*The first half-time, you forget it. Now you must impose yourself, physically, with your team.*
	Technique	*Don't hesitate to pull his arms.*
PLA-ENV (maintenance or elaboration of a shared situation awareness [Planning level])	Information	*He does not score as many goals as you think.*
	Comprehension	*Actually, we are in trouble only on Morgan's duels. Everything else is OK.*
	Anticipation	*Careful, they will play "Szeged" (name of combination).*
PLA-ACT-REPFCT (maintenance or setting of role allocation [Planning level])		*Cedric, you will change with Morgan.*
META (activities of metacooperation eliciting the different structures of knowledge used by the operators (general knowledge, task-related knowledge, knowledge about partners, etc.)		None
OTHER		

The verbal data were coded by both authors. The overall Kappa revealed a satisfying rate of agreement among the two coders ($k = .71$; $z = 25.94$, $p < .0001$). All the conditional coefficients were also high and significant.

Hypothesis

According to the Hoc and Amalberti model, we expect that the coach will adjust his cognitive control to the situational features. The relevant features of a handball match are the current phase (offensive or defensive), the score (in favor of the team or not), and the current period (first or second half-time). Cognitive control adjustments may concern the level of abstraction and the data used by the coach. Predicates refer to three different levels of abstraction (action, planning, and metacooperation levels). Arguments refer to different kinds of data evoked by the coach.
This general hypothesis will be proved using chi2 tests and the "transition" function of the MacSHAPA software, which enables the calculation of the frequencies and probabilities of transitions between predicates.

RESULTS

Results will be presented separately for the two half-times of matches and for the break period.

During Half-times

Predicate analysis

Predicates belonging to the metacognition level are almost non-existent. During the two half-times of each match, two predicates are predominant (cf. Fig. 2): CR-ITF (coach interference with players' actions) and PLA-ACT-PLAN (maintenance or generation of plans).

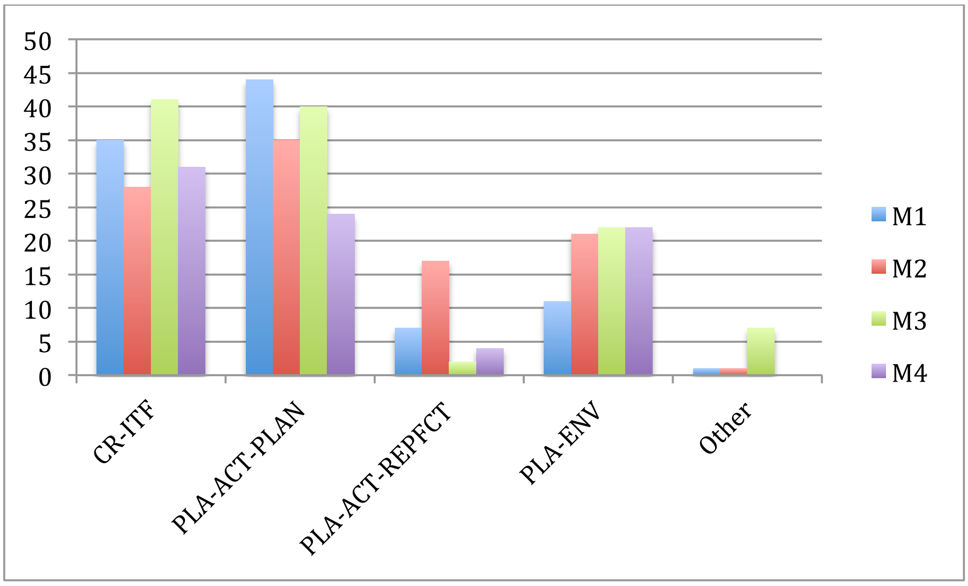

Figure 2. Breakdown of units into predicate categories during the two half-times

There are no significant relationships between their occurrence and the match period (first or second half-times). There is a significant link between the occurrence of the PLA-ACT-PLAN and the performance in only one match (M4). In this match, this predicate is more frequent when the team's score becomes unfavorable, $\chi 2$ (1, 81) =4.5, $p < .05$.

Above all, there are strong significant relationships – in all matches – between the occurrence of these predicates and the match phase (attack or defense). PLA-ACT-PLAN is far more frequent during offensive phases than during defensive phases ($p < 0.0005$ for three matches and less than 0.003 for the fourth one), whereas CR-ITF is far more frequent during defensive phases ($p < 0.0005$ for three matches and $p < 0.001$ for the fourth one).

The "transition" function of the MacSHAPA software confirms the importance of these two predicates. The probability that they initiate a sequence of several predicates is higher than 0.6. The probability for them to be present at the first, second or third place of a sequence is higher than 0.8. In contrast, the predicate PLA-ENV (related to situation awareness) appears to initiate the sequences with a probability lower than 0.3 and even, for some matches, lower than 0.2 (M1, second half-time of M2, first half-time of M3).

Argument analysis

Argument occurrence of the CR-ITF and PLA-ACT-PLAN predicates varies significantly according to the match period, in three of the four matches. In two matches (M1, M4), the plan deals mainly with structure in one half-time and mainly with players' involvement in the other. In two matches (M2, M4), the same may be said concerning interference.

In M2, there is a also a link between the "structure" argument and the score ($p < .05$). "Structure" is absent when the performance is balanced, and it becomes the main argument when the team's score is unfavorable. In this match, as in M1, there is also an obvious link between the score and one sub-argument of "structure": the defensive system becomes the predominant (and even the only) sub-argument as soon as the team's score is unfavorable.

As far as the situation awareness is concerned (PLA-ENV), one may notice differences depending on the match. They are related to predicate frequency (it is almost absent in the second half-time of M2 as well as in the first half-time of M1), to the object addressed (either own team or the opposite team), and to the data nature (information, comprehension, or anticipation). In most matches or half-times of matches, they deal mainly with the opposite team (M3, M4, first half-time of M2). They are mainly of "anticipation" type in the M4 two half-times as well as in the M2 and M3 first half-times.

During Breaks

Predicate analysis

During the break periods, three predicates are represented (Cf. Fig. 3): PLA-ACT-PLAN, PLA-ENV and PLA-ACT-REPFCT. In three matches (M2, M3 and M4), PLA-ACT-PLAN and PLA-ENV are the two predominant categories.

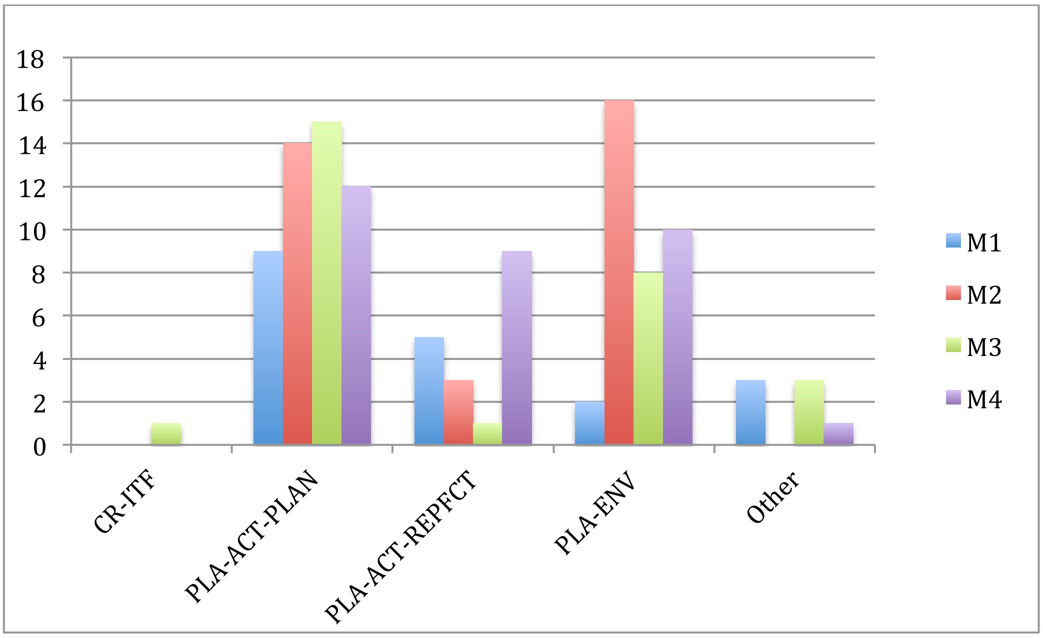

Figure 3 - Breakdown of units into predicates categories during the break

The probability – for a sequence – to begin with the predicate PLA-ACT-PLAN ranges between 0.4 and 0.55. The probability, for this predicate, to appear at the first, the second or the third place is higher than 0.9 for three matches (over 0.7 for M4). The probability, for the predicate PLA-ENV, to appear at the first or the second place of a sequence changes depending on the match: from a low value of 0.25 (M1) to medium and high value (0.5 for M3, 0.57 for M2 and 0.76 for M4).

Argument analysis

Arguments of the PLA-ACT-PLAN also vary depending on the match. They are mainly related to the team structure (depth) for M3, divided between the "structure" and "energy" categories for M4 and mainly related to energy for M1 and M2.

For M3 and M4, the arguments of PLA-ENV deal mainly with the opposite team, whereas they deal mainly with own team in M2.

DISCUSSION AND CONCLUSION

These results show two main adjustments in the cognitive control mode of the coach.

The coach alternates between two abstraction levels, depending on the match phase. During the offensive phase, both central defensive players are replaced by offensive players. The coach therefore takes advantage of each offensive phase to plan future actions for the central defensive players. In the same manner, the break period is mainly used to maintain or elaborate plans.

The cognitive control mode is adjusted according to the team performance.

In one match (M4), the level of abstraction rose (from action to planning) when the score was unfavorable. In this match, the score worsened in the second half-time; in the same time, the data used by the coach changed: the plan no longer concerned the players' involvement but only the team's structure. This kind of data is more "internal", since it relies on a mental representation of what a satisfying structure is. It also implies a more extensive information span (in causal and spatial terms) than data related to players' involvement.

Changes in the data used, according to the performance, are noticeable in two other matches. Structure becomes the main planning object in M1 when the score is unfavorable for the team. In this match, as well as in M2, it is – more precisely – the team's defensive system that is evoked by the coach.

The adjustment of the cognitive control mode may also depend on the opposite teams. In two matches (M3 and – to a lesser extent- M4), planning deals more with team structure than in the other matches. These two matches have several common features. In both cases, the opposite team was ranked lower than the coach's team. A victory was, therefore, possible. Furthermore, the stakes were high, since a victory would have ensured the team's presence in first division for the next season. The effect of the opposite team on the activity of the coach has been shown in a previous study (Debanne & Fontayne, 2012).

The consideration for the environment also varies depending on the match. The opposite team's features are taken into account mainly in M3 and M4 and are almost ignored in M1. These results may be explained in the light of the match stakes. Additionally, French teams (involved in M3 and M4) are better known to the coach than foreign ones (the opposite team is a foreign one in M1). This knowledge makes the situation awareness elaboration easier.

From a theoretical point of view, this study tends to show that the cognitive control mode migrates towards more abstraction and more complex data when the coach faces difficulties and/or high stakes. These results need to be confirmed by the analysis of data that have been recorded with two other coaches in eleven other matches.

These further studies should also enable the identification of the cognitive control modes that are the most satisfying according to the match main features.

From a practical point of view, these findings are expected to have an impact on coach training and, more generally, on the training of supervisors in dynamic and competitive situations. Following this study, we may suggest that coaches (a) should acquire, before each match, specific knowledge concerning the opposite team and opponent players, and (b) should learn to use routine procedures concerning the management of the defensive phase. They should make it a habit to make a systematic assessment of the defensive phase and plan future actions during the offensive phase.

ACKNOWLEDGEMENTS

We would like to thank the head coach, his assistant, and all the team players for allowing us to carry out this study.

REFERENCES

Amalberti, R., & Hoc, J. M. (1998). Cognitive activity analysis in dynamic situation. Why? How? *Le Travail Humain,* 61, 3, 209-234.

Bowes, I., & Jones, R. L. (2006). Working at the edge of chaos: Understanding coaching as a complex, interpersonal system. *The Sport Psychologist*, 20, 2, 235-245.

Debanne, T., & Fontayne, P. (2012). Communication from professional coaches to athletes during handball games according to the "Coaching Model" perspective. *Movements & Sport Sciences,* 76, 1, 11-23.

Fiore, S. M., & Salas, E. (2006). Team cognition and expert teams: Developing insights from cross-disciplinary analysis of exceptional teams. *International Journal of Sport and Exercise Psychology,* 4, 4, 369-375.

Hoc, J.-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations, *International Journal of Human-Computer Studies.* 54, 4, 509-540.

Hoc, J.-M. (2006). Dynamic situations. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors.* CRC Press.

Hoc J.-M., & Amalberti R. (1995). Diagnosis: Some theoretical questions raised by applied research. *Current Psychology of Cognition*, 14, 1, 73-101.

Hoc, J.-M., & Amalberti, R. (2007). Cognitive control dynamics for reaching a satisfying performance in complex dynamic situations. *Journal of Cognitive Engineering and Decision Making,* 1, 1, 22-55.

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29, 6, 861-878.

Rasmussen, J. (1986). *Information processing and human-machine interaction*, Amsterdam, North-Holland.

Sanderson, P. M., Scott, J., Johnston, T., Mainzer, J., Watanabe, L., & James, J. (1994). MacSHAPA and the enterprise of exploratory sequential data analysis (EDSA). *International Journal of Man-Computer Studies,* 41, 5, 633-681.

Silva, J. M. (1983). The perceived legitimacy of rule violating behavior in sport. *Journal of Sport Psychology,* 5, 4, 438-448.

A Naturalistic Decision-Making Investigation of Football Defensive Players: an exploratory study

Gilles KERMARREC[a] and Cyril BOSSARD[a]

[a] *Centre de Recherche sur l'Éducation, l'apprentissage et la Didactique (EA 3875), Université Européenne de Bretagne, U.F.R Sport et EP de Brest*
European Centre of Virtual Reality

ABSTRACT

Introduction: The aim of this study was to describe decision-making of football defensive player according to RPD model, and to investigate dynamic relationships between recognition processes and by-products used in naturalistic setting. Yet very few researchers have investigated the dynamic nature of decision-making, in sport real-world settings **Method**: Behavioural data was recorded from four high-level football players supplemented by verbal data collected during self-confrontation interviews. Seven critical defensive stages were studied. The data were analysed using a content analysis in five steps. **Results and Discussion:** Considering the dynamic of decision – making, 112 decisions were classified into three types of processes and 8 typical decisions for defensive stages. These findings are discussed from three perspectives: the RPD model highlighting the recognition processes of expert sport players; decision-making as an anticipatory thinking; the implications of these underlying mechanisms for training.

KEYWORDS

Decision making; recognition processes; anticipatory thinking; defensive stages; football players.

INTRODUCTION

A deep understanding of sport competition is essential for games because the success of coaches and players in such events is dependent on many qualitative and latent factors. Various studies in sport psychology over the past three decades have focused on decision-making in order to elicit "objective factors" of athletes' expertise (for a review, see Williams & Ward, 2007). Different aspects, such as visual search strategies (Williams, Hodges, North & Barton, 2006), memory processes (Zoudji, Thon & Debû, 2010), knowledge bases (McPherson & Kernodle, 2003), and the role of practice in decision-making (Kibele, 2006) have been examined. The decision-making context used in these studies was semi-experimental (schemata, slides, and videos).

The difficulty of setting out the "winning-factors" in games has led researchers to study activity in its complex and dynamical properties, and not only as an information-processing system (Sève, Saury, Ria & Durand, 2003; Lenzen, Theunissen & Cloes, 2009). Those studies have introduced theoretical models for studying the dynamic links between the athletes' cognitive resources and contextual constraints. Methods used were different recall techniques from a previously experienced case in natural competitive settings, verbal report and qualitative analysis. In reference to the theoretical approach of the course of action (Theureau, 1992), many studies concerned with sport have focused on the significant components of athletes' experiences: goals in sailing (Saury *et al*., 1997), knowledge building during tennis tables matches (Sève *et al.*, 2003), and emotions experienced in table tennis (Sève *et al.*, 2007). Inspired from the theory of situated action, Lenzen *et al.* (2009) studied the relationship between planning and action and the perception-action coordination in hand-ball. Mouchet *et al.* (2005) considered both first-person approach and the psycho-phenomenology propositions in order to analyze subjectivity of decision-making in rugby. In these games decision-making constitutes a practical fulfillment, situated in a particular context because circumstances are never identical (score, tiredness, players in presence, personal purposes). The assumption is that experts memorize "typical traces" for unusual, competitive, complex or difficult situations, i.e. situations involving "critical decisions" (Hoffman & Lintern, 2006).

In these psycho-phenomenological approaches of course of action, the researchers studied the situated experiences of expert athletes, but they didn't focus on the decision-making processes.

THE NATURALISTIC DECISION-MAKING FRAMEWORK

The NDM paradigm aims to improve the systems that help people to make choices in military, nuclear power or aviation settings. NDM examines the ways in which experts in real-world contexts, alone or in a team, identify

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013.* Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

and assess situations, make decisions and execute actions with consequences which are meaningful to both the actors and their environment (Lipshitz, Klein, Orasanu, & Salas, 2001). The features of the contexts in which experts' significant decisions are made relate to: ill-structured problems, ill-defined or competing goals, uncertain and dynamic environments, time constraints, high stakes, multiple event-feedback loops, multiple players, and organizational settings (Cannon-Bowers, Salas, & Pruitt, 1996). Parallels have already been drawn between working situations and sport situations (Fiore & Salas, 2006; Ward & Eccles, 2006).

The Recognition-Primed Decision Model

In NDM framework, the RPD model is an alternative way to the information-processing model in explaining experts' decisions under time pressure (Hoffman & Lintern, 2006). Klein (2008) refutes the idea that people in dynamic situations make decisions based on rational deductions or exhaustive analyses of expectancies. For Klein, experts use their experience to make decisions under time pressure. Experts confronted with dynamic situations are able to recognize typical situations and to associate an appropriate course of action (*i.e* pattern matching).

Conversely, the RPD model suggests three levels of experiencing the situation: *simple match*, *diagnosing the situation*, and *evaluating a course of action*. At the first level, the situation is quickly perceived as typical, allowing the expert to react directly by means of an action or sequence of actions. The second level occurs when the situation is perceived to be incongruous. The expert must clarify it by diagnosing it, focusing on its similarities with similar cases and hereby choosing an appropriate action. At the third level, the expert perceives the situation as typical but looks for a new solution and evaluates it through mental simulation before implementing a course of action. This process allows him to imagine how effective his action will be in the current situation.

In NDM perspective, experts use a holistic evaluation of the potential during the course of action. The first and most immediate option considered is therefore a plausible option (Lipshitz *et al*., 2001). For Klein (2008), both the assessment of the situation and decision-making during the action itself are supported by the recognition of spatiotemporal patterns. Ross, Shafer & Klein (2006) mentioned the "instantiation of a prototype" (p. 406). This prototype is a cognitive package that includes: information about the current typical situation (relevant cues), what to expect from this situation (expectancies), suitable goals, and typical action.

NATURALISTIC APPROACH OF DECISION MAKING IN SPORTS

RPD models have been used to study decision-making in various dynamic situations in work settings (for a review, see Ross *et al*., 2006), and are starting to be applied in sport settings (Omodei *et al*., 1998; Johnson & Raab, 2003; Macquet, 2009 ; Bossard *et al*., 2010, 2011). However, experimental procedure and quantitative analysis used by Omodei *et al.* (1998) and Johnson & Raab (2003) didn't describe accurately typical components or cognitive packages athletes use during course of action.

More recently, some researchers used RPD model for qualitative analysis of empirical data. Macquet (2009) has shown that expert volleyball players predominantly used a simple match (level one) to make decisions in competitive situations. The author explains that time pressure during certain tasks forces athlete to rely on the first level. In order to recognize a situation as typical and to react instantaneously (simple match), experts recognize typical components (opponents actions, trajectory of the ball, rules) in the course of action. This process allows them to retain meaningful and efficient actions. In the same perspective, Bossard *et al.* (2010) adapted RPD model to describe cognitive packages supporting decision-making during the counter-attack phase in ice-hockey. They identified ten "typical schemata" used by expert ice hockey players during successive stages of course of action. These schemata represent the underlying structures which link perceptive and cognitive elements and facilitate the recognition under time pressure. In this perspective, Bossard *et al.* (2011) used RPD model and schema theory in order to obtain a dynamic representation of experts' courses of action during a counter-attack training setting in football. These results are used to support participative simulation design for decision-making training.

These studies were based on the assumption of relatively stable cognitive structures are used to recognize situations. They didn't consider the dynamic aspect of decision-making, the refined assessment process, and the role of by-product in this process. Yet very few researchers have investigated the dynamic nature of decision-making, in sport real-world settings.

Purpose of the study

The purpose of this exploratory naturalistic case study was to elicit decision-making components and processes in a competitive football setting. Decision-making in football is a complex mechanism; particularly in defensive plays defenders must make their team-mates continually aware of potential problems regarding their own and the other team's locations and activities. This study focused on temporal analyses to describe course of action as successive decisions athletes have to make under time pressure and tried to propose a refined description of assessment mechanism and sub-products role in these process.

This qualitative exploratory study focuses on examining how defensive players contribute to the joint production of successful defense, and on identifying recurrent practices in the data, i.e. typical decision. In so doing, we will contribute to scant existing knowledge of the ways in which decision is typically achieved: which are the assessment processes, and the refined by-products / components experts used for decision making in critical situation.

METHOD

Participants and setting

Four international level football players from a professional football team in the French first League participated voluntarily in this study. The overall age was 25.5 (SD = 3,8 years) and they had played in this team for several months. Although the players did not ask to remain anonymous, they were given the following pseudonyms to guarantee some degree of confidentiality: Louis was left back, Robert was right back, Carl and Steeve were the central defenders. The players' decision-making was studied during a first league football match. Training, previous matches, feedback between players and coaches had developed team expertise for defensive stages, so that they were considered as one of the best defenders in the first league.

The team used zone defense or zonal marking. It is a tactic which is used in sports such as football where the players are made to guard a specific area of the field. In zone defense, if a defender is under pressure or in a critical situation their teammates must assist him, so that zone defense promotes or needs lots of interactions between defensive players.

Data collection

Two types of data were gathered: (a) observational data from continuous video recordings of the defenders actions from the beginning to the end of the defensive stage (from the loss of the ball till its recovering) and (b) elicited data from self-confrontation interviews.

The observational data were obtained by video recordings of the players' actions during the match. A digital camera was positioned high in the stand and was set for a fixed, wide-angle view, framing both the field and the players (see Figure 1).

Elicited data were obtained during self-confrontation interviews conducted after the match had been played. The interview techniques were derived from video recordings of players as they played; as the players viewed the video footage, they were asked to comment on the decisions made. The methodology of this interview is closed to the Critical Decision Method (CDM) developed to investigate experts' decisions in naturalistic settings. Self-confrontation interviews have been used by others researchers in natural sport settings (Lenzen *et al*., 2009; Macquet, 2009; Sève, Ria, Poizat, Saury & Durand, 2007).

The interviewer's prompts were related to description of the actions, thoughts, feelings and events as experienced by the player before, during and after each critical decision. According to RPD model, decision – making is a recognition process, and the recall process was facilitated by using some non-direct questions about feelings (how do you feel at this time?), perceptions (what are you looking for?) or focus (what is drawing your attention?), intentions (what do you want to do?), and thoughts (what are you thinking about?) associated with each decision. The researcher stopped the tape before each decision was made and asked the player to comment on his actions and the events leading up to the decision. Players could stop the video at any time to explain something that they deemed relevant.

The present self-confrontation interviews (duration, *M* = 45 min, SD = 5 min) were conducted the day after the game for each of the four players. To avoid potential bias, the coach agreed not to analyze the match on with the players until the interviews were over. The interviews were recorded in their entirety using a digital video camera and a tape recorder. This type of interview was supported by a formal contract of cooperation between the coach, the players and the researcher. All of the players and the coach had expressed a great deal of interest in the study. The players gave their permission for all recordings to be made available.

Data processing

The video of the match was viewed in order to choose critical stages for the study. Stages were identified as critical situations by the first researcher, another sport psychologist researcher and a football professional trainer. These stages were chosen because a) only the four defenders were between the player in possession of the ball and the goal; b) the player in possession of the ball drove it quickly to the defense area; or c) a long pass could put defenders under time pressure; d) finally, defenders had to stop opposite team attempt. Each defensive stage we chose was between ten and twenty seconds long, and achieved in production of successful defenses (i.e. the ball was intercepted). Finally, each four-defender team participated in seven critical defensive stages; they were characterized by the locus of the end of the adverse attack/the begin of the defensive stage.

Observable behaviors were systematically coded and organized into categories relating to the technical language of football, without making inferences about their intentions. The verbal exchanges between the players and the researcher during the interview were recorded and fully transcribed. These verbal data were processed in five steps: (a) generating defensive stages logs, (b) selecting and identifying decision making by-products, (c) analyzing decision making process thanks to previous coding (i.e. by-products), (e) identifying typical decisions, and (f) ensuring validity.

Generating defensive stages logs

This first step consisted of generating a summary table or log of the sectioned data for each seven defensive stages (Table 1). The main objective was to prepare the data for subsequent content analysis. For each participant, descriptions of the player's actions and comments were placed side-by-side in a three column table in chronological order for each situation studied. The first column references the situation in question, giving the participant's name and the reference number for the defensive stage, as well as the time at which the stage occurred in relation to the beginning of match. The second column lists the player's actions. A transcription of the player's verbalizations produced during the self-confrontation interviews can be found in the third column.

Table 1. *Example of Defensive Stage Log and Short Accounts*

Game highlights	Players' actions and behaviours	Verbal reports from self-confrontation interviews
First defensive stage		Robert – Right defender
	The ball was lost on the left side, about 40 meters away from the goal.	I see him losing the ball on the opposite side (I). Carl is guarding his opponent (I) I'm aligning on Steeve (A). (Level 1)
	The right winger passes to a midfield player (flick pass) and goes through the central defense.	We all guard (I), there is nobody behind me (I), I'm set in case of Steeve would be out of position (E), I check the alignement (I), I'm set to assist him (A). (Level 2)
	Carl intercepts the ball.	

Selecting and identifying decision making by-products

The second phase consisted of selecting data relating to players' decision-making. We used a category system derived from the RPD model (Klein, 2008) to code the salient features of decision-making. Defenders discourse, along with salient information, their own goals, the actions they have chosen and what they could expect, should enable us to identify the active cognitive packages (plausible goals, expectancies, typical action and relevant cues). We attributed a code for each of the salient features: *goals* (G), *action* (A), *information* (I), and *expectancies* (E).

Identifying recognition processes

Because we consider that the sports context could generate specific constraints, we decided to conduct empirical an inductive analysis of courses of action before comparing our results to the RPD model (simple match, diagnose and simulate). So, thanks to previous coding (i.e. decision making by-products), we looked for processes within the specific mechanisms the defenders used to assess successive situations and to make decision. We analyzed the verbal reports about each decision separately. In each short account, to clarify their decision, players established implicit relationships between by-products. They verbalized successively many components of a cognitive package, so that the chronology inside each short account could be considered as an indicator for identifying recognition process. For example a defender said: "I could see that the ball was on the right side (I), the attacker kicks a long pass (I) ; I run back (A)". In this example, relevant cues ("the ball on the right side" and "the attackers is …") are taken in account to assess the situation: so the defender decided to "run back". This approach aimed to elicit the dynamic of decision – making processes to be taken into account during the course of action. We noted/coded this recognition process: I(+I)---A. This kind of process, which included information perception led to action, was identified in 68 short accounts of decision-making. After each short account was analyzed in this perspective, we classified them into "typical processes" and compared the generated categories to those of the RPD model.

Identifying typical decisions

In this step, we conducted an empirical categorization of the data (Strauss & Corbin, 1998) on the different decisions made by the four defenders. We gathered together decisions made in the same way by different players during the seven critical defensive stages. The typical decisions were defined and named at the end of the analysis.

Ensuring validity of generated categories (i.e. types of processes and typical decisions)

The theoretical and empirical categorizations were validated by three researchers who had already coded protocols of this type in previous studies, had prior experience in team sport, and were familiar with the NDM framework. The reliability of the coding procedure was assessed using Bellack's agreement rate; the initial agreement rate was 80% for by-products coding, 90% for typical decisions and 90 % for types of recognition process.

RESULTS

The results will be presented in two stages: (a) the decision processes experienced by defenders; (b) typical decision of expert football defenders.

Processes and content of decisions experienced by expert football defenders

The analysis of the seven critical defensive stages suggested that players make tactical decisions based on combinations of significant information. By adopting a first person approach to dividing up the course of action, 112 successive decision-making situations were identified for the four players. We tried to elicit relationship between the significant by-products of the cognitive packages by considering the dynamic of the activity threw the chronology of self-confrontation interview. Considering the dynamic of decision – making, the 112 decision were classified into three types of processes. And, considering the content of each decision-making, we classified the 112 decisions into 8 typical decisions. The distribution of the processes related to what kind of decision was made is presented in table 2.

The first type of processes consisted in relevant cues invocation immediately conducted to action. In this case, critical information was perceived and the situation was assessed familiar to him, so that the player could rapidly choose an action. For example a defender (Paul) said: "I could see that Mark is following his opponent on the right side (I), I should cover back the central area (A)". Sometimes the process conducted the defender to explicit knowledge (" I have to protect central area; it's the most important to protect this zone; it's the best way to the goal"), so that it looks like a justification about decision-making. This process consisting in some kind of perception – action association, could led defenders to expected outcomes: "I could see that the ball was on the right side (I), the attacker could kick it for a long pass (I), I have to run back (A); my teammates should run back too, so that we'll stand on the same line to prevent a long pass (E)". In this example, it is remarkable that there is not only one relevant cue to be mentioned; the recognition process combined two relevant cues to active a course of action. These kinds of recognition process were coded: I---A; I+I--- A (table 2). This type of process, which included information perception associated to action, was identified in 68 short accounts of decision-making. According to the RPD model, if the player quickly assessed the situation to react immediately, the recognition process should be a simple match (level 1). Results in table 2 shows that this first type of process is the only process defenders used when the ball was the main relevant cue. Close to the ball and their opponent, they had to match immediately perception and action.

In the second type of process, the defender perceived one (or more) relevant cues that conducted him to attempt an evolution for the course of action (expectancies). In this case, information perceived was associated to expectancies, and then, the defender could see another relevant cue that conducted him to choose an action. For example a defender (Mark) said: "I could see that my opponent run behind me (I), I'm ready, I can run back to cut his way (E)". In this example, Mark's expectancy was a preparation to action. The players assess the situation in one way (but he wasn't sure it was the good option) and wait for significant information to confirm plausible options or goals. In some case, this first option could be invalided: "we've just lost the ball on the left side (I), far away from the goal (I), I should stand in my area because the attackers in possession of the ball couldn't become dangerous (E), "but, now, he's passing on the other side (I), so I have to run and to assist the central backs (A)". In these course of action the decision making process began when the defender performed an usual action ("I go back to my place" ; "I'm looking at my opponents" ; "I continue to be aligned") ; the defender pays attention for relevant cues ("I saw him in my back") which generated expectancies ("I feel he can run to the goal"). This type of process was identified in 18 short accounts of decision-making; it was coded I(+I)---E---A or I(+I)---G---A, and it could be considered as a diagnose process. According to the RPD model, we could consider that if the player can't assess the situation immediately, he has to diagnose, i.e. estimate the situation, and attempt for relevant cues before decide what to do (level 2).

The third type of processes consisted in identifying significant information which permitted to imagine the future of the ongoing situation. In this case, the defenders weren't directly concern with the ball; player perceived critical information, and the situation was assessed familiar so that the player could anticipate the future course of action. Knowledge led them to verbalize global assessment or anticipate the consequences of course of action so as they were more in observation rather than in action. For example a defender (Robert) said: "Long pass far forward (I); I know that Carl run really fast (K), he'll trap the winger in possession of the ball (E); so if I just assist the central backs, it's a good option for us (C)". It was coded I---K---E---C.

Sometimes the process conducted the defender to imagine a course of action, to get ready without any action to do: "I see the attacker alone in possession of the ball (I), we (the defenders) are aligned (I). If he (the attackers) challenges Carl (E), I'll have to help him, it's my role (K)! But on the other hand, he (the attacker) used to reverse the game I'll have to protect the goal (K); it's a critical moment for us (C)". This type of process, which included information perception associated to knowledge to generate expectancies, was identified in 26 short accounts of decision-making. According to the RPD model, if the player recognizes the situation and take time to verify it's a good option, he could simulate action (level 3).

Table 2. *Decision and types of decision-making process during defensive stages*

Main relevant cues	Types of processes	Types of process T1 I---A	T2 I---E---A or I---G---A	T3 I---K---E---C	Total
The ball, the attacker in possession of the ball	1. To track and intercept the ball (a pass or a cross)	8	-	-	**8**
	2. To block or track one attackers driving the ball	15	-	-	**19**
	3. To control and move quickly behind his attackers	7	7	-	**14**
One or more attackers	4. To move backward together and control a plausible long pass	8	6	-	**14**
	5. To look after his attacker and to stay quite close from his area	-	5	8	**13**
One or more team-mates	6. To assist a teammate (who had to challenge an attacker)	7	-	8	**15**
	7. To line up or get organised	15	-	6	**21**
	8. To move to the central defence and to protect the goal	8	-	4	**8**
	Total	68	18	26	**112**

DISCUSSION

This study aimed to extend current research by systematically examining relationships between underlying by-products and recognition processes used by expert when making decision; especially, it was expected that the dynamic of relevant features taken into account during the course of action could led us to decision-making processes. The results showed that elite football players' decision-making was based on three types of processes. These processes could be related to the RPD model recognition processes. If the situation were perceived as typical, the defenders used simple match when they had to a ball trajectory or to attackers 'movement. Sometimes (i.e. when the defenders were far away from the ball), typical situations were simulated in order to anticipate the consequences of course of action. If situations weren't perceived as typical, they were diagnosed and defenders paid attention for more information to generate expectancies. Complementarily, the results showed that elite football defenders' decision-making was based on eight typical decisions. The findings will be discussed from two perspectives: the recognition processes of expert sport players' decision-making and their consequences for training.

Consistency of findings to the three levels of the RPD model

The findings of the present study suggest that the players' decision-making was based on relationships between by-products and recognition processes of a typical situation. In previous study Macquet (2009) had compared the frequencies of the salient features considered for each level of the RPD model. In this study, association between salient features could be considered as criterions to help the identification of decision making processes. If the situations were perceived at first as typical they were simple matched (I---A), if not they were diagnosed (I---E---A or I---G---A). In level 3, the players recognized the course of action, but they used information and knowledge to generate expectations and to anticipate the consequences of action (I---K---E---C).

In previous works on expert decision-making in sport games (Macquet, 2009), recognition processes in these dynamical context are more often a simple matching process than a diagnosing or simulating process. Our results are in line with these findings. Therefore, our results are inconsistent with previous research considering the frequencies of each level of the RPD model. Previous studies in sport games have focused on attackers' decision-making in ice-hockey or football (Bossard *et al.*, 2010, 2011) or on decision-making in volleyball, which is a specific sport game because the players had to defend their playground in the same time (the same action) they have to attack their opponents' playground. In defensive stages in football, defenders have to quickly react to attackers' movement or to the ball trajectories; if they are in urgency, they used simple matching (level 1). But if defenders are far away from the ball, they have more time to assess the situation and to simulate the future of the course of action (level 3). In our findings, they evaluated the situation becoming dangerous for their teammates ("if he is overtaken, I should have to help him") or for themselves ("if he can kick a long pass in my back, I'll have to run back"). In sport game, the aim for the opponents is to be creative, to challenge defenders with low-probability events. While attackers have to initiate actions, defenders sometimes have to resist to typical decision and managing or handling higher levels of uncertainty (level 2).

Future research should aim to explore in more detail how defenders estimate how dangerous is a situation. In our findings, distances between defender and his opponent, distances from the line, from the goals, attackers speed, are precious relevant cues. The situation became particularly difficult for defenders when those spatiotemporal features weren't perceptible: "I can't see the ball, it's dangerous, I'm looking for my opponent and I'm late". In

these sequence the ball is a temporal marker and the line a spatial marker; they led defenders to "the feeling of urgency" for the situation assessment. In other words, the assessment of the situation and decision-making during the action itself seems to be supported by the recognition of spatiotemporal patterns (Klein, 2008).
Most of defenders recognition processes in this study are congruent with RPD model. According to Klein *et al.* (2007), these processes could be discussed toward "anticipatory thinking" and "sensemaking".

Decision-making as anticipatory thinking.

Anticipatory thinking is the process of recognizing and preparing for difficult challenges, many of which may not be clearly understood until they are encountered. Sensemaking often takes the form of explaining events and diagnosing problems, a retrospective process, and it can also take the form of formulating expectancies about future event. It is this future-oriented aspect of sensemaking that interests us here as an anticipatory thinking.
In previous studies about sport games, simple match was the main form of recognition process. Our results suggest that this first level of decision making process deals with two forms of anticipatory thinking: pattern matching and trajectory tracking. With pattern matching, the circumstances of the present situation bring out similar events and clusters of cues in the past: typical decisions could be "to move backward together", "to assist a teammate", "to line up", "to move to the central area and to protect the goal". In our studies, two types of trajectories can be seen as relevant cues. Pattern matching and trajectories tracking are used for the next immediate action they have to do: these processes could be related to "prediction" in Ward's proponents for underlying mechanisms of anticipation in soccer (McRobert *et al.*, 2009; Ward, 2003).
Moreover, we make distinction between tracking a ball trajectory and tracking an attacker, because of the additional complexity needed to track the possible trajectories of other actors in the decision space (Klein *et al.*, 2007). An attacker can change his own intention or/and his course of action. Defenders decisions are often building directly through typical experiences, and sometimes indirectly through patterns of meaning and response which can be used or blended to deal with an as yet unknown and unknowable future. In that way anticipatory thinking is aimed at potential events including low-probability high threat events, not simply the most predictable events. In this study, this diagnose process is used when defenders had "to look after an attacker and to stay quite close", "to control and move quickly", and "to move backward together" in order to get time to assess the course of action. This process could also be related to deep planning (i.e. considering potential alternative occurrences beyond the next immediate move; McRobert *et al.*, 2009; North *et al.*, 2011).
Then, anticipatory thinking relates to the "simulating process". This type of anticipatory thinking requires seeing the connections between events. It could be called conditional (Klein *et al*., 2007) and is supported by knowledge to generate pertinent expectancies. This is particularly consistent with some finding from the present study. Defenders used this kind of anticipatory thinking when they have to pay attention on the evolution of the course of action; they had "to look after his attacker and to stay quite close", "to assist his teammate", "to line up and stay organized", or "move to the central area and protect the goal". Instead of responding to a cue or to a ball trajectory, as in situational pattern matching, sometimes defenders need to appreciate the implications of different events and the consequences of course of action before choosing the most appropriate typical action. Thus the simulation process allows a positive, negative, or neutral assessment of the situation (Evaluation mechanism according to Ward, 2003).
Because experts in sport games have to be reactive and creative, and have to deal with both of high and low-probability events, expert performers in sport do not necessarily just recognize one option (Ward, Williams, & Ericsson, 2003). Thus, expertise is based upon the ability to predict and simple match, but also upon the ability to plan, to diagnose, to simulate and to evaluate, so that sport games should be a particularly interesting domain for studying underlying mechanism of anticipation and decision-making.

From experts' underlying mechanisms to training in decision-making

This study highlights the relationship between sub-products and types of processes in decision-making. According to the situated action paradigm and its related methods, these findings cannot be generalized to other populations and contexts in the form of a general model of decision making in game play. On the other hand, they reflect some of the underlying mechanism of expertise to be acquired for becoming a professional football player.
The results suggested that players' experiences led them to memorize typical decisions. The recurrence of units of meaning extracted from the experts experiences could be explain thanks to schema theory. Researchers in the field of NDM generally conclude that situation assessment is based on schema- or script-driven (Rentsch & Davenport, 2006). Our results showed that elite football defenders' decision-making was based on eight typical schemas. The implication is that training could be based upon these empirical evidences of expert performance Analysing the literature about decision-making training, Ward & al. (2008) regretted that researchers have used logical and tactical principles of action, rather than base training on expert data from the specific task and skill on which participants are to be trained (Ward & al, 2008). In this way, using defenders' schema and sub-products in order to implement a training regimen for defenders in soccer could be the first application of the expert performance approach to training in sport. In that way, findings about expertise or sport are used in a product-oriented approach of training (i.e. enhancing the content of training pratice, feedback, video simulation,

temporal occlusion training, with empirically derived data about expert performance). Training based on experts' performances provides large improvements in response time, and small benefits for decision accuracy (Ward & al, 2008).
Considering our findings, this products-oriented approach could be complemented thanks to a processes-oriented approach of training, including speed-processing (i.e. simple matching, predicting), but also in the aim to develop resistance in prior plausible options (diagnosing, deep-planning, simulations, evaluating).

ACKNOWLEDGMENTS

The authors would like to thank Yvan Bourgis for his help in collecting data.

REFERENCES

Cannon-Bowers, J. A., Salas, E., & Pruitt, J. S. (1996). Establishing the boundaries of a paradigm for decision research. *Human Factors, 38*, 193-205.

Fiore, S. M., & Salas, E. (2006). Team cognition and expert teams: Developing insights from cross-disciplinary analysis of exceptional teams. *International Journal of Sport and Exercise Psychology, 4*(4), 369-375.

Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In Ericsson, K.A., Charness, N., Hoffman, R.R. & Feltovich, P.J. (Eds.): *The Cambridge Handbook of Expertise and Expert Performance* (pp. 203-222). Cambridge: Cambridge University Press.

Johnson, J.G., & Raab, M. (2003). Take the first: Option generation and resulting choices. *Organizational Behavior and Human Decision Processes, 91*, 215-229.

Kibele, A. (2006). Non-consciously controlled decision-making for fast motor reactions in sports: A priming approach for motor responses to non-consciously perceived movement features. *Psychology of Sport and Exercise*, *7*, 591-610.

Klein, G., Snowden, D., & Pin, C. L. (2007). Anticipatory thinking. In K. Mosier & U. Fischer, (Eds.), *Proceedings of the 8th International Naturalistic Decision Making Conference* (pp. 120-127). CA: San Francisco State University.

Klein, G. (2008). Naturalistic Decision Making, *Human Factors, 50*(3), 456-460.

Klein, G., Ross, K.G., Moon, B.M., Klein, D. E., Hoffman, R.R., & Hollnagel, E. (2003). Macrocognition, *IEEE Intelligent Systems*, *18*(3), 81-85.

Lenzen, B., Theunissen, C., & Cloes, M. (2009). Situated Analysis of Team Handball Players' Decisions: An Exploratory Study, *Journal of Teaching in Physical Education*, *28*, 54-74.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Focus article: Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making, 14*, 331-352.

Macquet, A.C. (2009). Recognition within the Decision-Making Process: A Case Study of Expert Volleyball Players. *Journal of Applied Sport Psychology, 21*(1), 64-80.

McPherson, S. L., & Kernodle, M. W. (2003). Tactics, the neglected attribute of expertise. In Starkes, J.L. & Ericsson, E., (Eds.): *Expert performance in sports: Advances in research on sport expertise,* (pp.19-49). Champaign, IL: Human Kinetics.

McRobert, A. P., Williams, A. M., Ward, P., & Eccles, D. W. (2009). Tracing the process of expertise in a simulated anticipation task. *Ergonomics, 52*, 474-483.

North, J., Ward, P., Ericsson, K. A., & Williams, A.M. (2011). Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory, 19* (2), 155-168

Piegorsch, K. M., Watkins, K. W., Piegorsch, W. W., Reininger, B., Corwin, S. J., & Valois, R. F. (2006). Ergonomic decision-making: A conceptual framework for experienced practitioners from backgrounds in industrial engineering and physical therapy. *Applied Ergonomics, 37,* 587-598.

Rentsch, J. R., & Davenport, S. W. (2006). Sporting a new view: team member schema similarity in sports. *International Journal of Sport and Exercise Psychology*, *4*, 401-421.

Ross, K. G., Shafer, J. L., & Klein, G. (2006). Professional judgments and "naturalistic decision making". In Ericsson, K. A., Charness, N., Hoffman, R. R. & Feltovich, P. J. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 403-419). Cambridge: Cambridge University Press.

Sève, C., Ria, L., Poizat, G., Saury, J., & Durand, M. (2007). Performance induced emotions experienced during high-stakes table tennis matches. *Psychology of Sport and Exercise, 8*, 25-46.

Sève, C., Saury, J., Ria, L., & Durand, M. (2003). Structure of expert table tennis players' activity during competitive interaction. *Research Quarterly for Exercise and Sport*, *74*, 71-83.

Ward, P., Williams, A. M., & Ericsson, K. A. (2003). Underlying mechanisms of perceptual-cognitive expertise in soccer. *Journal of Sport and Exercise Psychology*, 25, 136.

Ward, P., & Eccles, D. W. (2006). A commentary on “team cognition and expert teams: Emerging insights into performance for exceptional teams”. *International Journal of Sport and Exercise Psychology, 4(*4), 463-483.

Williams, A.M., Hodges, N.J., North, J.S., & Barton, G. (2006). Perceiving patterns of play in dynamic sport tasks: Investigating the essential information underlying skilled performance. *Perception*, *35*, 317-332.

Williams, A.M., & Ward, P. (2007). Anticipation and Decision Making: Exploring New Horizons. In Tenenbaum, G., & Eklund, R. (Eds.), *Handbook of sport psychology* (pp 203-223). New York: Wiley.

Zoudji, B., Thon, B., & Debû, B. (2010). Efficiency of the mnemonic system of expert soccer players under overload of the working memory in a simulated decision-making task, *Psychology of Sport & Exercise,11(1),* 18-26.

Posters

Task Analysis of Coach Developers: Applications to The FA Youth Coach Educator Role

Andy ABRAHAM[a], Gareth MORGAN[a], Julian NORTH[a], Bob MUIR[a], Pat DUFFY[a]
Wayne ALLISON[b], Andy CALE[b] and Richard HODGSON[b]
[a]*Carnegie School of Sport, Leeds Metropolitan University, Leeds, UK*
[b]*Football Development Division, The (English) Football Association, Burton-upon-Trent, UK*

ABSTRACT
Introduction: This study attempts to understand the work of coach developers in the development of coaches. There is currently little understanding of what people in these roles do and, therefore, what their professional development needs are. **Method**: A total of 15 coach developer professionals were engaged in data collection. Data was collected through one or more of; interview, observation in one to one sessions, observation in workshop settings. **Results and discussion:** Using theoretical models to analyze data, an expert coach developer was defined through requisite professional skills, knowledge and typical behaviors in 6 domains. While this definition was achieved, caution is suggested since this definition is reflective of 'Frankenstein's monster' rather than reality. Due to the ill defined nature of the role 'coach educator', the defined domains should be used as benchmarks to measure personal practice against alongside the goals of the role and the context within which the role is undertaken.

KEYWORDS
Professional Judgment, Decision Making, Coaching, Coach Education, Coach Development, Professional Development.

INTRODUCTION

The work reported here reflects three separate but linked consultancy projects we were engaged in by Sports Coach UK, UK Sport and The (English) Football Association (The FA). The backdrop to this work in the UK is the emergence of the need to examine coaching as a profession both at a governmental level (Sports Coach UK, 2008) and sport federation level (The FA, 2010). While research in coaching has been around for about 40 years now, it has often only focused on identifying why expert coaches are expert or what makes an effective volunteer coach. This is not altogether surprising considering how relatively few well paid roles that there are in coaching, especially in comparison the huge numbers of voluntary coaches. As such, while there is a desire to improve the quality and professionalism of coaching there is relatively little research (especially in comparison to other similar domains such as teaching) to work from, however this is improving (e.g. Abraham & Collins, 2011; Côté, & Gilbert, 2009). However, despite this apparent change in capacity to understand coaching, there is virtually no research examining the role of the coach developer/educator. As such while research examining coaching expertise now exists, there isn't any evidence that this work is being delivered by coach educators to coaches.

The context of these projects therefore (drawn from the specific remits identified by the commissioning bodies) was to identify the demands of coach educators operating at a management and leadership level of coach development (the Sports Coach UK remit) and at the programme delivery and mentoring level (the FA Remit). The goal of both projects was to develop an informed view on the knowledge, skills and typical behaviors required to perform these roles. In addition to these core ideas this project reports back on how the differing roles within coach development may require differing application of skills and knowledge in actual behaviors.

Theoretical Framework – Professional Judgment and Decision Making (PJDM)

While there may be an argument to suggest that, with the relative immaturity of the coaching domain as a research milieu, especially in the area of coach education/development, a grounded approach may be necessary. However, given the demands of the commissioning bodies and the need for robust yet timely advice there was a clear need to avoid this long drawn out process and to borrow from existing theory. It is useful therefore that research in a broad range of occupations such as military, medicine and, importantly, sport (Abraham & Collins, 2011; Kahneman & Klein, 2009; Lipshitz, Klein, Orasanu, & Salas, 2001; Vickers, Reeves, Chambers, & Martell, 2004), has already highlighted that where judgment and decision making are crucial for effective performance, the key components of expertise are generally the same. Broadly speaking this research displays

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

that professionals make judgments that are rapid, intuitive and naturalistic in nature, often dictated by the time frame in which decisions have to be made, i.e. Naturalistic Decision Making (NDM). Furthermore the same research has shown that other judgments are slower, made with greater thought and deliberation and take a 'classical' approach to problem solving, i.e. Classical Decision Making (CDM). As such our assumption was that this theory of PJDM was relevant in this setting.

Supplementing the PJDM theories, other cognitive-behavioral and educational research has also been shown to fit this type of professional practice. Research from these areas has displayed that expertise is arrived at through the consistent application of metacognitive skills (i.e. thinking about how to think and learn more efficiently and effectively, Entwistle and Peterson, 2004) and mental skills (MacNamara, Button, and Collins, 2010).

The broad theories of PJDM, metacognition and mental skills bring a useful lens against which the role of the coach educator can be analyzed. However, providing greater definition is important if the work is to inform critical thought about coach development and professional development of coach developers. In order to bring greater definition to viewing the role of an expert coach developer, a model of coach development decision making (figure 1.) was developed by Abraham, Collins, Morgan & Muir (2009) that drew on similar work in coach decision making (e.g. Abraham, Collins & Martindale, 2006).

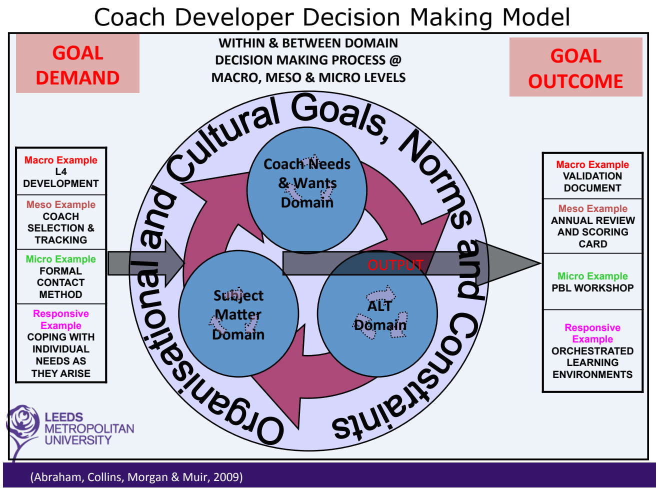

Figure 1. The Coach Development Decision Making Model (Abraham, et al 2009)

This model was developed based on a review of available evidence and literature and summarized six broad domains of understanding (i.e. the declarative knowledge) that should influence the decision making of coach developers. These six domains are summarized as:

1. DOMAIN - UNDERSTANDING CLUB AND FA CONTEXT, STRATEGY AND POLITICS – Understanding the culture of the situation that is being worked in and adapting behavior
2. DOMAIN - UNDERSTANDING THE COACH – Understanding the coach(es)'s motivations, needs and wants
3. DOMAIN - UNDERSTANDING ADULT LEARNING AND DEVELOPMENT – Understanding how to most effectively develop learning environments to support adult learning
4. DOMAIN - UNDERSTANDING COACHING CURRICULUM – Understanding the curriculum that will need to be delivered to support coaches in their development
5. DOMAIN – UNDERSTANDS SELF – Understanding own goals, strengths and limitations striving to improve when the opportunity exist
6. DOMAIN - PROCESS AND PRACTICE – Understanding the process and practice of coach development

Taken in combination these theories and ideas provide a solid foundation against which coach education practice could be unpacked, explored and therefore defined. As such this was the aim of the studies undertaken here.

Ultimately, improving a professional's ability to make and reflect on professional judgments is the overall goal of professional development, where professionalism is based in being able ground judgments in extensive theoretical, personal and practical expertise. Crucially this is true for both naturalistic and classical judgments. In essence, expertise in a professional judgment that is made in classical manner must be a result of much critical, often peer reviewed/informed thinking and planning emanating from a broad and deep knowledge base. If a professional judgment needs to be made intuitively and in a naturalistic setting then it must come from extensive practice that allows swift situational analysis, with the ability to pick up key cues in the environment (i.e. facial expressions, subtle changes in body language, subtle changes in performance) that align with ready made decisions. In both case expertise is informed by hours of previous critical thinking and reflection which will more often that not differentiate the lucky or recipe from the intuitive and adaptable professional.

METHOD

Participants

Three sets of related data from 3 different sets of participants were collected in this study. Participant group 1 were 8 (7 were male and 1 female) sport federation coach education managers drawn form a range of team and individual sports. All were experienced coach developers with responsibility for the development of coach education programmes for volunteer through to employed coaches. The second group were 3 (all male) coach developers engaged in formal mentoring roles in one to one development programmes with coaches from various sports. The third group were 5 (4 male 1 female) coach educators engaged in delivering workshops as part of overall programmes of coach development.

Process

Group 1 were engaged in an interview that 60 – 90 minutes. The questions used in these interviews focused on determining the goals of their role, how they defined coaches worth investing in, what factors impacted on capacity to support the development of coaches, what they thought a high performing coach was, what developmental practices are required to support coaches, what role assessment plays in the development of high performing coaches, how they know if their coach development practices are 'correct', and where has their knowledge of developing high performing coaches came from.

Both group 2 and 3 were observed in practice and field notes were made. The focus of the field notes was developed around the 6 domains (both individually and integratively) referred to earlier. For example, when observing the coach developers in mentoring roles, given that the coach educator would work with the coach during a coaching session it was interesting to watch when the coach educator chose to engage with coach and when they stood back. Similarly, when observing a coach educator leading a workshop, it was interesting to observe what prompted the coach educator to go 'off script' and when they chose to stay on script.

Three members of group 2 and two members of group 3 were engaged in post session discussions to draw out reasoning for their behavior where possible.

Whilst not initially intended, the approach taken has much in common with the ACTA methodology described by Gore and McAndrew (2009).

Data Analysis

Interview data from group 1 were content analysed using inductive and follow up deductive techniques as described by (Abraham et al., 2006). This approach allowed for all interviews to be analysed for statements that offer an insight or opinion regardless of underpinning theory. Follow up deductive techniques assigned these statements to categories based on the PJDM, metacognitive, and mental skills frameworks and the coach development model presented earlier.

Field notes taken from groups 2 and 3 were deductively content analysed with particular emphasis paid to naturalistic responses againsts the six domains of the coach development model. Additional explicit explanations for behaviour were used to add context and meaning to field notes. This was also used to confirm (or otherwise) interpretation of field notes.

RESULTS

In keeping with the suggestions of Gore and McAndrew (2009) the analysed data were written up as cognitive demands tables. In order to achieve this as meaningfully as possible for the final client (The FA) these tables were constructed around the two main concepts of CDM (identified as thoughtful mode in the table) and NDM (identified as intuitive mode). These concepts were then subdivided into professional skills and required knowledge. Further to this, due to the multiple demands placed on coach educators and the lack of consistency of how these roles are applied in different sports, additional suggestions for how these knowledge and skills could be applied based on the type of role being undertaken were added as 'typical behaviors' in three different yet connected roles. Finally, in keeping with the six domains identified earlier, six tables were constructed. Due to issues of space, edited versions of the full tables are shown below:

Table 1. Understanding the context

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Actively engages in working with relevant FA policy when implementing role. Conducts an informed analysis of organizational, group and individual strategy, politics and behavior.	Has a strong situational awareness of goings on in working environment Proactively and reactively recognizes and responds to opportunities to support and progress stakeholders toward achievement of nested goal
Required Knowledge	Works to an integrated mental model that encompasses a broad and deep knowledge base around relevant policy, strategic, emotional and political intelligence.	Has recourse to a rich set of critiqued experiences of working within complex relationship situations
Typical Leadership Behaviors	Actively engages and respects the opinions of key stakeholders in planning coach development programmes and interventions Defines boundaries and expectancies of coach development programme	Quickly identifies who main power brokers are and what their goals are Recognizes when the remit or position is being overstepped and withdraws to a 'safe' position
Typical Management Behaviors	Creates saleable message that can be delivered coherently and quickly Builds on relationships in order to sell more insightful and potentially complex ideas to increase shared understanding	Recognizes opportunities to reinforce message and sell ideas
Typical Coaching Behaviors	Develops plans for gaining trust and buy in with coaches and coach managers Plans how to respectfully challenge behavior and beliefs that are perceived as needing change.	Recognizes when coach is being threatened too much and backs off Recognizes when trust is being given in order to exploit opportunities offered.

Table 2. Understanding the coach

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Work with the coach to review current capabilities, set personalized goals and monitor, review and regulate progress toward set goals Build and maintain effective relationships with the coach	Has strong situational awareness of coaches' working environment and its demands Proactively and reactively recognizes and responds appropriately to moments of coach worry and/or stress when working with the coach
Required Knowledge	Draw on connections between life experiences and contemporary applied theories from social psychology, performance psychology and sociology to critically evaluate, understand and plan for changing coaches' behavior. Draw on rich mental model of what coaching is and how it changes at different levels of competence to facilitate goal setting and coach tracking	Has recourse to a rich set of critiqued experiences of working with coaches where cues in the environment are accurately connected to a limited set of 'correct' solutions
Typical Leadership Behaviors	Creates clear vision of what good coaching is Ensures the coach is making the required performance improvements and is displaying evidence of applying techniques independently	Quickly adapts predetermined developmental tasks to fit the performance of coaches Uses relevant verbal and non verbal interventions at relevant times in coach development activities
Typical Management Behaviors	Develops selection criteria for recruiting coaches into coach development Organizes regular updates to track progress	Where possible adapts formal course demands to in line with demands being placed on the coach Is flexible in application of formal recruitment methods on a case by case basis
Typical Coaching Behaviors	Shows empathy with and understanding of the change that coaches are going through Creates a personal connection or common ground with the coach to make coaching conversations more successful and able to overcome any barriers to having critical discussions	Recognizes when to push and when to support coach Recognizes the need for coach to put forward own opinions Keeps coaches focused on achieving developmental goals

Table 3. Understanding the adult learning

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Develop and monitor relevant learning environments, tasks and communication strategies to meet learning goals Design, deliver and evaluate meaningful learning opportunities and environments that meet the long-, medium-, and short-term learning needs of coaches	Has strong situational awareness of the quality of learning environments Recognizes uncertainty in everyday practice and selects relevant coping strategy
Required Knowledge	Works to an integrated mental model that encompasses a broad and deep knowledge base of learning theories and their application to classroom, workshop, on line, work-based, community and assessment learning opportunities.	Has recourse to a rich set of critiqued experiences within the domains of operation where cues in the environment are accurately connected to a limited set of correct solutions
Typical Leadership Behaviors	Works with a group of informed critical friends to develop constructively aligned learning programmes Identifies most effective learning strategies for achieving goals of coach development in line with available resources	Can offer justification for coach development programmes if challenged Foresees issues with coach development ideas in meetings
Typical Management Behaviors	Plans programmes of learning in line with available human, physical, financial, and learning resources and coach availability	Quickly identifies potential problems in programme delivery and identifies relevant solutions
Typical Coaching Behaviors	Plans development sessions with learning objectives that link to bigger development picture plan Creates a culture of coaches innovating, risk taking and developing new ideas	Responds to questions with insight using stories, analogies and examples Models risk taking and innovation in practice

Table 4. Understanding the coaching curriculum

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Design and/or understand developed coach development curricula that is aligned to FA coach development pathways/FA courses and to the needs of individual coaches Analyze best practice coaching to maintain currency in coaching curriculum	Has strong situational awareness of how well curriculum is being delivered. Also a strong awareness of how well curriculum is being received and work with by coaches Recognizes uncertainty in coaches relating to content to be learned and responds appropriately
Required Knowledge	Works to an integrated mental model that encompasses a broad and deep knowledge base relating to what good coaching is Draw on connections between life experiences and contemporary applied theories from coaching science, developmental psychology and performance psychology to form clear rationale for coach curriculum	Has recourse to a rich set of critiqued experiences within the domains of operation where cues in the environment are accurately connected to a limited set of correct solutions
Typical Leadership Behaviors	Works with coach developers and other relevant stakeholders to critique coach development plans Offers informed professional opinion on the development of curriculum	Provides a thorough rationale for all elements of practice when challenged Offers links between curriculum delivered and required improvements in practice
Typical Management Behaviors	Creates curriculum relevant to level of coach development Recruits and assigns relevant tutors to deliver Creates relevant documentation to support formal coach development programmes	Recognizes logistical and personnel issues and responds with quick and accurate solutions Provides a thorough rationale for all elements of practice when challenged
Typical Coaching Behaviors	Sets meaningful learning objectives for development sessions Plans tasks that are relevant to achieving learning objectives Sets progressive curricula that relates at macro, meso and micro levels of nested plan	Deliver verbal interventions with accurate information Quickly adapts predetermined curriculum to fit the performance of coaches

Table 5. Understanding of self

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Conducts critically informed, evidence-based self-analysis in order to examine, expose and challenge the congruence of intentions, assumptions and beliefs with practice. Works toward professional standards and values	Strives to recognize opportunities for self development and to work towards personal goals
Required Knowledge	Draw on contemporary concepts and applied theories of coaching expertise, reflection, social psychology, performance psychology and sociology to critically evaluate the reasoning and resources of own behavior and practice in order to generate development goals and action plans.	Has recourse to a rich understanding of self that recognizes strength and weaknesses in knowledge, skills and personal effectiveness Works to a mental model of personal effectiveness, excellence, professional values and ethics
Typical Leadership Behaviors	Plans for putting self improvement goals into practice Models expectancies in decision making integrity, honesty, sincerity, respect, risk taking and deprecating behaviors to break down barriers	Recognizes opportunities to model expectations in decision making, integrity, honesty, sincerity, respect, risk taking and deprecating behaviors
Typical Management Behaviors	Plans deliberately and objectively Keeps working towards plans	Ensures everyday practice stays to plan Recognizes and actions opportunities to work towards own management goals
Typical Coaching Behaviors	Respects and trusts the opinion of coaches Models integrity, honesty, sincerity risk taking and deprecating behaviors to break down barriers	Recognizes opportunities to model integrity, honesty, sincerity, respect, risk taking and deprecating behaviors Recognizes and actions opportunities to work towards own coaching goals

Table 6. Understanding process and practice

	Coach Developer - Thoughtful Mode	Coach Developer – Intuitive Mode
Professional Skills	Makes effective and informed decisions that reflect the big picture of coach development relating to the planning, implementation, monitoring, evaluation and regulation of nested goals and programmes of development Recognize and resolve problematic and atypical issues through the generation innovative strategies and solutions	Is aware of goings on in working environment Proactively and reactively recognizes and responds to opportunities in everyday work to support and progress stakeholders toward achievement of nested goal
Required Knowledge	Works to an integrated and explicit mental model of coach development that encompasses a breadth and depth of knowledge in the domains of; understanding process and practice of coach development, understanding the coach, understanding coaching, understanding adult learning and development, understanding context, strategy and politics, understanding self	Has recourse to a rich set of critiqued experiences within the domains of operation where cues in the environment are accurately connected to a limited set of correct solutions
Typical Leadership Behaviors	Negotiates goals with the coach and key stakeholders to create and manage expectations Develops constructively aligned learning programmes	Provides a thorough rationale for all elements of practice when challenged Uses relevant verbal and non verbal interventions at relevant times in leadership activities
Typical Management Behaviors	Develops plans that make most effective use of available resources Develops a milestoned tactical plan with built in contingency	Uses relevant verbal and non verbal interventions at relevant times in management activities Models integrity, honesty, sincerity risk taking and deprecating behaviors to set expectations
Typical Coaching Behaviors	Plans development sessions that link to bigger development picture plan Works to align needs established by benchmarking and wants established by coach	Quickly adapts predetermined developmental tasks to fit the performance of coaches Uses relevant verbal and non verbal interventions at relevant times in coach development activities

DISCUSSION

A cursory glance through the (few) job descriptions that appear for coach developers in job adverts suggests that the results presented in Tables 1 - 6 have gone much further to defining the role of coach developers. Typically these job descriptions display a lack of differentiation between knowledge, skills and typical behaviors. This leads to nebulous statements, often lacking definition with their focus being around experience, leadership and operational factors. Whilst not surprising or unusual in poorly defined/immature professional vocations, this does limit the capacity to recruit and/or develop emerging and even established professionals in this domain. Deductively analyzing the coach developer role through the lens of PJDM and the coach development model has offered a thorough insight into the role of a coach developer. Developing cognitive demands tables (Gore and McAndrew, 2009) has served a relevant and meaningful purpose by offering (relatively) a precise and concise overview of the role against which professional development approaches could be developed. As such, greater definition now exists.

Further examination of the results in Tables 1 – 6 does however, reveal a level of replication in the tables. This is not surprising as the tables (and the domains they reflect) are not meant to be orthogonal. They represent parts of one big picture of being a coach developer. Furthermore, creating the tables through multiple sources of 'data' means that they are informed through a wide peer group thus removing the chance of bias and ideas being missed. Creating tables from multiple sources also means that the tables are probably not reflective of reality, i.e., it is unlikely that coach developers exist that 100% reflect the skills and knowledge identified in all of the tables. In other words the tables created above are aspirational in nature. This is an important distinction if the tables are to be used to support the creation of professional development opportunities for coach developers. Given that coach developers are typically drawn from a pool of experienced coaches each coach educator/developer will have their own strengths and weaknesses across the six domains. Taking this approach therefore, the tables offer an initial benchmarking exercise against which the practice of recently employed and experienced coach developers can be mapped. In association with the goals these coach developers are working towards their professional development requirements can be critiqued and programmes created.

While progress has been made in defining the coach developer role, we do not claim to have a created a set of cognitive demands that is 'complete' and definitive. The tables above are intended to be neither – they are a well informed (i.e. informed by theory and practice) and succinct summary of the demands of a coach developer role. Sport coaching as a domain has previously been, and continues to be drawn to the behavioral competency approach to defining vocational roles because of the apparent control and objectivity that it offers (Thompson, 2000). However, as Thompson (2000) states, this approach has often only led to skilled robots rather than knowledgeable doers. Consequently, this has not been a competencising exercise; rather the tables display a set of ideas that attempt to capture the essence and complexity of coach development roles.

Further to these tables not being complete, we believe there remains much work that could and should be done within this domain. We are confident that the application of PJDM has brought much needed to focus to the role of coach developers. Indeed the explicit recognition of the thoughtful <u>and</u> intuitive demands of this role allows for more informed debate about the connection between theory and practice, thinking and doing, academic worlds and applied worlds. In short this sort of work allows artificial barriers to be broken down as it recognizes the skills and knowledge that all involved have. However, PJDM also strongly highlights the role of; perception, mental models, shared mental models, routines and recognition primed decision making in determining how people perform their roles. The content in the tables makes specific reference to all of these ideas however, it doesn't really unpack them, and viewed from this perspective we may have only scratched the surface.

For example, gaining trust (of the coach but also the club officials) was seen as being crucial for the coach developers within the FA. As such the statement 'Recognizes when trust is being given in order to exploit opportunities offered' was included in Table 1. However, we were not able to spend enough time with the coach developers on this specific topic so that we could unpack what perceptual markers were being identified that led to the coach developers feeling like they were being trusted. Indeed across all 6 tables, 'recognizes' 'perceives' 'identifies' are used frequently to express the importance of perception in making judgments. In most of these cases more work is required to unpack what these perceptual markers are. Indeed routines, mental models etc used in these roles all need further investigation.

CONCLUSION

The overall goal of any professional development must to improve practice in order to become more efficient and more effective in achieving goals. For coach developers therefore, any professional development should lead to them being better at developing better coaches. In order to do this work has identified that a, personalized, deepening of knowledge across the six domains should lead to improved awareness. However, both the deepening and awareness raising of knowledge <u>must</u> also lead to enhanced practical skills if there is to be an impact on practice and this must be obvious to the coach educators if they are to buy into the professional development. The completion of this work has allowed us to develop a bespoke postgraduate diploma course for the FA in order to facilitate the professional development of a selection of their coach education staff. Using the PJDM framework and aligning (even approximately) with the ACTA methodology has allowed this course to be evidence based and aligned with the needs and wants of the coach educators.

ACKNOWLEDGMENTS

This work has been made possible through the funding of Sports Coach UK, UK Sport and The Football Association.

REFERENCES

Abraham, A., Collins, D., & Martindale, R. (2006). The coaching schematic: Validation through expert coach consensus. *Journal of Sport Sciences, 24*(6), 549-564.

Abraham, A., Collins, D., Morgan, G., & Muir, B. (2009). Developing expert coaches requires expert coach development: Replacing serendipity with orchestration. In A. Lorenzo, S. J. Ibanez & E. Ortega (Eds.), *Aportaciones Teoricas Y Practicas Para El Baloncesto Del Futuro*. Sevilla: Wanceulen Editorial Deportiva.

Abraham, A., & Collins, D. (2011). Taking the Next Step: Ways Forward for Coaching Science. *Quest, 63*(4), 366-384.

Côté, J., & Gilbert, W. (2009). An integrative definition of coaching effectiveness and expertise. *International Journal of Sport Science and Coaching, 4*(3), 307-232.

Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behavior and influences of learning environments. *International Journal of Educational Research, 41*, 407–428.

Gore, J., & McAndrew, C. (2009). Accessing expert cognition. *The Psychologist, 22* (3), 218-219.

Kahneman, D., & Klein, G. A. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515-526.

Klein, G., & Militello, L. (2005). The knowledge audit as a method for cognitive task analysis. In H. Montgomery, R. Lipsitz & B. Brehmer (Eds.), *How professionals make decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Focus article: Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making, 14*, 331-352.

MacNamara, Á., Button, A., & Collins, D. (2010). The role of psychological characteristics in facilitating the pathway to elite performance: Part I: Identifying mental skills and behaviors. *The Sport Psychologist, 24*(1), 52-73.

Thompson, N. (2000). *Theory and practice in the human services*. Buckingham: Open University Press.

Vickers, J. N., Reeves, M.-A., Chambers, K. L., & Martell, S. (2004). Decision training: Cognitive strategies for enhancing motor performance. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research theory and practice* (pp. 103-120). London: Routledge

Perceiving and interpreting human activity: a normative multi-agent system

Fabien BADEIG[a], Benoît VETTIER[a] and Catherine GARBAY[a]
[a]*Laboratoire d'Informatique de Grenoble (LIG), AMA team, Université Joseph Fourier / CNRS, Grenoble, France*

ABSTRACT

Introduction: We present an approach for the follow-up of human activity in the complementary settings of health monitoring and collaborative work. **Method**: Rather than aimed at providing an accurate and complete description of this activity, our system is aimed at framing the perception or interpretation processes according to various universes of discourse and evolving requirements: the one of the application domain, the one of the human acting or observing this activity, the one of the technical system at hand. We further promote a unified designing approach grounded in normative multi-agent systems theory. **Results and discussion:** We discuss the approach potential with respect to (i) human activity monitoring, (ii) support to tangible collaborative activity.

KEYWORDS

Human activity follow-up; normative multi-agent systems; monitoring; remote collaboration.

PROBLEMS AND ISSUES

We consider the follow-up of human activity, in two complementary contexts: monitoring and remote collaboration. Grounding on activity theory (Engestrom, 1991), we propose a *situated* approach that accounts for various perspectives and universes of discourse. Our goal is not to provide a single "most-likely" interpretation result, but rather to collaboratively maintain an interpretation path within the bounds of varied and evolving frames. This is implemented by means of normative agents constructing and framing interpretation under various constraints (Vettier, Amate & Garbay, 2012; Garbay, Badeig & Caelen, 2012). As regards monitoring, the purpose is to track the activity of a person equipped with ambulatory sensors and to check whether it conforms to existing scenarios and personalized models (Vettier et al., 2012). Context sensitivity, personalization and pro-activeness are core properties for monitoring systems, together with their compliance to social conventions (Aarts & de Ruyter, 2009). The analysis process must imply several levels of abstraction, involving the global level of norms, functional requirements and goals (Weber & Glynn, 2006). For (Greenberg, 2001), context must be seen as a dynamic construct evolving with time, episodes of use, internal goals, or social interaction. Following (Klein, Phillips, Rall & Peluso, 2006), we approach interpretation as "a process of fitting data into a frame that is continuously replaced and adapted to fit the data". Therefore, we model the understanding process as interleaving focus, perception, interpretation, context modeling and anticipation activities. These activities are framed by mutually evolving norms, which represent requirements of three categories: application, actor- and system-dependent. A major breakthrough in the field of CSCW is to preserve the spontaneity and fluidity of human action while ensuring its consistency and proper coordination (Pape & Graham, 2010). Mutual awareness is central in this respect. We consider the case of actors working around remote interactive surfaces with no communication means except the moves of tangible objects on these surfaces. In this context, we approach collaboration as a conversational process involving several signs of dialog and threads of activity (Kraut, Fussell & Siegel, 2003). We further consider coordination as a compromise between the handling of implicit communication norms and affordant objects (Sire & Chatty, 1998). Dedicated tangible objects, called *tangigets*, are introduced to produce additional signs and traces to sustain conversational activity. Beyond information sharing, perceiving the other's activity must further be approached through the constraints and rules that ground social organization (Erickson & Kellogg, 2003). Informed virtual feedbacks are provided so that mutual awareness be approached as the sharing of norms that frame human activity (Garbay et al., 2012).

PROPOSED DESIGNING APPROACH

We propose an original designing approach for human activity follow-up, that we characterize as distributed (agent-based), situated (trace-oriented) and normative. Distribution is motivated by the heterogeneity and simultaneity of information threads to be managed at various levels of abstraction. Normative MAS are a class of Multi-Agent Systems in which agent behaviour is not only guided by their mere individual objectives but also regulated by norms specifying in a declarative way which actions are considered as legal or not by the group

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

(Castelfranchi, 2006). Agents acting in such system may be seen as “socially autonomous”. These norms are designed as condition-action rules; they are triggered by a dedicated engine and result in agent notifications (Boissier, Balbo & Badeig, 2010). Such design supports the separate modeling and exploitation of a variety of functional and non-functional requirements, from the application domain, singular actor/organization, or system at hand. The role of the norms is to express these requirements and frame agent’s processing accordingly. A norm has the following simplified expression: *N := <id, context, role, object>* in which *context* represents an overall evaluation condition, *role* represents the agent’s role concerned by this norm and *object* is a complex field, typically written as *launch (conditions, actions)* characterizing the conditional action attached to the norm. The trace is a multidimensional component reflecting human/system activity and its conformance to the requirements at hand. We distinguish between production, communication and coordination agents; depending on their types, they are provided with autonomous decision-making abilities as regards the enrichment of the trace (production), the feedback to the actors (communication) and the modification of the set of norms which frame interpretation (coordination), thus providing the system with pro-active behavior.

HUMAN ACTIVITY MONITORING

The goal is to provide a personalized follow-up of human activity, based on physiology and actimetry data (heart and breath rate, movement…) acquired by ambulatory sensors. Production agents are designed to manage concurrent and uncertain hypotheses at various abstraction levels: their role is to enrich a dynamic hypothesis population keeping track of human activity, both by verifying current hypotheses and by producing new hypotheses (prediction), to cope with the evolving input data (Vettier et al., 2012). The verification process results in the computation of hypothesis confidence degree (*low*, *medium* or *high*), depending on thresholds that may be dynamically adjusted by dedicated norms. Prediction may be performed by generalization (from a low to a higher level hypothesis), or by anticipation (transition between hypotheses at the same abstraction level), based on semantic networks expressing the relationships between hypotheses. The interpretation process is framed by a set of high-level requirements of three types: Application-Dependent (assessing the “normality” of features or situations…), Actor-Dependent (accounting for personnal profiles, expectations of the monitoring institution as regards what is alarming or not…) and System-Dependent (assessing the effectiveness, efficiency and informativity of monitoring). In the example below, the *Anticipation* rule activates agents whose hypotheses appear less than likely, so that successor agents, predicting new hypotheses, are generated. Some parameters such as confidence thresholds and minimum durations are necessary; they can be subjected to adaptation from the coordination agents.

Norm-anticipation:= <id, always, All, launch(cond1+cond2, create-successors) >
cond1 := [agent-confidence = medium/low]
cond2 := [agent-timer(medium/low) > minimum-duration-anticipation]
Norm-termination:= <id, always, All, launch(cond3+cond4, delete) >
cond3 := [agent-confidence = low]
cond4 := [agent-timer(low) > minimum-duration-termination]

These requirements are prone to evolve, in case of activites with varying constraints, modifications of surveillance priorities, or to ensure increased sensitivity and accuracy when faced with critical situations. In this design, *alarms* (situations which do not fit the *frame*) are considered as incentives to adapt the interpretation process' *frame*. Such adaption is managed by the coordination agents, which are launched by norms in charge of detecting such patterns of evolution. Two coordination norm examples are provided below. The *Parsimony* norm is only evaluated in a non-alarming context : its goal is to remove the least relevant agents when there are too many unlikely hypotheses, thus improving efficiency. The *reduce-openness* function will change the *medium/low* and *minimum-duration* thresholds (thus impacting the termination norm). The *Informativeness* norm is triggered when there are no likely hypotheses (an alarming situation from a non-functional viewpoint) : the *force-anticipation* function will result in *framing* the Production rules so that all current agents *anticipate*.

Norm-Parsimony:= <id, [no-alarms], Probe, launch(number-of-agents > N, reduce-openness) >
Norm-Informativeness:= <id, [always], Probe, launch(n-HighConfidence = 0, force-anticipation) >

Figure 1 shows an example coordination behaviour implying various kinds of norms and agents. A coordination agent (Probe Agent) is also running to check for the presence of any disorder (some requirements are violated) and frame the system activity accordingly. The green, yellow and red lines in this figure represent the likelihood of the agent's hypotheses (high, medium, low). The vertical arrows show the triggering of norms. The *A* arrows represent regular *Anticipation* notifications (when the input data changes, the agents reflect it by generating new hypotheses). The *P* arrows mark that the Probe Agent is activated to analyse more closely the situation, which eventually results in its forcing a Framing change (*F* arrow) so that all agents *Anticipate*. The Input Norms regularly notify the agents to feed them input data (no arrows shown for the sake of clarity).

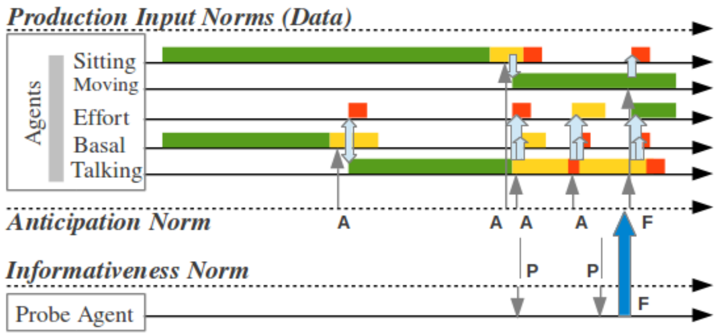

Figure 1. An example view of coordination behaviour for human activity monitoring.

COLLABORATIVE SUPPORT SYSTEM

The goal is to design a system ensuring smooth coordination between people working over remote interactive surfaces. The infrastructure, called *TangiSense* (Lepreux, Kubicki, Kolski & Caelen, 2011), makes use of RFID technology to locate/identify tagged tangible objects. Multicolor leds on top of each RFID antenna provide virtual feedback to local/remote users (assessing detection or conformity of tangible objects moves). No communication means is provided except the table and its tangible and virtual equipment. We discuss here an application to the RISK game, a strategic board game where players compete towards domination of the Earth. At start, each player is allocated an army and territories that his army occupies. An attack involves the designation of attacking and attacked territories and dice rolls to determine who is losing or winning this round. The proposed system involves production, communication and coordination agents communicating via multidimensional traces. Production agents ensure the follow-up of tangible object moves and perform application-dependent computations. Their activity is framed by norms of various kinds: application-dependent norms trigger production agents upon the detection of incoming moves; actor-dependent norms triggers higher-level production agents to follow events occuring over remote surfaces. We provide below an example of a norm ensuring the follow-up of two remote dice roll results, which results in the determination of a winner and loser and the launch of an agent whose role will be to update the traces accordingly.

Norm-dice-result:= <id, [step = fight], dice-result, launch(cond, win)>
cond := [trace.type(?t1) = dice] & [trace.value(?t1) = ?v1] & [trace.onTable(?t1) = ?x] & [trace.type(?t2) = dice] & [trace.value(?t2) = ?v2] & [trace.onTable(?t2) ≠ ?x] & [?v1 > ?v2]

The role of communication agents is to ensure the processing of events from the infrastructure layer and to provide informed virtual feedback (led display) to local/remote user. Their activity is framed by System-, Application- and Actor-dependent norms to regulate (i) the creation of traces (to keep track of events generated by object moves) and (ii) the launching of local/remote feedback agents (to provide informed follow-up of object moves, depending on conformity and privacy criteria). The role of coordination agents is to ensure the consistency of trace components (and therefore of overall activity) in a context where concurrent actions are performed by remote human and artificial actors. Updating the conformity and privacy fields of trace properties is of particular importance in this respect, since it results in modifying the trace local visibility and accessibility to remote actors. Application-dependent norms are used to this aim: they update the trace conformity and privacy, depending on the compliance of activity, to the task at hand, and to potential privacy tangiget moves. Coordination agents further ensure the adequacy of the set of norms in front of system-dependent, application-dependent or actor-dependent evolutions. Handling coordination tangigets is a way to indicate such evolutions. In the RISK game, a *stage* tangiget allows to follow gameplay evolution (e.g. from intialization to fight) and a *designation* tangiget to indicate the opponent area that the player wants to attack. We provide below an example of a coordination norm handling these coordination tangigets and launching a coordination agent whose role will be to modify the policy at hand by depositing the corresponding set of norms.

Norm-attack-policy:= <id, [step = fight], coordination, launch(cond, manage-norm-policy)>
cond:= := [trace.type(?t1) = coordination] & [trace.onTable(?t1) = ?tab1] & [trace.value(?t1) = “attack”] & [trace.type(?t2) = designation] & [trace.onTable(?t2) = ?tab1] & & [trace.value(?t2) = ?jd]

Inter-player coordination is finally ensured (i) by actor-dependent tangiget handling, (ii) by system-driven policies of action and (iii) by user-oriented visual feedback. This process (figure 2) is decomposed as follows: handling the coordination and designation tangigets modify the corresponding traces, which triggers the tangiget norm f1; this norm will launch AT1 on the attacking player local environment together with AT2 on the attacked player remote environment; it is then the role of AT1 and AT2 to deposit the policy they plan to follow for this new phase of the game, in the form respectively of communication norms f2 and f3 (these norms differ, depending on the table and player’s role); the role of f2 is to launch AI2, so that information about the attack under preparation be provided to the remote player; upon reception of this feedback, player 2 reacts by handling his own coordination tangiget, thus acknowledging reception of this action; the role of f3 is now to launch AI1, to transmit player 2's acknowledgement of reception.

CONCLUSION

We have proposed in this paper an original architecture for advanced ambient intelligence system design. Our proposal allows for a clear separation between two universes of discourse, that of the data, and that of the requirements (be it application-dependent, actor-dependent or system-dependent). It further ensures the proper coupling of these evolving universes, by means of situation-dependent evolving frames that act in turn to frame the perception and interpretation processes. We have considered two complementary applications, highlighting the unfolding of interpretation/collaboration processes. Following activity theory, framing processes do not act as a prerequisite to action. Rather, they allow to situate activity within a variety of socio-technical stances and operate as incentives to the process unfolding.

Work partly supported by the French DGA (REI SUPERCO n°2009 34 0013) and ANR (project IMAGIT 2010 CORD 01701).

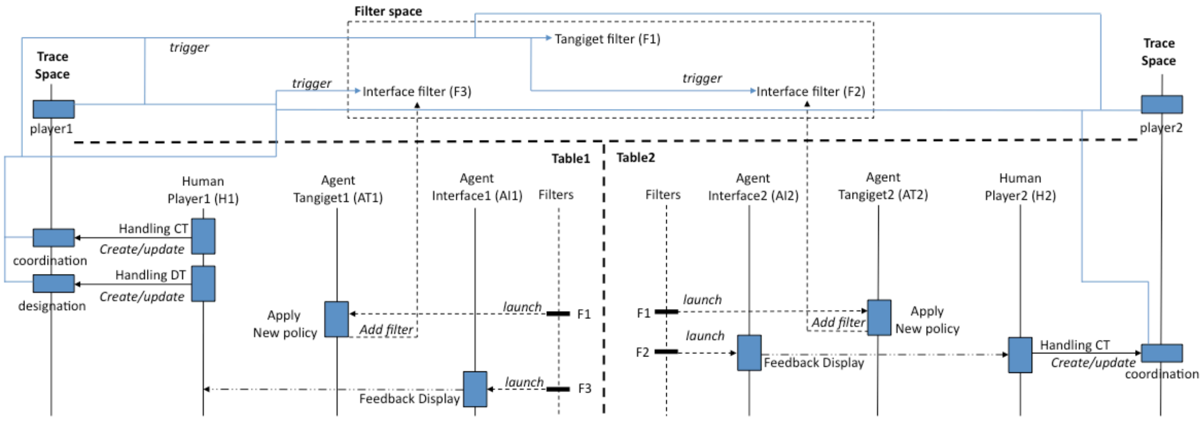

Figure 2. Inter-player coordination in the RISK game: the interleaving between tangiget, norms and agent handling may be observed in this short example.

REFERENCES

Aarts, E. & de Ruyter, B. (2009) New Research Perspectives on Ambient Intelligence, *Journal of Ambient Intelligence and Smart Environments*, 1, 5-14. doi:10.3233/AIS-2009-0001

Boissier O., Balbo F., and Badeig F. (2010) Controlling multi-party interaction within normative multi-agent organisations, *Proceedings of the MALLOW Workshop on Coordination, Organization, Institutions and Norms in Multi-Agent Systems*.

Castelfranchi, C. (2006) The architecture of a social mind: the social structure of cognitive agents, In Sun, R. (Eds.) *Cognition and Multi-Agent Interaction: from Cognitive Modeling to Social Simulation*, Cambridge University Press: Cambridge.

Engestrom, Y. (1991) Developmental work research: reconstructing expertise through expansive learning, In Nurminen, M. & Weir, G. (Eds.) *Human jobs and computer interfaces*, New York, NY: North-Holland.

Erickson, T. & Kellogg, W. (2003) Social translucence: Using minimalist visualizations of social activity to support collective interaction. In Höök, K., Benyon, D. & Munroe, A. (Eds.) *Designing Information Spaces: The Social Navigation Approach*, Springer-Verlag London.

Garbay, C., Badeig F. & Caelen J. (2012) Normative Multi-agent Approach to Support Collaborative Work in Distributed Tangible Environments, *Proceedings of the 2012 ACM Conference on Computer Supported Cooperative Work (Companion)*, Seattle, WA, USA.

Greenberg, S. (2001) Context as a dynamic construct, *Human-Computer Interaction*, 16, 257-268.

Klein G., Phillips J., Rall E. & Peluso D. (2006) A data/frame theory of sensemaking, *Proceedings of the 6th International Conference on Naturalistic Decision Making*, Pensacola Beach, Florida, USA.

Kraut, R. E., Fussell, R. & Siegel, J. (2003) Visual information as a conversational resource in collaborative physical tasks, *Human-Computer Interaction*, 18, 13-49.

Lepreux S., Kubicki S., Kolski C. & Caelen, J. (2011) Distributed interactive surfaces using tangible and virtual objects. *Proceedings of the Distributed User Interfaces CHI Workshop*, Vancouver, Canada.

Pantelopoulos, A. & Bourbakis, N. (2010) A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis, *IEEE Transactions on Systems, Man, and Cybernetics*, *Part C*, 40, 1, 1-12.

Pape, J. & Graham, T. (2010) Coordination policies for tabletop gaming. *Proceedings of Graphics Interface 2010*, Ottawa, Ontario, Canada.

Sire, S. & Chatty, S. (1998) Vers des interfaces à collaboration directe pour le travail de groupe, *Actes du Colloque GRESICO Communication, Société, Internet*, Vannes, France.

Vettier, B., Amate, L. & Garbay, C. (2012) A Multi-Hypothesis Monitoring Architecture: Application to Ambulatory Physiology, *Proceedings of the Sixth Starting AI Researchers' Symposium*, Montpellier, France.

Weber, K. & Glynn, M. (2006) Making sense with institutions: Context, thought and action in Weick's theory, *Organization Studies*, 27, 1639-1660.

Collective cognitive activity of fooball officials as it occurs

Simon BOYER, Géraldine RIX-LIEVRE and Michel RECOPE
Clermont Université, Université Blaise Pascal, ACTé, EA 4281, Clermont-Ferrand, France

ABSTRACT

Introduction: To investigate the cognition within a team of football officials, we scrutinized their inter-individual activities. **Method**: We studied two games in a professional football championship. We worked with two teams of officials. Games were filmed from the stands, and individual self-confrontation interviews were realised. Verbalization analysis and video description were combined. **Results and discussion:** The assistant and replacement referees develop a specific awareness about the doubt the referee express in decision making process. They evaluate the preoccupations of the referee in a timing enabling to make appropriate information manifest in order to assist the referee in his decision making. Their individual awareness toward referee behaviours enables to construct a mutual intelligibility of the opponent physical confrontations.

KEYWORDS

Decision Making; Coordination; Compatible Awareness; Self-confrontation interviews ; Football refereeing.

INTRODUCTION

We conducted an exploratory study about activity of teams of football officials in match by highlighting cognitive process between team members. These processes take place in the context of referee decision making which has all the characteristics of the environment as it is envisaged by the Naturalistic Decision Making community (Mascarenhas, Collins & Mortimer, 2005).

Rix-Lièvre et al. (2011) proposed a non-normative perspective and methodology to understand referee decision making: referee makes judgment acts which shows to players and impose upon them what is acceptable during a match; judgment acts are bodily or linguistic manifestations embedded in context (Rix-Lièvre et al., 2011). The rules are not applied but are an essential resource for refereeing. However these results do not take into account that the referee is a member of a team. This research aims to investigate the cognitive and interactive processes between officials which enable the judgment act process.

Bourbousson et al. (2011) studied the collective activity of a basketball players' team throughout a match. The authors combine an individual activity analysis perspective with an ergonomic one. They highlighted the mutual intelligibility (Salembier & Zouinar, 2004) of game situations constructed by players in order to coordinate the construction of a collective performance: players understand when their partners have the same understanding of the game situation than themselves. Players have tacit strategies to communicate their shared understanding of the situation by acting on field. The authors stressed that mutual intelligibility is founded both on sharing relatively stable and pre-existent knowledge and on sharing the present situation through sharing contextual information.

According to Heath et al. (2002) staying aware of the reciprocal behaviors of other team members is fundamental for mutual coordination even though the members may be involved in distinct activities.

In football, the members of the officials' team are involved in different tasks: it's a heterogeneous team. So this research takes into account the organized partition of tasks (Cooke & Gorman, 2006) and its organizing role for coordination (Salembier & Zouinar 2004). In football, the referee is the director of the game: he makes the decisions. Two assistant referees are responsible for signaling the outs of play – throw-ins, corners – and the off sides in their zone. But they also have the duty of helping the referee to control the match. The replacement referee manage the substitute players. More, partition of tasks determine official movements and subsequently their perspectives on game: the referee can move everywhere he wants on the field; the assistant referees move on a half length field on the edge of the pitch along the throw-in line; and the replacement referee stay in a zone situated at the middle of a edge of the field. Others specific procedures regard communicating and coordination.

To investigate referee judgment acts and interactive process between officials we work on the meanings the participants give to their inter-individual behaviors in situation. This exploratory study seeks to highlight a mutual intelligibility of the situation and mutual understanding between members. We investigate what is manifest (Salembier & Zouinar, 2004) and meaningful for each member of the officials' team in situation. We

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

aim to study how each team member contributes to produce judgment acts and what coordination actions enable this contribution, especially how each official contribute to this production.

METHOD

Participants

We investigated two games in the French national professional football in the first and second divisions (Ligue1, Ligue2). We studied the activity of two teams of officials, each comprising four members: a referee, two assistant referees, a replacement referee. Referees and assistant referees are members of the elite category. During the games officials were equipped with audio communication systems - authorised by the French Football Federation capable of communicating in real time -.

Protocol

For each match the game videotaped from the stands was collected from the networks broadcasting it.
After each game, individual self-confrontation interviews (Theureau, 2003) were realised : each participant was re-immersed in his own lived game experience. The officials were asked to describe their activity to tell the researcher what was meaningful in the course of their action in a particular lived-experience. To a great extent, this meaning is implicit and/or tacit (Kahneman & Klein, 2009) so we help each participant to formulate it (Vermersch, 1999). The match was reviewed throughout its length at the speed the participant considered to be most suitable for reliving into his experience (Lorains, Ball, & MacMahon, 2013). So doing, we identified what was meaningfull for the officials all along the match. But the reseacher especially points some particular manifest events of communication between officials.
The interviews were transcribed. Verbalizations about judgment acts and interindividual communications were selected. Descriptions of the corresponding moments were realized. The description data and the officials verbalisations were then combined in order to relate the role of the information delivered by each member in judgment act.
The processing consisted in understanding what is meaningful for each participant in the verbal or non-verbal behaviors of other team members, i.e. the mutual understanding.
More, we focused on mutual intelligibility. The processing also consisted in grasping the way each participant make manifest his understanding of a game situation – i.e. the way he shows that to other members. Conversely, the processing aims to highlight the way team members percieved the manifest information concerning a partners' understanding of the game situation. We highlighted their understanding of manifest information and the use they made of this information to produce a judgment or a judgment act.
So we are thereby able to reveal the mutual understanding between team members and the mutual intelligibility of the officials concerning the judgment of game situations. Thus, we determined which official(s) was/were involved in each judgment act. We have determined how each one takes part in the construction of what is imposed on the players and how each official construct the way he takes part in it.

RESULTS AND DISCUSSION

The awareness of the assistant referee and the replacement referee

Referee's colleagues are aware of cues related to the referee's doubts while he is engaged in the process of constructing what he imposes on the players. The referee's partners focus on indications of hesitations in referee's behavior – the way he moves on the pitch or the way he declares a foul or not. For example, the assistant referees are focused on: (1) the way (speed) the referee brings the whistle to his mouth; (2) his posture and his ground supports that is whether the referee is ready to make his usual movements to follow the play or hesitates in continuing his movements. The assistant referees and replacement referee evaluate the doubt expressed by the referee judging the game situation. For them, the doubt is a sign of the referee's need for information. But detecting the referee's need is not enough. Assistant and replacement referees have to assist the director of the game in producing judgment act. So, they make manifest "helping" information: information within the framework of the referee's preoccupations *hic et nunc*, i.e. the way in which the referee constructs the situation of play.
To make manifest "helping" information adequately, the assistant and replacement referees have to determine – in the timing of the referee judgment act construction – (1) the compatibility of raising the level of manifestness of an element of context with the referee's preoccupations *hic et nunc* or (2) the compatibility of their own judgment with the referee's preoccupations *hic et nunc*.
The referee partners make manifest information in two ways: (1) so that it is a complementary meaningful element to judge because it can support referee's understanding of the game situation; (2) so that it will become a salient meaningful element because it has the potential – for the referee – to be decisive in producing a judgment act, i.e. to show and impose the acceptability of the game situation. For example, the assistant referee may make manifest an element of the context of play that the referee has not been able to see and so which he could not have taken into account in constructing what is imposed on the players.

The assistant and the replacement referees seem to develop a complex awareness (Endsley & Garland, 2000) about the referee duty: to judge *hinc et nunc*. They take into account simultaneously the context of play and the referee doubts behavior.

The assistant and replacement referees – artisans of mutual intelligibility

The assistant and replacement referees acquire an understanding of the manner the referee judge opponents' physical confrontations in the first minutes of the match. Indeed, in situations involving shoulder to shoulder duals or tackles on the ball they "evaluate" the referee level of acceptability for the intensity of physical contacts between players. Each partner of the referee constructs a way to value contact intensity according to this level of acceptability – in order to assist the referee for the rest of the match. In this way assistant and replacement referees seem respectively to construct an awareness based on the referee judgment of physical confrontations. Awareness of each official is about the same item – the intensity of the contacts – but it is constructed from their different perspectives. The referee partners try to make their own judgment match with the referee's judgment framework. It's like a calibration of their judgment (Unkelbach & Memmert, 2008).

To assist the referee in producing judgment act, assistant and replacement referees make their judgment manifest. In return, the effective use of this manifest information by the referee in constructing his judgment act validates the information relevance delivered by his partners *hic et nunc*. So officials construct mutual inteligibility: the referee enables the assistant and replacement referees to shape their judgment according to his. As a consequence, the manner the referee uses his partners' manifest information makes enable the progressive construction of their awarness compatibility (Salmon, 2009) with his judgment.

In this way, the referee's judgment acts enable the awareness of his colleagues to be checked and regulated, so as to maintain the mutual intelligibility of game situation. This mutual intelligibility is thus progressively and locally constructed by episodes (Bourbousson, et al., 2011) of implicit validation (or invalidation) of the relevance of information the officials deliver *hic et nunc*. Mutual intelligibility is a co-construction for the control of game situations.

CONCLUSION

This research highlights the specific awareness the replacement and assistant referees develop in order to assist the judgment act process of the referee, i.e. when he doubts. This allows them to deliver information which is consistent with the dynamics of the referee activity.

The referee partners construct collectively their individual understanding of game situations in order to control the contact intensity between opponents. The officials' respective understandings are (1) partly shared by each official and (2) partly distributed and compatible with each other – because of their specific roles and perspective on players' actions.

But, if the compatibility is constructed to judge contact intensity between opponents according to the same values, the way assistant and replacement referees construct this compatibility from their perspective remain to be precisely investigated.

More, if the perceptive activity of referee partners is focused on cues about doubt occurrences, this study has to bring some insight to the way the assistant referee can understand the referee preoccupations precisely, i.e. what are the referee's doubts.

ACKNOWLEDGMENTS

This study was carried out within the framework of the PERF Arbitrage activities and with the support of La Poste Group and the French Football Federation.

REFERENCES

Bourbousson, J., Poizat, G., Saury, J., & Seve, C. (2011). Description of dynamic shared knowledge: an exploratory study during a competitive team sports interaction. *Ergonomics, 54*(2), 120-138.

Cooke, N. J., & Gorman, J. C. (2006). Assessment of team cognition. In P. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (pp. 270-275). London: Taylor & Francis Ltd.

Endsley, M. R., & Garland, D. J. (2000). *Situation Awareness: Analysis and Measurement*: Taylor & Francis Group.

Heath, C., Sanchez Svensson, M., Hindmarsh, J., Luff, P., & Lehn (vom), D. (2002). Configuring awareness. *Computer Supported Cooperative Work, Special issue on Awareness, 11*, 317-347.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise. *American Psychologist, 64*(6), 515-526.

Lorains, M., Ball, K., & MacMahon, C. (2013). Expertise differences in a video decision-making task: Speed influences on performance. *Psychology of sport and exercise, 14*(2), 293-297.

Mascarenhas, R. D., Collins, D., & Mortimer, P. (2005). The Accuracy, Agreement and Coherence of Decision-Making in Rugby Union Officials. *Journal of Sport Behavior*, *28*(3), 253-271.

Rix, G., Boyer, S., & Récopé, M. (2011). *Referee cognition as it occurs: different kinds of judgment acts.* Paper presented at the 10th International Conference on Naturalistic Decision Making (NDM 2011), Orlando, FL, USA.

Salembier, P., & Zouinar, M. (2004). Mutual intelligility and shared context, Conceptual inspirations and technological reductions. *@ctivités, 1*(2), 64-85.

Salmon, P. M. (2009). *Distributed situation awareness: Theory, measurement and application to teamwork.* Alder: Ashgate.

Theureau, J. (Ed.). (2003). *Course-of-action analysis and course-of-action centered design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Unkelbach, C., & Memmert, D. (2008). Game Management, Context Effect, and Calibration: The Case of Yellow Cards in Soccer. *Journal of Sport & Exercise Psychology, 30*, 95-109.

Vermersch, P. (1999). Introspection as practice. *Journal of consciousness studies, 6*(2-3), 17-42.

Issues of Grounding and Team Coordination in Asynchronous Mission Control-Space Crew Interactions

Ute FISCHER[a], Kathleen MOSIER[b], and Judith ORASANU[c]
[a]*Georgia Institute of Technology*
[b]*San Francisco State University*
[c]*NASA Ames Research Center*

ABSTRACT

Introduction: During long duration missions space-ground communications will involve delays up to 20 minutes one way, a reality that poses a formidable challenge to team communication and task performance. In the present paper we examined how transmission delays impacted the interactions between mission controllers and space crews and their joint performance during routine and off-nominal tasks. **Method**: Four teams of NASA mission controllers and astronauts participated in a space simulation study involving two 1.5-hour scenarios with transmission delays of 50 sec and 300 sec. Team communications were transcribed and coded. Analyses focused on communication problems as well as identified communication strategies that may have helped the mission controllers and space crews establish and maintain common ground. **Results and discussion:** Inefficiencies in team communication pertained to structural aspects (turn taking) and the content of communications (missing identifications by speakers and ambiguous listener feedback). Strategies supportive of grounding processes were also identified.

KEYWORDS

Common ground; coordination; manned space operations.

INTRODUCTION

Effective and efficient communication between Mission Control and space crews is essential for successful task performance and mission safety. The importance of team communication is heightened when unforeseen problems arise, such as system failures that are time-critical and require extensive coordination and collaboration between space and ground crews. However, fractious interactions between space crews and mission control personnel have been observed during past missions and space simulation studies and posed a risk to mission success (Kanas & Manzey, 2008). These problems are likely to be exacerbated as missions travel further from the Earth. During long duration missions and missions beyond Low Earth Orbit, space-ground communications will involve delays up to 20 minutes one way, a reality that poses a formidable challenge to team communication and task performance.

Team communication requires the collaboration between speakers and addressees. Conversational partners need to coordinate the communication process (e.g., when to speak) as well as its content (e.g., speakers present information and addressees have to confirm their understanding or request clarification) to ensure that the information becomes part of their common ground—that is, is accepted as mutually understood, accurate and relevant to shared goals (Clark 1996). To do so effectively, partners need to adapt their behavior to the opportunities and constraints associated with different communication situations and media (Brennan & Lockridge, 2006). During face-to-face interactions conversational partners are co-located and thus may presume as mutually known information that is in their shared visual field. Turn-taking between partners is rapid and in sequence, and partners may rely on gestures and facial expressions to direct the other's attention and provide feedback on their understanding. Remote partners that communicate synchronously—e.g., air traffic controllers and pilots—lack a shared visual flied and visibility; however, turn-taking can be rapid, with messages received almost instantaneously, and their order easily determined. Co-present partners are also able to indicate understanding and agreement as messages are produced. In contrast, remote collaborations that involve time delays in team members' communications come with a considerable "cognitive overhead" (Olson, G. & Olson, J., 2003). The timing of turns is challenging, and individual contributions may be out of sequence, making it difficult for team members to follow the thread of a conversation and thus to develop shared situation models (Olson, G., & Olson, J., 2003). Grounding is more effortful and misunderstandings more likely in asynchronous communication due to a lack of immediate feedback.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Investigations of asynchronous communication in domains such as telemedicine have identified communication delays as a primary impediment to effective telesurgery, and have prescribed faster transmission technology (e.g., asynchronous transfer mode) as the solution (e.g., Eadie, Seifalian, & Davidson, 2003). However, given the current limitations of earth-space transmission technology, it is essential to explore solutions that focus on communication processes per se, rather than on transmission speed.

Our project involves several studies with the overall goal to develop and validate protocols supporting mission control–space crew communication and collaboration under time-delayed conditions. In the present paper we report initial findings from a space mission simulation study. Analyses examined how transmission delays impacted the interactions between mission controllers and space crews and their joint performance during routine and off-nominal tasks. Specifically, analyses focused on communication problems as well as identified communication strategies that may have helped the mission controllers and space crews establish and maintain common ground (i.e., mutual task and situation awareness).

METHOD

The present study analyzed communications between mission controllers and space crews that were collected as part of the Autonomous Mission Operation (AMO) study conducted by Frank et al. (2012). The AMO project addressed the allocation of responsibility among flight crew, ground crew and automation given time delay between the space vehicle and earth. These issues were examined during simulated anticipated off-nominal (procedural), and unanticipated off-nominal (ill-defined) situations involving four space crew and mission control teams.

Participants

Four teams of NASA mission controllers and astronauts were recruited for the study. Each team consisted of eight mission controllers and four space crew members.

Procedure

The teams participated in six simulated space missions over two days. Simulations took place in NASA's Deep Space Habitat. Each day consisted of one training session and two experimental sessions. The first day of the experiment, the Baseline condition, involved present day equipment and space crew and mission control (MC) communicated via voice-loop. The second day, the Mitigation condition, introduced new automation and communication equipment (texting) to support space crew – mission control collaboration. Both experimental days included one scenario in which space-ground communications were delayed by 50 sec (representing missions to Near Earth Objects, NEO) and one in which the dealy was 300 sec (representing missions to Mars). Each experimental scenario lasted for 1.5 hours and required teams to complete 12 activities. One of these was either a medical emergency or a system failure (= ill-defined task); the remaining 11 activities were routine maintenance (= procedural) tasks. The present analysis considers only the team communications that occurred during the two experimental sessions on day 1 (i.e., when team members communicated by voice loop).

Communication Coding

Communications between mission controllers and space crews were transcribed and subsequently coded. Our analysis of team members' communications addessed both structural and content variables. Coding categories were adopted from past research on team communication, in particular from studies of distributed but co-present teams, such as air traffic controllers and pilots (Fischer, McDonnell, & Orasanu, 2007; Morrow & Fischer, 2013). The analysis of structural aspects concerned the timing and sequence of turns. In particular we examined whether there were turns that were out of sequence (i.e., related turns by partners did not follow each other as one partner inserted a turn before addressee could respond to the initial contribution), and we looked for instances in which mission controllers and space crew members talked over each other (i.e, step-ons). We also noted the presence (and absence) of strategies that facilitate turn taking, such as the use of specific phrases (i.e., *over*) to mark the end of one's turn. The analysis of communication content focused on participant identification, the presentation of information and on addressee feedback. As communications between mission controllers and space crew members had one designated channel, the identity of a speaker and his or her intended addressee could be ambiguous. Conventions, such as call signs that are used in Air Traffic Control (ATC) – pilot communications could mitigate against potential confusions. We also examined whether speakers and addressees employed strategies supportive of mutual understanding. In particular we looked for instances in which speakers structured complex information into concise units, or repeated critical pieces of information ; and we classified how addressees provided feedback on their understanding—that is, whether they simply acknowledged receipt, or gave more specific indications of what they understood. Instances of misunderstandings were noted, as well as the presence or absence of repair attempts.

RESULTS AND DISCUSSION

Presently the communications of three (of the four) mission control – astronaut teams have been transcribed and analysed. Accordingly, we will present here only a descriptive analysis of instances of communication problems and of strategies that likely facilitated common ground and team coordination as illustrations of our analytic approach.

Structural issues in asynchronous communication

Inefficiencies in turn taking

Disruptions in the turn sequence were observed under both time delay conditions, and involved contributions that were out of sequence or occurred simultaneously. Figure 1 depicts an exchange between Mission Control (MC) and a flight engineer (FE-2) and illustrates how a delay of 50 sec can disrupt the orderly progression of individual contributions. In the example FE-2 requested input from MC concerning step 4 in a procedure. As he did not hear back from MC in time, he proceeded with the procedure just to encounter a new ambiguity in the next step, and thus turned to MC for help. However, MC answered his initial request before they received his second request, and, apparently because they did not hear any acknowledgment from FE-2, they repeated their by now superfluous answer instead of responding to FE-2's second request. Meanwhile FE-2 repeated his second request, which ultimately got answered 4 minutes after it was posed.

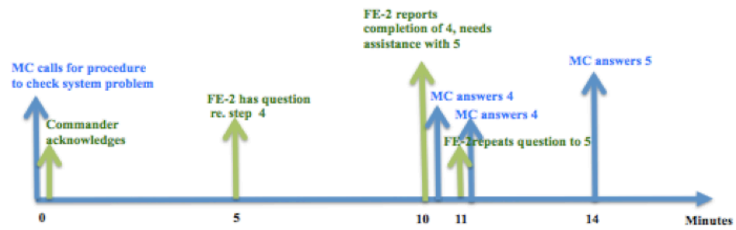

Figure 1. Depiction of an exchange in which turns are out of sequence

The turn sequence could also be disrupted by simultaneous transmissions from MC and the space crew. Step-ons occurred when a team member—for instance a space crew member—was speaking as a communication from MC came in that they had transmitted 50 sec or 300s before. Step-ons may not be easily resolved as partners could concurrently request a repeat from one another resulting in a second step-on.

Strategies to facilitate turn taking

In one of the teams we analysed, MC and space crewmembers announced a specific time (e.g., *we will have step 3 to copy in five seconds*) at which their partner could expect further communication from them. This strategy has not been observed in past research on synchronous team interactions, presumably because co-present partners can immediately respond to one another and thus an orderly progression of turns is rather effortless. On the other hand, when team members communicate asynchronously, they do not know when their partner will talk to them. Setting a time for one's communication eliminates this uncertainty at least during routine and procedural tasks and my thus mitigate both out of sequence communications and step-ons.

Mission controllers, in particular, as well as some space crewmembers tended to mark the end of their turns with phrases such as *over*; *that's all*; or *Thank you*. In so doing they explicitly handed the floor to their partner, a strategy that may support the sequencing of turns, even in non-routine situations.

Issues concerning the content of communication

Inefficient collaborations on content

Space crew members occassionally did not identify themselves when they called MC. This omission required MC to infer the identity of the caller from the content of the message. While dropped identifiers apparently did not hamper MC-space crew communications in our sample, this behavior could potentially impair mutual understanding as it creates ambiguity concerning the identity of the speaker. In time-critical or high workload situations, addressees may fail to make the correct inferences and thus may mistake the identity of the speaker and ultimately may misunderstand the intended meaning.

Listener feedback was at times not optimal as space crews and mission controllers provided minimal or ambiguous evidence of their understanding, or failed to respond altogether to a partner's communication. Minimal and ambiguous responses, such as *we copy all*, or *we copy your last* (after several transmissions by the same speaker), are short acknowledments with which addressees indicate receipt of a message and their belief that they correctly understood. However, these responses do not convey what addressees understood and thus deprive speakers of the opportunity to verify that their message was understood as intended. Read-backs are standard operational procedure in space operations and are intended to catch misunderstandings before they lead to incorrect actions. Missing responses by addressees also introduce ambiguity as they could indicate that addressees did not hear, are too busy to respond, or disagree. They likely increase speakers' workload and could result in frustration and miscommunication. For example, in one situation the flight surgeon had to instruct a space crew member (FE-3) on how to conduct an ultra sound for a bladder scan. As FE-3 did not respond to the surgeon's communications, she apparently got concerned that there was a transmission problem and finally requested: *Make sure that you copy after you received this message, please. I would like to have an understanding that you are hearing me correctly*.

Strategies to support mutual understanding

Addressees can facilitate mutual understanding by repeating what they heard and understood. In so doing, speakers can verify that their message was understood as intended and if necessary, can correct any problems. Mission controllers and space crew members predominantly followed standard operational procedures and provided proper read-backs. The value of read-backs was apparent in an instance in which a flight engineer (FE-3) misunderstood one value to be used in configuring a robotic camera. Fortunately, MC caught the error and corrected it right away. Even though there was a transmission delay of 50 sec, MC's correction reached FE-3 just after 2 min and thus in time before he could start the survey with an incorrect configuration.

Mutual understanding was likely aided by team members presenting information in a well-structured manner. Mission controllers and space crew members frequently prefixed complex messages with a summary, akin to a subject header in emails, such as *Houston FE-3 with a status on procedure 8 decimal 1*. In so doing, team members grounded their contribution in their ongoing discourse and thus likely facilitated comprehension. This strategy seems particularly helpful when communication is asynchronous and team members need to keep track of individual contributions and their relationship over an extended period of time. Likewise, mission controllers supported grounding by packaging complex instructions into meaningful chunks. A typical example of this behavior is the follwing communication by MC: *We have DPC comments for you today. First off with respect to the fluid transfer we have a supply tank level of 81% and atrium trank level of 19 % and a desired tank level of 90%. The comma value that we will use in step 2 is 39 minutes. That takes care of it for the fluid transfer. With respect to the vehicle survey in procedure 8 decimal 1 we like you to give us a heads up after you completed step 1 decimal 4. … And that's all we have from Houston.*

CONCLUSIONS

Team communication requires the collaboration of conversational partners to ensure common ground. Our analysis of mission control – space crew interactions identified some problems that distributed teams may encounter when their communications are asynchronous. We also characterized strategies that these professionals employed, apparently in an effort to overcome some of the challenges associated with remote collaborations. However, the present study is limited by its small sample and the fact that there was no synchronous condition included in the experimental design. To determine the impact of time delay on team communication we are currrently conducting a lab experiment that involves a large sample and examines team communication under synchronous and asynchronous conditions.

ACKNOWLEDGMENTS

Support for this research was provided by Cooperative Agreement (NNX12AR19G) between NASA Johnson Space Center, Georgia Tech, and San Francisco State University.

REFERENCES

Brennan, S. E., & Lockridge, C. B. (2006). Computer-mediated communication: A cognitive science approach. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics, 2nd Edition (pp. 775-780)*. Oxford, UK: Elsevier Ltd.

Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.

Eadie, L. , Seifalian, A. , & Davidson, B. (2003). Telemedicine in Surgery. *The British Journal of Surgery*, *90*(6), 647-658.

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space and Environmental Medicine, 78*(5), II, B86-95.

Frank, J. et al., (submitted to IEEE). *Autonomous Mission Operations*.

Kanas, N., & Manzey, D.(2008). *Space psychology and psychiatry* (2nd edition). El Segundo, CA: Microcosm Press.

Morrow, D.G. & Fischer, U.M. (2013). Communication in socio-technical systems. In J. D. Lee & A. Kirlik (eds.), *Oxford Handbook of Cognitive Engineering*. New York, NY: Oxford University Press.

Olson, G. M., & Olson, J. S. (2003). Mitigating the effects of distance on collaborative intellectual work. *Economics of Innovation and New Technologies*, *12*, 27-42.

Divided-attention task on driving simulator: comparison among three groups of drivers

Chloé FREYDIER[ab], Julie PAXION[a], Catherine BERTHELON[a]
and Mireille BASTIEN-TONIAZZO[b]
[a] French Institute of Science and Technology for Transport, Development and Networks (IFSTTAR) – Salon de Provence
[b] National Center for Scientific Research (CNRS), UMR 6057 – Aix en Provence

ABSTRACT

Introduction: Driving is a complex and dynamic task that requires performing simultaneously several sub-tasks, as traffic management and vehicle control. Driving involves both automatic and controlled processing depending on situation met and drivers' experience. **Method**: Three groups of drivers with different driving experience were submitted to a divided-attention task in order to assess the interference linked to a secondary task on driving behaviour. The main task was a car-following task and the secondary task was a number identification task which could appear on central or peripheral vision. **Results and discussion:** Results showed that driving performances increase with experience. Indeed, novice drivers, compared to more experienced drivers, took more time to brake and had more difficulties to maintain a stable position in the lane. This task allowed to differentiate driving behaviour depending on experience and could be used in training of novice drivers.

KEYWORDS

Secondary task; simulator; driving experience

INTRODUCTION

Many factors can negatively influence driver's behaviours and cause road accidents. Among them, the lack of experience is recognized as a main factor contributing to road accidents. Indeed, young drivers are widely overrepresented in road accidents so that, in France, it is the first cause of death among those under 25 years (Page, Ouimet & Cuny, 2004). Young drivers are particularly sensitive to distraction effect and have a higher risk of road accident linked to distraction (Hosking, Young & Regan, 2009). According to researchers, the lack of experience may be an additional reason that novice drivers are more easily distracted (Pradhan, Divekar, Masserang, Romoser, Zafian, Blomberg, Thomas, Reagan, Knodler, Pollatsek & Fisher 2011). Otherwise, attention is one of the most important factors and a primary cognitive requirement for safe driving. According to studies, attentional defaults are involved in about 22% to 50% of road accidents (Hendricks, Fells, & Freedman, 2001 ; Klauer, Dingus, Neale, Sudweeks, & Ramsey 2006 ; Lee, & Strayer, 2004 ; Van Elslande, Perez, Fouquet & Nachtergaele, 2005). Divided-attention is widely involved in driving activities because it consist in various subtasks. For example, driving implies maintain a stable trajectory while being careful with traffic. Each task requires an amount of attention. According to Kahneman (1973) cognitive system has a limited amount of attentional resources, and it's the amount of available resources engaged in the task that will determine quality, efficiency and deep of cognitive processing. Many studies using a driving simulator show that performing a secondary task deteriorates driving performances and increases reaction time (Andersen, Bian & Kang 2011 ; Atchley & Chan 2011 ; Bian, Kang & Andersen, 2010 ; Brookhuis & De Waard, 1994 ;; Horberry, Anderson, Regan, Triggs & Brown, 2006 ; Hosking & al, 2009 ; Lamble, Kauranen, Laakso & Summala, 1999). For example, using a mobile phone during a car following task increases the time-to-collision , the response time to brake (Lamble & al,1999) and delays the reaction time to headway changes (Brookhuis & De Waard, 1994). Driver's distraction by visual secondary task leads to an increase of mistake production (Young, Salmon, & Cornelissen, 2012). This deterioration of performance linked to an additional task is confirmed by a study carried out on real-environment (Blanco, Biever, Gallagher & Dingus, 2006). When drivers perform simultaneously several tasks, they are placed in a divided attention situation. The secondary task draws the driver's attention away from the driving task and the driver has to divide its attentional resources between driving and the secondary task. Studies linked to driver's distraction by an secondary task were the object of experiments related to phone and assistance device and experiments on ocular movements. It seems to be very interesting to complete these works with an approach in terms of attention resources.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

In this context and according to Rasmussen (1983) Three behaviour levels are distinguished : 1. Skill based behaviour where the risk is related to routine errors, inattention errors. This level involves majoritarily automatic processing which requiere few attentional resources ; 2. Rule-based behaviour which is activated if it is required by the situation. At this level, the risk is the activation of rules which are maladjusted to the situation ; 3. Knowledge based behaviour, in the case where no rules allow to solve the problem. The operator have to improvisate and find a solution based on anterior knowledge. The risk is not to find the solution. This level involves controlled processing which requiere a lot of attentional resources. This model is relevant to explain differences between novice and experienced drivers. Indeed, experienced drivers use majoritarily skill-based behaviour, while novice drivers, whose behaviours are not totally automatized, use more often a knowledge-based behaviour.

Our main hypothesis is that, due their lack of skill-based behaviour, young drivers will be less competent than experienced drivers to manage a dual-task.

Our second hypothesis is that the divided-attention capacity is gradually obtained and depends of the driving experience.

METHOD

Participants

Thirty-seven volunteers, divided up into three groups, participated in the study. The first group was consisted in 13 novices drivers aged 18 and had their license for less than 4 months. The second group was made up of 12 young drivers more experienced, i.e. with 36 months of driving license and aged 21. The third group was composed of 12 drivers more aged and experienced, i.e. with at least 8 years of driving license. All experienced drivers had their own vehicle and traveled over 20,000 km.

Simulator

The driving study was carried out on the SIM^2-IFSTTAR fixed-base driving simulator equipped with an ARCHISIM object database. The driving station comprised one quarter of a vehicle. The image projection (30Hz) surface filled an angular opening that spans 150° horizontally and 40° vertically. The vehicle had an automatic gearbox and was not equipped with rearview mirrors.

Procedure & Statistical analyses

Participants performed training before the beginning of the experiment. The main task consisted in following a vehicle while keeping a constant distance with this vehicle. The lead vehicle speed varied with sixteen accelerations and sixteen decelerations either with high or low amplitude. The driver was placed in the middle of three-lane road, in a way that the environment was perfectly symmetric. The secondary task consisted in the number identification, even or odd. A three digit number appeared each 1.5 seconds to 2.5 seconds during 400 milliseconds on a central or peripheral visual field. Participants had to activate the right control of the steering wheel if the target was even or the left control if the target was odd. Before the dual task and in order to obtain a reference measure of driving performance, volunteers performed a car-following task alone. The secondary task was also performed alone in order to ensure that it lead to a similar cognitive cost for novice and experienced drivers. The order of presentation of these two simple tasks was counterbalanced during the experiment. Then, drivers perform a dual-task: simultaneously a car following task and an identification task of the number parity in central and peripheral visual field. The interference related to the secondary task was computed and compared with reference measures. This divided-attention task seems relevant because it use the same perceptual channel (visual) for the two tasks and according to Wickens (2002), it is a good way to highlight an interference between two tasks. Moreover, we have specifically chosen the secondary-task apart from driving context in order to avoid possible learning effect linked to driving experience.

The dependent variables were divided into four categories: 1. Speed: index of speed regulation (time taken by drivers to reach the same speed of the lead vehicle) and reaction time to press on the brake; 2. Distance: mean and standard deviation of inter-vehicular distance; 3. Position in the traffic lane: mean and standard deviation of lane position (SDLP) and 4. Performance on the secondary task: reaction time and accuracy (percentage of correct, incorrect responses and omissions). The independent variables were driving experience (three groups) and task (single vs dual). Statistical analyses were performed using Statistica. The data was tested for significance using repeated measures analysis of variance ($p \leq 0.05$). Results from the divided-attention task were compared to performances in the reference task. Bonferroni post hoc tests were subsequently used for pairwise comparisons.

RESULTS

Only significant results are presented in this paper. ANOVA revealed a significant main effect of the dual-task. The analyses highlighted an impairment of performance in dual task compared to the single-task (car-following) on the time necessary to reach the same speed of the lead vehicle and on the mean inter-vehicular distance. In dual-task, the percentage of correct responses decreased and the percentage of omissions increased compared to those in single-task (number's identification) (See Table 1).

Table 1. Performances depending on task (single vs. dual)

Dependant variable	Single-task	Dual-task	F
Index of speed regulation (sec)	m= 11.04 (sd= 1.66)	m= 10.34 (sd= 2.24)	F(1,31) = 8.85 , p ≤ 0.005
DIV mean (m)	m= 54.92 (sd= 14.37)	m= 50.08 (sd= 8.57)	F(1,31) = 4.61 , p ≤ 0.05
Response correct (%)	m= 87 (sd = 0.11)	m= 91.2 (sd= 0.07)	F(1,31) = 24.87, p ≤ 0.001
Omission (%)	m= 2.9 (sd = 0.04)	m= 6.1 (sd= 0.1)	F(1,31) = 13.27, p ≤ 0.001

An interaction between task and number's location (central vs peripheral) on the response accuracy (correct responses F(1,31) = 12.77 , p <0.005 and omissions F(1,31) = 11.33, p < 0.005) showed that these impairments in dual task occurred only when the numbers appeared in peripheral vision (See Figure 1 & 2).

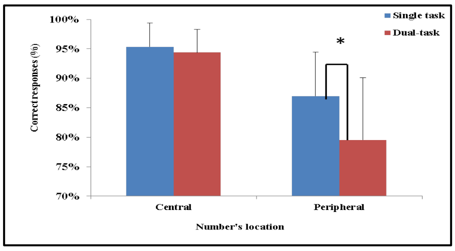

Figure 1. Correct responses' percentage depending on task (single vs. dual) and number's location (central vs. peripheral)

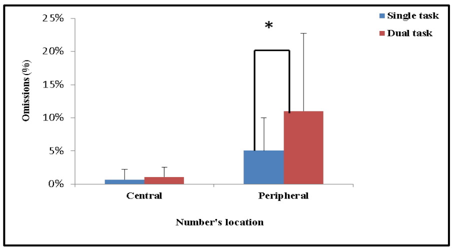

Figure 2. Omissions' percentage depending on task (single vs. dual) and number's location (central vs. peripheral)

ANOVA also revealed a significant decrease of time to brake (F(2,31) = 4.8, p ≤ 0.05) and of standard deviation of lateral position (SDLP) (F(2,31) = 8.22, p ≤ 0.005) with driving experience. Conversely, the percentage of correct responses (F(2,31) = 4.91, p ≤ 0.05) increased with driving experience (See Table 2).

Table 2. Performances depending on driving experience

Dependant variable	Novice	Experienced (36 months)	Experienced (8 years)	F
Time to brake (sec)	m= 12.39 (sd= 3.12)	m= 10.45 (sd= 3.73)	m= 8.55 (sd= 3.67)	F(2,31) = 4.80 , p ≤ 0.05
SDLP (m)	m= 62.82 (sd= 10.72)	m= 55.65 (sd= 15.55)	m= 45.78 (sd= 92.99)	F(2,31) = 8.22 , p ≤ 0.005
Response correct (%)	m= 86.2 (sd = 11.2)	m= 90 (sd= 8.4)	m= 91.4 (sd= 7.6)	F(2,31) = 4.91 , p ≤ 0.05

DISCUSSION

The aim of the present research was to explore the divided-attention abilities of young novice drivers compared to experienced drivers, and to determine whether young novice drivers were particularly affected by the interference linked to secondary task.

Globally, compared to a single-task, a divided-attention task affects negatively both driving performance and identification abilities on the secondary task. These results are congruent with the literature concerning a dual-task. Performing two tasks simultaneously causes an impairment of performance compared to the realization of each task separately. In accordance with the limited capacity theory of attention (Kahneman, 1973) performing a divided-attention task causes an overload and task demands can exceed driver's attentional resources.

Interaction effects between number's locations and performances indicate that impairment linked to the secondary task occurs only when numbers appear in peripheral vision. Results are consistent with eye-movement's studies while driving which indicate that when situation complexity increases, drivers focus their glances on central visual field (Underwood, 2007).

Globally, novice drivers brake later than experienced drivers with 36 months or 8 years of driving experience. The skills necessary to rapidly brake would be therefore acquired from 36 months of driving. Otherwise, novices and experienced from 36 months, had higher SDLP than drivers with 8 years of driving license. As SDLP reflects the skills of lane-keeping, these later don't seem completely acquired even after 36 months of driving license. Finally, novice drivers' percentage of correct responses is fewer than experienced ones, which is directly linked to driving inexperience. In terms of an amount of available resources, the main task (car-following) is not automatic yet for this group of drivers, and thus requires knowledge-based behaviour (Rasmussen, 1983) and many cognitive resources. Performing a secondary task has thus for consequence a mental overload and an inadequate attentional resource distribution. For novice drivers, attentional resources are highly mobilized by sensori-motor sub-tasks to the detriment of cognitive sub-tasks. These outcomes are consistent with previous

studies on distraction among young novice drivers (Stutts, Reinfurt, Staplin & Rodgman, 2001 ; Metz, Schömig & Krüger, 2011).
Moreover, the divided-attention task is a relevant choice because, conversely to a single-task, it allows experienced drivers with 36 months of driving to be distinguished of those who have 8 years of experience. It shows that even with 36 months of practice, all resources necessary to safe driving are not available.
In summary, as expected, experienced drivers have greater driving performance and use mainly skill-based behaviour. Young drivers have not completely acquired these skills, even after 3 years of driving experience. This study shows a progressive acquirement of skills necessary to safe driving and it could be improve by adequate training in driving simulator.

REFERENCES

Andersen, G. J., Bian, R. N. & Kang, J. (2011). Limits of spatial attention in three-dimensional space and dual-task driving performance. *Accident Analysis and Prevention,* 43, 381-390.

Atchley, P. & Chan, M. (2011). Potential benefits and cost of concurrent task engagement to maintain vigilance : a driving simultaor investigation. *Human Factors*, 53, 1, 3-12.

Bian, Z., Kang, J.J. & Andersen, G. J. (2010). Changes in extent of spatial attention with increased workload in dual-task driving. *Transportation Research Board of the National Academies*, Washington, D.C., 8-14.

Blanco, M., Biever, W.J., Gallagher, J. P. & Dingus, T.A. (2006). The impact of secondary task cognitive processing demand on driving performance. *Accident Analysis and Prevention,* 38, 895-906.

Brookhuis, K. & De Waard, D. (1994). Measuring driving performance by car-following in traffic. Ergonomics 37, 427-434.

Hendricks, D. L., Fells, J. C., & Freedman, M. (2001). The relative frequency of unsafe driving acts in serious traffic crashes. Washington (DC) : National Highway Traffic Safety Administration. Http://www.nhtsa.dot.gov/people/injury/research/UDAshortrpt/documentation_page.html.

Horberry, T., Anderson, J., Regan, M.A., Triggs, T. J. & Brown, J. (2006). Driver distraction : The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis and Prevention*, 38, 185-191.

Hosking, S.G., Young, K.L. & Regan, M.A. (2009). The effects of text messaging on young drivers. *Human Factors, 51*. Récupéré du site de la revue : http://dx.doi.org/10.1177/0018720809341575

Kahneman, D. (1973). *Attention and effort*: Englewood Cliffs, NJ: Prentice-All.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. & Ramsey, D. J. (2006). The impact od driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. National Highway Traffic Safety Administration, Washington DC.

Lamble, D., Kauranen, T., Laakso, M. & Summala, H. (1999). Cognitive load and detection thresholds in car following situations : safety implications for using mobile (cellular) telephones while driving. *Accident Analysis and Prevention*, 31, 617-623.

Lee, D. L., & Strayer, D. L. (2004). Preface to special section on driver distraction. *Human Factors, 46* (4), 583-586.

Metz, B., Schömig, N. & Krüger, H.P. (2011). Attention during visual secondary tasks in driving: Adapatation to the demands of the driving task. Transportation Research Part F, 14, pp. 369-380.

Page, Y., Ouimet, M.C. & Cuny, S. (2004). An evaluation of the effectiveness of the supervised driver-training system in France, Annual Proceedings of the Association for the advancement of automotive medicine.

Pradhan, A.K., Divekar, G., Masserang, K., Romoser, M., Zafian, T., Blomberg, R.D., Thomas, F.D., Reagan, I., Knodler, M., Pollatsek, A. & Fisher, D.L. (2011). The effects of focused attention training on the duration of novice drivers' glances inside the vehicle. *Ergonomics, 54*:10, 917-931.

Rasmussen, J. (1983). Skill; rules, knowledge:signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions on Systems, Man & Cybernetics, 13*, 257-267.

Stutts, J.C., Reinfurt, D.W., Staplin, L., & Rodgman, E.A. (2001). The role of driver distraction in traffic crashes. Washington, DC: University of North Carolina.

Underwood, G. (2007). Visual attention and the transition from novice to advanced drivers. *Ergonomics, 50*:8, 1235-1249.

Van Elslande, P., Perez, E., Fouquet, K. & Nachtergaele, C. (2005). De la vigilance à l'attention : déclinaison des problèmes liés à l'état psychophysiologique et cognitif du conducteur, et analyse de leur influence sur les mécanismes d'accidents. Rapport de convention INRETS/DRAST, INRETS/RE-05-918-FR.62p.

Young, K. L., Salmon, P. M. & Cornelissen, M. (2012). Distraction-induced driving error: An on-road examination of the errors made by distracted and undistracted drivers. Accident Analysis and Prevention. In press.

Transforming Knowledge: Capturing Engineers' Cognitive Expertise

Julie GORE
University of Surrey, UK

ABSTRACT
Introduction: This paper reports the initial stages of an on-going project with leading engineering experts working with the manufacture of petroleum additives. The project aims to provide transformative, effective methodological solutions to Knowledge Management transfer. **Method:** Engineering experts are being trained to use Applied Cognitive Task Analysis (ACTA) techniques in order to document the cognition of their expert peers. **Results and discussion:** Results to date have had high face validity and otherwise undocumented knowledge has been elicited. A key challenge here is to ensure the practitioners' capture cognition and continue to provide transformative innovations, and continuous knowledge transfer within this highly intelligent workforce.

KEYWORDS
Applied cognitive task analysis, engineering, expertise, situation-assessment, scenario planning

INTRODUCTION

Capturing, documenting and utilising professional cognitive expertise where technological excellence is fundamental to success is a critical challenge for global organisations. To date the naturalistic decision making (NDM) community and related communities of practice have successfully reported the merits of ACTA and associated CTA techniques (Hoffman & Militello, 2008; Militello, Wong, Kirschenbaum & Patterson, 2011). More recently these techniques have also begun to appear in other research areas of organisational behaviour and management practice (Gore and McAndrew, 2009; McAndrew & Gore, 2012; Osland, 2010) however, reports which focus upon the training of practitioners to adopt such techniques is less well documented. This work therefore also examines the importance of the role of academics translating methodological research developments for explorations *of* and *in* professional knowledge management practice (Anderson, 2007).

The researcher was asked to explore (within a much wider organisational project on knowledge management) how best expert cognition in engineering expertise could be elicited, documented and utilised. This poster shares the process of training transfer and the illustrative results of a practitioner CTA. The expert cognition associated with managing uncertainty is highlighted (Lipshitz & Strauss, 1997) and aspects of hot/sensory based cognition explored.

The task illustrated here is from an expert process engineers analysis of the key cognitive demands involved in the task of completing an "initial manufacturing plant trial start-up". When new petroleum additives are developed, teams of research engineers with professional expertise in chemistry and physics run pilot or 'start up' trials within the manufacturing plant. This process has a high degree of risk associated with it and can be expensive and time-consuming, in terms of both health and safety and continued product innovation success. Process and product development are also subject to unique legal confidentiality agreements in this area of engineering. The task of setting up and monitoring a plant trial therefore involves a high degree of macro-cognitive complexity within a NDM type framework (Orasanu & Connolly, 1993). The macro-cognitive complexity i.e. "the cognitive adaptation to complexity" (Klein, et al 2003; Klein & Hoffman, 2008) can involve several experts managing risk and uncertainty, making sense of their dynamic environment, with high stakes and shifting, ill-defined goals under time-pressured situations.

METHOD

A one day (7 hour) briefing about the use of ACTA techniques was provided for a small group of professionals with different areas of engineering expertise. The cognitive areas ACTA and CTA seek to address are: difficult judgments and decisions; attentional demands; critical cues and problem-solving strategies. During a second day the researcher trained 3 engineers to use a selection of the ACTA techniques (Militello and Hutton, 1998 recommend that three to five subject matter experts usually exhaust the domain of analysis).

First, the researcher completed a task diagram and knowledge audit in order to illustrate the interview techniques associated with stage one and two of ACTA. This process was stopped and re-started in order for the engineers

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

to ask questions and clarify the process. The first stage of ACTA a production of a *task diagram*, provides the interviewer with a broad overview of the task. This interview helps identify areas requiring complex cognitive skills which can be examined in depth in stage 2 of the process: *the knowledge audit*. In order to identify the type of tasks which were seen to be essential by the expert engineers, task diagrams were completed for key areas of engineering work which involved cognitive complexity. It is this type of work the organisation recognised was not currently documented meaningfully in training procedures. The professionals involved in the knowledge management project were mindful that areas of expert cognition were not being transferred to novice engineers.
Second, the engineers practiced knowledge audit techniques with each other and documented their understanding of complex cognition. Again, a stop – start approach was adopted to facilitate the question technique and the documentation of knowledge elicited. *The knowledge audit* focuses upon a cognitive sub-task elicited from the task diagram and is well documented in the research literature in expert-novice differences (Crandall et al, 2006). Table 1 illustrates key questions and probes.
An optional third stage, *the simulation interview* assists the understanding of participants' cognition within the context of a challenging scenario. This can be useful for developing training recommendations and is an area of ongoing work with the organisation. Finally, a cognitive demands table was completed by the engineers, providing an analytical summary of data elicited.

Table 1: Knowledge Audit probes.

1. **Past and future** : Experts know how the situation developed and know where the situation is going (de Groot 1946/1978; Endsley, 1995; Klein & Crandall, 1995; Klein & Hoffman, 1993): Is there a time when you walked into the middle of a situation and knew exactly how things got there and where they were headed?

2. **Big picture**: Experts understand the whole situation and understand how elements fit together (Endsley, 1995; Klein, 1997): Can you give me an example of the big picture for this task? What are the major elements you have to know and keep track of?

3. **Noticing**: Experts can detect cues and see meaningful patterns (de Groot, 1946/1978; Klein & Hoffman, 1993; Shanteau, 1985): Have you had experiences where part of a situation just 'popped' out at you; where you noticed things going on that others did not catch? What is an example?

4. **Tricks of the trade/Job smarts**: Experts can combine procedures and do not waste time and resources (Gore, 2004; Klein & Hoffman, 1993): When you do this task, are there ways of working smart or accomplishing more with less i.e. tricks of the trade – that you have found particularly useful?

5. **Improvising / opportunities** : Experts can see beyond standard operating procedures and take advantage of opportunities (Dreyfus & Dreyfus, 1986; Shanteau, 1985): Can you think of an example when you have improvise in this task or noticed an opportunity to do something better?

6. **Self-monitoring** : Experts are aware of their own performance and notice when performance is not what it should be and adjust to get the job done (Cohen, Freeman & Wolf, 1996; Glaser & Chi, 1988): Can you think of a time when you realized that you would need to change the way you were performing in order to get a job done?

7. **Anomalies**: Experts can spot the unusual and detect deviations from the norm (Klein, 1989; Klein, 1997; Klein & Hoffman, 1993): Can you describe an instance where you spotted a deviation from the norm, or knew something was amiss?

8. **Equipment difficulties**: Experts know equipment can mislead and do not implicitly trust equipment as novices might (Cannon-Bowers, Salas & Converse, 1993): Have there been times when the equipment pointed in one direction, but your own judgment told you to do something else? Or when you had to rely on experience to avoid being led astray by the equipment?

Adapted from Militello & Hutton (1998)

RESULTS

At first the engineers found the process of interviewing and being interviewed using the CTA techniques a little challenging. They quickly however, became very comfortable using the knowledge audit probes and found the structured approach very rewarding. Each of the engineers reported that the knowledge elicited, including key cues for improving situation awareness and scenario planning had rarely been documented in such a pragmatic way previously. The results of an independently completed cognitive demands table completed by the experienced CTA researcher was comparable to the engineers' interpretations (See Table 2). Here we see clear areas of cognitive complexity for the expert plant trial engineers: *effective communication*; *planning ahead* (responding to action-feedback loops); and *noting key technical cues and strategies*. By eliciting the macro-cognitive processes involved in this challenging task, the experts have noted why these three areas can be difficult (especially for novices), this includes time-pressure, using insightful information rather drowning in data (the engineers can request many reports to check the development of the new petroleum substance and be easily overloaded); and assessing the situation in order to be flexible. The engineers then go onto make explicit suggestions about how to improve and sharpen the cognition required to complete a successful trial, for example always over sample progress reports and recognise there is no need to analyse all of them under time-pressure.
In addition, an important feature which emerged to the surprise of the engineers was the importance of hot/sensory based cognition. Documenting cues associated with emotions and sensory based cognition with

specific tasks may then lead to improvements in situation awareness, especially concerning hazards and potential errors in the manufacturing process. For example several engineers described noticing peculiar smells at 2 am in the morning which resulted in adjusting the manufacturing process before the new petroleum additive was destroyed, making significant economic savings and avoiding potential hazards.

Table 2. Illustrative Cognitive Demands Table

For Plant Trials/Start Up

Difficult Cognitive Element	Why difficult?	Common errors	Cues & strategies used
Effective communication	• Maximising information sources requires effort and maintenance • Sorting important information from nice-to-have • Possible language barrier	• Excessive focus on one direct contact • Neglecting applications engineers	• Get involved in operator training • Get out and about, talk to operators and analytical chemists • Regular e mail summary update to customers/interested parties
Planning ahead / sticking to plan	• Time consuming & tedious • Misplaced fear of corrective feedback to published plan • Concern about "planning for failure • Plant may want to complete trial to return to regular production	• Inadequate plan • Fail to document possibility of unexpected events • Tendency to panic when faced with a "surprise" • Inadequate sampling schedule • Not knowing when to be flexible	• Publish and agree in advance • Include "what if" scenarios and plan for off spec product • "Phone a friend" – use all available resources, collaboration • Over sample, no need to analyse all of them • Send some samples to own lab • Apply full testing schedule to 1st batch, be flexible later if appropriate
Technicals • Compare plant samples with lab programme • Test Property trend • hallenge analytical data • DCS constraints • Commissioning	• Too much control room time • Separating important data • Embarrassment factor • Believe briefing	• Don't visit unit/lab • Neglect logical analysis • Only look at numbers, not the samples • Underestimate complexity and risk • Wrong rotation direction	• Get out of the control room • Plot or tabulate data • Diplomacy • Get face to face contact • Ask to meet • Be present for the commissioning testing

CONCLUSION

Whilst this work is on-going it aims to be original in its application as few studies document such applied innovation including practitioners with the co-construction of knowledge. The elicited scenarios will aim to assist novice engineering professionals raise their situation awareness in relation to specific tasks, with areas of cognitive complexity being clearly defined in an organisational based repository of training scenarios. Further, more detailed work is currently being completed in this area which should support knowledge management development within the organisation. In addition, further work needs to be completed to assess if all of the professional engineers can easily utilise the ACTA techniques, assisting organisational learning in order to provide transformative innovations to knowledge management and support macro-cognitive awareness.

ACKNOWLEDGMENTS

Thank you to Infineum UK, for their continued support with this on-going work.

REFERENCES

Anderson, N. (2007). The practitioner-research divide re-visited: strategic-level bridges and the role of IWO psychologists. *Journal of Occupational and Organizational Psychology*. 80, 2,175-183.

Gore, J. & McAndrew, C. (2009). Accessing expert cognition, *The Psychologist*, 22, 3, 218-219.

Hoffman, R.R., & Militello, L.G. (2008). *Perspectives on cognitive task analysis: Historical origins and modern communities of practice*. New York, NY: Taylor & Francis.

Klein, G., Ross. K.G., Moon, B.M., Klein.D.E., Hoffman, R.R., Hollnagel. E. (2003). Macrocognition. *IEEE Intelligent Systems*, 18, 3, 81-85.

Klein, G., & Hoffman, R. (2008). Macrocognition, mental models, and cognitive task analysis methodology. In J. M. Schraagen, L. Militello, T. Ormerod, & R. Lipshitz, (Eds.). *Naturalistic decision making and macrocognition.* (pp. 57-80). Ashgate: Hampshire, U.K.

Lipshitz, R. and O. Strauss (1997). Coping with Uncertainty: A Naturalistic Decision-Making Analysis. *Organizational Behavior and Human Decision Processes,* 69, 2,149-163

McAndrew, C. & Gore, J. (2012). Understanding preferences in experience-based choice: a study of cognition in the wild. *Journal of Cognitive Engineering and Decision Making*, on line in advance.

McAndrew, C. & Gore, J. (2010). "Convince Me..." An Inter-Disciplinary Study of NDM and Portfolio Managers. In Mosier, K.L. & Fischer, U.M. (Eds.) *Informed by Knowledge: Expert Performance in Complex Situations.* Psychologist Press , 353-369.

Militello, L.G. & Hutton, R.J.B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for

understanding cognitive task demands, *Ergonomics*, 41, 11, 1618-1641.

Militello, L.G., Wong, W., Kirsechenbaum, S.K., and Patterson, E. (2010). Systematizing discovery in cognitive task analysis. In Mosier, K.L. & Fischer, U.M. (Eds.) *Informed by knowledge: Expert performance in complex situations*: 23-40. New York, NY: Taylor and Francis.

Orasanu, J. & Connolly, T. (1993). The reinvention of decision-making. In: G.A. Klein, J. Orasanu, R. Calderwood & C.E. Zsambok (Eds.), *Decision making in Action: Models and methods:*3-20. Ablex, Norwood, NJ.

Osland, J. S. (2010). Expert cognition and sense-making in the global organization leadership context: A case study. In Mosier, K.L. & Fischer, U.M. (Eds.) *Informed by knowledge: Expert performance in complex situations*: 23-40. New York, NY: Taylor and Francis.

Schraagen, J.M.C., Militello, L., Ormerod, T., & Lipshitz, R. (2008) (Eds), *Naturalistic decision making and macrocognition*. Aldershot, UK: Ashgate Publishing Limited.

How is the Diagnosis Made? The Observation of Paramedics Performance in Simulated Competition Task[1]

Jitka GURŇÁKOVÁ[a] and Hana HARENČÁROVÁ[a]
[a]*Institute of Experimental Psychology, Slovak Academy of Science, Bratislava*[2]

ABSTRACT

Introduction: We examined the process of decision making related to diagnosis in paramedic teams in an international competition. **Method**: Observation of 28 paramedic teams in selected task was compared with objective medical evaluation of their performance in the whole competition. **Results and discussion:** The real process of examining the patient and establishing the diagnosis by the paramedics is not in accordance with the prescribed procedures. Paramedics show a tendency to make assumptions about the case from early steps of dealing with it, which has a strong influence on the subsequent process of examination of the patient and establishing diagnosis.

KEYWORDS

Problem solving; health; simulation; paramedics; making the diagnosis.

INTRODUCTION

In this paper we examined the process of making the diagnosis in paramedic teams. Situations which paramedics usually have to deal with can be classified as a complex problem solving, as they are dynamic, time-dependent and complex (Quesada, Kintsch & Gomez, 2005).

As stated by Dekker (2005), doctors are 7 500 times more likely to kill somebody by mistake then gun owners are. The problem of human error is even more important in the field of emergency medicine. In the study of Najaf-Zadeh, Dubos, Pruvost, Bons-Letouzey, Amalberti and Martinot (2011) the most common alleged misadventures were the diagnosis-related error (47%). Moreover, Rittenberger, Beck and Paris (2005) found frequently occurring deviations from the prescribed protocol.

However, in complex real world situations it is very complicated to objectively evaluate the performance of paramedics. One possibility how to overcome this difficulty is to observe the performance of paramedic teams at a competition consisting of a wide range of tasks with different level of complexity. This form of data collection enables also to compare different teams in the same situation and to gain access to objective assessment of quality of team performance from the point of view of different independent judges – experts in the field of emergency medicine. Understanding how the rescue teams deal with real time problems can provide us with useful insights into how the trainings or medic tools should be designed to support the successful resolving of the situations.

Aim

Aim of this study was to assess the frequency and nature of deviations from a standardized treatment protocol of paramedic teams in the selected competition task.

METHOD

The paramedic teams were observed and videotaped at one selected task at the international competition „Rallye Rejvíz 2010“. The competition had a strict time schedule where the teams had to solve 13 tasks of different level of difficulty in the course of 24 hours. The task we have selected for the analysis was part of the first 12-hours day period. The competition tasks were modelled as close to the real problems that paramedics can encounter as possible. The competitors were not just under time pressure, they also had to cope with unclear instructions, lack of feedback and physical fatigue. Contrary to real life situations they had to deal with the subjectively perceived pressure of public assessment by the judges and occasional onlookers.

[1] The study was supported by Grant Agency VEGA no. 2/0095/10
[2] Authors' email addresses: jitka.gurnakova@savba.sk, hana.harencarova@savba.sk

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Participants

We observed all 28 teams in the competition (46 men, 10 women). The teams consisted of two paramedics without physician. 20 teams were from the Czech Republic and 8 teams came from Slovakia. We obtained additional information from 24 participants from the total number of 56 participants (42.86%). Anyway, we found neither statistically significant difference among these participants and the rest of group in their level of performance, nor significant relationship between their performance, age, years of practice or number of participations at previous competitions.

Procedure

In the selected task the team had to execute simple basic examination of a healthy, communicating man. After arriving at the position, they received the following instruction: „*Emergency dispatch centre received an emergency call and sent you to: Call from wife: husband at home after night shift, sounds "strange" on the phone, unusual speech, reluctant. Confirmation phone call home – no response. Your goals are: assess situation and patient, examine patient and recommend care and treatment.*" The instruction was followed by detailed description of possibilities of the next transport of the patient into three hospitals with different facilities and different distance from the spot. The timer started after handing the instruction back to the judge; the limit was 15 minutes.

Analysis

The quantitative analysis of videotaped materials of all 28 paramedic teams was conducted in order to determine the chronology and selection of employed diagnostic procedures. On the basis of the assessment of judges in the course of the whole competition we selected 10 teams for the transcription and a detailed qualitative analysis. The selected teams represented specific cases of solving the task.
We divided all teams according to the total score they achieved in the competition into three performance groups and compared them in the variables related to the performance of the team.

RESULTS

Processing of initial information from emergency dispatch centre

We created two catergories of dealing with written instructions representing a usual call from emergency dispatch centre. The first represented the way of reading the instruction, the second was related to further information processing (e.g., selection of the key information, sense making about the situation on the spot). Further information processing slightly correlated with higher evaluation of approach towards the patient from the patient himself ($r_s = 0.342$; $p = 0.041$). Moreover, despite no significant differences, there was some tendency of more successful teams to study and process information in the instruction more thoroughly.
Qualitative analysis of special cases showed that some of the most successful teams, by reading aloud, were simultaneously selecting key information about the patient, formed an appropriate image of the situation on the spot and also selected the best possibilities for further most probable treatment of the patient. Interaction with team colleague regarding the instruction enabled them to correct the primary misunderstandings. In some teams from the first and second group according to performance at least one of the members read carefully the possibilities of next transport of the patient.
On the contrary, the least successful strategy was not paying full attention to the instructions – reading it on the go, ignoring the possibilities of patient treatment, inadequate understanding of situation or fixation on the anamnestic detail – "he was speaking strangely". Inefficient was to create hypotheses in advance about the patient if they negatively influence the approach to the patient (bias, suspicion of mental disorder from which disrespect for the patient resulted).

Anamnesis and examination

An ideal solution of the task was careful and systematic taking of the anamnesis, thorough examination: physical examination, measuring non-invasive values, establishing glycaemia, 12-lead ECG and complete basic neurological examination. Even the most successful teams did not perform the task without mistakes. The procedures differed in how many times they were used, some of the procedures were rare or totally absent (Table 1.).

Table 1. Examinations used according to their frequency

% of teams	Number of teams	
75-100%	21-28	examination – blood pressure, oxygen saturation, pulse, glycaemia, pupils, ECG (11correct /11 incorrect), neurological examination – squeeze of the hand anamnesis – medicaments, treatment
50-74%	14-20	anamnesis – actual problem, actual state, previous examination, allergies, pain, occupation, what happened, alcohol, injury, accident
25 – 49%	6-13	examination – breathing, neurological examination – tongue, sticking out the tongue, arms stretched forward, finger-to-nose test; temperature, stiffness of the neck, hearth, stomach, head, legs anamnesis – pain in a chest, food, fluids, workload, diabetes, previous illnesses, breathing, medicaments

The teams with a narrow focus just on the application of their examination procedure often overlooked the important anamnestic information spontaneously provided by the patient. These teams were also forgetting to ask about possible connections with actual state of patient's health. A complete opposite were teams that, although listening actively to the patient, relied completely on his information and did not verify his state by examination. According to our observation nobody did a complete examination. Three categories of teams according to their performance in the whole competition showed significant differences in the number of examination procedures (K-W = 9.435; p = 0.009; Table 2), but despite some descending tendency not in the number of anamnestic data obtained in selected task (Table 2.).

Table 2. Number of examination procedures and anamnestic data according to performance group

	N	Mdn	M	SD	Mdn	M	SD
		number of examination procedures			number of anamnestic data		
1. performance group	9	15	14.44	5.83	15	12.67	4.42
2. performance group	10	12	10.90	2.56	11	12.10	4.31
3. performance group	9	8	7.89	2.47	11	11.22	1.92

Linear regression analysis validated that, how teams approached this task represented to a certain extent how they worked in the whole competition. The relation between the number of gained anamnestic information in the selected task and the ranking in the whole competition was not significant. But the amount of examinations executed in the selected task explains 26% of variance in overall ranking of the team (adjusted R square; F = 10.49; p = 0.004). After adding the variable "assessment of the approach to the patient from judges" the value of explained variance rose to 36% (adjusted R square; F = 8.46; p= 0.002) which is more than what explains the total score of the teams in the selected task (23%, F = 9.09; p = 0.006).

The correlation of the number of gained anamnestic information and the number of examinations conducted was on the border of statistical significance ($r_s = 0.316$; p = 0.051).

Successful procedure in this phase of dealing with the task was systematic examination from head to toe, with simultaneous taking of the anamnesis, perceiving the relevant information spontaneously obtained from the patient and their verification by examinations. It was also helpful to summarize the gained information after some intervals and their verification, an effort to understand the situation – why the wife called, how the patient reacted and why. An alternative was also a combination of the mostly systematic procedure of anamnesis taking with a parallel examination, while at the same time verifying the preliminary hypotheses, perceiving information from the patient and deriving new plausible hypotheses, however, if the paramedic was still able to return to the systematic procedure.

Diagnosis

Although 23 teams came to the right conclusion that the patient is healthy, this conclusion was not well supported by the conducted examinations. Some teams showed nervousness, uncertainty or helplessness, as the number of examinations with good results increased. Two teams had a tendency to stick to any kind of minor discrepancy in the examination results and use it as an argument for the next examination by the doctor. One team did not arrive to the final conclusion in the limited time. Two teams refused to accept the idea, that the patient was healthy despite all good results and they concluded that the patient has to be intoxicated or a psychiatric case. They presented this as a reason for forced transport. Their way of thinking is illustrated by the statement: "We did not find anything, but there has to be something wrong with you when your wife said that you were weird." Apart from that, there were also incorrect conclusions, e.g., suspicion of ventricular fibrillation or mental disorder.

The task should have ended with the attempt to call the doctor on the spot (doctor could not come) and then use consultation by phone for the purpose of signing the against medical advice form. The correct solution to this task was to conclude that the patient had no medical problems, that he was healthy, just tired. He reacted irritated to his wife's phone calls because she repeatedly woke him up after the night shift and so he switched the phone off.

DISCUSSION

Results both from quantitative as well as qualitative analysis together with other resources support the following findings: Misunderstanding of initial information lead to more frequent conflicts with the patient, which resulted in errors in the further process. Similarly Balla, Heneghan, Goyder and Thompson (2012) claim that the cognitive biases developed at the initial framing of the problem relate to errors at the end of the process. In our case there were false conclusions about the patient's state (e.g., mental health, intoxication). This can be explained by the fixation errors, which occur when the practitioner concentrates solely upon a single aspect of a case to the detriment of other more relevant aspects (Fioratou, Flin & Glavin, 2009). Fixation errors or cognitive lockups have been reported as a unique type of performance failure in dynamic work environments (Xian & MacKenzie, 1995).

Creating an assumption about a forthcoming event can be the natural part of mental preparation for the action (Xiao, Milgram & Doyle, 1997; Dominiguez, 2001). Although this preparation can be useful, it is not immune to mistakes. We suppose that paramedics incline to create a small number of successively arising hypotheses about the condition of the patient. Less successful teams consider their plausibility just on the basis of a minimal number of examinations and sometimes even against a direct evidence of results. Narrow focus of attention to their own goals resulted, in less successful teams, in indifference towards important anamnestic information spontaneously provided by the patient. Keyser and Woods (1990) see the failure to revise situation assessment as new evidence comes in as a major source of human error in dynamic domains. According to Balla et al. (2012) deliberate practice of looking for warning signs may be a potential method of professional development to reduce error through reflection in action.

CONCLUSION

The real process of examining the patient and determining the diagnosis by the paramedics in the selected simulated situation was not in accordance with the prescribed procedures. The tendency of paramedics to create preliminary hypotheses about the state of the patient and situation can be helpful just in the case when the team is able to stay open towards the alternative development of the situation. They can flexibly adapt to new conditions, they are able to maintain a systematic procedure, they do not refuse contradictory evidence, and do not stop to look for information after gathering the minimum evidence of the validity of their preliminary assumptions. In the opposite case the team risks misunderstandings, conflicts and mistakes not just in contact with the patient but also in establishing the diagnosis or deciding about the next process.

The next research should broaden the range of observed real and simulated situations and compare it to the subjective statements of the participants. Our study replicated the results according to which paramedics do not adhere to the prescribed procedures and that they commit mistakes in the diagnosis process. We also brought description of the most and the less often used procedures of investigation of Slovak and Czech paramedics.

REFERENCES

Balla, J., Heneghan, C., Goyder, C., & Thompson, M. (2012). Identifying early warning signs for diagnostic errors in primary care: a qualitative study. *BMJ open, 2*(5), 1-9.

Dekker, S. W. A. (2005). *Ten questions about human error: A new view of human factors and system safety.* Mahwah, NJ: Erlbaum.

Dominiquez, C. O. (2001). Expertise in laparoscopic surgery: anticipation and affordances. In E. Salas, G. Klein (Eds.), *Linking expertise and naturalistic decision making,* Lawrence Erlbaum Associates, Inc., 287-301.

Fioratou, E., Flin, R., & Glavin, R. (2009). No simple fix for fixation errors: cognitive processes and their clinical applications*. Anaesthesia, 65*(1), 61-69.

De Keyser, V., & Woods, D. D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In A. G. Colombo, A. Saiz de Bustamante (Eds.), *Systems reliability assessment*. Netherlands: Springer, 231-251.

Najaf-Zadeh, A., Dubos, F., Pruvost, I., Bons-Letouzey, C., Amalberti, R., & Martinot, A. (2011). Epidemiology and aetiology of paediatric malpractice claims in France. *Archives of disease in childhood, 96*(2), 127-130.

Quesada, J., Kintsch, W., & Gomez, E. (2005). Complex problem-solving: a field search of a definition? *Theoretical Issues in Ergonomics Science*, *6*(1), 5-33.

Rittenberger, J. C., Beck, P. W., & Paris, P. M. (2005). Errors of omission in the treatment of prehospital chest pain patients. *Prehospital Emergency Care, 9*(1), 2-7.

Xiao, Y., & Mackenzie, C. F. (1995, October). Decision making in dynamic environments: Fixation errors and their causes. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 39*(9), SAGE Publications, 469-473.

Xiao, Y., Milgram, P., & Doyle, D. J. (1997). Capturing and modeling planning expertise in anesthesiology: Results of a field study. In C. E. Zsambok, G. Klein (Eds.), *Naturalistic decision making.* Mahwah, NJ: Lawrence Erlbaum, 197–205.

Naturalistic Decision Making on the Ship's Bridge

Catherine HARVEY[a], Pengjun ZHENG[b], and Neville A. STANTON[a]
[a]University of Southampton, UK
[b]Ningbo University, China

ABSTRACT

Introduction: With the maritime industry having been the subject of relatively little Human Factors research, an exploratory study was conducted to investigate decision making by captains on the ship's bridge. **Method**: Ningbo University's ship bridge simulator was used to assess the decision making, communications and training requirements of novice and expert teams. **Results and discussion:** Interview transcripts were mapped onto the Recognition Primed Decision (RPD) model and compared to communication and video data, indicating that novice participants may struggle to provide an accurate account of events. This may be due to the pressure associated with their position and cultural factors relating to perceived loss of face. The potential for a decision-making oriented tool to support training is also discussed.

KEYWORDS

Ship's bridge; Maritime industry; Critical Decision Method; Recognition-Primed Decision Model.

INTRODUCTION

The maritime industry appears to be lagging behind other transport domains in terms of Human Factors research (Hetherington et al., 2006). This is despite the fact that the maritime domain is extremely complex and, according to much of the literature, highly prone to 'human error' (e.g. Chauvin and Lardjane, 2008; Gregory and Shanahan, 2010; Hetherington et al., 2006). Human issues including fatigue, stress, time pressure, communication, environmental factors and long working hours all affect the performance of decision makers on the ship's bridge (Hetherington et al., 2006). These issues contribute to the loss of an average of two ships every day in the global shipping industry, and although mechanical failures are decreasing, the overall number of vessel incidents is not (Gregory and Shanahan, 2010). In order to reduce risk in the maritime industry, the behaviour of human operators must be understood: only then can systems be designed to mitigate error.

This study was designed as a preliminary exploration of behaviour on the ship's bridge. A descriptive approach was taken to examine a range of phenomena, including decision making, communications, training, and the performance of novice and expert teams. In this paper, we describe our initial findings with the aim of highlighting issues which are specific to decision making in the Maritime domain and which can be taken forward for further investigation.

METHOD

Participants

Twelve participants (all male) were split into four groups each consisting of three team members: 1 captain, 1 watchman and 1 helmsman (the roles were allocated to participants randomly). 6 participants (two groups) were students on Ningbo University's Maritime Navigation degree course: these were classified as 'novice' participants. The novice groups had only experienced classroom- and simulator-based training and had no experience on a real ship's bridge. The two novice captains were aged 21 and 22. The other six participants (2 groups) were experienced sailors: these were classified as 'expert' participants. The two expert captains were aged 51 and 36, and had 15 and 12 years' experience on the ship's bridge respectively. All participants were native Chinese speakers but had English as a second language (the tests were carried out in English).

Equipment

The tests were conducted in Ningbo University's ship's bridge simulation facility, shown in Figure 1. The bridge interface is made up of 9 display screens, simulating the complete range of information on a standard ship's bridge, including Electronic Chart Display and Information System (ECDIS) and radar. High resolution projectors offer a 260 degree field of view. Weather, sea, and tide conditions, along with the movements of other vessels are completely controllable. Communications from the Vessel Traffic Service (VTS) and other vessels reach the bridge via a two-way radio and communications from the engine room come via a telephone located on the bridge. Communications from the VTS, Chief Engineer and other vessels were made by a single 'actor'

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

located away from the bridge simulator, in a separate control room: this person was an experienced navigator and trainer who had detailed knowledge of the scenario.

Figure 1. Novice group operating Ningbo University's ship bridge simulator

Procedure

The captains from each team were asked to complete a demographic questionnaire to collect information on age, gender and experience. Each team completed the same scenario, lasting approximately one hour. During each scenario, three incidents were programmed to occur in a fixed order:

1. Failure of the ship's main engine
2. Visibility becomes reduced due to fog
3. Another vessel passes in close proximity to the ship, presenting a potential collision risk

The ship bridge teams were instructed to navigate into Ningbo Port and to interact and communicate with each other in the way that would be expected on a real ship's bridge. They were not made aware that any incidents would be simulated during the scenario. The team also had to communicate with the Ningbo VTS, the ship's chief engineer and captains of other vessels in the port and surrounding area. Each team's performance was recorded (video and audio) during the scenario and all communications were transcribed following the tests. Notes were also made during the scenarios to assist the transcription process and to capture extra detail. Communications worked on an 'announce and address' basis, so that for every communication it was clear who was speaking and who they were speaking to.

Semi-structured interviews were constructed around the CDM cognitive probes and were conducted post-test to elicit information from the captains about their decision making during each incident in the scenario. The interviews were conducted in English; however the CDM probes were converted into Mandarin and asked by an interpreter where necessary. The interviews were audio-recorded and later transcribed.

Data Analysis

The CDM data was compared to the communications transcripts in order to evaluate the validity of the interviews. The CDM data was coded (using QSR nVivo) according to the stages of the Recognition-Primed Decision (RPD) model, in order to investigate the nature of decision strategies between novices and experts.

RESULTS AND DISCUSSION

Comparing the CDM and Communications Data

Comparisons between the interview data and the communications transcripts highlighted discrepancies in response reporting by novices. For example, describing the engine failure incident, one novice captain stated:

'First, I must send information, like err 'pan pan', second I must communicate with the Ningbo VTS, third I must call other vessels, fourth I take some measures like stop engine and check vessels crossing.'

Later in the interview, the captain reports a different sequence of events for the same incident:

'When the accident is dangerous, first I must stop my engine right now, second I will give orders to the helmsman or the watch officer, third I communicate with the other vessels to let them know my vessel is having problems and that I am dangerous, so to keep clear of me.'

The order of actions performed by the novice captain and his colleagues is not clear from the CDM information; it was therefore necessary to examine the videos and communications transcripts for this incident, in order to find out exactly what happened. According to these sources, the captain first sent a radio message to all ships in the vicinity, initially reporting 'I have a problem with engine.., with steering'. He then called the Ningbo VTS and (incorrectly) reported a problem with the ship's steering gear. Later, the VTS had to request position information from the captain, rather than him providing it as a matter of course. This behaviour is contrary to the two descriptions provided retrospectively by the captain, for example, the captain did not mention his confusion over whether the problem related to the engine or steering gear at all in the inteview. This may have been caused by the captain forgetting what he had done and this is a fairly common problem with retrospective techniques, however, the interviews were conducted immmediately following the simulations, so it is unlikely that memory would have been the problem here. An alternative explanation is a lack of confidence on the part of the novice captain, which may have resulted in him responding to the probes in a way that he thought the interviewer expected, rather than providing an honest account of what happened in the scenario.

Populating the RPD

To further explore decision making, the CDM data was coded according to the different stages of the RPD model. The coded extracts were mapped on to the RPD model diagram: this is illustrated in Figure 2, which shows an example for an expert captain dealing with the passing vessel incident. This examplifies the serial descision making strategy identified by Klein et al. (1989), as the captain seemed to quickly recognise that it would be easy for him to maintain course and immediately asks other vessels to allow him to do so. However, when the other vessel refuses to change her course, the captain re-evaluates his decision and changes course. There was significant time pressure involved in the decision as both vessels were moving closer to a collision, so the captain would not have had time to evaluate alternative strategies (i.e. concurrent option evaluation) before deciding on and declaring his course of action. Klein et al. (1989) suggested that by focussing on non-routine events, the CDM can elicit tacit knowledge from operators and the mapping exercise highlighted evidence of this. Figure 2 shows that the captain evaluated the likelihood of success of his decision (see 'Will it work?) based on tacit knowledge about the difficulties in communicating with the captains of smaller ships. Although the CDM – RPD model mapping was very clear for the expert captain illustrated in Figure 2, interpretation of the novice captains' CDM interviews was more difficult. The RPD identifies mental simulation of action as a stage in the decision making process; however, for the novice captains it was impossible to know from the CDM data if the participants were doing this during the incident or whether they were just constructing a retrospective simulation of actions. The CDM probe 'Did you imagine the possible consequences of this action?' is designed to elicit this information; however, in this study the novice captains seemed to interpret this as a requirement to perform some retrospective walk through of the situation rather than describing what they actually did at the time of the incident. Furthermore, with the novice participants there seemed to be a concern that they should answer this question in a certain way, i.e. by providing a number of possible actions; however, it seemed likely that they were creating these alternatives in response to the interview. This problem needs to be considered particularly for novice participants, who are likely to feel under more pressure to give answers that they think the interviewer expects and that reflect positively on their own performance. Furthermore, there may also be cultural factors which influence the information provided by participants, for example the issue of loss of face is particularly significant in Chinese culture and the novice captains may have been affected by the perceived pressure of being 'assessed' by higher level colleagues (Chow et al., 1999).

CONCLUSIONS AND FURTHER WORK

This exploratory study of decision making on the ship's bridge highlighted some problems with using retrospective interviewing as a method for eliciting decision making information, for example, the pressure felt by participants to give a positive, but not necessarily accurate, account of an incident. In this case this behaviour seemed to be influenced by experience level and cultural factors.

Our initial findings suggest that mapping the CDM interview responses onto the RPD model is a useful way of modelling the decision making strategies of novice and expert captains. Part of this exploratory study focussed on the training needs of future captains as the Ningbo simulator is used extensively as a training facility. The RPD model appears to have potential value as a training tool to encourage students to consider how they make critical decisions. For example, students could go through each stage of the RPD for a case study task and, rather than learning rules, this would encourage them to explore their own decision making strategies and to identify the knowledge that they need to acquire in order to make the transition from novice to expert captains. Decision making 'games' have also been suggested as a training tool by Chauvin et al. (2008). This is recommended as an area for future work in the maritime domain.

ACKNOWLEDGMENTS

The authors would like to thank the Faculty of Maritime Engineering, Ningbo University, China for providing the simulation facility, and to the participants for giving their time to take part in the tests. The research was supported in part by the National Natural Science Foundation of China (No. 61074142) and Disciplinary Project of Ningbo under grant B01342104900.

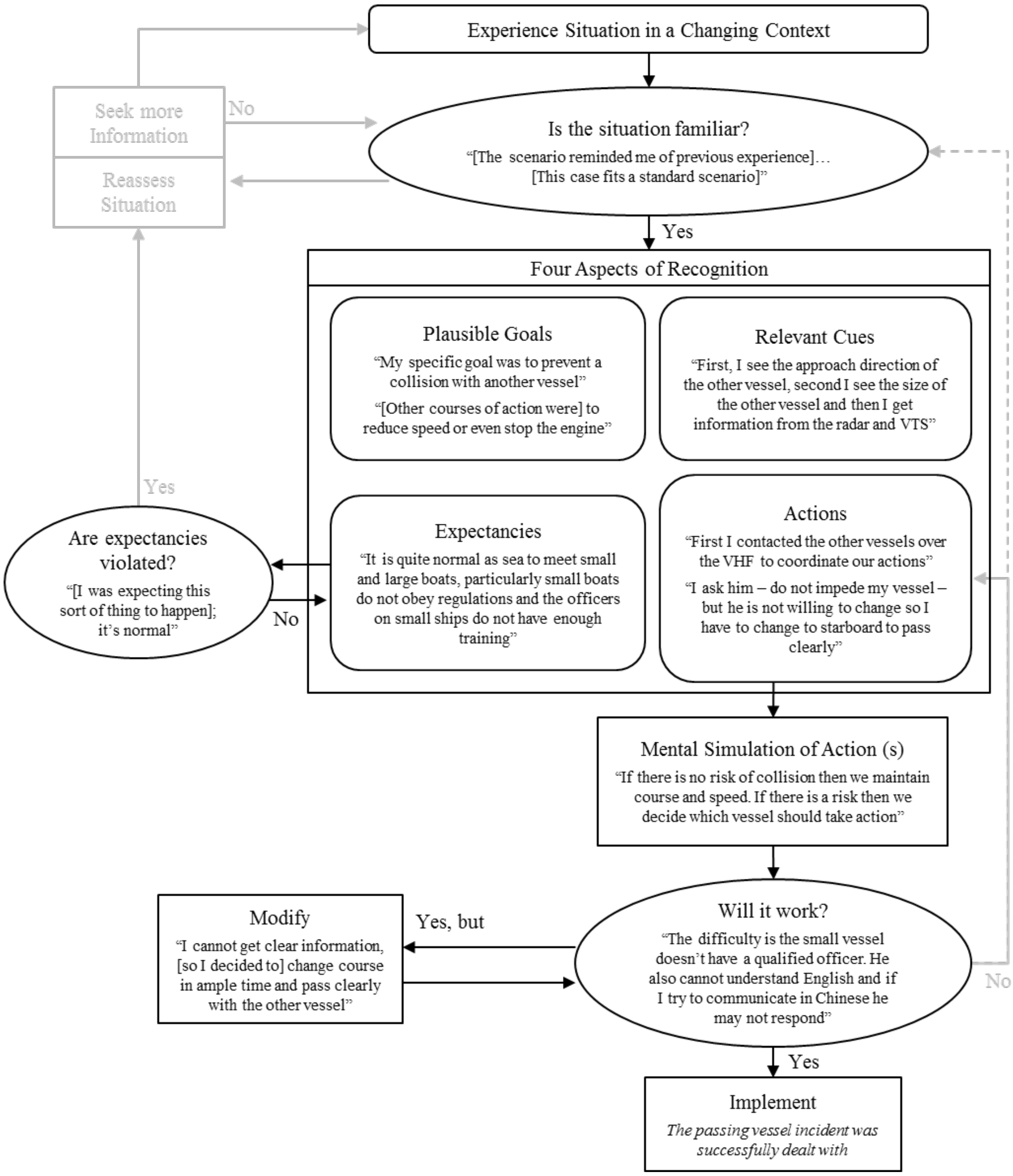

Figure 2. CDM interview extracts mapped onto the RPD model

REFERENCES

Chauvin, C., Closterman, J.P. & Hoc, J-M. (2008) Situation awareness and the decision-making process in a dynamic situation: avoiding collisions at sea, *Journal of Cognitive Engineering and Decision Making*, 2, 1, 1-23.

Chauvin, C. & Lardjane, S. (2008) Decision making strategies in an interaction situation: collision avoidance at sea, *Transportation Research Part F: Traffic Psychology and Behaviour*, 11, 4, 259-269.

Chow, C.W., Harrison, G.L., McKinnon, J.L. & Wu, A. (1999) Cultural effects on informal information sharing in Chinese and Anglo-American organizations: an exploratory study, *Accounting, Organizations and Society,* 24, 7, 561-582.

Gregory, D. & Shanahan, P. (2010) The human element: a guide to human behaviour in the shipping industry, *London, UK Maritime and Coastguard Agency.*

Hetherington, C., Flin, R. & Mearns, K. (2006) Safety in shipping: the human element, *Journal of Safety Research*, 37, 4, 401-411.

Klein, G., Calderwood, R. & MacGregor, R. (1989) Critical decision method for eliciting knowledge, *IEEE Transactions on systems, man, and cybernetics*, 13, 3, 462-472.

The "Cry Wolf" Effect and Weather-Related Decision Making

Susan JOSLYN and Jared LECLERC
University of Washington, Seattle, WA, USA

ABSTRACT

Introduction: Poor compliance with weather warnings might be due partly to residents' prior experience with false alarms, leading them to distrust future warnings. **Method**: In a computer-based task, participants used forecasts to make weather-related decisions. Some participants were given explicit advice from an automated decision aid, comparable to weather warnings. We systematically manipulated the false alarm rate and whether the forecasts included uncertainty estimates. **Results and discussion**: Lower false alarms increased participants' decision quality and compliance, but the effect was small. Adding an uncertainty estimate to the forecast was more beneficial. These results carry important implications for weather risk communication.

KEYWORDS

Decision making; earth and atmospheric sciences; risk communication; decision support aid.

INTRODUCTION

General public end-users have vast experience with weather predictions and their subsequent outcomes: daily weather forecasts are readily available and widely used. Evidence suggests that, as a result, end-users have garnered a fairly sophisticated understanding of the relative accuracy of such forecasts and the uncertainty involved (Joslyn & Savelli, 2010). This is one domain in which most of us have at least some degree of expertise. Nonetheless, many people make what appear to be remarkably poor decisions when it comes to severe weather situations. Despite recent improvements in lead time and weather forecast accuracy, weather-related injury and death remain a serious problem. There is growing consensus that public response to warning forecasts, or lack thereof, is a significant contributing factor (e.g., Nagele & Trainor, 2012; Riad et al., 1999).

There are numerous possible reasons that people ignore weather warnings. Extreme weather events require vulnerable residents to make a choice between two general courses of action: 1) take precautionary action, e.g., evacuation, which is often regarded as costly because of the inconvenience, the risk of looting, etc. (Blendon, 2008; Cutter & Smith, 2009 ; Smith & McCarty, 2009); or 2) take no action and risk a severe loss, e.g., injury or death. In situations such as this, in which both alternatives involve potential losses, research suggests that people tend to be "risk seeking" (Kahneman & Tversky, 1979), preferring to take the risk rather than incur the cost of precautionary action. Furthermore, survey research suggests that people understand that extreme events are rare (Joslyn & Savelli, 2010) so they might underestimate the risk of an actual threat.

In addition, people might distrust weather warnings (Dow & Cutter, 1998; Morss & Hayden, 2010). Prior experience with false alarms, warnings that were perceived as unnecessary, may lead to a reluctance to comply with future alarms, known as the "cry wolf" effect (Breznitz, 1985). Because of the high potential loss that can result from extreme weather events, as well as the need to prepare several days in advance, weather warnings are often given when the probability of the event at any given location is low. Thus, the predominant error for warning forecasts tends to be a false alarm (Joslyn & Savelli, 2010; Barnes et al., 2007). While there has long been concern about the impact of false alarms on trust in weather warnings, the psychological effects remain unclear (Barnes et al., 2007). Among survey studies that investigate the impact of false alarms in natural settings, there is some evidence suggesting that people are fairly tolerant of such errors (e.g., Baker, 1991), while other evidence suggests classic false alarm effects (e.g., Atwood & Major, 1998). However, in survey studies, there are multiple uncontrolled variables, not the least of which is respondents' prior exposure to false alarms. There is some evidence that the cry wolf effect may not be apparent after a single false alarm but arises only after several false alarm experiences (Carsell, 2001). For that reason, an experimental approach may be required to obtain direct evidence for false alarm effects.

The experiment reported here is the first direct experimental evidence of which we are aware for a cry wolf effect in weather-related decision making. It was obtained by systematically manipulating the false alarm rate in a winter weather task to determine the impact on trust and compliance. The practical motivation for this research was to reproduce the effect in the laboratory and then explore factors to attenuate it. For instance, there may be a rate of false alarms that is tolerable to users and does not reduce trust or compliance (Roulston & Smith, 2004).

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

On the other hand, there may be ways to communicate warnings to reduce negative effects of false alarms and increase compliance so that warning thresholds need not be altered. Here we tested warnings that included probabilistic uncertainty estimates. Warnings that acknowledge the uncertainty up front may be judged less wrong when they fail to verify. Indeed, there is a growing literature that suggests that people have more trust in forecasts and make better decisions when uncertainty estimates are included (Joslyn et al., 2007; Nadav-Greenberg & Joslyn, 2009; Roulston et al., 2006). Thus, the research reported here also sought to explore whether uncertainty forecasts attenuated decrease in decision quality due to false alarms.

METHOD

Participants performed a computer-based task in which they made a series of decisions about applying salt brine treatment to the roads to prevent icy driving conditions the following morning. Participants were told that treatment should be applied beforehand if sub-freezing temperatures were anticipated that night. On each of 60 trials, simulating two winter months, participants chose between spending $1,000 to apply the salt treatment or risking a $5,000 penalty if they chose to withhold treatment and a freezing temperature was observed. They were given a monthly budget of $35,000 and paid cash commensurate with ending balance. Thus, the goal of the task was to maximize the budget. The decisions, which were a simplified version of the real-world task, were based on overnight low temperature forecasts and the advice of a computerized "decision support aid." After each trial, participants were informed of the observed temperature and any budget adjustments. The 60 forecasts and observed temperatures included a realistic range of temperatures for Washington State, where the experiment was conducted, as well as reliable probabilities where applicable.

The economically rational strategy in this situation (Murphy, 1977) was to apply salt treatment whenever the probability of freezing was 20% or higher. That is the point at which the expected value of the penalty (Bernoulli, 1954), obtained by weighting the $5,000 penalty by the probability that it will be incurred, and the cost of salting are equivalent: $5,000 (possible penalty) * 20% (probability of freezing) = $1,000 (cost).

On each trial, some participants had the advice of the "decision support aid" (DSA), which was described as taking into account the forecasted temperature and uncertainty, as well as the cost of salting and the penalty for not salting. In fact, it employed a rule based on the probability of freezing. In one condition, it was the economically rational rule, recommending salting whenever the probability of freezing was 20% or higher. Thus, following the advice of the decision support aid in this condition led to optimal performance, although, as with warnings for actual severe weather, it exposed participants to many false alarms, i.e., trials on which salting was recommended but freezing temperatures were not observed. Our own past research using this experimental paradigm and threshold suggested that people become reluctant to follow the advice after a few trials (Joslyn & LeClerc, 2012). Roulston & Smith (2004) suggested that raising the threshold to a level greater than the economically rational threshold might increase user compliance by lowering the false alarm rate (which, however, also increases the miss rate). Building upon this idea, we systematically manipulated the false alarm rate by creating three conditions in which the thresholds for recommending salt were 10%, 20% and 30% chance of freezing. The lower the threshold, the greater the proportion of false alarms and the smaller the proportion of misses. It is important to note, however, that the forecasted temperature, observed temperature and the economically rational threshold for salting (20% chance of freezing) were identical in all three conditions. The only thing that changed was the probability of freezing at which the decision support aid advised applying treatment.

In addition, in order to determine whether adding an uncertainty estimate to the forecast attenuates the cry wolf effect, we manipulated forecast format. One group of participants received the nighttime low temperature forecast and decision advice. Another group received the nighttime low temperature forecast, decision advice, and the probability of freezing. A final group received only the nighttime low temperature forecast and served as an overall control condition. Thus, there were 7 between-groups conditions.

Although this experimental task differs in many respects from real-world decisions, it shares several critical features with weather warning situations. Participants make decisions with outcomes that affect them directly, represented in the task as cash rewards. The decision task involves a choice between expending resources for protection and risking a potentially costly loss, as is the case when preparing for real-world weather events. Moreover, because of the high cost-loss ratio, precautionary action is required at a fairly low threshold, also true of many warning forecasts. Finally, it is a task in which a common error is risk seeking, mirroring the non-compliance with weather warnings.

RESULTS

We explored two dependent measures: mean expected value of participant decisions and compliance with the advice. Expected value was assessed by assigning each decision a theoretical value. A decision to salt was assigned the cost of salting, -$1,000. A decision against salting was assigned the penalty (-$5,000) multiplied by the probability of freezing on that trial (the likelihood that penalty would be incurred). A mean expected decision value was calculated for each participant. Compliance was summarized as a ratio of the number of times that the participant followed the advice provided and the control baseline rate (the number of advice-compatible decisions when no advice was provided), with higher values indicating greater compliance.

The false alarm rate had a significant effect on both expected value and compliance, suggesting a cry wolf effect. Participants experiencing the highest false alarm rate (DSA 10% condition) had significantly lower mean expected value than those with the lowest false alarm rate (DSA 30% condition) $t(64.66) = 2.56$, $p = .01$ (see table below). Moreover, compliance with the advice decreased as false alarms increased. Participants in the DSA 10% condition were significantly less compliant than participants in the DSA 30% condition, $t(69) = 2.25$, $p = .03$. However, performance and compliance were similar between participants in the DSA 20% and DSA 30% conditions, suggesting that lowering the false alarm rate slightly from the optimal rate is not effective.

Experimental condition	Mean expected decision value	Compliance ratio
DSA 10%	-$1,078.12	1.01
DSA 20%	-$1,049.55	1.10
DSA 20% + Freeze Probability	-$988.04	1.20
DSA 30%	-$1,033.36	1.12

In fact, the most beneficial manipulation was adding an uncertainty estimate to the forecast, which dramatically increased both compliance and mean expected decision value. Participants with forecasts that included both decision advice and the freeze probability had significantly higher mean expected decision values than participants with the advice alone in both the DSA 20% condition, $t(55.13) = 4.02$, $p < .001$, and the DSA 30% condition, $t(71) = 3.52$, $p = .001$. The same pattern of results was found for compliance.

CONCLUSION

This experiment provides evidence for both the cry wolf effect and the benefit of uncertainty forecasts. First, the decision advice with the highest false alarm rate led to worse decisions and lower compliance than did the advice with the lowest false alarm rate. However, a slight reduction of false alarms from the economically optimal threshold (DSA 20% condition) to the 30% threshold did not have a significant effect. Additional research is necessary to determine whether an even greater reduction in false alarms would be effective or whether the corresponding increase in miss rate would counteract any benefits.

Second, and even more importantly, advice that included probabilistic forecasts led to better decisions and higher compliance than advice alone. As noted in earlier studies (e.g., Joslyn & LeClerc, 2012), inclusion of uncertainty estimates might make forecasts seem more plausible, compensating any negative impact of false alarms, and allow individuals to tailor their decisions to their own risk tolerances.

Thus, despite the longstanding concern about cry wolf effects in warning forecasts, reducing false alarms to any practical level may not be effective in increasing compliance. This evidence suggests that a much more effective solution is to leave the false alarm rate where it is and provide users with an explicit uncertainty estimate.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under Grant No. 1023354.

REFERENCES

Atwood, L. E., & Major , A. M. (1988). Exploring the "cry wolf" hypothesis. *International Journal of Mass Emergencies and Disasters,* 16, 279–302.

Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters,* 9, 287–310.

Barnes, L. R., Gruntfest, E. C., Hayden, M. H., Schultz, D. M., & Benight, C. (2007). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting,* 22, 1140 –1147.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22, 23-36.

Blendon, R. J. (2008). High-risk area hurricane survey. [Available at www.hsph.harvard.edu/hurricane/topline.doc.]

Breznitz, S. (1985). False alarms: Their effects on fear and adjustment. *Issues in Mental Health Nursing,* 7, 335-348.

Carsell, K. (2001). Impacts of a false alarm : The January 29, 2000 Ventura, California experience. M.S. thesis. Dept. Of Geography and Environmental Studies, University of Colorado at Colorado Springs, 86 pp. [Available from Dept. Of Geography and Environmental Studies, University of Colorado at Colorado Springs, P.O. Box 7150, Colorado Springs, CO 80933.]

Cutter, S. L., & Smith, M. M. (2009). Fleeing from the hurricane's wrath: Evacuation and the two Americas. *Environment,* 51, 26 -36.

Dow, K., & Cutter, S. L. (1998). Cry wolf: Repeat response to hurricane evacuation orders. *Coastal Management,* 26, 237–252.

Joslyn, S., & LeClerc, J. (2012). Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied,* 18, 126-140.

Joslyn, S., Pak, K., Jones, D., Pyles, J., & Hunt, E. (2007). The Effect of probabilistic information on threshold forecasts. *Weather and Forecasting,* 22, 804–812.

Joslyn, S., & Savelli, S. (2010). Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteorological Applications,* 17, 180–195.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica,* 47, 263-292.

Morss, R. E., & Hayden, M. H. (2010). Storm surge and "certain death": Interviews with Texas coastal residents following Hurricane Ike. *Weather, Climate, & Society*, 2, 174-189.

Murphy, A. H. (1977). The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Monthly Weather Review*, 105 (7), 803-816.

Nadav-Greenberg, L., & Joslyn, S. (2009). Uncertainty forecasts improve decision-making among non-experts. *Journal of Cognitive Decision Making,* 2, 24-47.

Nagele, D. E., & Trainor, J. E. (2012). Geographic specificity, tornadoes, and protective action. *Weather, Climate, & Society*, 4, 145–155.

Riad, J. K., Norris, F. H., & Ruback, R. B. (1999). Predicting evacuation in two major disasters: Risk perception, social influence, and access to resources. *Journal of Applied Social Psychology,* 29, 918-934.

Roulston, M. S., Bolton, G. E., Kleit, A. N., & Sears-Collins, A. L. (2006). A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting,* 21, 116-122.

Roulston, M. S., & Smith, L. A. (2004). The boy who cried wolf revisited: The impact of false alarm intolerance on cost–loss scenarios. *Weather and Forecasting,* 19, 391–397.

Smith, S. K., & McCarty, C. (2009). Fleeing the storm(s): An examination of evacuation behavior during Florida's 2004 hurricane season. *Demography,* 46, 127-145.

Evaluative Feedback Spaces for Cultural Mental Models

Gary L. KLEIN[a], Craig A. BONACETO[a], Alison A. DINGWALL[a], Jill L. DRURY[a], and Mark S. PFAFF[b]
[a]*The MITRE Corporation*
[b]*Indiana University Indianapolis*

ABSTRACT
Introduction: Feedback from either social context or personalized information can result in more effective changes in mental models and consequent behavior changes. Techniques for collecting mental models on a group or societal scale could improve communications for social change or collaborative decision making. **Method**: A social media application is being developed that will allow users to personalize a cultural model to closely match their network of beliefs regarding some event, action, or entity. A graphical *evaluative feedback space* will enable real-time views of how their model compares with socially significant groups, or exchange information that is tailored to others' mental models.

KEYWORDS
Practical Application; Macrocognition; Mental Models; Cognitive Social Psychology; Cultural Network Analysis.

INTRODUCTION

Cultural models are shared prototypical concepts drawn from distributed members of a population known as a cultural group (Sieck, Rasmussen and Smart, 2010). Sieck and colleagues developed a synthesis of multi-disciplinary techniques called Cultural Network Analysis (CNA) to model culture as networks of ideas. As formal descriptions of the knowledge possessed by members of particular groups, cultural models describe and represent how the world is understood by the members of these groups. A key premise is that cultural knowledge comprises many networks of causally-interconnected ideas. These mental models become activated within particular situations to drive thinking and decision making, and can change under suitable conditions. Cultural models could therefore be especially useful for sharing understanding among collaborative decision makers, as well as anticipating effects in communications campaigns (Sieck et al., 2010).

Amjad and Wood (2009) present a compelling example of how even weakly held beliefs can lead people down a path to extremist behavior. They found that among their college student participants, beliefs held about the acceptability of aggression toward Jews strongly predicted whether participants would be willing to join an extremist anti-Semitic organization. However, even a brief educational intervention of attending a short lecture on the positive historic relationship between Jews and Muslims caused significant changes in beliefs and a dramatic 94% reduction in the number willing to join an extremist group.

By extending the CNA method, we plan to show that, beyond the canonical cultural models, personalized cultural models of individuals can be collected that reflect an individual's own network of beliefs regarding some event, action, or entity. In this way, the responses of individuals can be viewed as specific, varied instantiations of the canonical model. As such, this extension results in a type of crowd-sourced exploratory modeling (Bankes, 1992) of the canonical model. Therefore, the results can be viewed in much the same way as the frequentist formatted decision spaces demonstrated by G. L. Klein[3] et al. (2011) for comparing the distributions of outcomes generated by exploratory computer modeling. In addition, these individual models can be mined to identify relationships between people's beliefs and the outcomes they expect from actions or events. Such a rich set of data at the individual level should enable better understanding among collaborative decision makers, and enable effective targeted communications aimed at specific individual variations of the canonical cultural model. Collecting mental models through a widespread social media venue, and visualizing this data to enhance collaboration and targeted communication, constitute a novel synthesis of tested sociocultural research processes and new technological capabilities.

[3] Both Gary A. Klein and Gary L. Klein are working in the decision making area. We distinguish the work of each by using their middle initials in citations.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

METHOD

As described above, we will implement a two-phase process to extend the CNA method. As a case study, we will assess the mental models of employees regarding a new corporate policy.

The first phase is a customization of the CNA methods, as described by Sieck et al (2010). In this phase we will determine existing cultural models by interviewing a representative cohort of individuals to obtain and validate a set of cultural models regarding their expectations of a new corporate policy. Guidance for these interviews will be based on a set of employee issues that the corporate policy makers want to better understand. Interviewers will then elicit "critical incidents" from each individual in the cohort in terms of issue-related tasks or work conditions that the interviewees believe will be affected by the policy. Based on these incidents, interviewers can engage in more abstract discussions with each individual regarding the causal chain the interviewee believes will lead to those effects. Coding of the interviews will be done by two independent coders and then validated by the analysis team. From these coded interviews, the analysis team will identify the cultural models shared by cultural groups within the cohort. We plan to represent these cultural models as directed acyclic graphs (DAGs), although the exact form will be dependent upon the actual models derived from the interviews.

In the second phase, these cultural models can then be used as a basis for a computer-automated survey of a broader sampling of the corporate population. Using visualizations of these cultural models, such as a DAG, individuals can select a cultural model that comes closest to matching their network of beliefs, similar to the Nearest Neighbor technique used by G. A. Klein and Militello (2001). For their selected model, individuals can assign personal values to any or all of the edges connecting the vertices in the DAG representation of the cultural model, as illustrated in Figure 1. The sign of the numbers indicates whether the association between the vertices is positive or negative, and the value indicates the strength of the association. The alphabetic labels of the vertices are just arbitrary substitutions for what will be in actual use labels for real-world states or actions.

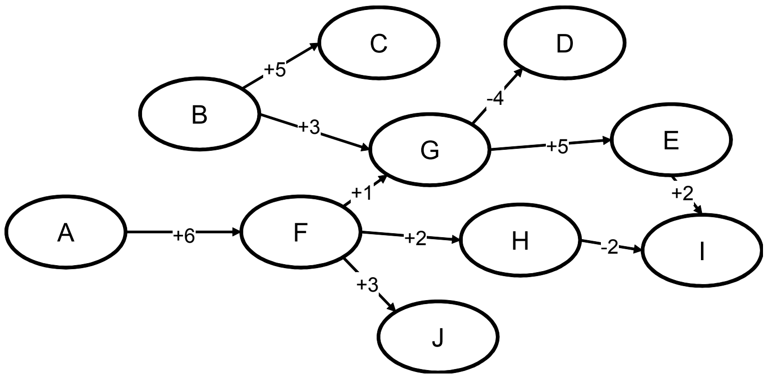

Figure 4. Example individual evaluated DAG representation of a generic cultural model

When the cultural models are so evaluated by the population of individuals, a number of analyses and functions become possible. For instance, the distance between different individuals can be calculated with regard to the values they assigned to any subset of *n* edges. Based upon this calculation, an evaluative feedback space can be generated that allows an individual to see their relationship to the members of significant social groups who have also evaluated the edges of the model, such as the different employee seniority levels (e.g., 3, 4 or 5) in the evaluative feedback space illustrated in Figure 2. Notice that on the evaluative measure, models of members of XYZ team differ substantially across seniority levels. Within seniority levels 3 and 5 there are even greater differences among the XYZ team members.

The evaluative feedback system will be able to maintain permanent anonymous links between individuals and their models. These links will allow corporate communications to send messages directly to individuals while maintaining anonymity. For example, anonymous individuals can be selected by cultural model, any combination of edge evaluations, or social group (e.g. team or seniority) to receive specific messages crafted to address their beliefs. In this way, messages can be tailored to address individual beliefs without identifying specific individuals. In addition, it is possible to perform a controlled evaluation of the effectiveness of different messages. Variations on messages can even be sent to subgroups to compare their effects on beliefs expressed in the models.

DISCUSSION

This poster will describe a prototype of the evaluative feedback space system and discuss results from its initial use in support of a corporate policy initiative.

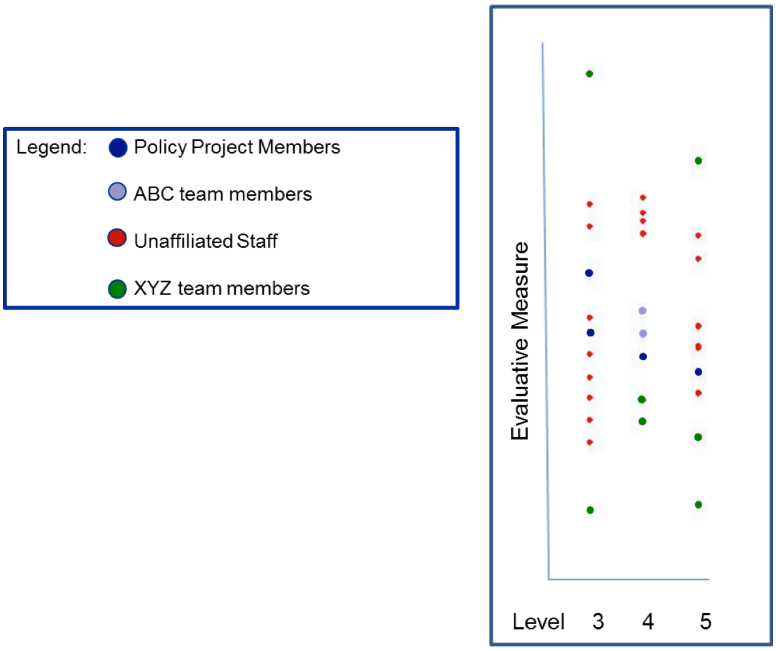

Figure 5. An evaluative feedback space showing individual mental model evaluations by project group and seniority level

ACKNOWLEDGEMENTS

This work was funded by MITRE Corporation projects 51MSR603-AA and 51MSR605-BA.

REFERENCES

Amjad, N. & Wood, A.M. (2009). Identifying and changing the normative beliefs about aggression which lead young Muslim adults to join extremist anti-Semitic groups in Pakistan. *Aggressive Behavior*, 35, 514–519.

Bankes, S. C. (1992). Exploratory modeling and the use of simulation for policy analysis, Santa Monica, CA: The Rand Corporation.

Klein, G. A. & Militello, L. G. (2001). Some guidelines for conducting a cognitive task analysis. In E. Sales (ed.), *Human/Technology Interaction in Complex Systems, Vol. 1.* Oxford: Elsevier Science Ltd (pp. 161-97).

Klein, G. L., Drury, J. L., & Pfaff, M. S. (2011). Providing an option awareness basis for naturalistic decision making. *Journal of Cognitive Technology*, Vol 16(2), pp. 10 - 19.

Sieck, W. R., Rasmussen, L. J., & Smart, P. R. (2010). Cultural Network Analysis: A Cognitive Approach to Cultural Modeling. In D. Verma (Ed.), Network Science for Military Coalition Operations: Information Extraction and Interaction. Hershey, Pennsylvania, USA.: IGI Global.

Intraoperative Surgical Decision Making – A Video Study

Lucy MITCHELL[a], Rhona FLIN[a], George YOUNGSON[a], Momin MALIK[b], Irfan AHMED[b]
[a] *School of Psychology, University of Aberdeen, UK*
[b] *Aberdeen Royal Infirmary, UK*

ABSTRACT

Aim: This study investigated the differences between trainee and consultant general surgeons in their intraoperative risk assessment, risk tolerance and decision making. **Method:** A sample of 13 consultant general surgeons and 14 trainee general surgeons from 3 Scottish hospitals were interviewed while observing three videotapes of gall bladder surgery (laparoscopic cholecystectomy). Decision making and risk assessment were examined by requesting them to verbalise their cognitive processes at two decision points when the videotape was paused. They also completed a short risk tolerance and skills judgement questionnaire.

Results and Discussion: There were no significant differences between consultant surgeon and trainee risk ratings (1-7) at the first or second decision points in any of the 3 videos. Consultants' and trainees' risk-taking preference scores were not significantly different. The interview results are currently being analysed and will be presented in the poster at the meeting.

KEYWORDS

Decision making;surgeons; expertise; risk judgements; intra-operative

INTRODUCTION

Surgeons' cognitive processes, especially situation awareness and judgmental ability are key non-technical skills (Yule et al, 2008) contributing to effective performance. Most decision making occurs when the surgeon diagnoses the patient's condition and works out a plan for the operation, and so the surgical literature mainly focuses on pre-operative decisions. However, intraoperative decision making can be important during routine operations where the operative site is not as anticipated (e.g. anatomic variations) or the procedure does not go to plan. Moreover, during emergency surgery, there can be less time for preoperative planning and proportionately more decision-making may be required in the intraoperative phase.

The few studies on surgeons' intraoperative decision making have examined differences related to expertise, cognitive style or personality. A comparison of conversion preferences during observations of videorecorded laparoscopic cholecystectomy, found that 50% of residents and 40% of staff surgeons would convert the operation at some point, while other surgeons preferred to continue with the laparoscopic approach (Dominguez, 2004). This raised several questions around cue recognition and differences in risk tolerance. Intraoperative risk management requiring consideration of risk levels and personal competence, as well as the surgeon's risk tolerance may also be important factors during conversion decision making (Pauley et al, 2011). Injury to the extrahepatic bile duct occurs at a rate of 0.1-0.5% seemingly regardless of seniority or experience (Way et al, 2003). Bile duct misidentification is reported as being the causal factor in 86% bile duct injury and since no surgeon would divide a structure without prior identification of the anatomy nor knowingly transect the common bile duct, this suggests that inaccurate visual perception is the major cause of such injuries (Dekker & Hugh, 2008). Visual misperception (i.e. seeing what you believe), cue ambiguity and underestimation of risk were important factors in misidentification in previously reviewed cases (Massarweh et al, 2009). More recent research has suggested that surgeons develop an intuitive sense of when they need to 'slow down' in order to make the right decisions intraoperatively (Moulton et al, 2010), and that experienced surgeons will identify critical information at an earlier stage of diagnosis than junior colleagues (Abernathy & Hamm, 1995). Experts develop the ability to make decisions rapidly and intuitively, often without effortful processing (Klein, 1993) and on that basis, it was anticipated that trainee surgeons' comments on their decisions might indicate a more analytical style, explicitly comparing optional courses of action, while consultants' comments might indicate a more intuitive style, stating identification of a single preferred course of action. Although Pauley et al's (2011; under review) interview studies with general and ophthalmic surgeons show the use of both methods and surgeons' reliance on intuitive or analytical decision styles probably depends on both personal preferences and situational constraints.

To test intra-operative surgical decision making, video vignettes of common operative procedures were shown to students, novice and proficient urologists, and differences were found in their knowledge and surgical judgement, through analysis of their 'think aloud' cognitive processes (Chatterjee et al, 2009). The current study

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

investigated the differences between trainee and consultant general surgeons in their intraoperative risk assessment, risk tolerance and decision making styles using a similar method using video tapes of a common procedure. Decision making and risk assessment were examined by requesting surgeons to verbalise their cognitive processes at decision points whilst watching videos of laparoscopic cholecystectomy operations.

METHOD

Ethical approval was granted for the study from the [University of Aberdeen, School of Psychology and the North of Scotland NHS Research Ethics Committees (refs: 2/032/10; REC 10/S0802/83)]. Surgeons were allocated a participant number which enabled their ratings and questionnaire data to remain anonymous.

Sample Surgeons (n=27, 6 female)) were individually interviewed in response to an invitational email. There were 13 consultant general surgeons (n=13; age *M*=49years; experience *M*=13years, SD 8.46, range 2-28 years) and 14 trainee surgeons; (age *M*=35years, experience *M*=6 years, SD=1.86, range 3-9 years) from three Scottish teaching hospitals.

Procedure Recordings of three different laparoscopic cholecystectomy operations were shown in a randomised order, each displaying differing degrees of risk and complication. The video sequences were sourced through commercial contacts and pre-existed this study.
Each video was paused at two judgement points (selected by authors IA and MM) and the same set of questions posed to each participant. The questions were as follows:

i) What do you think is happening here?
ii) What would you do at this point and why?
iii) What cues or information are you attending to at this point?
iv) Do you have any concerns here? If so, what are they?
v) What do you think could happen next?
vi) Are you reminded of any previous experiences?

This was to investigate the information/ cues being utilised, the assessment of the situation and possible actions, e.g. what options were available to proceed. Each surgeon also rated the degree of risk at the two decision points (1=low risk to 7=high risk).
Following the interview, a six item questionnaire on risk taking preference was completed, adapted from the Jackson Personality Index (Jackson, 1994 ; Pearson et al,1995). Risk-taking preference was scored on a 5-point ordinal scale with the highest point value in questions one, three and five, being assigned to the most affirmative responses ("Not at all like me" = 1 point, "very much like me" = 5 points). In the other three questions, scoring was reversed. Surgeons were also asked to report how many times in the previous 24 months they had made a decision which could have resulted in an adverse event. Finally, two questions adapted from an aviation study (O'Hare, 1990) required surgeons to rate (1-7) their own skills and judgement compared with their peers.

Analysis The risk ratings were analysed using PASW Statistics Version 18. (2011). Responses to questions at the pause points were transcribed, de-identified and are being analysed using NVivo 8 software (2008). These are being coded by categorising responses with emerging themes and cues being mentioned in response to each of the questions

RESULTS

There were no significant differences between consultant surgeon and trainee risk ratings (1-7) at the first or second decision points in any of the videos. A series of t-tests comparing surgeons' (n=27) risk ratings at each decision point showed that in video **A**, surgeons rated the risk in the first decision point significantly higher than at the second decision point; $t(26) = 4.22$; $p<.0$. There was no significant difference between the decision points in video **B**; $t(26) = -.75$, ns and in video **C**, surgeons rated the risk higher in the second decision point than in the first; $t(25) = -5.13$, $p<.01$.
Consultants' and trainees' mean risk-taking preference scores on the questionnaire were not significantly different (*M*=14.1; SD=3.7; *M*=16.9; SD=5.6, respectively).
Preliminary results from the analysis of the surgeons' responses at the decision points will be given in the poster. A sample of analysis from video B is given below (C = consultant surgeon, T = trainee surgeon):

Video **B**

At the first decision point, more consultants expressed criticism of the techniques being used than trainees (C=7; T=2). More trainees' statements suggested they were unclear of the anatomy at this point than consultants' (T=8; C=0). More consultants specified the structures they were identifying than trainees (C=8; T=4) and converting to an open procedure was considered as an option by a minority (C=3; T=3) at this decision point. At the second decision point, converting to an open procedure was an option expressed by four consultants; no trainees mentioned conversion at this point but three stated that they might ask for assistance, and one consultant mentioned asking for a second opinion.

DISCUSSION

Consultants' and trainees' risk-taking scores were not significantly different and preliminary analysis of the interview responses does not indicate major differences. The trainee surgeons in the study were almost all at an advanced stage of their training and in many cases are showing similar responses to the consultants and so may have been too experienced to reveal significant expert novice differences for this type of procedure which is frequently carried out by general surgeons. As the sample sizes of the two groups are not large, then one option for anlaysis will be to combine the groups and to conduct the remaining analyses without the expertise variable, focussing instead on individual differences in approach. As in our previous study (Pauley et al , 2011), there are appear to be a wide range of responses across the 26 surgeons to a given situation on the video taped procedure, as individual sugeons tend to prefer different techniques and technical solutions for the same procedure. The intention is to explore these differences in relation to the risk and skill judgement ratings.

ACKNOWLEDGEMENTS

We are grateful to all the consultant and trainee surgeons who took part in the study. Thanks also to Dr Keryn Pauley for her contribution to the design and to Mr Simon Paterson-Brown for advice. Funding was from Scottish Funding Council, Strategic Research Development Grant to Scottish Patient Safety Research Network.

REFERENCES

Abernathy C. & Hamm R. (1995) *Surgical Intuition. What it is and how to get it*. Philadelphia: Hanley & Belfus.

Chatterjee S., Ng J., Kwan K. & Matsumoto E. (2009) Assessing the surgical decision making abilities of novice and proficient urologists. *The Journal of Urology*, 181, 2251-2256.

Dekker S. & Hugh T. (2008) Laparoscopic bile duct injury: understanding the psychology and heuristics of the error. *Australian J. of Surgery*, *78*, 1109-1114.

Dominguez C. (2001) Expertise in laparoscopic surgery: Anticipation and affordances. In: *Linking Expertise and Naturalistic Decision Making,* Salas E., Klein G.(eds). Mahwah, NJ: LEA, 287-301.

Dominguez C., Flach J., McDermott P., McKellar D. & Dunn M. (2004) The conversion decision in laparoscopic surgery: Knowing your limits and limiting your risks. In: *Psychological Investigations of Competence in Decision Making,* Smith K., Shanteau J., Johnson P. (eds). Cambridge University Press: Cambridge, 7-39.

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

Flin R., Yule S., Paterson-Brown S., Maran N., Rowley D. & Youngson G. (2007) Teaching surgeons about non-technical skills. *The Surgeon,* 5, 2, 107-110.

Jackson D. (1994) *Jackson Personality Inventory-Revised.* Sigma Assessments Systems, Inc: Port Huron, MI.

Klein G. (1993) A recognition-primed decision (RPD) model of rapid decision making. In: *Decision making in action: models and methods*, G. Klein, J. Orasanu, J. O. R. Calderwood, C. Zsambok (eds). Ablex: Norwood, New Jersey, 138-147.

Massarweh N., Devlin A., Gaston Symons R., Broeckel Elrod J. & Flum D. (2009) Risk tolerance and bile duct injury: surgeon characteristics, risk taking preference, and common bile duct injuries. *American Journal of Surgery*, 209, 17-24.

Memon M., Muhammad G., Arshad S., Jat M., Bozdar A. & Shah S. (2011) Study of open conversion in laparoscopic cholecystectomy. *Gomal Journal of Medical Sciences*, 9,1, 51-54.

Moulton C., Regehr G., Lingard L., Merritt C. & Macrae H. (2010) 'Slowing down when you should': Initiators and influences of the transition from the routine to the effortful. *J. Gastrointestinal Surgery*, 14, 1019-1026.

O'Hare D. (1990) Pilots' perception of risks and hazards in general aviation. *Aviation Space and Environmental Medicine*, 61, 599-603.

Pauley K., Flin R. & Azuara-Blanco, A. (under review) Ophthalmic surgeons' intraoperative decision making.

Pauley K., Flin R., Yule S. & Youngson G. (2011) Surgeons' intraoperative decision making and risk management. *American Journal of Surgery*, 202, 375-381.

Pearson S, Goldman L., Orav E, Guadagnoli E., Garcia T., Johnson P. & Lee T. (1995) Triage decisions for emergency department patients with chest pain: do physicians' risk attitudes make the difference? *Journal of General Internal Medicine*, 10, 10, 557-564.

Spencer F. (1978) Teaching and measuring surgical techniques: the technical evaluation of competence. *Bull. American College of Surgeons*, 63, 9-12.

Way L., Stewart L., Gantert W., Liu K., Lee C., Whang K. & Hunter J. (2003) Causes and prevention of laparoscopic bile duct injuries. *Annals of Surgery*, 237, 460-469.

Yule S., Flin R., Paterson-Brown S., Maran N., Rowley D. & Youngson G. (2008) Surgeons' non-technical skills in the operating room: Reliability testing of the NOTSS behaviour rating system. *World Journal of Surgery*, 32, 548-556.

Decision making strategies used by experts and the potential for training intuitive skills: A preliminary study

J. O. OKOLI[a], G. WELLER[b], J. WATT[b], B.L.W. WONG[b].
[a]*School of Science and Technology, Middlesex University. JO547@live.mdx.ac.uk*
[b]*School of Science & Technology, Middlesex University. G.Weller@mdx.ac.uk; J.Watt@mdx.ac.uk W.Wong@mdx.ac.uk*

ABSTRACT

Experts do better than novices- not only because they possess more perceptual and cognitive skills but also because they are able to organize and apply those skills better than their counterparts. However, experts find it difficult to express what they know. This is mainly because their knowledge is highly resistant to surface articulation, even by the experts themselves. Thus, it is evident experts need help telling what they know and do. In this study, expert firefighters were interviewed using the critical decision method (CDM) procedure. Results from the investigation revealed certain tacit (procedural) knowledge and cognitive skills they used in performing complex tasks in time pressured environments. It is hoped that systematically organizing this knowledge and skill into a framework will enhance the design of a well informed instructional curriculum for training less experienced officers.

KEYWORDS

Decision Making; education and training; expertise; critical decision method; tacit knowledge; intuitive skills.

INTRODUCTION

During any crisis, civilians whose lives and properties are at stake usually expect a lot from incident command teams. Hence, responding to more dangerous and un-predictable crises will definitely call for the skills of more experienced personnel (Zsambok, 1997). Managing real-world crises poses numerous challenges to professionals because they operate under time pressure, uncertainty, dynamic and changing conditions, ill-defined goals, high stakes etc. (Lipshitz, Klein, Orasanu & Salas, 2001). The goodnews however, is that experienced decision makers still carry on despite these challenges, and they perform reasonably (and sometimes exceptionally) well under these conditions (Orasanu & Connolly, 1993).

Nevertheless, there is compelling evidence to suggest that experts are not fully aware of about 70% of their own decisions and mental analysis of tasks (Clark, Feldon, van Merrienboer, Yates & Early, 2006). In other words, they sometimes find it challenging to fully explain how they arrive at their judgement especially when such information is required to support the design of training, assessment or decision aids (Hannabuss, 2000; Dane & Pratt, 2009).

The objectives of this research therefore are: (1) to investigate how real experts make critical decisions in performing complex tasks (2) to develop a well structured instructional curriculum (educational framework) using the information generated from objective 1 above. This framework will facilitate the learning process and the transfer of relevant knowledge, skills and competence (KSC) to novices.

This current paper presents a preliminary result on some key insights generated from the pilot study, which is part of the ongoing research process. This result will subsequently inform part of the design of a training curriculum using the four component instructional design (4C/ID) framework.

EXPERTS, DECISION MAKING AND THE NATURALISTIC ENVIRONNMENT

Shanteau (1992) defined experts as "those who have been recognized within their profession as having the necessary skills and abilities to perform at the highest level". The school of thought that favours experts' cognitive competence has generally been referred to as naturalistic decision making (NDM); where complex and challenging tasks are performed in naturalistic (real-life) settings (Gore, Banks, Millward & Kyriakidou, 2006; Lipshitz, Klein, Orasanu & Salas, 2001). In contrast to the normative model which prescribes how decisions should be made, NDM describes how people actually make decisions using their experience (Zsambok, 1997; Shanteau, 1992). This way, NDM researchers make prescriptions based on descriptive models of expert performance. As Kahneman and Klein (2009) puts it:

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

"A central goal of NDM is to demystify intuition by identifying the cues that experts use to make their judgments, even if those cues involve tacit knowledge and are difficult for the expert to articulate. In this way, NDM researchers try to learn from expert professionals".

There is growing evidence from the literature to suggest that decision making within the naturalistic setting involves more than one reasoning strategy (cf Gigerenzer & Gaissmaier, 2011).Within the field of cognitive science, two main modes of thinking have been suggested: the intuitive and analytical strategies (Klein, 2003:64), elsewhere known as system 1- effortless, intuitive and automatic, and system 2 - effortful, analytical and deliberate (Kahneman, 2011). However, the intuitive strategy (tacit system) is usually the "default" operational procedure, and the deliberate (analytical) system is only invoked when the former cannot solve the problem at hand or when there is need to make a conscious decision- such as planning what to do next.

Unfortunately, despite arguments that individuals may benefit from switching between intuitive and analytical approaches, little research agreement has emerged concerning the preferred sequence by which individuals should employ these approaches e.g. should people take stock of their intuition first and then engage in analysis? Or, should they expect intuition to play a key role after engaging in an analytical decision-making process? It is therefore not entirely clear whether intuitive and deliberative thinking actually represents two different modes of thinking or whether they are end points of the same dimension.

METHODOLOGY

Critical Decision Method

The critical decision method (CDM), which is a semi-structured interview process, was specifically used to elicit expert knowledge (cognitive strategies and mental models) in this study. Critical decision method is a retrospective interview strategy that applies a set of cognitive probes to actual non-routine incidents that required expert judgment or decision making (Klein, Calderwood & MacGregor, 1989)

This method was preferred to others within the cognitive task analysis family for two reasons: First, the CDM has a strong theoretical basis and it is highly instrumental in defining the emerging field of Naturalistic Decision Making (Hoffman, Crandall, & Shadbolt, 1998). Second, it has been widely used in different ways to elicit expert knowledge across different domains, hence, the validity and reliability of this method has been proven (Wong, 2000). CDM (being a semi-structured protocol with cognitive probe) has also been successfully used to overcome the effects of memory limitations- which has proved to be a major concern regarding retrospective verbal reports in the knowledge elicitation process (Clark et al., 2006)

Participants

30 experienced fire-fighters (n=30) selected across various fire stations in the UK and Nigeria makes up the sample size for the wider study. Currently, six experienced firefighters (n=6) in the UK have been interviewed as part of the piloting process. The participants were carefully chosen based on their rank/position as well as through peer nomination- as a proof of their expertise. This is to ensure that expertise is verified and not assumed. All participants had personally been involved in managing real-life fire incidents in which they made critical decisions, and have at least operated as an incident commander.

Procedure

Participants were first asked to recall and 'walk-through' a memorable fire incident that tasked their cognitive skills. Thereafter a timeline was constructed jointly with the participants, and decision points were identified. This was then followed by probing the decision points using some set of cognitive probes- at this point, we are interested in unmasking the tacit knowledge and decision making strategies behind experts' performance. Each interview lasted between 1hr-2hr and was recorded with the consent of each participant.

RESULTS AND DISCUSSION

Data from the CDM procedure were carefully transcribed. A total of 19 decision points were identified from the six interviews, and these decision points formed the basic unit of analysis (Figure 1). Data analysis was conducted using both the critical decision analysis and the emergent themes analysis (ETA) processes (Wong, 2004).

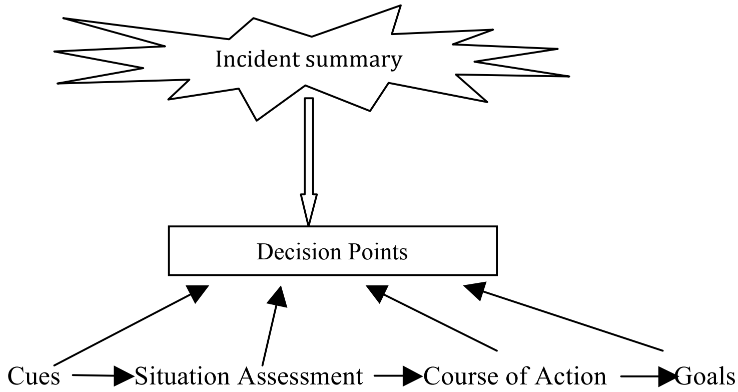

Figure 1. CDM data analysis process

Specific themes from the CDM data analysis were then collated across the six different incidents in order to identify commonalities. Five main themes emerged from this preliminary study and are briefly discussed below:

Complex Tasks

For a task to be judged as complex they must be performed using both automated and explicit knowledge, and will in addition usually extend over many hours or days (Clark et al., 2006). To demonstrate this, interviewees reported certain difficulties they encountered when managing the various incidents which they narrated. These difficulties include: working over a long duration of hours usually without breaks or refreshments, the need to manage and coordinate resources effectively (crew members, fire engines and appliances), the need to continuously monitor the situation and develop plans based on changing conditions, and the possibility of encountering novel situations etc. Interviewees explained some complexities associated with their job tasks:

> *"This incident was very challenging for me because we were working against very difficult atmospheric conditions- very dry, severe drought and windy conditions in a densely populated area- with houses all over" (participant 6)*

> *"It was very unusual for me because the woman involved was psychiatric and threatened to burn down the building (Arson) if we left. So the incident turned out to be a welfare issue" (participant 2)*

If we therefore aim to achieve a more efficient process of transferring intuitive skills from experienced officers to novices, it then becomes necessary to understand and capture how these experts perform such complex tasks (van Merrienboer, Clark & de Croock, 2002)

Cue Identification

Interviewees reported relying heavily on their senses- eyes, ears, and nose- as useful aids in formulating their decisions. In other words, they made decisions based on identified cues, such as the colour, intensity and smell of smoke. Generally, the level of risk that expert firefighters are willing to take depends largely on whether or not the environment presents favourable cues (Kahneman & Klein, 2009). When asked about the cues they looked for, interviewees responded thus:

> *"I used information I gathered from people around, I used my eyes, my ears, and a bit of common sense"* (Participant 5)

> *"I think seeing is believing; so I will take all I can see as my basis of information"* (participant1).

Situation Assessment

Interviewees reported that they always conduct a 360^0size up, also known as situation assessment immediately they arrive at the scene of an incident. They also reported that past experiences, in addition to their existing knowledge were crucial factors in maintaining accurate situation assessment.

> *"The only way I can describe it is that those incidents contribute to a template. You may only have 5 or 6 templates perhaps, but most of the incidents you go to will fit into one of those templates"* (participant 3)

However, interviewees explained that the task begins to place more cognitive demand on them when the current situation fails to match any of the pre-stored patterns in their memory. This is the point where they normally shift from rule-based to a more creative decision making strategy, which sometimes might contradict the SOPs (Klein, 2003)

> *''Yes we don't normally get people down through a ladder because it is risky, but I had to take that risk and it turned out that I was right''* (Participant 1)

Table 1. Analysis of oil storage fire incident showing shifts in Situation Assessments

Situation assessment 1	
Cues	Very large fire involving oil storage; collapsed roof; site of incident very close to residential houses
Expectations	Very intense fire with high potential to spread further
Goals	Get access to the building; get enough water to attack the fire; contain the fire
Decision-Point 1	Ask for reinforcement (Requested 15 additional pumps)
Decision-Point 2	Exterior attack- it is too dangerous to commit crew
Situation assessment 2	
Cues	Fire growing bigger; arrival of 15 additional pumps
Expectations	Presence of additional workforce will result into better control
Goals	Get access to the seat of the fire; resort to another option since the initial option of water attack is not working; safety of crew members
Decision-Point 3	Get specialist appliance to climb higher in order to see the actual seat of the fire
Situation Assessment 3	
Cues	Fire still burning due to involvement of petrol at the seat of the fire; water unable to put out the fire; pollution of water courses.
Expectancies	Fire may remain uncontained and burn out itself unless a more rigorous strategy is employed
Goals	Reduce environmental pollution from the flames as much as possible; clear the road for road users to get to work as soon as possible.
Decision-Point 4	Decision to request specialist appliance (foam attack)

Decision Making Strategy

Interviewees agreed that they employ both intuitive and analytical strategies in making most of their decisions. However, they stated that many of such decisions have to be intuitive. Even the analytical decisions still have to be made as quick as possible. According to them, time is a critical factor which is never sufficient when managing fires- thereby limiting the possibilities of conducting any thorough analysis. The breakdown of the 19 decision points based on decision making strategies and the estimated decision time is shown on table 2:

Table 2. Analysis of the decision making strategies from the six incidents

Decision making strategy	Estimated decision time by participants	No of Decision points
Intuitive	≤ 1minute	10
Analytical	>10 minutes	2
Intuitive + Analytical	≤ 10 minutes	7
		Total =19 decision points

Previous research has shown that professionals could actually draw on the repertoire of patterns that they had compiled during more than a decade of both real and virtual experience to identify a plausible option, which they considered first (cf Klein, 2003). This is what the recognition primed decision (RPD) model developed by Klein and his colleagues illustrate (Klein et al., 1989). According to the RPD model, the patterns stored in the memory of the decision maker highlights the most relevant cues, provide expectancies, identify plausible goals, and then suggest appropriate types of reactions (i.e. using recognized patterns to solve current problems).

Table 3. Analysis of the problem solving strategies from the six incidents

Problem solving strategy	Description	No of Decision points/Percentage
Standard	Generally agreed by all officers as the most appropriate option	5 (26.3%)
Typical	Modifications to the standard operating procedures to meet the requirements of the situation	9 (47.4%)
Creative/constructed	No standard solution exist; typically requires creative problem solving	5 (26.3%)
		19 (100%)

It can therefore be suggested that, depending on the complexity of the problem at hand and the ease of recall to memory, officers usually confront challenges using any of the 'standard', 'typical' or 'creative' problem solving strategies (see Table 3). To validate this, interviewees were asked if the situation they narrated and their subsequent decisions fit into the scenario they were generally trained for. They reported thus:

> "*Normally you have a set of options that you pick, and they will be right- one way or another. You might start up with option A, and a subtle change to option B will be successful. You adapt the procedures for the job to do something unusual*" (Participant 3)

> "*It was an unusual incident, but something inside you takes over- where you go into a mode of professionalism- and it comes because you have been doing it for that long*" (Participant 5)

Training Requirements for Less Experienced Officers

Interviewees ascribed their decision making ability to both the quality and quantity of the training they have undergone over their years of practice. They were able to identify some of the training they have but which their non-expert counterparts did not have (or which they have at just a basic level). These include: command and control training, breathing apparatus training, human resource management training, training involving acetylene cylinders and asbestos, training on how to use hydraulic platforms, first aid trainings etc. They however explained that this training could be merely theoretical if officers do not have the opportunity to apply them in real life incidents.

> *"Most firefighters are taught to do the techniques that extinguish fire. But an incident manager needs to be trained to manage the incident, needs to be trained to command, and that is quite a long and extensive period of training"* (participant 3)

> *"As a manager as well, you have management training; how to deal with incidents, how to deal with people, and how to deal with staff"* (participant 2)

Thus, to be able to attain automaticity and expertise, novices must be given opportunity to practice their job procedures deliberately and continuously until processing becomes more effective and efficient (Ericcson, Prietula & Cokely, 2007).

CONCLUSION AND FURTHER WORK

One of the ways of improving the overall level of human performance in a task is by understanding how proficient individuals actually perform the task. This preliminary study, using the critical decision method (CDM) has begun to reveal some of the skills, knowledge and competencies inherent in expert firefighters we interviewed.

However, at the advanced data collection stage of this overall study we hope that data from subsequent CDM process will be analyzed further to provide detailed and in-depth information about the skills, cognitive strategies, sequence, action scripts, mental models, prerequisite knowledge, and rules required for complex skill learning. This learning process will then be coherently organized for ease of teaching through the four component/instructional design (4C/ID) model.

REFERENCES

Clark, R.E., Feldon, D.F., van Merriënboer, J.J.G., Yates, K., & Early, S. (2006). Cognitive Task Analysis. A final draft.

Dane, E. & Pratt, M.E. (2009) *chapter 1* conceptualizing and measuring intuition: a review of recent trends *International Review of Industrial and Organizational Psychology, 2009, Volume 24, pp.1-40.*

Ericsson, K.A., Prietula, M.J., Cokely, E.T. (2007) The making of an expert. *Harvard Business Review* 85:114–121

Gigerenzer, G., & Gaissmaier, W. (2011) Heuristic Decision Making. *Annual Review of Psychology* Vol. 62: 451-482

Gore, J., Banks, A., Millward, L., and Kyriakidou, O. (2006). Naturalistic decision making and Organizations: Reviewing Pragmatic Science. *Organization studies* 2006; 27(7) 925-942

Hannabuss, S. (2000). Narrative Knowledge: eliciting organizational knowledge from storytelling. *Aslib proceedings*, Vol.52, No.10, pp.402-413

Hoffman, R.R., Crandall, B., & Shadbolt, N.R. (1998). Use of the Critical Decision Method to elicit expert knowledge: A case study in the Methodology of Cognitive Task Analysis. *Journal of Human Factors and Ergonomics Society.* June 1998, vol. 40, no.2, 254-276

Kahneman, D., and Klein, G. (2009) Conditions for Intuitive Expertise. A failure to Disagree. *Journal of the American Psychological Association.* Vol 64, no 6: 515-526

Kahneman, D. (2011) *Thinking, fast and slow*. England: Allen Lane Publishers

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics, 19*, 462-472.

Klein, G (2003) *The power of Intuition: How to use your gut feelings to make better decisions at Work*. New York: Doubleday

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Focus article: Taking stock of Naturalistic Decision Making, *Journal of Behavioural Decision Making,* 14, 331-352

Orasanu, J. and Connolly, T. (1993) The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood and C. Zsambok (eds) Decision Making in Action. NewYork: Ablex

Shanteau, J. (1992). Competence in experts: the role of task characteristics. *Organisational Behaviour and Human Decision Processes,* 53, 252-266.

van Merriënboer, J. J. G., Clark, R. E., and de Croock, M. B. M. (2002). Blueprints for complex learning: the 4C/ID model. *Educational Technology Research Development.,* 50(2), 39–64.

Wong, B.L. William. (2000). The Integrated Decision Model in Emergency Dispatch Management and its implications for Design. *Australian Journal of Information Systems, 2*(7), 95-107.

Wong, B. L. W. (2004). Data analysis for the Critical Decision Method. In D. Diaper & N. Stanton (Eds.), Task Analysis for Human-Computer Interaction (pp. 327-346). Mahwah, NJ: Lawrence Erlbaum Associates.

Zsambok, C. E. (1997). Naturalistic decision making: Where are we now? In C. E. Zsambok & G. A. Klein (Eds.), *Naturalistic decision making* (pp. 3–16). Mahwah, NJ: Erlbaum.

Acceleration to Expertise in Healthcare: Leveraging the critical decision method and simulation-based training

Mary PATTERSON[a], Laura G. MILITELLO[b], Regina TAYLOR[c], Amy BUNGER[c], Derek WHEELER[c], Gary KLEIN[d] and Gary GEIS[c]
[a]*Akron Children's Hospital*
[b]*Applied Decision Science, LLC*
[c]*Cincinnati Children's Hospital Medical Center*
[d]*Macrocognition, LLC*

ABSTRACT
Introduction: This study at a Children's Hospital applied the critical decision method (CDM) to identify pediatrician's expertise in recognizing sepsis. **Method**: Fourteen CDM interviews were conducted with pediatric residents and faculty. **Results:** We identified critical cues used by expert and novice physicians to recognize sepsis. Using contextual information from critical incidents, the critical cues articulated by expert physicians were integrated into simulation-based training scenarios for novice physicians to accelerate the prompt recognition and early treatment of sepsis.

KEYWORDS
Expertise; Health; Decision Making; Cognitive Task Analysis; Situation Awareness/Situation Assessment

INTRODUCTION

Naturalistic Decision Making methods and models are being used at Cincinnati Children's Hospital Medical Center (CCHMC) to drive the development of simulation-based training for resident physicians. New duty hour restrictions aimed at reducing sleep deprivation, along with increases to resident supervision, have the unintended consequences of limiting exposure to certain cases and closed loop learning within one case. Simulation-based training offers a potential strategy to increase resident exposure to critical illnesses, such as sepsis, over a range of presentations. Sepsis can present in relatively subtle ways and is easily confused with other less urgent and more common conditions. Unless detected and treated quickly, sepsis can be deadly.

Innovation

Critical decision method (CDM) (Crandall, Klein, Hoffman, 2006; Klein, Calderwood, & MacGregor, 1989) data are being used to advance the development of realistic simulation-based sepsis scenarios, leveraging a high-fidelity medical simulation facility that allows for short-duration, high-impact training experiences. This is a novel application of CDM data to generate training scenarios for healthcare simulations. The CDM was designed to aid experts in articulating tacit knowledge, aspects of expertise that are not easily described. In this study, we conducted CDM interviews with both experienced (emergency room and critical care) and novice pediatricians about patients with sepsis. The aggregate set of critical incidents provided realistic cues, cue clusters, and contextual detail to create novel simulation-based training for sepsis recognition. This research extends an earlier study focusing on sepsis cues in neonates (Crandall and Getchell-Reiter, 1993) to include all of pediatrics.

METHODS

Preparation

The principal investigator and core research team were pediatricians and qualitative researchers who were familiar with high reliability methods but were not experienced in CDM interviewing. In preparation for the data collection, three pediatricians and two qualitative researchers attended two half-day workshops conducted by experienced CDM interviewers.

Participants

Fourteen paediatricians, including 8 resident, 3 emergency and 3 critical care pediatricians, participated in individual CDM interviews conducted by trained study team members.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Interviews

Interviews were conducted by interview dyads containing one experienced pediatrician and one qualitative researcher. Interviewers asked each interviewee to describe a challenging incident involving sepsis. Each interview lasted approximately one hour, was audio recorded and then transcribed.

Coding

The objective of the coding was to identify cues and strategies to be incorporated into simulation-based training. We used a card sorting technique to organize and prioritize the potential coding categories. The resulting coding scheme included seven large categories and subcategories within each. One experienced emergency pediatrician and three qualitative researchers coded the remaining data. For each code, we noted whether the category was mentioned as well as the type of cue (increased/decreased suspicion of sepsis, confusing/misleading) and the type of judgment (normative, ipsative or cue discrepancy).

FINDINGS

A total of 23 patient-based sepsis incidents were discussed in the 14 interviews, 12 incidents relayed by residents and 11 by faculty. Table 1 depicts a subset of the coding categories most relevant to simulation design.

Table 1. Sample coding categories

Categories	All interviews		Residents		Faculty	
	Freq.	% of interviews	Freq.	% of interviews	Freq.	% of interviews
Classic Indicators of Sepsis:						
Consensus Criteria:						
Fever or Hypothermia	18	**78.26%**	8	**66.67%**	10	**90.91%**
Tachycardia or Bradycardia	20	**86.96%**	10	**83.33%**	10	**90.91%**
Tachypnea or Bradypnea	10	43.48%	5	41.67%	5	45.45%
WBC abnormalities (high, low)	6	26.09%	4	33.33%	2	18.18%
Experienced-based Criteria:						
Distal perfusion	15	65.22%	7	58.33%	8	**72.73%**
Mental status changes	19	**82.61%**	11	**91.67%**	8	**72.73%**
Ill appearance	13	56.52%	7	58.33%	6	54.55%
Hypotension	13	56.52%	8	**66.67%**	5	45.45%
Other lab abnormalities	14	60.87%	7	58.33%	7	63.64%
Rhythm disturbances	2	8.70%	1	8.33%	1	9.09%
Patient's interaction w/parent	4	17.39%	1	8.33%	3	27.27%
Risk factors for sepsis:						
Medical history						
Obvious source of infection	12	52.17%	6	50.00%	6	54.55%
Chronic medical illness	15	65.22%	8	**66.67%**	7	63.64%
Age less than 2 months	1	4.35%	1	8.33%	0	0.00%
Indwelling plastic	7	30.43%	4	33.33%	3	27.27%
Lack of immunizations	0	0.00%	0	0.00%	0	0.00%
Fever of unknown origin	0	0.00%	0	0.00%	0	0.00%
Source of Information:						
Family or caregiver concerns/anxiety	11	47.83%	7	58.33%	4	36.36%
Non-physician healthcare provider concerns	6	26.09%	3	25.00%	3	27.27%
Strength of data	3	13.04%	2	16.67%	1	9.09%
Personal trust/respect	4	17.39%	2	16.67%	2	18.18%
Disparity between trainee and faculty impression	5	21.74%	3	25.00%	2	18.18%
Cue clusters	16	**69.57%**	7	58.33%	9	**81.82%**

The left column contains coding categories. Frequency counts and percentages for each category are depicted in the columns to the right. If a category was mentioned in an incident, it received a count. For example, *Fever or Hypothermia* was mentioned in 18 out of 23 incidents, or in 78.26% of the incidents. Frequency counts and percentages are further broken down by experience level with residents being the least experienced participants and faculty representing the experts. Items highlighted in red were mentioned by at least two thirds of the interviewees. . These critical cues became the focus of the scenario design. We have not yet analysed the data for expert-novice differences.

Classic indicators/Consensus Criteria

As expected, the classic indicators of sepsis appeared frequently in the critical incidents related by interviewees. These indicators are traditionally associated with sepsis and can be considered the "textbook signs" of sepsis. Signs of fever and tachycardia or bradycardia were the most frequent classic indicators, with tachypnea or bradypnea and white blood cell count abnormalities appearing less frequently.

Experience-based Criteria

Experience-based criteria require more judgement and intuition to recognize the clinical importance of the cue. For example, abnormal distal perfusion was described frequently in the critical incidents, in a range of different ways. Some physicians described color change as a cue while others mentioned cool or mottled extremities. Figure 1 provides a list of descriptors used to describe important changes in distal perfusion.

Distal Perfusion Descriptors			
Skin color	Extremities	Temperature	Other
Pale	Mottling, especially lower extremities	Cold	Delayed cap refill
Pale-ish gray	Hands were mottled	Mottled + warm	Decreased peripheral perfusion
Pasty	Extremities were cold	Sweating	Poor peripheral pulses
Pallor (African Amer.)	Pale extremities	Perfusion was warm	Vaso-constricted
Yellowish	Nose was yellowish		Pulses were thready
No nice flush on cheeks			
Mottled			
Flushed			
Purple			
Reticulated pattern			

Figure 6. Distal perfusion descriptors

Mental status changes were another frequently mentioned experience-based cue. These can be subtle and often depend on a comparison to a "normal" or baseline state for that patient. Mental status is particularly difficult to assess in patients who have comorbidities or may be on medications that alter mental status and activity level.
Ill- appearance was mentioned frequently by first-year residents. In this category, we included any mention of an impression or gestalt of well, ill, toxic, or unable to explain/non-descript "not normal."

Medical History

Chronic medical illness is a risk factor for developing sepsis that appeared consistently in the critical incidents. Often this was in the context of a discussion that children with chronic medical conditions are more likely to have infections, represent a population where it is more difficult to discern altered mental status due to medications or developmental delays, and routinely exhibit some of the consensus criteria such as difficulty breathing or rapid heart rate, which could be used to reject sepsis as a possible diagnosis.

Source of information

Interestingly, the least experienced physicians in our sample, first year residents, frequently mentioned *family or caregiver concerns* as an important cue in the critical incidents they discussed (75% of interviews). Parent identification of a change in baseline, that the child "just isn't himself", often cued a resident to begin to consider more serious conditions such as sepsis.

SCENARIO DESIGN

Scenarios

Critical cues identified in the CDM interviews drove the design of five scenarios presenting challenging sepsis cases, including one garden path scenario. The intent was to design scenarios that would accelerate the recognition of sepsis in novice physicians. Junior residents often observe and provide support functions; they are rarely the primary team leader. As a result, they generally gather information based on assessments and interpretations made by other healthcare professionals rather than experiencing the cues directly. The simulation facility represents a safe environment in which we would be able to place residents in that leadership role where they would have to seek, interpret, and make meaning from the cues available without the help of a more senior physician.
We chose to focus on critical cues and cue clusters frequently mentioned in the interviews as important indicators of sepsis. In more traditional scenarios, it is common to provide the participant with findings from radiographs, ECGs, or distal perfusion by simply reporting the results. For this project, we wanted to present the

patient with the raw data to support the development of perceptual skills and pattern recognition in the context of a challenging incident. Rather than having an actor provide radiographic findings over the phone, we presented the participant with the image so that s/he could interpret them first-hand. Some cues required innovative strategies such as placing the feet of the mannequin in ice prior to training to simulate cold extremities. To improve the fidelity of mental status cues, we used a voice modulator that allowed adult actors to speak through the mannequin in the voice of an adolescent male, a 2-year old girl, 6-year-old boy, depending on the requirements of the scenario. Videos and screen shots were used to depict changes in skin appearance and capillary refill and required the physician to interpret the image and independently draw their conclusions.

Cognitive artifacts such as medical history and lab results were incorporated into the simulation. Scripted comments from support staff and family members were included to simulate cues identified in the interviews. We also incorporated verbatim comments from the critical incidents into scenarios to make the actors playing the role of support staff more realistic. Scripted telephone responses from consulting physicians and radiology techs were embedded to increase realism while limiting the number of actors required to conduct the simulation.

Another innovative approach integrated into the scenarios was more realistic timing and urgency. Often scenarios are focused on emergent crisis situations, such that the patient has impending or actual cardiorespiratory failure leading to a medical code. Several of the CDM incidents included situations in which a patient's condition slowly deteriorated, making it easy to miss the sepsis cues. In order to address this, we designed scenarios that are more representative of the way a situation is likely to unfold on the unit..

To illustrate, one scenario involves a 6-year-old boy with developmental delays who is recovering from surgery for a broken arm. Initially, the child has a normal temperature, but develops labored breathing and an increased heart rate (tachycardia). His eyes are closed and he cries intermittently. He is slightly blue around his lips and chin. As the pediatrician examines the child, the child becomes agitated and inconsolable. Capillary refill is delayed. As the scenario unfolds, the patient's condition changes depending on the intervention(s) performed by the pediatrician. However, even if the pediatrician intervenes quickly and appropriately, signs of sepsis become more evident (i.e., heart rate increases and blood pressure drops, respiratory rate goes down, level of consciousness decreases). In this scenario we incorporated a number of cues mentioned in critical incidents including a chronic medical illness (developmental delay), delayed capillary refill (visible on video), and changes in activity level and mental status in addition to standard consensus criteria (tachycardia and bradypnea).

CONCLUSIONS & NEXT STEPS

CDM interviews revealed a set of critical cues and cue clusters that were successfully integrated into scenarios used in a simulation environment. The rich detail elicited in the CDM interviews drove the use of innovative strategies for presenting raw data to residents in the simulation so that they would be challenged to notice and interpret the cues without the help of more senior personnel.

At the time of this submission, we are testing the scenarios to distinguish novice and expert differences in recognition of sepsis. Using the scenarios developed based on critical incidents, a cohort of pediatric residents and faculty were recruited to participate in the study. Measures include an adapted version of the situation awareness global assessment technique (SAGAT) (Endsley & Garland, 2000) to assess each participant's understanding of the situation at three points in each scenario. Following the testing process, a full simulation-based curriculum for the recognition of sepsis will be implemented for all first year pediatric residents.

ACKNOWLEDGMENTS

This research was funded by the Agency for Healthcare Research and Quality.

REFERENCES

Crandall, B. Klein, G., Hoffman, R.R. (2006). *Working Minds: A practitioner's guide to cognitive task analysis.* MIT Press, Cambridge.

Crandall, B. and Getchell-Reiter, K. (1993). Critical decision method; A technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Science,* 16, 42-51.

Endsley, M.R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT, 131-157. In M.R. Endsley & D.J.Garland (Eds.). *Situation awareness analysis and measurement*. Erlbaum, Mahwah.

Klein, G.A., Calderwood, R., and MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics,* 19, 3, 462-472.

Militello, L.G. and Lim, L. (1995). Patient assessment skills: Assessing early cues of necrotizing enterocolitis. *Journal of Perinatal and Neonatal Nursing*, 9, 42-52.

The role of the Perceptual Cycle in teams

Katherine L. PLANT[a], Neville A. STANTON[a] & Catherine HARVEY[a]
[a]Transportation Research Group, Faculty of Engineering and Environment, University of Southampton, Southampton, SO17 1BJ, UK

ABSTRACT

Introduction: The Perceptual Cycle Model has been previously applied to help explain decision making, but this has only been from the perspective of one individual and not a team. This paper explores the team perceptual cycle process. **Method**: Four crew members from a helicopter team were interviewed about an incident using the critical decision method. Deductive thematic analysis was employed to analyse the transcripts. **Results and discussion:** The data were modelled into a team perceptual cycle to demonstrate how individual perceptual cycles contribute to the overall team process. The intentions of future work are discussed.

KEYWORDS

Theory and modelling; decision making; transportation (aviation); Perceptual Cycle Model

INTRODUCTION

The Perceptual Cycle Model (PCM; Neisser, 1976) is based upon the idea of a reciprocal, cyclical, relationship between operator and environment. The PCM models interaction between person and world, with heavy emphasis on the role of schemata (knowlegde structures based on similar past experiences). Neisser presented the view that human thought is closely coupled with a person's interaction in the world, both informing each other in a reciprocal, cyclical relationship. Existing knowledge (schemata) leads to the anticipation of certain types of information (top-down processing); this then directs behaviour (action) to seek out certain types of information and provides a way of interpreting that information (bottom-up processing). The environmental experience (world) can result in the modification and updating of cognitive schemata and this in turn influences further interaction with the environment.

Recent research has applied the PCM to a variety of incidents which involved a high decision making component, including the Ladbroke Grove rail crash (Stanton and Walker, 2011), the Stockwell shooting (Jenkins et al., 2011) and the Kegworth plane crash (Plant and Stanton, 2012). The model only provides a representation of information processing for one individual. Contemporary perspectives on decision making acknowledges the importance of distributed decision making, i.e. multiple individuals and technological agents working together to make decisions (Stanton and Bessel, *in press*). The PCM has been used to underpin theories of team-processes. For example, the theory of Distributed Situation Awareness uses the PCM to explain why teams achieve compatible, rather than shared, situation awarness (SA; Salmon et al., 2009). The PCM emphasises the role of schemata in perception and decision making; Schema Theory argues that no two individuals will ever posess exactly the same schema because they will have had slightly different experiences, therefore team SA will only ever be compatible and not shared.

Despite the fact that the PCM has been used in decision making research, where team decision making is increasingly becoming the focus of research, there is little research that explores the perceptual cycle of a team. This paper presents a case study of a Search and Rescue (SAR) helicopter team when dealing with a critical incident to explore a team perceptual cycle. The four crew members who were present in the helicopter: pilot flying (PF), pilot not flying (PNF), winch operator (WO) and winch man (WM), provided accounts of the incident. The data was qualitatively analysed using a coding scheme developed from the PCM. The resulting data has been modelled into a team-PCM to demonstrate how the world information, schemata held and actions performed for each individual team member amalgamate to form an overall perceptual cycle for the whole team. At this exploratory stage future research is also discussed.

METHOD

Critical decision method

The CDM (Klein et al., 1989) elicits decision making data through the use of cognitive probes as a tool for reflecting on strategies and reasons for decisions. The method was originally designed to understand decisions during non-routine incidents. Usually the CDM consists of four phases: incident identification, timeline construction, deepening probes and 'what if' queries. In this study, phase four was omitted, Klein and Armstrong

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

(2005) have acknowledged that it is acceptable to adapt the method to suit the needs of individual research projects. The cognitive probes were used to collect decision making information when the crew of a SAR team had to deal with a fault with the engine oil temperature (EOT) during a training exercise. A selected example of the probes are provided in Table 1. Interested readers are directed to additional texts for the full CDM procedure (e.g. Crandall et al., 2006; Klein and Armstrong, 2005).

Table 1. Examples of the cognitive probes in the CDM

Area	Probe
Goals	What was your specific goal during the scenario?
Decisions	What was the primary decision that you made?
	What features were you looking for when formulating your decision?
Information	What information did you use to make that decision?
	Was there additional information that you might have liked to assist your decision making?
Experience	Was the decision made comfortably within your experience (why/why not?)
	Did your experience influence the decision that you made?
Expectations	Did you hold any expectations that influenced the decision making process?
Options	What other courses of action were considered/available to you?
	How was one option chosen and the others rejected?

Participants

All four crew members were male. The PF was 48 years old with 5500 hours of experience (600 on type). The PNF was 55 years old with 9000 flying hours experience (1500 on type). The WO was 65 years old with 6000 hours of relevant experience (500 on type) and the WM was 60 years old with 7500 hours of relevant experience (850 on type).

Procedure

The four crew members were interviewed seperately at their helicopter base. Each participant provided a high level overview of the incident and the cognitive probes were asked in relation to the decision making made during the incident. The interviews were audio recoreded and later transcribed. The interviews took place six months after the critical incident occured.

Data analysis

Deductive thematical analysis has been employed in this study whereby data was classified into meaningful themes which were generated from existing theory (Boyatzis, 1998). In accordance with the objectives of this paper, the coding scheme was based on the three categories of the PCM (see Table 2). The text segments obtained from the CDM transcripts were coded for instances of the themes identified in the coding scheme (Table 2). The data was used to understand individual perceptual cycle processes and to explore how a team PCM can be represented.

Table 2. PCM coding scheme

	Code name		
	Schema	Action	World
Definition	Mental structures held by individuals that organise their representations of the world.	The process or statement of doing something, or the intention to do something.	Externally available information in the world (environment).
Description (for coding)	Statements relating to the use of prior knowledge and experience, i.e. things based on experience, expectation or 'knowing' things (this could be implied information through the discussion of training and/or standard operating procedures)	Statements of doing an action or discussion about potential actions that could be taken.	Statements relating to potential or actual information existing in the world (environment). These can be physical things, conditions or states of being.
Example	"*my expectation was that the engine would take a while to start in the rain*"	*"I turned on the engine"*	*"Caution light came on"*

A PERCEPTUAL CYCLE REPRESENTATION OF TEAM PROCESSES

The incident occurred when flying an AW139 helicopter as part of a routine SAR winch training exercise over a vessel. The WM was out of the aircraft and over the vessel when the pilots were altered to a problem with the EOT, the high tempterature meant the crew had one minute of flight time before the engine had to be shutdown. The WM had to be rapidly returned to the helicopter. Transitioning from the hover into forward flight caused air to circulate and naturally cool the engine oil which gave the pilots thirty minutes flying time so the decision was made to return to base which was ten minutes away.

As part of the CDM procedure the incident was broken down into seven distinct phases: Briefing, dummy run, live run, onset of critical incident (EOT problem), immediate actions (training aborted), diagnostics and return to base. Due to the space constraints only the team's percptual cycle for dealing with the onset of the critical incident will be considered. Figure 1 and the associated discussion exemplifies the perceptual cycle that each individual crew member engaged with at this phase of the incident.

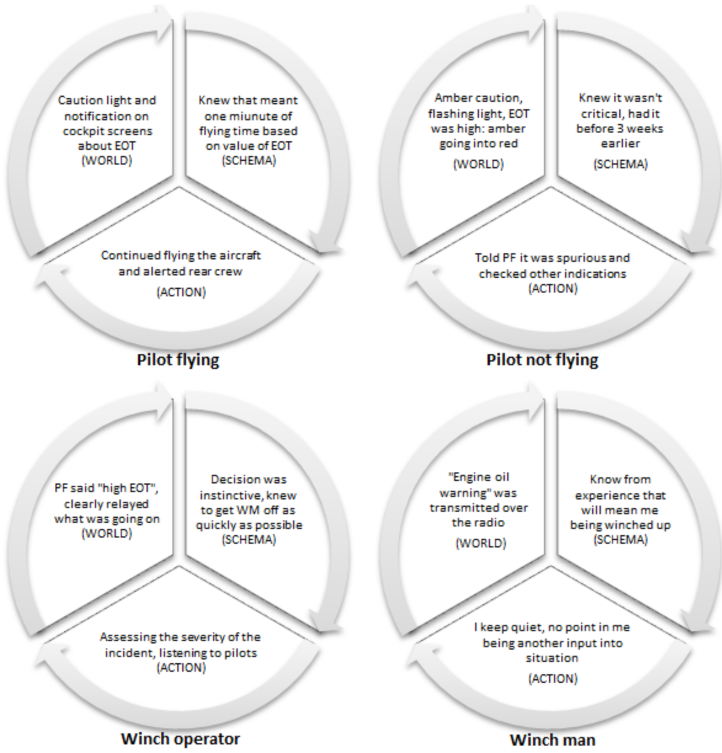

Figure 1. Individual PCMs for each crew member during the onset of the incident (high EOT)

Both pilots were alerted to the fact there was a problem with the EOT by indications in the cockpit (flashing amber caution light and text appearing on Primary Flight Display). The WO and WM were not privey to this information so they were alerted to the problem from communications made by the PF; the actions of the PF becomes world information for the WO and WM. It is clear that each individual posesses a schema for the situation that is relevant to their situation. For example, when the incident occured the WM was deployed on the winch so his schema, based on similar past experiences, is that he will be winched up and his resulting action is to remin quiet so that the crew can deal with the situation. Similarly, the world information of hearing “high EOT” activates what the WO described as an instinct, i.e. he knew what had to be done based on the world information he was exposed to. The most interesting difference between crew members lies between the two pilots who, for the same situation (EOT problem), have very different schemata. For the PF the onset of the problem activates a generic training schema, i.e. the level the EOT had reached meant that they only had one minute of flight time before an engine had to be shut down. However, the PNF had experienced the same problem three weeks earlier and therefore his activated schema was that this was not a critical fault and was most likely to be a spurious indication, which he verbalised to the PF.

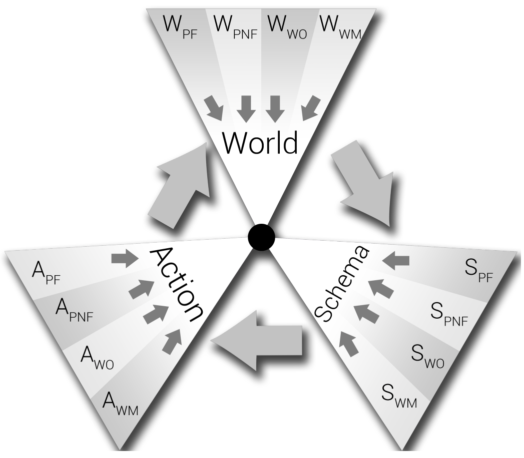

Figure 2. Team Perceptual Cycle Model

The PF and PNF would have been exposed to very similar, if not identicle, training. This example therefore demonstrates the unique, individual, nature of schemata and the potential influence of schemata in decision making: due to the activated schemata, the PF was in a 'react now' frame of mind, whereas the PNF was more relaxed about the situation as his schema led him to believe it was spurious. If the PNF had been PF that day (and therefore would have had ultimate responsibility for the decisions made) he would have still reacted immediately because standard operating procedures are designed to prescribe a clear way of reacting. However, it is still interesting to so clearly see differences between crew members in terms of the perceptual cycles each are engaged with and it demonstrates that the potential exists within a team to make inappropriate decisions based on the viewpoint of one team member. Figure 2 exemplifies the PCM from a team point of view. Each individual team member makes a contribution to the central part of the model which represents the overall team perceptual cycle; the EOT problem activated existing schemata that were devloped as a result of previous training and experiences and the actions taken to deal with this resulted in the training being aborted.

CONCLUSION

This paper intended to demonstrate the importance of considering the whole team in the context of the PCM, which is a model that has been previously used in decision making research but with little consideration to team processes. As expected, with exploratory work, this case study has highlighted further areas to investigate. A basic illustration of the approach is provided in Fig.2; highlighting how each team member contributes to an overall team perceptual cycle. However, future work is intended to further refine this, both diagrammatically and in understanding the processes at play. For example, as part of a model to explain schema-driven everyday activity, Norman and Shallice (1986) discussed the process of contention scheduling. This is the process by which competing schemata are prioritised based on strength of activation and motivations of the individual. In a team situation new elements are added: motivations of the team and individual team member's roles and dominances. Future work intends to explore how contention scheduling works in teams and this will be particularly relevant in ambiguous or spatially remote teams where a clear leader is not apparent. Other authors, for example, Reason (1990) have proposed additional schema-based processes such as similarity matching and frequency gambling. Again, it will be interesting to see how these processes manifest at the team level. Much has been researched and documented about the processes involved at the level of individual micro-cognition. What is less well known is how these processes translate to team behaviour at a macro-level. This paper has looked at this in the context of the perceptual cycle and future work intends to answer some of the questions the research has highlighted.

ACKNOWLEDGMENTS

The authors would like to thank the pilots who were interviewed and Mike Howell who assisted with the illustrations.

REFERENCES

Boyatzis, R. E. (1998). Transforming Qualitative Information: Thematic Analysis and code development, Sage, Thousand Oaks, CA.

Crandall, B. Klein, G.A. & Hoffman, R.R. (2006). Working minds: A practitioner's guide to cognitive task analysis, MIT Press, Cambridge, MA.

Neisser, U. (1976). Cognition and Reality, W.H.Freeman and Co., San Francisco.

Norman, D.A. & Shallice, T. (1986). Attention to Action. In: Davidson, R.J., Schwartz, G.E. & Shapiro, D. (eds.) Consciousness and Self-Regulation, Plenum Press, New York, 1-18.

Jenkins, D., Salmon, P. M., Stanton, N. A., Walker, G. H. & Rafferty, L. (2011). What could they have been thinking? How sociotechnical system design influences cognition: a case study of the Stockwell shooting, *Ergonomics,* 54, 2, 103-119.

Klein, G. A. & Armstrong, A. (2005). Critical Decision Method. In*:* Stanton, N. A., Hedge, A., Brookhuis, K., Salas, E. & Hendrick, H. (eds). Handbook of Human Factors and Ergonomics Methods*.* CRC Press, Boca Raton, FL, 35.1-35.8.

Klein, G. A., Calderwood, R. & Macgregor, D. 1989. Critical Decision Method for Eliciting Knowledge, *IEEE Transactions on Systems, Man and Cybernetics,* 19, 3, 462-472

Plant, K.L. & Stanton, N.A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model, *Safety Science*, 50, 300-315.

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

Salmon, P.M., Stanton, N.A., Walker, G.H. & Jenkins, D.P. (2009). Distributed Situation Awareness. Theory Measurement and Application to Teamwork, Ashgate, Farnham.

Stanton, N.A. & Bessell, K. (*in press*). Team decision making: eliciting the structure of interdependencies when returning to periscope depth. Submitted for the proceedings of NDM11.

Stanton, N.A. & Walker, G.H. (2011). Exploring the psychological factors involved in the Ladbroke Grove rail accident, *Accident Analysis and Prevention*, 43, 1117-1127.

A Model for Firefighting Managers Making Decisions in Emergencies

Agoston RESTAS, PhD
Department of Fire Prevention and Rescue Control
National University of Public Service, Budapest, Hungary

ABSTRACT

Introduction: firefighting managers have a typical profession, during which making recognition-primed decisions is regularly required. This article gives a thorough review of the topic, than shows a simple and a complex model, created by the author. **Method**: one of them was an essay analysis, the second one was a word association test, specially created for this research. **Results and discussion**: the author created a simple and a complex model for firefighting managers making decisions, taking into account time pressure, the limited capability of processing information and also a mechanism complementing the recognition-primed decision.

KEYWORDS

Firefighting manager, RPD - recognition primed decision, model for making decisions in emergencies.

INTRODUCTION

The background of recognition of a special decision-making mechanism in the focus of this article was given that, in some cases, no sufficient time is available, necessary for classic decision-making. Therefore, strategists sought to design and plan the details of military operations in advance, however, the application of different decision support instruments in live situations, designed for optimal decisions, failed many times. Decisions made in reality, often not harmonized, could not be harmonized, considering the circumstances, with the pre-formulated strategies, mostly because there was not enough time needed to achieve them. In the article, author illustrates the limits of the possibilities of analytical decision-making, presents the general operating mechanism of recognition-primed decision-making, elaborate its special model relevant to firefighting managers, as well as explore and systemize the factors that facilitate (catalyze) the processes.

An important element of the activities of emergency responders is that they cannot or only to a very limited extent can modify the terms of the task, improve them as desired. Despite the differences of environment, indications of the *complexity* of the situation, the possibility of the *volatility* in the given situation, *uncertainty* and *ambiguity* of the information available can be recognized and well identified (VUCA environment). The peculiarities of each specialized branch can be illustrated through the examples of several authors: Klein (1989) dealt with the analysis of the decision circumstances of the military, Killion (2000) took examples from the navy, Bruce (2011) shows his own medical case, Johansen (20007) simplifies difficult circumstances.

The most limiting factor from the above is *time*, proven also in author's own studies. This provides a framework impossible to burst and a forced drift, a *pressurized channel* for the decision-maker, entangled in which one can no longer break free.

DECISION-MAKING MECHANISM OF A FIREFIGHTING MANAGER

Author refers, at the general model of recognition-primed decisions, mostly to Klein's work (1989; 1999), which is analyzed by Cohen with others from the direction of critical thinking (Cohen et al., 1996). Killion supplements and combines with his multi-aspect decision-making model, (Killion, 2000). Based on Klein's work, the essence of recognition-primed decisions is that the decision-maker, through his previous experience, has several different solution schemes in his mind, which he is capable of recalling in a new situation from memory. The decision-maker immediately applies the first pattern that matches the typical features of the given problem of, that is to say, makes decisions fast as a result of previous experience. We know from Miller's researches also that the *short-term memory* of the vast majority of people can only process simultaneously 7±2 units of information (Miller, 1956). This information, of course, can be quite different, e.g. a characteristics of fire, the capacity of the response unit, a number, or even the absence of information searched.

Author has proven by essay analysis how complex the tasks of emergency responders are (Restas, 2013); this shows that in several cases, simultaneously, there is or would be a need to process many more units of

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

information than the capacity of our short-term memory would allow. *The maintenance of our decision-making capability, i.e. our short-term memory, based on the above, clearly requires that we should omit analyzing and evaluating decision-making processes protracted and use the recognition-primed decision-making procedure, based on previous experience.*

Author wishes to create a model element to demonstrate the decision-making mechanism of firefighting managers, which takes into account the limits of the simultaneous processing of information. Since the information units may be qualitatively independent of each other, author choses the simplest *graphical representation of the unit-based discrete difference* to separate them from each other. A model element must be such, which can graphically demonstrate the schemes based on earlier experience, the characteristics of different fires, and the interlocking of the former as the application of the scheme, which represents the technically correct solution of the task, i.e. effective decision. The model refers, at the general model of recognition-primed decisions, mostly to Klein's work (Klein, 1989; 1999).

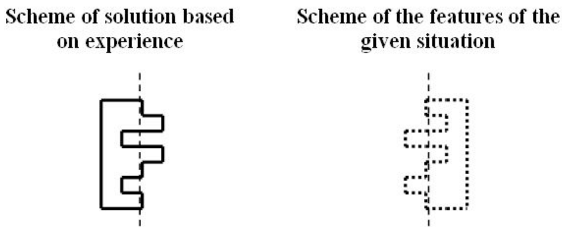

Figure 1. Graphic representation of the empiric scheme of recognition-primed decisions matching a given situation

The schemes in *figure 1* represent 7 graphical discrete values each, which are marked by positive or negative protrusions and their "center line"; these values indicate the amount of simultaneous decision-making capacity. Thus, the "negatives" of the schemes can be matched as a given situation and the solution necessary therefor. As an integration of above processes, decision mechanism functions as follows: an experienced firefighter has performed the elimination of a large number of and different fires. Despite the fact that as far as the parameters each fire is different from another, some characterizing features can be well conceived (*figure 2*).

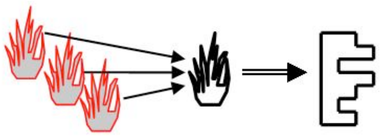

Figure 2. Evolvement of the scheme on fire

The characterizing features of identical types of fires are crystallized by experience, and are fixed in our *long-term memory*. Similarly, to the characteristics of a fire, the characteristics of successful extinguishing, the facilitating decisions are also fixed (*figure 3*); just as the mistakes desired to be avoided and the unsuccessful procedures and failures. Experience gained through many years, based on the features of fires, formulate the system of schemes, behind which we can find actions (decisions) efficiently applicable to eliminate them.

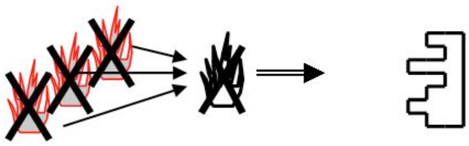

Figure 3. Evolvement of the scheme on the lessons learnt from extinguishing a fire

If another incident has almost the same circumstances as one already many times successfully eliminated by a firefighting manager previously (*model of positive confirmation*), he will attempt to use the same ones in the procedures. Therefore, another fire, quasi bearing the typified properties of previous similar fires, a decision-maker involuntarily immediately recalls the typified decisions in his conscience.

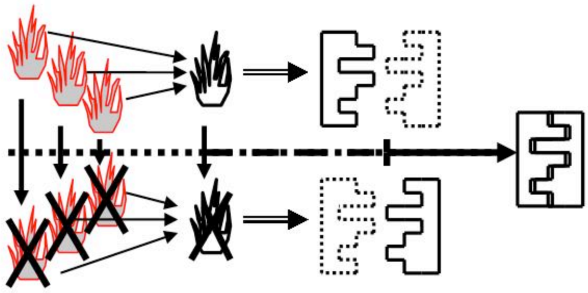

Figure 4. Aggregated scheme on fire and the evolvement of the lessons learnt from extinguishing it

The properties of a fire and of previous successful extinguishing operations, based on the above, are closely interlinked; they are each other's "reflections" (*figures 4*). *Author proved with association studies that the above, i.e. the characteristics of a fire and the thoughts directed towards its extinguishing, the schemes of response, in the case of firefighters, are very closely connected in a complex way.*

When a firefighting manager identifies a fire, he imagines what would happen if he applies the usual tactics to fight it. If the scheme of solution matches, he accepts it, if not, he rejects it and thinks of the next most typical action. Thus, it is a recognition-primed, model-matching process, which can be followed by a quick and almost automatic decision.

The long-term memory of a firefighting manager, through practical experience, has the schemes of both different fires and their extinguishing characteristics. During another alert, information available and collected on a fire automatically generates the recollection of the scheme necessary to solve it, based on which a firefighting manager defines the firefighting tactics necessary. *However, the results of association studies clearly point in the direction that at a given fire (problem) managers do not focus on the fire as a problem but rather on its immediate solution.* From this, author makes the conclusion that a decision-maker will not follow the change of the characteristics of a fire, but the validity of solution scheme, that is, the dynamics of the implementation of the extinguishing process. This does not mean a contradiction with the previous, but rather a difference in views, the shift of emphasis of the focus of attention.

The difference in views, that is, the shift of emphasis means that a firefighting manager does not focus on the change of characteristics of a fire, but rather on the expected evolvement and dynamics of the scheme selected, i.e. extinguishing tactics. The thought sequence fire–characteristics–solution is attractively logical, however, the decision capacity of our memory is facilitated if it manages and reduces the necessary information in the simplest possible way. Since the schemes of characteristics relating to a fire exist together with the schemes of solution, there is no real need for it to appear in our short-term memory. Thus, the function appearing is modified to the simplest and shows the format fire–solution.

The above do not contradict Klein's model, they rather complement it. Klein, in his model, evaluates (imagines what will happen) the results of matching schemes by the decision-maker prior to performing action version, however the aftermath of the decision, in author's opinion, is much more significant in case of firefighting managers. Since the problem immediately and automatically generates both the direction of the solution and start of the action version, rather the process itself is important in terms of efficiency, which is caused by the decision. *The schemes based on experience certainly contain the information on the dynamics of the process of fire, so if it meets the expectations, we do not have to modify the original firefighting tactics.* However, if the dynamics of the process does not suit the expected, the change is inevitable in the performance of efficiency. *Based on the above, the recognition-primed decision is not just an individual act before extinguishing the fire, but it is also the continuous accompaniment as needed.* By doing this, author shares the view that the experienced decision-maker perceives the problem together with its solution, *furthermore, author extends the continuous co-existence of the problem and of the whole process of solution of an emergency (firefighting and technical rescue).*

THE COMPLEX MODEL OF DECISION-MAKING OF FIREFIGHTING MANAGERS

If not enough time is available for analyzing and evaluating decision-making, recognition-primed procedures receive a greater role. Critical thinking uses recognition procedures, during which the decision-making process can be accelerated or analyzed with the help of a quick test and depending on the time available. The quick test, considering the circumstances, hinders recognition-primed decision and prefers critical thinking. However, when the circumstances are inappropriate for critical analyzing thinking, the quick test allows immediate reply.

Despite the limited decision capacity, thanks to recognition-primed mechanisms, in most of the occasions, correct decision are made by firefighting managers. Time limit precludes the possibility for the firefighting manager to carry out analyses necessary for the classic model, therefore, the selection of the optimal possibility is objectively not attainable by the decision-maker. The decision-maker is not striving to achieve ideal results, as a response to the difficulties of collecting information and reducing costs in relation, but depending on the circumstances, he is satisfied with the its satisfactory solution.

By reducing the time available for decision-making and for maintaining decision-making capacity, a firefighting manager applies the management (decision-making) method based on exceptions in numerous situations. Its essence is that several moments of interventions proceed protocol-like, thus, they need not be controlled all the time; on the other hand, not all the phases of the processes require direct management decision.

Based on the study of creativity, author has concluded that there is no such a feature characteristic of the working circumstances of firefighting managers that would not be advantageous to perform efficient work in a VUCA environment. Therefore, it is sure that the creative capabilities of firefighting managers can be explicitly advantageous to facilitate the professionally correct decisions on firefighting and rescue tasks.

Heuristics are not random-like errors or specific distortions facilitating our everyday activities. These are the results of simplifying mechanisms, through which decision-makers can make difficult tasks manageable for themselves. Besides the benefits of heuristics, the greatest challenge for a firefighting manager can mean the inherent erroneous distortions, which surely often help, but their uncritical acceptance, in certain cases, can end up in fatal dangers.

The declared objective and sense of the decisions of firefighting managers is the efficient implementation of emergency interventions. It is symbolized by the principles of firefighting with structured division, on the top of which we clearly find the saving of human lives (Restas, 2013).

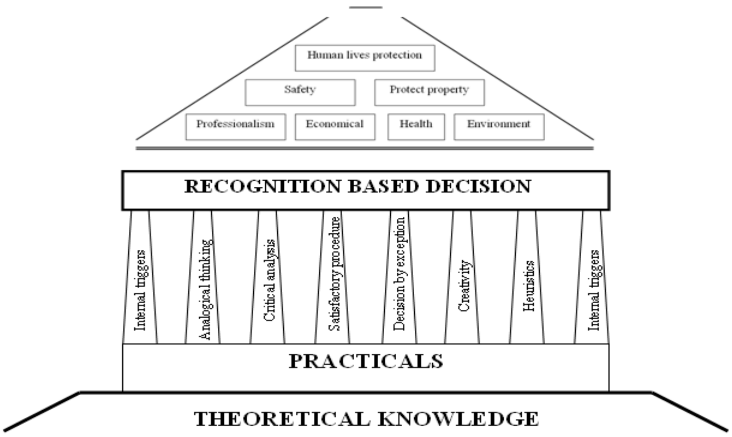

Figure 5. Complex model of decision-making of firefighting managers in emergencies

Firefighting managers certainly have less time to make their decisions compared to the time interval of classic decisions, so, their decision mechanism is strongly based on recognition procedures due to the peculiar environment (VUCA), and the limited process possibility of simultaneous pieces of information. *The competence of firefighters is based on the unity of theoretical knowledge and practical experience. Building on practical experience, the different mechanisms like analogical thinking, critical analysis, satisfactory procedure, decisions based on exceptions, creativity and heuristics, together with the internal triggers, hold as pillars and make recognition-primed decision procedure of firefighting managers operational.* Author illustrates the above as a complex system of emergency decision-making of firefighting managers in *figure 5.*

REFERENCES

Bruce, E. [2011] A Picture is Worth a Thousand Words – at Least; Pentington Media Inc. USA

Cohen, S. M., Freeman, J.T., Thompson, B.B. [1996] Integrated Critical Thinking Training and Decision Support for Tactical Anti-Air Warfare; Report, Cognitive Technologies, Inc., Naval Air Warfare Center Training System Division, Contract No. N61339-96-R-0046.

Johansen, B. [2007] Get There Early: Sensing the Future to Compete in the Present. San Francisco, CA: Berrett-Koehler Publishers, Inc.. pp. 51–53. ISBN 9781576754405

Killion,T.H. [2000] Decision Making and the Levels of War; Military Review, United States Army Combined Arms Center, Fort Leavenworth, Kansas, 2000 novemer-december

Klein, G. A. [1989] Strategies of decision making , Military Rewiev, No.5

Klein, G. A.: [1999] Sources of Power: How People Make Decisions Cambridge, MA: MIT Press 1999 ISBN 0262611465

Miller, G. A. [1956] The Magic Number 7 Plus or Minus 2; Some Limits on our Capacity for Processing Information, Psychology Review, Vol. 63

Restas, A. [2013] Principles of Decision-Making of Firefighting Managers, Based on Essay Analysis; Poster, The 11th International Conference on Naturalistic Decision Making, 22-24 May 2013, Marseilles, France

Principles of Decision-Making of Firefighting Managers, Based on Essay Analysis

Agoston RESTAS, PhD
Department of Fire Prevention and Rescue Control
National University of Public Service, Budapest, Hungary

ABSTRACT

Introduction: this article focuses on determining the principles of decision-making of firefighting managers in emergencies. **Method**: the author used a freely compiled essay analysis, made by firefighters with several years of practical experience. **Results and discussion**: the research ascertains that firefighters clearly know their special decision situations, and that they are forced to make them under time pressure, and also that the cornerstones of the practical experience are highlighted in the essays. Based on relevant statutes and the cornerstones elaborated in the essays, the author logically arranged them and determined the principles of decision-making of firefighting managers in emergencies.

KEYWORDS

Firefighting manager, recognition primed decision, principles of decision-making, essay analysis.

INTRODUCTION

It is well known that it is difficult to define the genre of an essay. Those attempting to do so will certainly fail (Kerner, 2009). Bakó's opinion is that "*it is easier to define what an essay is not than to compact it in one sentence*" (Bakó, 2002). Due of the nature of the genre, essays have several divisions, groupings, however, the reflection is a common feature of all (Gyergyai, 1984), that is, its author freely explains his own views on the subject, which can shift even in the direction of self-expression and self-confession (Szabó, 2001).
According to Samuel Johnson (1709-1784), an essay is no other than "*The free flow of thoughts of the mind, a single raw work, it is not an orderly, organized work.*"[4] The analysis of the free flow of ideas is obviously not easy; however, there was a need for this spontaneously appearing expression of opinions for author's researches. He was curious about the information on qualitative processing, in which firefighters with more or less experience mention the difficulties of their own decisions and of their companions, and their specialty. Author wanted to draw up characteristics from them, which are able to generally demonstrate and express the difficulties, the circumstances of one of the specific types of emergency decision-makers, of firefighting (rescue operations) managers.
Applying the method of analyzing essays, author shares Thorn's (2000) view that a researcher following the qualitative method "*does not always exactly know what and how to ask, but entrusts himself to the characteristics of the phenomenon researched, lets it control the research according to its own internal logic*" (Juhász, 2008). It is not the undemanding state or lack of interest qualifying the preparation of the research, but on the contrary, it is a predictive factor, which allows to ensure the extent of freedom, which the researcher needs (in this case the author), for the sake of achieving the goals of the study. With response to this thoughtfulness, research wanted to take advantage of the freedom of opinion of firefighters by ensuring the extent of freedom provided by the essay during the research, and based on the above, and by the possibility that is proven suitable, researched by others.

APPLIED METHOD OF ANALYSIS OF THE RESEARCH HYPOTHESES AND THE ESSAY

Research Hypotheses

Using the well-known regularities of author's own experience, as well as of the adherence to practical activities, repeated many times, research had an assumption: *based on hypothesis, the cornerstones of practical experience will be outlined from the content of the essays and the factors will be formulated that affect or limit their*

[4] http://hu.wikipedia.org/wiki/Essz%C3%A9

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

decisions in the most disadvantageous way. Author regards it as a clear result, if emergency or the related notion, *limited time* in a commonly understood sense, appears in the papers predominantly and in large numbers.

Method Applied for the Analysis of an Essay Freely Compiled

In order to perform the analysis, 13 students of the *Firefighting specialization*, having several years of practical experience, submitted essays freely compiled. The least experienced person had been a firefighter for five years, the most experienced one for 24 years. Most of them, based on engagement, had gained significant experience in the management of firefighting before. Their task was to formulate a short essay (minimum 3, maximum 10 pages) on their experiences and lessons learnt relating to their decisions made during interventions, everything or anything that pops up in their minds, freely without limits. The fact that author did not give them any other restrictions or guidance, research intended to ensure the broadest possible framework for their free and spontaneous expression of opinion.

They received a one-month deadline to write their essays. 12 papers were sent on deadline; 10 could be evaluated by their quality; 2 were obviously completed with minimum targets in mind necessary to meet the requirements of the semester. The shortest content without compulsory elements (front page, contents, declaration on own sources) was expressed in 3 pages (2 essays), but there was an essay exceeding the maximum (1 essays with 11 pages of content in merit). The average extent of the content in merit of the papers was 6.4 pages. All essays, as a compulsory element, also contain a declaration, by which the students certify that the material was compiled by themselves.

Research referenced marking with at least one but no more than two codes (e.g. DPF, D5C) during analyses, after each core or specific findings, as justification, but in cases where they are acceptable and easy to retrieve from the papers, clearly as the opinion of the vast majority of students (reference to e.g. time pressure or professional experience), research did not see the need.

RESULTS OF THE ANALYSIS OF ESSAYS FREELY COMPILED

Author used basically two guiding principle for evaluation; one of them was the matching of responses to the accepted tactical elements used by the fire fighting tactical handbook (Bleszity, 1989), the other was the *expert estimation method* based on the author's own 12 years firefighting experience.

The content of the papers is a good lesson learnt for a variety of reasons. Students refer, in almost all the cases, to classic, founding psychological, sociological, management theory, decision theory knowledge, studied earlier. Thus, the role of *Maslow's hierarchy of needs* (DPF) and of *stress* (D5C, SX9) influencing decision is available in field of psychology, the "*great man*" and "*trait*" theories as the 1920's *feature theory* (DPF, BYM) characterizing the topic of management theory, however, from a developmental point of view, by this it ends vividly with the *style theories* (leadership styles) associated with the names of Lewin and Likert in each case! One paper referred to author's earlier elaboration (Restás, 2001) relating to the topic (D5C). Based on the above, it shows from the essays that the training of firefighters in subjects like management (leadership) or decisions did not keep pace with the teachings of later schools; they end with the *leadership styles* preferred in the 60's.

Along with granting constitutional rights, however, the law[5] imposes obligations on the firefighting manager, whose stringent performance is inevitable to meet the technically proper implementation of the intervention. Therefore, special requirements are clearly spelled out in the papers against a firefighting manager. These show that *students attach primacy to professionalism*, to which the link *the hands-on experience* (SX9, BYM) or explicitly or implicitly express it the context of their message (BJS, C84). This primacy can be attributed to two factors fundamentally interrelated, one of which is the *constraint to decide*, emphasized in almost all cases, the other is the *time factor*, declared in all the papers, without exception. They set out counter-examples consistently conflicted to emphasize experience, in which they present the lack of professionalism (BJS, C84), or refer to its impact causing inability to decide (G87, PDS).

They show the effects of the start of an intervention, influencing the efficiency of the entire firefighting in different ways, but all mention professional experience and time factor:

- *„I think this is the moment when the success or failure of the elimination of the incident is decided. Most depends on the first 3-5 minutes”* (BJS)
- *„there is no time or possibility for compromises, one has to act immediately”* (RWL)
- *„ (a firefighting manager) has only minutes, moments to make his decisions.”* (RWL)
- *„we feel the decision is burden on us as we approach the irreversible moment”* (PVR)
- *„ (the personnel) having arrived at the site faces the facts. The majority of our static decisions are discarded, dynamic decision processes commence.”* (DPF)
- *„it is essential that (a firefighting manager) assesses the situation in a short time and not be protracted in issuing instructions”* (SX9)
- *„the chief fire officer on site has only seconds and in many cases little or misleading information to make his decision.”* (D5C)

5 Hungary, MoI Decree No. 1/2003. on the Rules of Firefighting, Annex, item 32.

They search for solutions in the essays to clarify and explain the features of decision-making without any request. Thus the attempt to establish a static division of decisions (e.g. drills) before an intervention or dynamic division at the incident sites can be found (DPF, SX9), the static arrangement of decision support tools (BJS), the scrutiny of the conceptual differences between decision and selection, keeping in mind time factor (PVR, SX9), the gradual learning process (G87), but also the evolvement of programmed and routine decisions (RWL, FEG).
Summarizing the above, author ascertain that firefighters clearly know their special decision situations, that they are forced to make them under time pressure even if its expression or appearance can be quite different in each paper. Despite this fact, or exactly because of the significant differences, it is obvious, however, that they are not able to create a uniform image or provide satisfactory explanation to its real background.
The activity of firefighters, the primary aspects the circumstances of intervention are quite distinct in the essays and can be expressed as follows:
As the primary aspect of interventions, all unambiguously identify *lifesaving*, but *safety* (C84, D5C) and *professionalism* (RWL, FEG) are also present. There is a firefighter, who pairs safety with *special knowledge*, *experience* and *routine* (SX9), and there is another one, who attaches a greater role to avoiding *stress* (C84, D5C).
It is unambiguous that they regard as important the role of the *knowledge of the site* (BYM, G87) and *„clairvoyance"* as well, the ability with which they comprehend the site and the given situation (D5C, RWL). It is remarkable how important role they attach to calmness by using different expressions like *„laid back"* (BYM), *„cool head"* (DPF), which obviously refer to the danger of their opposite as well. To explain it, author could list his own experience as well, but the statement may be experience, in which one expressed: *„It is important that it is not events which should drift the personnel, but the person in charge should control and direct the activities on site."* (G87)
The inevitability to comply with the law is mentioned in 7 essays, 6 out of them tried to emphasize it by copying the rights and obligations of a firefighting manager stipulated by law word for word. This strict compliance with the law, also justifiable by other researches, is obviously characteristic of all order-controlled organizations (e.g., Mezey, 2009). Despite the above fact, we can see in the essays some examples of forced but conscious and necessary infringements of rules:

„One may not order something that is against the law. My personal opinion is that some exceptional cases, however, may give reason not abide by the rules. These may be cases when all the experience and individual skills of a leader are necessary in a given situation to be able to find a solution, which legislators could not have in mind or it is not properly regulated." (BYM)

As a summary of the above, author ascertain that, the cornerstones of the practical experience are highlighted by the content of the essays, and all the factors conceived that influence or limit the decisions of firefighting managers the worst, is fulfilled. Author's assumption, according to which the role of time is sensibly predominant, even by expressing the concept of constraint in different ways, can be justified in each essay.

As a circumstance influencing a decision or defining its professionalism other cornerstone can be conceived, e.g. the *primacy of lifesaving* overriding everything, the importance of *safety* and professionalism, or preserving calmness by avoiding *stress*, i.e. the continuous maintenance of decision capacity.
The latter, i.e. the maintenance of decision capacity clearly ensures the fulfillment of the previous cornerstone; ranking and arranging them, and the aspects of the target function of interventions logically supplemented, the principles of firefighting (rescue) can be ascertained.
Another issue also emerges from the papers that a fire or incident appears in the thoughts of the students in the form of images very clearly. They regard as unambiguously important the role of *knowing the site* (BYM, G87) and also that they be able to immediately see through, assess and comprehend a given situation (D5C, RWL). Author draws the conclusion from this that situation exercises at different locations and the detailed analysis, "*scrutiny"* of previous cases, can also significantly increase the efficiency of later decisions. To facilitate it, an excellent opportunity is provided by plotting boards (sand tables) designed and set up for different situations, but even the use of 3D imagery IT programs. Through these assets one could illustrate the movement of responders and vehicles step by step, the progress of events, the different opportunities, and just as well the emerging dangers.
Based on the essays, author ascertains that students not only have general organizational and managerial knowledge, but they aware of their importance as well. They regard the anomalies between the stipulations put down in conventional education materials and the mechanisms experienced during decision-making under the pressure of time as job-specific features. They link the classic knowledge mostly usable in the periods between incidents to the maintenance of the daily operational order of an organization, which means, as concluded by the author, that the operational mechanism of a fire brigade, from several aspects, is identical to that of other organizations.

CONCLUSIONS - THE PRINCIPLES OF FIREFIGHTING

Research has not found an example of the *declared* ascertainment of the principles of firefighting. Despite this fact, it can be justified both by elaborating the essays and the relevant statutes[6] and accepted without dispute that the first and foremost principle cannot be other than saving and protecting *humans* lives. This is defined in the Rules[7] as a primary role, but any firefighter's immediate response can only be the same, obviously; this is clearly justified by the elaboration of the essays. Accepting lifesaving as the basic principle, putting the following tasks in a logical order, we come to a hierarchic structure.

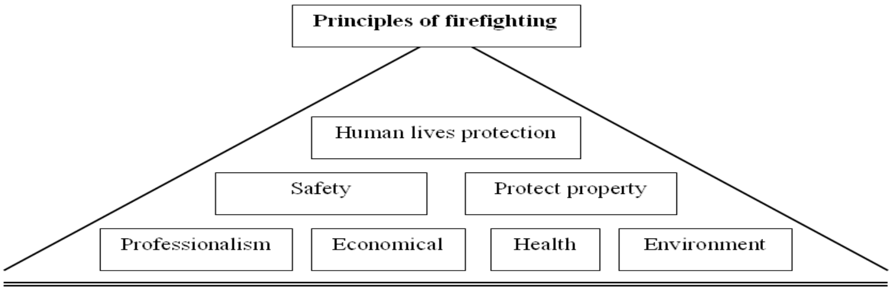

Figure 1 The principles of firefighting

Based on the above, the first principle is the *protection of human lives.* In the hierarchic structure, the most important task after lifesaving is to maintain *safety*, which includes both subordinates and the victims of an incident. Looking at the orientation of the organizational objectives, the performance of the mission, the efforts made to *protect property* should be regarded as of equal value. The maintenance of safety can be ensured by *professionalism*, which, logically, must also certainly fulfill the expectations of *being economical*. The efforts made to protect property may not incur the disproportionate deterioration of *health* or *the environment*, so, these concepts, in my opinion, should also be regarded as principles. The principles hierarchically built are in interaction with each other; however, the horizontal order is not fixed on an identical level. Therefore, *professionalism* during the *protection of property* is also authoritative, just as *health* or *environment protection* belongs to *safety*. Author illustrated his recommendation on the hierarchic division of the principles of firefighting in figure 1.

REFERENCES

Bleszity, J. & Zelenák, M. [1989] A tűzoltás taktikája. Tankönyv, BM Könyvkiadó, Budapest, Hungary

Bakó, E. [2002] Az esszéíró Kiss Tamás. Verseghy Ferenc Electronic Bibliothek, ESŐ Irodalmi lap V./2. downloaded: 08. March, 2012. http://vfek.vfmk.hu/00000013/06.html

Gyergyai, A. [1984] Esszé az esszéről. In: *Védelem az esszé ügyében.* Szépirodalmi Könyvkiadó, Budapest, Hungary, pp 7–25.

Juhász, V. [2008] Egy internetes honlap, az iwiw szegedi felhasználóinak szociolingvisztikai vizsgálata, különös tekintettel a nemre és a korra; PhD dissertation, PTE Alkalmazott Nyelvészeti Doktori Program, Pass: 09. May, 2008.

Kerner, A [2009] Esszé a középiskolában. A filozófiai esszé; Anyanyelv-pedagógia, Pedagógiai folyóirat, Magyar Nyelvtudományi Társaság, Magyartanári Tagozat, dowloaded: 08. March, 2012. http://www.anyanyelv-pedagogia.hu/cikkek.php?id=164

Mezey, Gy. [2009] Döntés és kockázat; Monography, Szent István Egyetemi Kiadó, Gödöllő, Hungary ISBN 978-963-269-099-5

Restás, Á. [2001] A tűzoltásvezető döntéshozatali mechanizmusa; Védelem, VIII./2., Budapest, Hungary, pp. 28-30, ISSN: 1218-2958

Szabó, B. I. szerk. [2001] Irodalmi fogalmak kisszótára, Korona Kiadó, Budapest, Hungary, ISBN 9638153974

Thorne, S. [2000] Data analysis in qualitative research; Evidence-Based Nursing, 2000/3 pp. 68-70. downloaded: 10. March, 2012 http://bmj-ebn.highwire.org/content/3/3/68.full

[6] Hungary, Act XXXI of 1996 on the protection against fire, on technical rescue and on the fire service.
[7] Hungary, MoI Decree No. 1/2003. on the Rules of Firefighting.

Exploring Compatible and Incompatible Transactions in Teams

Linda J. SORENSEN[a, b] and Neville A. STANTON[a]

[a] *Faculty of Engineering and the Environment, Civil, Maritime and Environmental Engineering and Science Unit, University of Southampton, Highfield, SO17 1BJ, United Kingdom.*

[b] *Institute for Maritime Technology, Operations and Innvoation, Maritime Human Factors, Vestfold University College, 3103 Tønsberg, Norway.*

ABSTRACT

Introduction: Transactions of information and situation awareness are a critical commodity in team decision making and the development of distributed situation awareness. Schema theory explains production of behaviour and argues that different schemata guide interpretation of external information. Schema theory has been given most attention at an individual level, this paper therefore seeks to consider the manner in which schema influence team interaction, coordination and distributed situation awareness. **Method**: An exploratory study was devised in which the communication of team members were analysed. **Results and discussion:** Compatible and incompatible transactions were explored in light of schema theory. It was found that compatible transactions were associated with a slightly higher efficacy compared to incompatible transactions. The transactions explored were associated with a range of schemata and these early findings indicate that schemata guide not only individual but also team behaviour.

KEYWORDS

Schemata; team decision making; distributed situation awareness; teamwork.

INTRODUCTION

Stanton, Salmon, Walker & Jenkins (2009a) pointed out that situation awareness (SA) transactions are a critical commodity in the development of Distributed SA in teams. As yet, little is known of the nature of compatible and incompatible transactions and the role these may play in Distributed SA. Transactions which take place in a team may not be used by team members in the manner it was intended and, therefore, could play a role in SA breakdown. This paper seeks to shed light on the manner in which compatibility and incompatibility of SA transactions manifests itself in teams and the role of schemata in the establishment of overall team strategies. An exploratory analysis was performed to reveal the manner in which compatible and incompatible SA transactions contribute to the regulation of teams' behaviour and contribute to the development of Distributed SA.

In order for teams' behaviour to be guided in a coordinated manner the individual team members must agree a joint approach for acting. This necessitates that the team members understand what is required of them and is brought about by the spread of information within the team.

Schema theory, based on the work of Bartlett (1932), explains the production of behaviour as an organisation of experience which are drawn when dealing with a current situation (Stanton, Salmon, Walker & Jenkins, 2009b). Stanton et al. (2009b) explained that the schemata held by a person combines with the goals they hold, tools they use and the situations they find themselves in to generate, or blend, new behaviour. Individuals gain different experience and as a result no individual may hold different schemata.

Grasser and Nakamura (1982) argued that schemata are generic knowledge structures which serve to guide interpretation of external information. Marshall (1995) explained that these knowledge structures can be represented as a network of associations. Schemata have been described as "hierarchically organised sets of units describing generalised knowledge about an event or scene sequence" (Mandler, 1984, p.14).

Actions are specified only at the highest, abstract, level and activation of a higher-order schema leads to the activation of lower level schemata to complete a sequence of behaviour (Norman, 1981). Norman and Shallice (1986) defined the higher order schemata 'source schema' and lower-level schema 'component schema'. Component schema, when activated through the source schema, become source schema in their own right as a person runs through the sequence of actions required for performing some task. As an example, "making a stew" may be a source schema which triggers a number of component schemata such as "preparing beef" which in turn become a source schema for "cutting meat", and so on. Schemata are therefore structured in a hierarchical manner (Plant & Stanton, 2012). Graesser and Nakamura (1982) differentiate between mental models and schemata by the example "restaurant eating schema". They state that this schema is generic for any restaurant a

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

person might visit, whilst a mental model would have to be related to individual restaurants and the specific time at which the restaurant is visited (Plant & Stanton, in press). An individual's schemata will be combined with the goals they possess and the situation they find themselves in to develop new types of behaviour (Stanton et al., 2009b).

Norman and Shallice explained that "when numerous schemata are activated at the same time, some means must be provided for selection of a particular schema when it is required. At times, however, there will be conflicts among potentially relevant schemata and so some sort of conflict resolution procedure must be provided" (p.4). In many areas of teamwork one course of action must be chosen and agreement within the team must be established if a common goal is to be met in a timely manner. This poses the question of how teams resolve a conflict between opposing ideas or views on what the right course of action may be. Norman and Shallice (1986) described this resolution of conflict as "contention scheduling".

This paper presents an exploratory study which considers the information passed around in the team. In particular, the information is explored to reveal whether the information trigger a number of alternative courses of action from which one must be chosen, explained as contention scheduling (Norman & Shallice, 1986).

Norman (1981) described situations in which the wrong schemata were selected as a means of describing different types of human error. He suggested that three basic types of schemata account for most errors: activation of the wrong schemata (as described in contention scheduling similar triggering conditions may lead to the wrong schemata being activated), failure to activate appropriate schemata (e.g. lack of attention to the triggering conditions which could have activated the schema) and a wrong triggering of schemata (e.g. triggering of schema at the inappropriate time).

The Role of Compatible SA

Salmon, Stanton, Walker, Jenkins, Baber &McMaster (2008) explained that each agent may hold different SA for the same situation. The individual is governed by their specific team role, tasks and goals in the manner in which they perceive the situation as it evolves (Stanton et al., 2009b; Salmon et al., 2008). This is closely linked with Schema Theory, as described above, which argues that each individual holds different schemata (as the sum of their experiences) and that no schema will be identical between two individuals (Stanton et al., 2009c; Stanton et al., 2009d). This is also closely linked to the idea that it is not necessary for the whole team to know everything (Salmon et al., 2008; Hutchins, 1995a). Successful team performance depends on knowing who knows what to access information, not knowing everything. Given the difference between individual team member's schemata and interdependent tasks awareness is not shared (Salmon et al., 2008). One team member's SA could therefore be different but remain compatible as their SA will be required to ensure that the team can perform successfully (Salmon et al., 2008). This was argued by Stanton, Salmon, Walker & Jenkins (2006) who asserted that team members have unique but compatible portions of awareness. In other words, that the team requires separate awareness but also compatible awareness whilst working towards a goal (Salmon et al., 2008).

The diverse but related literature described here point to a number of pertinent questions which may shed light on the role of schemata in team decision making and compatible SA in the team's development of Distributed SA. The exploratory research presented here was therefore guided by the following research questions:

1. Do teams exhibit the use of source schemata and component schemata (Norman & Shallice, 1986; Grasser & Nakamura, 1982; Plant & Stanton, 2012)?
2. Are conflicts of schemata, such as that described as contention scheduling, observed in team communications (Norman & Shallice, 1986)?
3. Do the team members exhibit transactions of information which are either compatible or incompatible and associated with a component schema (Stanton et al., 2006a; Stanton et al., 2009b; Salmon et al., 2008)?

METHODS

Research Design

A qualitative approach was chosen to explore the data and shed light on the three research questions detailed above. The research utilised two qualitative approaches; a top-down approach were the game rules were used as a guide to identify schemata and a bottom-up process where content analysis was utilised to explore compatible and incompatible transactions observed in communications.

Experimental Tasks

A strategy game was developed in which a chess board was used with players of four different colours; blue, yellow, green and red. The blue pieces signified friendly players and were controlled by the experimental team. Yellow players were unknown, while green were neutral and red players were enemy or opponents pieces. The rules of the game were as follows:

- The aim of the game is to take as many red players as possible
- Each Blue player has one move per turn, however, each player can give their move to another player on a turn-by-turn basis

- Each player can move in any direction but not through another player
- Moving through another player constitutes taking
- Blue players have to outnumber a red player before they can take it
- Blue must not take blue, green or yellow players
- Red must move away from blue if a blue player gets to within one space of red
- If red players outnumber the blue players they must move towards them and try and take them
- In two games the opponent players move
- In two games the opponent players are disguised as yellow and will only reveal their true colour (e.g. red or green) if a blue is next to it.
- Changing colour is considered a move (the player cannot immediately be moved after colour change). After revealing the colour the player cannot change back to yellow.

A military SME verified the game as reflecting those strategy games used in command training. The four games played were:

- Static game: The opponent players do not move. All opponent players are shown to the experiment team in their true colours (e.g. red is shown, yellow is shown and green is shown)
- Moving game: The opponent player's move after the experiment team has moved. All opponent players are shown to the experiment team in their true colours (e.g. red is shown, yellow is shown and green is shown)
- Static and disguised game: The opponent players do not move. All opponent players are shown as yellow (e.g. red and green are disguised as yellow) so that the experiment team must reveal what the true colour of the opponent players are (i.e. green, red or yellow).
- Moving and disguised game: The opponent player's move after the experiment team has moved. All opponent players are shown as yellow (e.g. red and green are disguised as yellow) so that the experiment team must reveal what the true colour of the opponent players are (i.e. green, red or yellow).

All teams performed four experimental tasks of a strategy game. The team collaborated to achieve the aim of the game which was to take as many red players as possible whilst at the same time avoiding taking yellow, green or blue (e.g. other team members) players. Collaboration in the teams was ensured through communication.

Data Reduction and Analysis

Communications were explored using content analysis to identify compatible and transactional information elements, using a similar approach to that applied by Stanton et al. (2009a; 2009b) among others (Salmon et al., 2008; 2009a). Using Norman's (1981) and Norman and Shallice's (1986) description of source and component schemata the most prevalent source schemata were identified. Norman and Shallice (1986) defined source schema as a "highest-order control" mechanism which organises a set of learned action sequences. A component schema can therefore be seen as a lower-order schema which achieves some part of the actions which the higher-order, or source, schema initiates.

The communications were further explored for source and component schemata and for schemata which were in conflict with each other as described by contention scheduling. The transactions identified were linked to the component schemata they originated from. The transactions observed were depicted in state-space diagrams. Sanderson, Verhage & Fuld (1989) described the use of state-space diagrams as a means of exploring process control as a dynamic problem solving task. Using state-space diagrams they showed how the operator handled a set of problems and moved from point to point within the state space as they did so. The state-space diagrams were constructed to show how the understanding of team members changed as new information was provided and how it conflicts with existing assumptions, or schemata. These are described as either assimilation or accommodation, where assimilation reflects instances where the incoming information fits with the schema and where accommodation reflects that new schemas had to be developed (Piaget, 1961). In this way compatible and incompatible transactions are shown.

RESULTS

A common source schema, or a 'super-source' schema, observed in all communications was "win the game" from which all other schemata appeared to originate. The component schemata observed from this schema and the manner in which compatible and incompatible transactions between team members developed the schemata are explored here.

Taking a red

Figure 1 illustrates the manner in which teams' compatible transactions were passed around the team with regards to taking red.

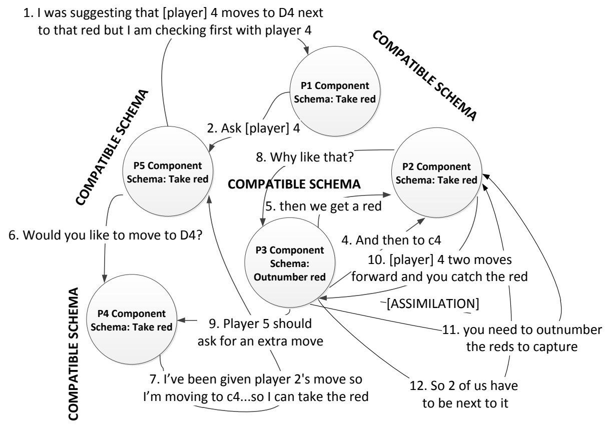

Figure 7. State-space diagram showing compatible transactions associated with the schema "take red"

As seen in Figure 1, Player 5 suggested to Player 1 that Player 4 moves next to a red player (*"I was suggesting that [player] 4 moves to D4 next to that red"*). Player 3 made a suggestion to Player 2 in terms of another move which would take another player next to red (*"And then to c4"*). This was followed by a statement which asserted that doing so would enable the taking of a red (*"then we get a red"*). This prompted Player 2 to ask why this was necessary (*"why like that?"*). Player 3 appeared to have a component schema for taking red which differed slightly from the other team members, namely that in order to take red the red must be outnumbered first, by there being at least two blue players to every red (e.g. the schema "outnumber red"). This was seen in the transaction from Player 3 to Player 2 where this game rule was explained (*"you need to outnumber the reds to capture" and "so 2 of us have to be next to it"*). Player 3's transaction to Player 2 appeared to have triggered the activation of a further component schema, namely "giving away moves" which Player 3 transacted to Player 4 (*"Player 5 should ask for an extra move"*). Whilst this was a different schema to that held by the other team members this was not incompatible and originated from the source schema "take red".

Making Moves

Figure 2 illustrates the state-space diagram for "making moves".

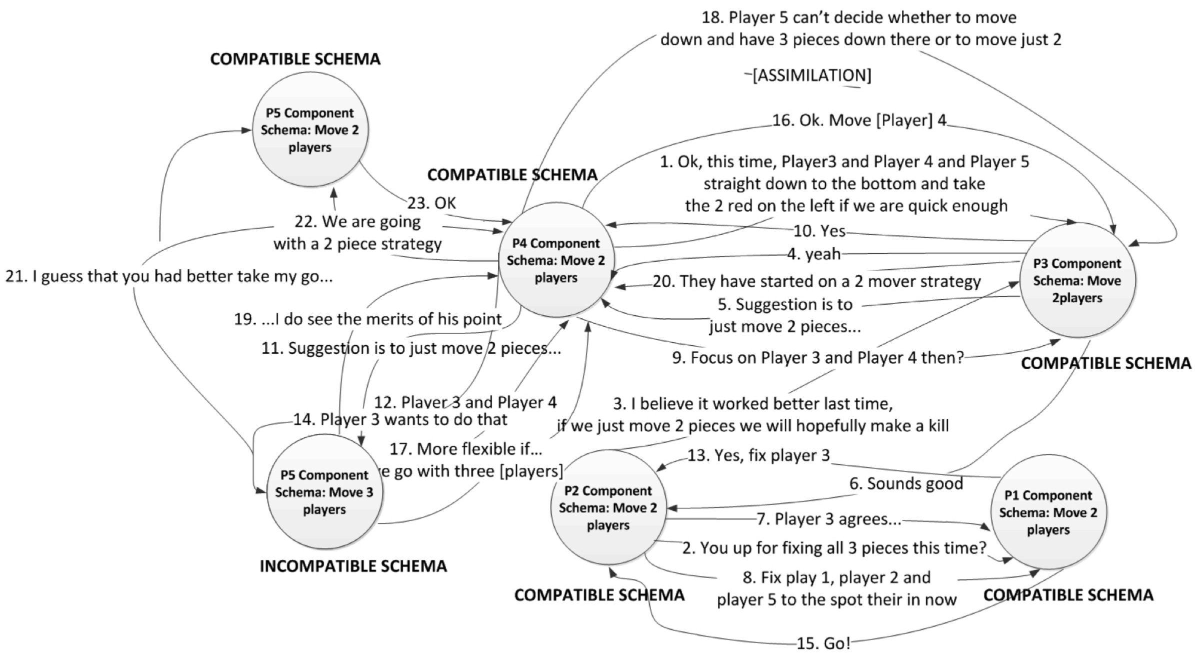

Figure 8. State-space diagram showing adjustment of schemata related to moving players

The component schema of "moves" was observed in the communications, however, contention was observed between the need to move two or more players in order to be effective. Figure 2 shows the state-space diagram developed for the component schema "moves". The schema "moves" has here taken the role of a source schema triggering two different component schema "move two players" and "move three players".

A contention can be seen in the team communications with Player 5 and Player 4 beginning the game with an active schema for "moving 2 players" whilst Player 3, Player 2 and Player 1 have an active schema for "move 2

players". The first transaction, passing between Player 4 and Player 3 (*"Ok, this time, Player 2 and Player 4 and Player 5 straight down to the bottom…"*), appears to arise from Player 4's schema "move 3 players" and was incompatible with Player 3's schema "move 2 players". This resulted in a transaction from Player 3 to Player 4 (*"Suggestion is to just move 2 pieces"*) where Player 3's active schema for moving only two players is conveyed. Through a process of accommodation Player 4 then adapts the original schema for moving three players to two players. Player 5, like Player 4, held a conflicting schema to that of the other team members (*"move three players"*) which is adjusted to "move two players" through the transactions received from Player 4. Player 4 therefore; after having had their schema changed, goes on to initiate accommodation of Player 5's schema. As can be seen in Player 4's transaction to Player 5 where the same message as that Player 4 received from Player 3 was passed on to Player 5 (*"Suggestion is to just move 2 pieces"*). Player 5 argued against the proposed strategy initially (*"more flexible if we go with three [players]"*) but relents and, seen in the reply, *("I do see the merit of his point"*) adjusted his schema to that held by the majority of the team (e.g. "move two player").
Figure 3 shows the compatible transactions being assimilated into the team member's schemata "move towards red". Here, no conflicts are observed and each team member's transactions aligned with the schemata.

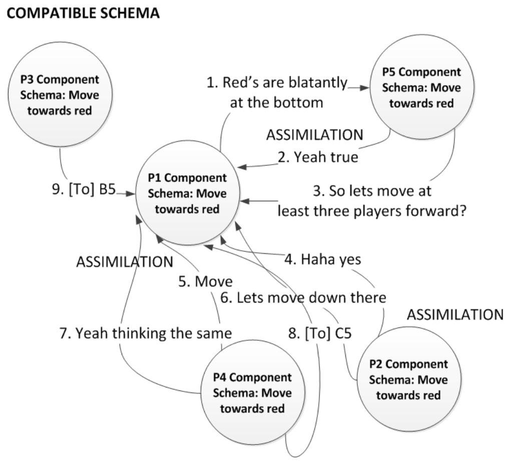

Figure 9. State-space diagram showing assimilation of compatible transactions in relation to moving players

Summary of Results

The transactions explored were associated with a range of schemata, where two were in direct contention with other component schemata, as illustrated in Figure 4 below. In the following section these exploratory findings are discussed.

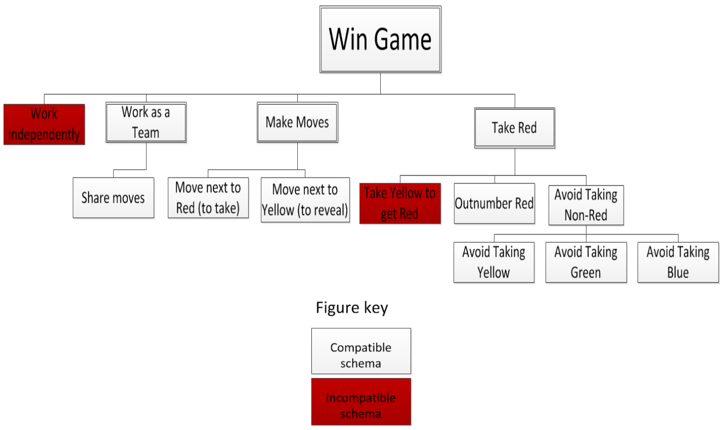

Figure 10. Summary of compatible and incompatible schemata activated from the source schema "win game"

DISCUSSION

Teams are interdependent entities from which Distributed SA emerges through interactions between team members (Stanton et al., 2009a; Stanton et al., 2009b; Salmon et al., 2008). The team's interdependence means that each team member performs separate but related tasks to enable the team to achieve an overall goal. Understanding the role of transactional and compatible SA in holding different parts of a system, or team, together is important to further the theory of Distributed SA. This exploratory research sought to shed light on the manner in which compatible and incompatible transactions support the regulation of team behaviour and the development of Distributed SA.

Three research questions guided the exploratory analyses conducted for this chapter. The first asked whether the teams exhibited use of source and component schema. The findings presented here showed that all teams exhibited the activation of source schemata which in turn triggered the activation of component schemata, as described in the literature (e.g. Norman, 1981; Norman & Shallice, 1986).

The second research question asked whether the teams exhibited contention between schemata, as described by Norman (1981). The findings revealed that whilst the triggering of component schemata was mostly appropriate for the context of the game variant played, the triggering of subsequent schemata clearly made the team members vulnerable to activation of inappropriate schemata. The findings highlighted one example of "wrong triggering of schemata" (Normal, 1981) where the team activated a schema which was inappropriate at that time but which could potentially have been appropriate at another time (e.g. in a different type of game). As was seen in the team which held conflicting schemata concerning team working strategies (e.g. work as a team or work independently). Salmon et al. (2008) argued that deficiencies in one agent's SA can be compensated by another. This was exemplified when a team member who displayed the wrong schema adjusted it via accommodation whereby information that conflicted with the original schema was used to develop a new schema. Similarly, in discussing taking a red player the team members supplemented each other's understanding of the manner in which red was to be taken (for instance, by being outnumbered).

Norman and Shallice (1986) explained that individuals may not possess schemata for novel tasks. In such instances no schema will be available for selection and a new schema must be developed. An agent draws on existing experience and knowledge whilst interacting with the world to form a new schema appropriate for the novel task. This may in turn lead to wrong schemata being developed as the interpretation of the new task may not be entirely fitting. This was exemplified in the team communications where a team member had developed a schema for taking yellow as a means by which red players could be got to. The application of previous experience and schemata, which may not be appropriate, may be a rational means by which the teams instigate behaviours. Whilst the schema may be incorrect for a particular game variant expressing it means it becomes possible for the team to adapt it in light of conflicting transactions made by other team members. This is in line with the explanation offered by Bartlett (1936), and early Schema Theory, that the production of behaviour arises from an organisation of experience which are being drawn on in dealing with a situation (Plant & Stanton, 2012; Plant & Stanton, in press). The sequence of activation of a source schema and associated component schemata that were evident in the team's transactions also showed that the team quickly adapted their behaviour to the context, once it was understood, and this led to a triggering of further schemata and acts relating to those. Stanton et al. (2009b) argued that schemata support individuals to proficiently deal with situations in the production of appropriate responses. This was exemplified in the extract of communications where taking a yellow was discussed. Player 1 had an active schema for taking a yellow and expressed this to Player 3. Player 1's schema was therefore transacted to Player 3 who had the same schema triggered. Player 1 then appeared to have checked the game rules whilst Player 3 checked the board and found neither that taking a yellow was allowed by the game rules nor gave any advantage in terms of movement on the board. The "taking of a yellow" schema was then dismissed and a new schema activated. The players went on to discuss making moves around the yellow.

The third research question asked whether contention scheduling was observed in the team communications. The communication extracts presented here did show a degree of conflict between different team members' opposing schemata. Norman and Shallice (1986) explained that when several schemata are activated at the same time selection between these is required. A conflict resolution procedure must then be provided and it would appear that transactions, in conveying what an agent knows, has a 'conflict scheduling' (Norman & Shallice, 1986) function in the teams. Compatible and incompatible transactions, through a process of assimilation and accommodation (Piaget, 1961), appeared to enhance and develop the schemata of other team members, thereby resolving the contention.

CONCLUSION

Given the exploratory nature of this research limited conclusions can be drawn from this study with respect to contention scheduling in teams. It appears that when a conflict existed between team members (as where a yellow player was considered taken and a team member insisted that taking this course of action would be wrong) a resolution was found. It may be that in teams, like the ones studied here, conflict resolution occurs through the schema of member with high status being given higher 'activation threshold' in the team, resulting in this schema being triggered when in conflict with a "lesser" team members' conflicting schema. Such scenarios are commonly found in military C2 and in hierarchical teams where one leader is in charge. The activation threshold value given to team members schemata could, perhaps, be reduced or increased by aspects such as whether their schemata have been appropriate for other situations before (i.e. dependent on team members experience) and therefore build on trust and cohesion. It is also possible that where a more democratic team structure exists, the schema which is held by most team members will be given the highest activation value and thus is selected for team behaviour. This is supported by the finding that compatible transactions were associated with slightly higher efficacy when contrasted to incompatible transactions. As such, more moves were made with fewer transactions than for the incompatible transactions. The absence of contention scheduling between component schemata held by different team members may explain why fewer transactions were required. In these instances the teams' attention was focused on making the moves rather than establishing the appropriate schema.

ACKNOWLEDGMENTS

This work from the Human Factors Integration Defence Technology Centre was part-funded by the Human Sciences Domain of the UK Ministry of Defence Scientific Research Programme.

REFERENCES

Bartlett, F.C. 1932. *Remembering: A Study of Experimental and Social Psychology*, Cambridge, Cambridge University Press.

Grasser, A.C. & Nakamura, G.V. 1982. The impact of schemata on comprehension and memory. In: BOWER, G.H. (ed.) *The psychology of learning and motivation: advances in research and theory.* New York: Academic Press.

Hutchins, E. 1995a. *Cognition in the Wild*, Cambridge, The MIT Press.

Mandler, J.M. 1984. *Stories Scripts and Scenes: Aspects of Schema Theory*, Hillsdale, NJ, Lawrence Erlbaum Associates.

Marshall, S.P. 1995. *Schemas in problem solving*, Cambridge, Cambridge University Press.

Norman, D.A. 1981. Categorization of action slips. *Psychological Review,* 88, 1-15.

Norman, D.A. and Shallice, T. 1986. Attention to Action. *In:* Davidson, R.J., Schwartz, G.E. and Shapiro, D (eds.) *Consciousness and Self-regulation. Advances in Research and Theory.* New York: Plenum Press, 1-18.

Piaget, J. 1961. The Genetic Approach to the Psychology of Thought. *Journal of Educational Psychology*, 52, 275-281.

Plant, K.L. & Stanton, N.A. 2012. Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science*, 50, 300-315.

Plant, K.L. & Stanton, N.A. in press. The explanatory power of Schema Theory: Theoretical foundations and future applications in Ergonomics. *Ergonomics*.

Salmon, P.M., Stanton, N.A., Walker, G.H., Jenkins, D.P., Baber, C. & McMaster, R. 2008. Representing situation awareness in collaborative systems: a case study in the energy domain. *Ergonomics*, 51, 367-384.

Salmon, P.M., Stanton, N.A., Walker, G.H., Baber, C., Jenkins, D.P., McMaster, R. & Young, M.S. 2009a. What is really going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, 9, 297-323.

Sanderson, P.M., Verhage, A.G. & Fuld, R.B. 1989. State-space and verbal protocol methods for studying the human operator in process control. *Ergonomics*, 32, 1343-1372.

Stanton, N.A., Salmon, P.M., Walker, G.H. & Jenkins, D.P. 2006. Distributed situational awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics*, 49, 1288-1311.

Stanton, N.A., Salmon, P.M., Walker, G.H. & Jenkins, D.P. 2009a. Genotype and phenotype schema and their role in distributed situation awareness in collaborative systems. *Theoretical Issues in Ergonomics Science*, 10, 43-68.

Stanton, N.A., Salmon, P.M., Walker, G.H. & Jenkins, D.P. 2009b. Genotype and phenotype schemata as models of situation awareness in dynamic command and control teams. *International Journal of Industrial Ergonomics*, 39, 480-489.

Doctoral Consortium

Decision inertia: Defining its meaning and dichotomising its context as a latent and emergent variable

Nicola POWER[a] , Laurence ALISON[a] , Jason RALPH[b]
[a]Critical and Major Incident Psychology, University of Liverpool, UK
[b]Electrical Engineering & Electronics, University of Liverpool, UK

ABSTRACT

Introduction: Decision inertia is a failure to execute an important, irrevocable decision resulting in non-optimal consequences. Such decisions are high-stakes, irreversible and with far reaching consequences which may outlast the event; typical of a critical or major incident. This research aims to explore decision inertia in three ways. **Method**: Firstly, a detailed literature review will define *what* decision inertia is and *how* it manifests behaviourally (to inform a theory-driven coding dictionary assisting later analysis). Secondly, as a latent variable, it will be tested by exploring verbal and behavioural manifestations of decision inertia in practice. Thirdly, as an emergent variable, the impact of interoperability and new technologies on its likelihood will be tested. Data will be analysed using mixed methods both quantitatively and qualitatively. **Results:** Findings will extend the theoretical understanding of decision inertia as a concept and provide practical application to assist emergency service critical incident decision making.

KEYWORDS

Decision Making; General and miscellaneous; Decision inertia; Critical incidents; Interoperability; New technology

INTRODUCTION

When individuals make risky decisions involving high stakes and under conditions of uncertainty, the consequences of their choice can have large, irrevocable immediate and long term consequences. When under such uncertain and pressurised conditions, decision makers have been found to act in an avoidant manner, favouring inaction over decision implementation (van den Heuvel, Alison & Crego, 2012). Rather than commit to choice, decision makers will continually delay their engagement with the decision process (Eyre, Alison, McLean & Crego, 2008). This is defined as decision inertia: the topic of this research.

Research into decision making has advanced through disciplines of psychology, human factors and management. Yet the research on inaction or non-decision making has had less attention. According to Anderson (2003), decision avoidance is 'a tendency to avoid making a choice by postponing it or by seeking an easy way out that involves no action or no change' (p. 139). Anderson (2003) provided a useful introduction to this topic, yet failed to explicitly measure how it manifests in context (as a latent variable) or what appears to influence its likelihood (as an emergent variable). The present research will address this gap by: (i) explicitly defining decision inertia; (ii) outlining how it manifests in context; and (iii) exploring the likelihood of decision inertia in relation to the endogenous or exogenous characteristics of the critical incident environment.

Research will be conducted in the context of critical incident decision making with emergency 'blue light' responders (i.e. police, fire and rescue, ambulance). Alison and Crego (2008) defined critical incidents as highly risky and uncertain events with the potential to produce outlasting negative effects on those directly involved and the wider community. As such, naturalistic decision making paradigms have been a useful tool for conducting such research. Moreover, naturalistic research methodologies have produced mutual benefits between practitioners and academics in utilising immersive simulated learning environments to assist applicable training whilst generating ecologically valid psychological results (Alison, van den Heuvel, Waring, et al., in press). With this in mind, critical incident decision making was chosen as a fruitful research environment for studying decision inertia as a general concept, and also as conclusions may be usefully applied to assist policy and training for emergency service decision making.

METHODOLOGICAL APPROACH

There are three distinct themes to this thesis plan. Research will be mixed methods to generate empirically valid conclusions within an ecologically valid setting by utilising both a priori defined quantitative methods with post-hoc qualitative analysis of verbal and behavioural indicators.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

RESEARCH QUESTION ONE: WHAT IS DECISION INERTIA?

Research will firstly define decision inertia as a distinct concept. Although it may be refined in response to experimental results, the first step to defining decision inertia will be via a detailed and exhaustive review of current literature on this topic - a recommended process for PhD research (Dunne, 2011). A systematic literature review will be conducted to provide a thorough and transparent basis for theory building (Rhoades, 2011). By following guidelines on how to conduct an objective systematic review (Neely, Magit, Rich, et al., 2010), a search protocol will be developed consisting of keywords, inclusion and exclusion criteria. The chosen search strategy is electronic search engines. In addition 'grey literature' relating to practitioner guidelines will be included to see how decision inertia is currently addressed in practice. Papers will then be identified, screened and selected before evaluating and extracting key information and synthesising findings. Results will assist in the preliminary generation of a theoretical, deductive coding dictionary of decision inertia as a latent theme.

RESEARCH QUESTION TWO: HOW DOES DECISION INERTIA MANIFEST IN CONTEXT?

The second research question will address *how* decision inertia manifests in the real-world via observational data from a simulated multi-agency training event and semi-structured interviewing.

Participants:

Participants will be strategic decision makers attending training for a multi-agency CBRN (chemical, biological, radiological, nuclear) event. Participants will be in the position of Gold command and involved in the Strategic Coordinating Group (SCG) meeting where strategic decisions are made.

Procedure:

Trainees from multiple agencies, including police, fire and rescue and ambulance, will respond to a live simulated CBRN incident as they would in the real-world. Measures of decision making are outlined below.

Measures:

Decision logs:

Decision logs will be utilised to identify the key decisions made during the incident. Key decisions will be verified in post-event interviews.

Videoed SCG meetings:

SCG meetings will be recorded and transcribed. Key decisions (generated from decision logs) will be used to guide the selection of discussion frames about these key decisions. A thematic analysis (Braun & Clarke, 2006) will be conducted to identify what key themes appear to be representative of decision inertia. This process will involve both 'bottom up' and 'top down' processing. Top down processing will be based upon the theoretical coding dictionary generated from the literature review to guide initial coding of data. In addition, the data will also be explored 'bottom up' for indicators of decision inertia not identified in the a priori coding dictionary. A full coding dictionary will be synthesised and data set returned to for full coding of identified themes.

Post-task semi-structured interview:

Semi-structured interviews with Gold commanders involved in SCG meetings will also be conducted. Interviews will be supplemented with video footage of SCG meetings to aid recall. Commanders will be asked about their perceived key decisions during the event and further asked about their experience of decision inertia. This is to further unpack decision inertia as a latent concept to inform its semantic definition. Interviews will be transcribed and thematically analysed following the procedure above.

RESEARCH QUESTION THREE: WHAT FACTORS DURING CRITICAL INCIDENT RESPONSE CAN PREDICT THE LIKELIHOOD OF DECISION INERTIA?

Research will also explore *why* decision inertia may occur as an emergent variable by analysing its relationship with predictor variables. This research will be conducted in a more experimental and controlled manner during two simulation studies.

Study one: Does the presence of multiple agencies during interoperable emergency response increase the likelihood of decision inertia?

Participants:

Participants (n=60) will be trainees from the Fire and Rescue service who will complete a laptop based simulation scenario. They will range in rank from Crew up to Area managers and randomly split into one of two experimental groups. Both groups will complete the same simulation however, crucially, group A will be given additional information informing them that other agencies (police and ambulance) are also in attendance whereas group B will be told that the fire service are the only agency at the scene.

Procedure:

All participants will be told that they are Incident Commander (IC) and will individually complete a computer based 30 minute scenario though PowerPoint using headphones and a keyboard. The scenario involves an incident at a transport tunnel which initially appears to be a road traffic collision but escalates to become a potential terrorist attack. A key strategic decision will be the declaration of a major incident. The scenario will be developed in accordance with a UK based Fire Service Training and Development Academy and subject matter experts used to validate realism and provide a 'gold standard' response. The only difference between groups during the scenario is the initial framing of whether other agencies are in attendance at the scene.

The scenario will consist of six auditory information feeds. Following each information feed the scenario will freeze and they will answer questions on: situation assessment; risk perception; information requests; actions taken. They will also score measures of their uncertainty using a Likert scale. Following the scenario, they will answer a questionnaire about whether they declared a major incident and their reasons for this. The questionnaire will also identify, via a series of set questions, whether they experienced decision inertia and if so in what form.

Measures (Independent variables):

- Individual differences prior to the experiment will be taken by filling in a questionnaire to measure:
    - Work based locus of control
    - Trust in teams (both intra- and inter- agency based trust)
    - Organizational stressors (perceived work-based constraints; coping tendencies; psychological work-related strain)
    - Experience
- Experimental manipulation (multi-agency environment)

Measures (Dependent Variables – by comparison to 'Gold Standard'):

- Situation Assessment
- Risk perception
- Information requests
- Actions taken (including 'key' declaration of major incident)

Main Hypothesis:

H1: Decision inertia (to declare a major incident) will increase when other agencies are at the scene. This effect will be greatest in (and moderated by) those with:

- External locus of control
- Low inter-agency trust
- High organizational stressors
- Less experience

Future application:

It is hoped that this study can be extended to both ambulance and police to assess the multi-agency effects on decision inertia.

Study two: Can the use of autonomous systems in critical incident response reduce decision inertia?

Participants:

Participants from a UK Fire and Rescue Service will test the effects of a prototype autonomous quadrotor being designed in conjunction with the Engineering Department. They will be split into two groups: one with use of the quadrotor and one control group.

Procedure:

Participants will respond to a live simulated event. Their decisions and responses will be logged by an observer during the task. Participants will later be asked to attend a short semi-structured interview to measure their recalled situation assessment following the event.

Measures (Situation assessment – recall and recognition task):

Prior to being interviewed, participants will be asked to fill in a short questionnaire to measure their recognition. They will be asked closed questions about the event and must select the appropriate options from a given list. During their interivew, they will be asked to recall the event as they remember it. The observer logs will be used to measure accuracy.

Measures (Decision inertia):

During the semi-structured interview, they will also be asked about their decision inertia experienced during the task. Responses will be transcribed, coded and results compared between groups.

IMPLICATIONS

This research will contribute to the theoretical understanding of decision inertia. Although findings are generated from critical incident decision making, it is hoped that the coding dictionary can be extended to future testing in other domains. Research will also provide practical recommendations for assisting emergency service crisis response by assessing how decision inertia can be affected by interoperabiltiy and new technologies. The coding dictionary of decision inertia indicators will also be useful for assessing training.

REFERENCES

Alison, L. and Crego, J. (2008). Policing Critical Incidents: leadership and critical incident management. Devon, England: Willan Publishing.

Alison, L., van den Heuvel, C., Waring, S., Power, N., Long, A., O'Hara, T., & Crego, J. (in press). Immersive simulated learning environments for researching critical incidents: a knowledge synthesis of the literature and experiences of studying high-risk strategic decision making. Journal of Cognitive Engineering and Decision Making.

Anderson, C.J. (2003). The psychology of doing nothing: forms of decision avoidance result from reason and emotion. Psychological Bulletin, 129(1), 139-167.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.

Dunne, C. (2011). The place of the literature review in grounded theory research. International Journal of Sociological Research Methodology, 14(2), 111-124.

Eyre, M., Alison, L., McLean, C., & Crego, J. (2008). Decision Inertia: the Impact of Organisations on Critical Incident Decision Making. In L. Alison & J. Crego (eds). Policing Critical Incidents: Leadership and Critical Incident Management. Devon, England: Willan Publishing.

Neely, G.J., Magit, A.E., Rich, J.T., Voelker, C.C.J., Wang, E.W., Paniello, R.C., Nussenbaum, B., & Bradley, J.P. (2010). A practical guide to understanding systematic reviews and meta-analyses. Otolaryngology – Head and Neck Surgery, 142(6), 6-14.

Rhoades, E.A. (2011). Commentary: Literature Reviews. The Vola Review, 111(3), 353-368.

van den Heuvel, C., Alison, L., & Crego, J. (2012). How uncertainty and accountability can derail strategic 'save life' decision in counter-terrorism simulations: a descriptive model of choice deferral and omission bias. Journal of Behavioral Decision Making, 25(2), 165-187.

Managing Disruptions in Complex Adaptive Systems A Case for Resilience in Airline Operations Control

Kenneth E. IGBO[a]
[a]Swinburne University of Technology, Melbourne Australia

ABSTRACT

Introduction: Studies in naturalistic decision-making and resilience engineering have made significant advances towards characterising emergent cognitive behaviours in highly dynamic and distributed work settings. This doctoral project draws on such advances to examine disruption-management processes that involve multiple players spanning multiple functions, expertises and differing levels of authority in the way each player influence the overall outcome. **Method**: This research integrates ethnographic and grounded theory methods in examining distributed decision-making, negotiations and coordination processes in airline operations control. **Results and discussion:** Preliminary findings suggest that complex networks of interdependencies—relating to resources and performance variables—across units involved in airline operations control promote reciprocity and the adoption of cooperative strategies. The findings further suggest that such emergent cognitive behaviours must be examined in context of the specific work setting studied. Such insights provide a framework for delineating how decision-makers collectively adapt formalised procedures *in-flight* while managing many-to-many mappings across conflicting goals.

KEYWORDS

Distributed decision-making; resilience; reciprocity; common ground; sense-making; airline operations control.

INTRODUCTION

Airline operations control, given its highly dynamic and distributed nature, presents a fitting 'natural laboratory' to re-examine propositions developed while studying cognitive behaviours in other complex and highly adaptive systems. Control activities in airline tactical operations are often characterised by goals that are dynamically changing or locally adapted across multiple centres of control. Regulating interactions across these centres of control presents interesting challenges because each centre possesses 'partial authority, partial autonomy and partial responsibility' (Ostrom, 1990) in relation to the extent they can adapt overall operational goals and activities. Complex interdependencies between key resources controlled by the different centres further exacerbate the challenge to adapt planned operations, particularly in the event of unforeseen perturbations (see Abdelghany, Abdelghany, & Ekollu, 2008; Clausen, Larsen, Larsen & Rezanova, 2010; Jarrah, Yu, Krishnamurthy & Rakshit, 1993; Teodorović & Guberinić, 1984; Yan & Yang, 1996).

Clearly, managing such interdependencies, often under severe economic pressures and time restrictions, necessitates that decision-making protocols reflect the intrinsic complexities of interactions across multiple centres of control. This premise is typically exemplified in most accounts of airline operational recovery in the wake of a disruptive event (Yu & Qi, 2004). The following section briefly reviews theoretical perspectives that underpin this research. Thereafter, a brief description of the methodological approach, including the unit of analysis, the research carried out so far and future fieldwork plans, is presented. Preliminary findings and expected theoretical contributions are documented afterwards. In the concluding section, this paper argues for the originality of its contributions.

THEORETICAL PERSPECTIVES THAT UNDERPIN THIS RESEARCH

This doctoral research integrates notions, metaphors and theories in resilience engineering and naturalistic decision-making research to develop a theoretical framework with which to investigate disruption management in complex adaptive systems. From a naturalistic decision-making perspective, this study examines the research on distributed decision-making to ascertain how shared understanding and common grounds are constructed *in-flight* in rapidly changing environments. Notions that emerged from this stream of research provide one part of the theoretical foundation on which this doctoral research is grounded. These include joint predictability, sense-making, team situation awareness, and other aspects of macrocognition (see Klein, Orasanu, Calderwood & Zsambok, 1993; Montgomery, Lipshitz & Brehmer, 2005, Mosier & Fischer, 2011; Rasmussen, Brehmer & Leplat, 1999; Schraagen, Militello, Ormerod & Lipshitz, 2008). On the other hand, this research also draws from

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

insights developed within resilience engineering perspective that revolve around basic tradeoffs in human adaptive systems. Of particular interest is the advancement of a proposition to consolidate multiple tradeoffs observed in human adaptive systems into a few fundamental ones (e.g., Hoffman and Woods, 2011; Woods and Branlat, 2011). Along this frontier, scholars have begun to explore a set of critical properties of human adaptive systems, as well as fundamental architectural principles, in regards to managing tradeoffs across multiple centres of control (see Woods and Branlat, 2011).

Drawing on these insights, this research advances the legitimization of a unified conceptual framework of tradeoffs. It examines disruption-management processes in airline operations control through the lens of the unifying framework proposed by Hoffman and Woods (2011), and further discussed in Woods and Branlat (2011). The notion of disruption is used here to represent varying degrees of perturbations to a complex system processes and operations (see Westrum, 2006), as opposed to interruption of an individual's cognitive task processes that has become a major research focus in the fields of human-computer interaction and cognitive psychology (e.g., Latorella, 1998 ; Suchman, 1987, Wickens, 2002). The intention of study is to delineate a set of contextual variables that allow controllers to regulate interactions across multiple centres of control. In particular, this research uses its investigation of decision-making processes in airline operations control to shed light on the roles that pertinent informational and organizational dynamics play—including architectural principles, communication systems, work culture and the associated social organizational factors—in managing tradeoffs and conflicting goals across multiple centres of control. Further, it draws on interesting phenomena observed in various airline operations control centres (OCCs) to speculate on the underlying mechanisms that influence the construction of common ground, sense-making, reciprocity, and the choice of recovery strategies in most highly adaptive and distributed work settings.

METHODOLOGICAL APPROACH

This research integrates a variety of theoretical orientations and qualitative research traditions. Whilst the methodological approach adopted here draws heavily from both ethnography and the grounded theory perspectives, it also integrates insights from deductive qualitative approaches (Atkinson, Coffey, Delamont, Lofland & Lofland, 2001; Charmaz, 2006; Denzin & Lincoln, 2011; Glaser & Strauss, 1967; Strauss & Corbin, 1990). The mix in methodological and theoretical perspectives underscores the philosophy that investigations into contextual factors in complex adaptive systems stand to gain from the strengths of a range of methodological choices available, while minimising their latent limitations. The fieldwork was designed to be undertaken in four stages, each successive stage builds on the findings of the earlier stages (Charmaz, 2006; Denzin & Lincoln, 2011). Initially, the study aim was broadly defined in line with the grounded theory approach (Glaser & Strauss, 1967). This approach allowed the researcher to progressively narrow the study focus around key issues and lines of inquiry as these were identified in the field (Straus & Corbin, 1990). Participant observation, audio-recorded interviews and documentary analysis were used for data collection purposes.

Unit of analysis

The unit under analysis is the *Airline Operations Control* system spanning both international and domestic operations. This system encompasses all processes and interactions geared towards the execution of planned schedule of services, as well as all schedule recovery processes in the wake of a disruptive event. The chief actors include the operations controllers (*ops controllers*), crewing operations, flight dispatch, maintenance operations, customer services and shift managers. However, *ops controllers* and shift managers occupy the central position in this research. Hence, this research focuses mainly on their interactions with other actors within and across the system boundary during negotiations, decision-making and coordination processes. Whilst air traffic controllers (ATCs), airport authorities and station operations (ramp, gates, passenger and baggage handling, etc.) are only considered as agents within the broader air transportation environment, the negotiation processes between these agents and the OCCs are considered part of the operational system under analysis.

Participant Characteristics

Ethnographic study was undertaken in four airline OCCs (see Atkinson, Coffey, Delamont, Lofland & Lofland, 2001). Two of the airlines operate large, multi-hub networks with extensive international connections and large regional feeder services. The third airline operates medium-sized international networks with extensive domestic and regional feeder services for its affiliated airlines in other geographic locations. The fourth airline operates extensive domestic and regional networks and code-shares some regional feeder services with its parent company. All four airlines are based within Asia-Pacific and the Middle East regions.

Field research

The first-stage visits to three OCCs served as a preliminary study to allow the researcher get acquainted with disruption-management practices as they are currently adopted at the sites visited. The visits lasted approximately three hours for each site. During these visits, the researcher adopted the role of a mentee and sat with *ops controllers* to learn how they interact with other functional units involved in the operations control processes. During periods of minor perturbations, *ops controllers* and/or shift managers explained certain domain specific terms and intimate the researcher on the state-of-affairs, their recovery options and rationale behind their choices. Evidence gathered was initially documented as field memos. The memos were discussed with randomly selected participants to ensure that the documented observations accurately reflect meanings that are held by participants within this work setting. The revised memos were later used in conjunction with deductively derived themes from the literature to design a structure for the next stage of study. Further, the researcher identified critical operational points of interest and formal interview sessions were arranged with selected participants in charge of some roles of interest to this research.

The second-stage visits were undertaken in only one OCC and lasted approximately 18-hours over three weeks. Here, the researcher explored, in greater depth, the themes put together from field memos and literature reviews in several 45-mins interview sessions. In addition, a decision was made to extend the discussions beyond phenomena recorded in the memos to accommodate 'memorable' disruptive events that are particularly significant to the participants within six months prior to the interview dates. This decision is imperative given that it is not guaranteed that the researcher could have observed all requisite phenomena due to access and time limitations. Reflecting on past events provided opportunity for participants to discuss events that pushed their operations outside the planned envelope, but which were not observed directly. Hence, participants were allowed to talk freely about scenarios they had witnessed and their experiences, but were brought at some point to address selected themes of interest. Data collected was used to develop a more structured framework with which to compare practices and processes in that particular OCC against the other sites participating in this study.

The third-stage visits will be undertaken in the other three airlines, and will involve mainly participant observation and retrospective protocols. The purpose here is to determine how the phenomena observed so far is reflected across the four participating sites. Fourth-stage visits, if necessary, will entail clarifications and further validations, and will involve selected participants across all four airlines. Overall, ideas generated from these fieldworks and the ones deduced from literature reviews will be integrated to form a 'mid-range' theory (Glaser & Strauss, 1967; Charmaz, 2011) of disruption management in airline operations control. This could be extended to the broader human adaptive systems to validate the framework developed in this project against findings in other complex adaptive systems.

PRELIMINARY FINDINGS

So far, evidence from one airline OCC suggests a reciprocal relationship across centres involved in airline operations control. This phenomenon is linked to the tight-coupling of operational activities across these centres. The participants appear to share a common belief that it is hard to extricate the performance of one centre from the global system's performance. This belief in itself, coupled with a complex network of resource interdependencies, promotes reciprocity and the adoption of cooperative strategies.

On the definition of resilience, participants unanimously contend that resilience depends on what a company wishes to achieve as their operational goals. Using complex mappings between location, size of operation, company policies, political and cultural values, etc., participants argue that various airlines have very divergent perspectives in relation to what resilience means to their operations. In particular, a senior network-control manager suggests that to understand resilience in context of a specific airline's operations, one must 'dig deep' to unveil not only the economic drivers, but also cultural, political and social driving forces behind key decision considerations or criteria deployed in that company. Of course, it is broadly acknowledged in the literature that such factors do influence the adoption of either cooperative, autonomous or defensive strategy. However, it is not yet clear how precisely they influence these strategies in airline operations control. Perhaps subsequent stages in this study will elucidate on this phenomenon.

Other interesting lines of findings include the effect of co-location, the driving forces behind the adoption of'command-and-control' in airline OCCs, the necessity to validate information emanating from multiple centres, and how loosely *ops controllers* and shift managers hold onto their 'illusion of control'. Regrettably, due to space limitations, these findings would not be documented here but will be presented in the consortium.

EXPECTED THEORETICAL CONTRIBUTIONS

It is expected that the final thesis findings will produce a framework for comparing contextual mechanisms in complex adaptive systems, including bush-fire fighting, large-scale commercial laundry services and social-services systems. Comparing driving forces and contextual variables across these complex work domains has the potential to delineate variables that are shared across the broader human adaptive systems and those that are context specific. Furthermore, it is expected that this research will contribute towards the formalization of a unified framework of tradeoffs by characterising more definitively how the phenomena observed in airline operations control relate to previous studies of other complex adaptive systems. Such characterisations could lead to the development of more comprehensive understanding of the notions studied as they relate to complex adaptive systems. Lastly, this paper contributes to the validation of theories developed in other domains by testing them empirically against the specifics of airline operations control.

ORIGINALITY OF CONTRIBUTIONS

With regards to originality, this research integrates advances in two theoretical perspectives: resilience engineering and naturalistic decision making, in explaining the driving forces behind distributed decision making, reciprocity and the construction of common ground. Specifically, our study highlights that autonomous and protective modes of adaptation are particularly significant in airline operations control. This phenomenon is indicative of the necessity to implement a course of action under severe time constraints, while managing complex network of interdependencies relating to resources and performance variables controlled by different units. As strange as it sounds, there is a compelling evidence to suggest that autonomous and protective strategies are especially useful in speeding up decision making processes in highly dependent but distributed systems. Furthermore, the findings underscore the need to maintain an autonomous regulation strategy between safety-critical units and other units in the broader system. Given that safety is important in aviation, this move will ensure an internal regulation mechanism that monitors the position of a system in relation to its boundary of acceptable performance. Overall, we surmise that different modes of adaptation are best suited for managing different forms of trade-offs under different circumstances. The challenge for both researchers and practitioners, therefore, is to ascertain what contextual mechanisms that support the effectiveness of specific modes of adaptation when managing specific trade-offs in varied contexts.

REFERENCES

Abdelghany, K.F., Abdelghany, A.F. & Ekollu, G. (2008) An integrated decision support tool for airlines schedule recovery during irregular operations, *European Journal of Operations Research*, 185, 825–848.

Atkinson, P., Coffey, A., Delamont, S., Lofland, J. & Lofland, L. (2001) Handbook of Ethnography, Sage Publications, London.

Charmaz, K. (2006) Constructing grounded theory: A practical guide through qualitative analysis, Sage Publications, London.

Clausen, J., Larsen, A., Larsen, J. & Rezanova, N.J. (2010) Disruption management in the airline industry – concepts, models and methods, *Computers and Operations Research*, 37, 809–821.

Denzin, N. K. & Lincoln, Y. S. (2011) The Sage handbook of qualitative research, Sage Publications, London.

Glaser, B. G. & Strauss, A. L. (1967) The discovery of grounded theory: strategies for qualitative research, Aldine, Chicago, IL.

Hoffman, R. R.& Woods, D. D. (2011) Simon's Slice: Five fundamental tradeoffs that bound the performance of human work systems, Tenth International Conference on Naturalistic Decision Making, Orlando, FL.

Jarrah, A. I. Z., Yu, G., Krishnamurthy, N. & Rakshit, A. (1993) A decision support framework for airline flight cancellations and delays, *Transportation Science*, 27, 266–280.

Klein, G.A., Orasanu, J., Calderwood, R. & Zsambok, C.E. (1993) Decision making in action: models and methods,Ablex Publishing Corporation, Norwood, NJ.

Latorella, K. A. (1998) Effects of modality on interrupted flight deck performance : Implications for data link, *HFES*, 87-91.

Montgomery, H., Lipshitz, R.& Brehmer, B. (2005) How professionals make decision, Lawrence Erlbaum Associates, Mahwah, NJ.

Mosier, K. L.& Fischer, U. M. (2011) Informed by Knowledge : Expert performance in complex situations, Taylor and Francis, New York, NY.

Ostrom, E. (1990) Governing the commons: The evolution of institutions for collective action, Cambridge University Press, New York, NY.

Rasmussen, J., Brehmer, B. & Leplat, J. (1999) Distributed decision making: Cognitive models for cooperative work, John Wiley & Sons, West Sussex.

Schraagen, J. M., Militello, L. G., Ormerod, T. & Lipshitz, R. (2008) Naturalistic Decision making and macrocognition, Ashgate Publishing Limited, Hampshire.

Sheffi, Y. (2005) The Resilient Enterprise: Overcoming vulnerability for competitive advantage, MIT Press, Cambridge, MA.

Straus, A. L. & Corbin, J. M. (1990) Basics of qualitative research: Grounded theory procedures and techniques, Sage Publications, London.

Suchman, L. (1987) Plans and situated actions : The problem of human-machine communication, Cambridge University Press, New York, NY.

Teodorović, D. & Guberinić, S. (1984) Optimal dispatching strategy on an airline network after a schedule perturbation, *European Journal of Operations Research*, 15, 178–182.

Westrum, R. (2006) A typology of resilience situations, In E. Hollnagel, D. D. Woods & N. Levenson (eds), Resilience engineering : Concepts and precepts, Ashgate, Farnharm, UK, pp. 55-65.

Wickens, C. D. (2002) Multiple resources and performance prediction, *Theoretical Issues in Ergonomic Science,* 3(2), 159-177.

Woods, D. D. & Branlat, M. (2011) How human adaptive systems balance fundamental tradeoffs: Implications for polycentric governance architectures, Fourth International Symposium on Resilience Engineering, Sophia Antipolis, France.

Yan, S. & Yang, D. (1996) A decision support framework for handling schedule perturbation,*Transportation Research*, 30, 405–419.

Yu, G. & Qi, X., (2004) Disruption management: frameworks, models and applications, World Scientific Publishing, Singapore.

Investigate naturalistic decision-making of a workgroup in dynamic situation. From the modelling to the design of a training virtual environment

Yohann CARDIN[a,c,d], Cyril BOSSARD[a,c] and Cédric BUCHE[b,c]
[a]Centre de Recherche sur l'Éducation et la Didactique (EA 3875), Université Européenne de Bretagne, U.F.R Sport et EP de Brest
[b]Laboratoire en Sciences et Techniques de l'Information, de la Communication et de la Connaissance (CNRS-UMR 3192), École Nationale d'Ingénieurs de Brest
[c]Centre Européen de Réalité Virtuelle
[d]STDI

ABSTRACT

This thesis aims to rely on a work of activity analysis to develop a virtual training platform for firefighters (SécuRéVi). The use of this type of simulation is more and more common in the field of training, but often suffers from a lack of credibility in terms of learning content and method. To solve this problem, this project aims to model the collaborative work of firefighters during training sessions in order to provide assistance to the development of SécuRéVi. The activity analysis of these group works, relying on the EAST (Event Analysis of System Teamwork) methodology and self-confrontation interviews, is expected to highlight the particular "know-how" and to develop pedagogical scenarios essential in the design of such a training platform.

KEYWORDS

Decision making; collaborative work; firefighter; virtual environment.

INTRODUCTION

This PhD work is at the intersection of the fields of ergonomics, occupational didactics and virtual reality. Its originality lies in the facts it etablishes a link between the field of activity analysis and the computerized design of a virtual environment for training.

All this work is part of a larger project between the European Center of Virtual Reality (ECVR) and STDI company, specialized in the production of books in any field and being open for the past few years to new learning methods based on multimedia technologies. It aims to formalize a generic methodology and operational design of virtual environment for human activity (VEHA) using the Mascaret platform while pursuing three main objectives: i) placing the training specialists (trainers, educationalists, teachers, experts) at the center of the design of the VEHA, ii) developing authoring tools allowing these professionals to bring their expertise without the intervention of a computer engineer iii) rationalizing the design process of the VEHA in order to reduce costs and capitalize on the developments of a project for another.

VIRTUAL ENVIRONMENT FOR HUMAN ACTIVITY (VEHA)

Simulation in virtual environment is emerging as a major and innovative technology in the field of training. However, these systems are considered complex because they involve a variety of components, structures and a diversity of interactions (Tisseau, 2001).

Among these systems, there are particularly VEHA which are defined as environments that interact human and artificial agents, and providing access to training resources in order to promote the construction of knowledge by the learner. This type of environment often relies on the implementation of pedagogical scenarios whose role is to provide instructions to individuals to ensure the smooth progress of the learning session. An example using this type of tool is a training virtual simulator of F-16 airplane engine maintenance (Fonseca, Paredes, Rafael Morgado & Martins, 2011). This environment has the characteristic to train team coordination through the use of sub-scripts (corresponding to each member's tasks) and a management component of all interpersonal relationships.

In the field of firefighters, a multitude of training simulators have been developed, and this in different topics. For example, Omodei, Elliott and Walshe (2004) conducted a critical review of literature on wildland firefighting simulators. Thus, they presented seven softwares for which they made a brief description and drew

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

up the advantages and disadvantages of each one. Nevertheless, this review was more particularly based on the range of scenarios they are able to generate and all the control possibilities that they offer to the trainer.
SécuRéVi (Querrec, Buche, Maffre & Chevaillier, 2003) is another example of virtual platform using pedagogical scenarios, this in order to train firefighters to the procedures and operational management against fires with chemical risks. A distinctive feature of this simulator, developed in the ECVR and which the results of this thesis are intended for, is to rely on a generic "scenario creation" tool, independent of the type of task or the learning area (Marion, Querrec & Chevaillier, 2009).
Beyond any aesthetic aspect, the primary objective of such an environment is learning and knowledge acquisition by the learner. From this perspective then, Hall (2011) has attempted to measure the effects of computer-based simulation training on fire ground incident commander (IC) decision-making. The purpose of this study was to establish a correlation between the training program on the simulator called FLAME-SIM and increases in the decision-making efficiency of fire ground ICs. However, the improvement in the participants decision-making being only assessed on the simulator and finally, the interest of learning transfer from VEHA to the real field was missing from this study.
Another important point to consider in the perspective of a learning is that a simulator must have adequate levels of fidelity to both physically and psychologically (Omodei et al., 2004). Physical fidelity thus refers « to the extent to which the simulation is successful in reproducing the physical aspects of the environment » like the sounds, the fires' behaviours or the equipment available. For its part, the psychological fidelity refers to « the degree to which the simulation captures the functional and cognitive aspects of the performance domain » (Entin, Elliott & Schiflett, 2001). More precisely, a virtual environment in order to be credible have to place the same cognitive demands on participants' decision-making and thought processes as occurs in real situation. Thus, the scenarios generated by the program must provide sufficient physical fidelity to allow the required high-level psychological fidelity.
In this way, in this type of learning platform, collaboration between humans and/or virtual entities can be undermined when the task of modelling the collaborative activity is left to the sole discretion of the computer engineer (Tchounikine, Baker, Balacheff, Baron, Derycke, Guin, Nicaud & Rabardel, 2004). To resolve this problem, this thesis has the ambition to try to analyze, describe and modelize human collaborative activity in natural situations with the purpose of providing assistance to the design of training devices using new technologies.

DECISION-MAKING ANALYSIS IN THE FIELD OF FIREFIGHTERS

For over twenty years, models and concepts from the approach of Natural Decision-Making gained some influence as a method to explain the decision-making in many complex situations. NDM approach has accumulated observations in most of the major risk situations (driving, military, nuclear, medical emergencies, etc.)..
Thus, in 1986, Klein & Brezovic have developed the Recognition-Primed Decision (RPD) model in order to modelize the decisions of firefighter's officers in the management of serious incidents in a control center. From interviews with 26 experienced firefighter's officers, the authors analyzed the decisions made during 32 critical interventions by collecting verbal and behavioural data. This data analysis then showed that 80% of the decisions made by the officers were based on a process of recognition of typical situations. Expert officers then did not compare several options but perceived situations encountered as "typical cases" to which they associated certain types of actions (or sequence) appropriate and commonly used with success. In addition, they highlighted the fact that the first option considered by these experts often intuitively was usually satisfactory.
Other authors adhering to the NDM approach followed in the study of decision-making of firefighters. Thus McLennan and Omodei made various studies in this area, leading them to put forward some reflections on the use of the RPD model for unusual and complex events (McLennan, Pavlou & Omodei Klein, 2005). By the setting up of interviews of 40 experienced controllers, they also highlighted the different processes of anticipation used to better manage wildfires: "recognition", "extrapolation" and "construction". Processes that can be directly linked to the three forms of anticipation issued by Klein, Snowden & Pin (2007): "pattern matching", "trajectory tracking" and "convergence". In addition to these results, their studies have implemented many methodological innovations such as the introduction of cameras in the firefighter helmets to improve self-confrontation interviews (McLennan et al., 2005) or the creation of a Human Factors Interview Protocol (HFIP) guide (McLennan & Omodei Reynolds, 2005) to facilitate the detailed recall of decision-making process in fire management.
For his part, in a more theoretical approach, Rahman (2009) took an interest in the effect of emotion in decision-making under life threatening conditions thus highlighting the concept of emotion-primed NDM. This theoretical concept enabling to study the NDM under threat is based on "constructive" and "destructive" modulations to explain the enhancement or reduction of the probability of making a decision leading to a satisfying outcome under these conditions. In this way, he provided a framework for designers and trainers to better take into account the effect of emotion in the design of different training systems for decision-making.
Finally, in order to synthesize and bind all this findings, Pardue (2009) wrote a review of literature on risk behaviour and decision-making among firefighters with the aim to highlight all the possible ways of

improvement in pedagogical methods for initial training and continuing education in professional development. After identifying all external and internal factors that influence the decisions firefighters make on the ground, he insists on developing more exercises in unfamiliar situations of danger (most strongly associated with firefighter injuries and fatalities on the ground) setting up through pedagogical scenarios and enabling the protagonists' ratings in the various areas of learning. This type of training enabling to distinguish the combinations of fire ground factors that are only slightly dangerous from those considered deadly contrary to usual exercises in which the emphasis is on strategy and tactics.

However, all these studies focused more particularly on the decision-making of some individuals and how to train them. They did not really take into account the collaborative aspect of the firefighter job and the influence that can have on the decision-making of each one in dynamic situation. This problematic point is thus underlined by Fern, Trent and Voshell (2008) who base their opinion on a critique of the RPD model. For them, the latter only concerns the activity of a single person, not a whole team of which the dynamics should be taken into account. In addition, this model provides only how firefighters make their decisions but finally does not describe the specific cognitive work and the decisions required in the domain that can help for the design of training systems. To do this, they built their study on the Cognitive Task Analysis (CTA) which is not a single method of analysis but a set of techniques and methods classified as « formal analysis », « empirical techniques » and « computer models ». In this way, the authors rely especially on a functional goal decomposition of urban firefighting outlining the functions, decisions and information requirements of the firefighters on the ground. This led them to consider the difficulties of urban firefighter decision-making and the importance of improving communication systems within intervention teams.

In this purpose, a recent metholodogy called EAST (Event Analysis for Systemic Teamwork) emerged to study the group activity (Walker, Stanton, Stewart, Jenkins, Wells, Salmon & Baber, 2009).

EVENT ANALYSIS FOR SYSTEMIC TEAMWORK METHOD

EAST is a macro-ergonomic method developed by the HFI DTC (Human Factors Integration Defence Technology Center) and used for studying the command and control sociotechnical networks. It is a descriptive method which extracts large scale systems level data on the emergent properties of a sociotechnical network arising from interaction between its different components (humans or technical devices).

EAST method rests on the Distributed Situation Awareness (DSA) which is a systemic concept unlike the Situation Awareness concept which is only focused on the individual (Endsley, 1995). In a crisis situation, the performance of the network is directly linked to the quality of the DSA (Stewart, Stanton, Harris, Baber, Salmon, Mock, Tatlock, Wells & Kay, 2008). Concretely, the DSA takes into account the fact that there can be as many points of view of a situation as agents within it and particularly the relatively invariant element which is "information" and its distribution (implicit or explicit) by its different agents across the entire system (Stanton, Stewart, Harris, Houghton, Baber, McMaster, Salmon, Hoyle, Walker, Young, Linsell, Dymott & Green, 2006).

All of this allows EAST to model three networks which are the aim of the method (Figure 1) : Task, Social and Propositional (or Knowledge) networks. The task network is a chronological representation of performed operations showing "who" is performing "what", "when".The social network represents a modeling of the relationships between actors allowing to know who are the central agents within the system. Finally, the propositional network describes the interaction between agents and the knowledge objects to characterize the flow of information required to decision-making within the system.

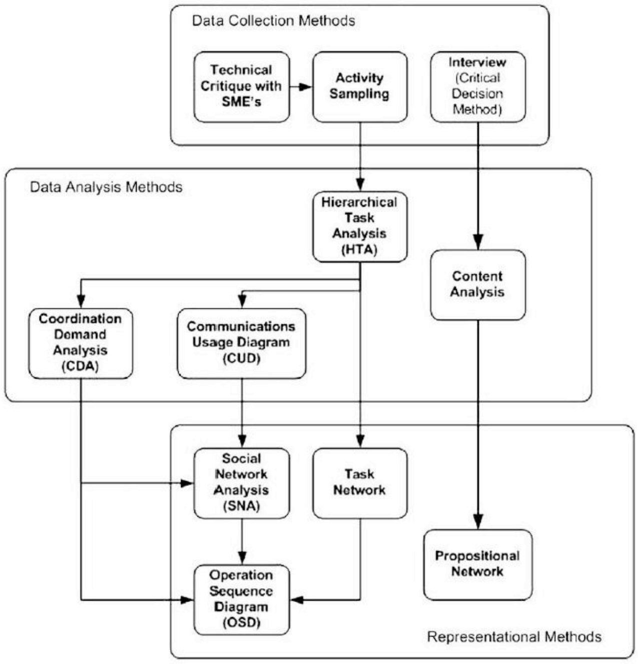

Figure 11. Structure of the EAST method (Walker et al., 2009)

This EAST method has already been used in various fields (fire service, air traffic control, military, police and so on) and allowed to obtain numerous information about the decision-making of individuals within collaborative situation. Thus, in the framework of our project this could represent an interesting method to set up (probably after some modifications to define in relation to our study) in order to collect essential data.

TOWARDS A COMBINED METHODOLOGY

As we have seen above, deal with the problem of agent activity within a group is very complex and needs to be able to rely on two approaches: psychological and systemic. Following this principle, the concepts of DSA and schema were discussed jointly by Stanton, Salmon, Walker and Jenkins (2009) so that it is more suitable for the study of situation awareness in collective environments. Thus, they studied an example taken in the UK energy distribution domain by linking the concept of genotpe and phenotype schemata with the DSA.

After that, several recent studies have used this type of methodology combining systemic and psychological approaches especially for accident analysis in different fields. Thus, Salmon, Read, Stanton & Lenné (2013) studied an incident with a semi-trailer truck that occured at rail level crossings on investigating systemic and psychological factors. The systemic analysis permited the study of interactions between the different composants involved in the situation. For that, they relied on the building of an "Accimap" (Rasmussen, 1997) summarizing all the decisions, actions and mistakes made within the different system levels (technical and operational management, physical processes and actor activities, local government, etc.). Their psychological approach relied on the schema theory and the Neisser's perceptual cycle (1976) to highlight the processes at the individual level. They used the Norman's taxonomy of schema-related slips errors (1981) which proposes that errors are caused by problems in the activation of schema that controls action sequences. In this way, they could find that the incident were caused by a truck driver mistake (more precisely, a faulty activation of schema error according to the Norman's taxonomy) maybe brought by some ambiguous environmental elements.

A lot of studies used the same methodology in others fields like aviation (Plant & Stanton, 2011) or public health incidents (Cassano-Piche, Vicente & Jamieson, 2009) for example. However, for their systemic approach, all of these studies used the Rasmussen's Accimap contrary to Stanton et al. (2009) who tried to link the concepts of genotype and phenotype schemata with a propositional networks on which is based the EAST method. In our project, the use of a propositional network, which can be moreover decomposed by indivual, seems to be more suitable for the transfer of agents' decision-making underlying processes on the ground to a virtual environment.

METHODS

The activity analysis during this thesis will take place during the training exercises for firefighters of Morbihan. This work will focus more particularly on one or several exercises performed in collaborative situation and requiring a global coordination and a decision-making from each member of the team under high time pressure.

Thus, this procedure might combine several methods to collect data. One of them is the EAST method described above. This method could give all the information used by the team members during the exercice and the dynamic of their flow during each phase of the training exercise. Particularly, this method will provide a propositional network which represents all the knowledge used by the agents on the ground (Figure 2).

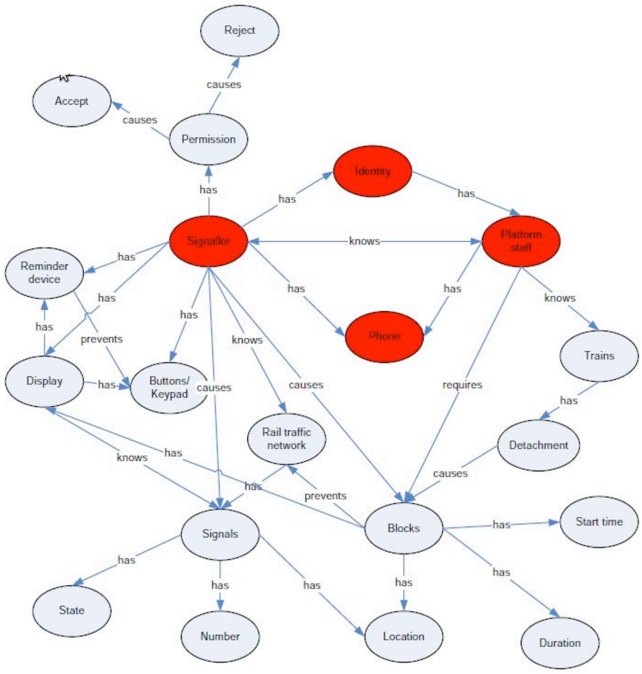

Figure 12. Example of Propositional Network (Gibson, Walker, Stanton & Baber, 2004)

This kind of network has the advantage to detail precisely and to link environmental cues and knowledge taken into account by agents to act, moreover in a dynamic way. Another big advantage is that these networks can be decomposed individually for each person involved in the exercise describing in this way the situation awareness of the individuals working in the system (Stanton et al., 2009).

This method could be completed by the setting up of self-confrontation interviews (Theureau , 2010) with each member in order to obtain verbalization data on their own decision-making. Thus this data, by a qualitative analysis based on the RPD model, might show us "how" the information is used by them at each step and if it forms a part of typical situation recognized.

These data could also be supplemented by information provided by trainers about their expectations regarding the knowledge acquired by the protagonists and their learning methods.

EXPECTED RESULTS AND PROSPECTS

The activity analysis of these work groups should contribute to highlight the particular "know-how". The latter should be able to define the autonomy area of each member of a team, the "possible violations" and acceptable practices in terms of safety. In all the studies, the management by individuals of their own resources is at the heart of the management of dynamic and collaborative situations. Here, the resources in question are not only "extrinsic" resources of the individuals (material and technical resources, organizational resources, human resources outside the system, etc.). Resources managed in natural situations are also "intrinsic" resources of the system operators. These immaterial resources cover the know-how, skills, cognitive resources, job rules, meta-knowledge, etc. (Falzon & Teiger, 1995). In the end, it is the extraction of these internal resources which should enable to contribute to the development of SécuRéVi and make all the interest of this work.

Indeed, firstly with a pedagogical point of view, the individual propositional networks obtained by the EAST method will supply all information (environmental cues or knowledge) that the VEHA users have to take into account at each phase of the exercise. Thus, we can imagine a virtual tutor which could automatically provide the information or highlight an environmental cue of which one user needs at a particular time. Secondly, with a computing point of view, the fact that propositional networks can be decomposed for each team member could normally allow to develop the processes which would direct all the virtual agents in the environment. Finally, the study of « how » information is used by each team member on the ground could help to enhance the artificial intelligence of virtual agents in order to be more credible.

By relying on all these results, we hope to design a credible VEHA that could be a real complement of the firefighters or incident commanders initial training.

REFERENCES

Cassano-Piche, A., Vicente, K., Jamieson, G. (2009) A test of Rasmussen's risk management framework in the food safety domain: BSE in the UK, *Theorical Issues in Ergonomics Science* 10 (4), 283-304.

Endsley, M. (1995) Toward a theory of situation awareness in dynamic systems, *Human Factors* 37 (1), 32–64.

Entin, E., Elliott, L. & Schiflett, S. (2001) DMT-RNet: An internet-based infrastructure for distributed multidisciplinary investigations of C2 performance, Proceedings of the 6th International Command and Control Research and Technology Symposium, Annapolis, MD.

Falzon, P. & Teiger, C. (1995) Construire l'activité, *Performances Humaines & Techniques, Hors Série (septembre)*, 34-39.

Fern, L., Trent, S. & Voshell, M. (2008) A functional goal decomposition of urban firefighting, *Proceedings of Information Systems for Crisis Response and Management (ISCRAM 2008),* Washington DC, May 4-7, 2008.

Fonseca, B., Paredes, H., Rafael, J., Morgado, L. & Martins, P. (2011) A Software Architecture for Collaborative Training in Virtual Worlds: F-16 Airplane Engine Maintenance, *In Proceedings of CRIWG*, 102-109.

Gibson, H., Walker, G., Stanton, N. & Baber, C. (2004) EAST Methodology for Railway Data: Protection Scenario, *Human Factors Integration Defence Technology Centre*.

Hall, K. (2010) The effect of computer-based simulation training on fire ground incident commander decision making, PhD dissertation, University of Texas at Dallas.

Klein, G. & Brezovic, C. (1986) Design engineers and the decision process: Decision strategies and human factors literature, *In Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting*, numéro 2, pages 771-775.

Klein, G., Snowden, D. & Pin, C. (2007) Anticipatory thinking. In K. Mosier & U. Fischer, (Eds.), *Proceedings of the 8th International Naturalistic Decision Making Conference,* 120-127, CA: San Francisco State University.

Marion N., Querrec R. & Chevaillier P. (2009) Integrating knowledge from virtual reality environments to learning scenario models. A meta-modeling approach, *International conference of Computer Supported Education,* pages 254--259, Lisbonne, Portugal.

McLennan, J., Pavlou, O., Klein, G. & Omodei, M. (2005) Using video during training to enhance learning of emergency incident command and control skills, *The Australian Journal of Emergency Management,* 20, 3, 10-14.

Mclennan, J., Elliot, G. & Holgate, A. (2009) Anticipatory thinking and managing complex tasks: Wildfire fighting safety and effectiveness, *In Proceedings of the Industrial & Organisational Psychology Conference,* pp. 90-95, Sydney, Australia.

Neisser, U. (1976) Cognition and Reality: Principles and Implications of Cognitive Psychology, Freeman, San Francisco.

Norman, D. (1981) Categorisation of action slips, *Psychological Review* 88 (1), 1-15.

Omodei, M., Elliott, G., Walshe, M. (2004) Development of computer simulated wildfire scenarios for the investigation of unsafe decision making, Bushfire Cooperative Research Centre Report 2: 2004. Melbourne, Victoria, Australia. 56 pp.

Pardue, J. (2009) Pedagogical reforms in firefighter education. Fire and Emergency Training Institute. http://www.feti.lsu.edu/archives/2009/OnlinePedogogicalReformsForFireFighterEducation.pdf.

Plant, K., Stanton, N. (2011) Why did the pilots shut down the wrong engine? Explaining errors in context using schema theory and the perceptual cycle model, *Safety Science* 50 (1), 300-315.

Querrec, R., Buche, C., Maffre, E. & Chevaillier, P. (2003) SécuRéVi : virtual environment for fire-fighting training. Dans Richir, S., Richard, P., et Taravel, B., éditeurs, *5th virtual reality international conference (VRIC'03)*, pages 169–175, Laval, France. ISBN : 2-9515730-2-2.

Rahman, M. (2009) Understanding naturalistic decision making under life threatening conditions, *Proceedings of the 9th International Conference on Naturalistic Decision Making* (pp. 121-128), Swindon, U.K.: The British Computer Society.

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

Salmon, P., Read, G., Stanton, N. & Lenné, M. (2013) The Crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings, *Accident Analysis and Prevention* 50, 1278-1288.

Stanton, N., Stewart, R., Harris, D., Houghton, R., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M., Linsell, M., Dymott, R. & Green, D. (2006) Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology, *Ergonomics* 49(12–13), 1288–1311.

Stanton, N., Salmon, P., Walker, G., Jenkins, D. (2009) Genotype and phenotype schemata as models of situation awareness in dynamic command and control teams, *International Journal of Industrial Ergonomics* 39, 480-489.

Stewart, R., Stanton, N., Harris, D., Baber, C., Salmon, P., Mock, M., Tatlock, K., Wells, L. & Kay, A. (2008) Distributed situation awareness in an airborne warning and control system: application of novel ergonomics methodology, *Cognition Technology and Work,* 10 (3), pp. 221–229.

Tchounikine, P., Baker, M., Balacheff, N., Baron, M., Derycke, A., Guin, M., Nicaud, J. F. & Rabardel, P. (2004) PLaton-1: Quelques dimensions pour l'analyse des travaux de recherche en conception d'EIAH, Rapport technique, Rapport de l'action spécifique : Fondement théorique méthodologique de la conception d'EIAH, Département STIC du CNRS.

Theureau J. (2010) Les entretiens d'autoconfrontation et de remise en situation par les traces matérielles et le programme de recherche 'cours d'action', *Revue d'Anthropologie des Connaissances,* Volume 4, n° 2, 2010/2, 287-322, http://www.cairn.info/revue-anthropologie-des-connaissances-2010-2.html.

Tisseau, J. (2001) Réalité Virtuelle – autonomie in virtuo, Habilitation à diriger des recherches, Université de Rennes I.

Walker, G., Stanton, N., Salmon, P., Jenkins, D., Stewart, R. & Wells, L. (2009) Using an integrated methods approach to analyse the emergent properties of military command and control, *Applied Ergonomics,* 40 (4), pp. 636–647.

Situation Awareness in Offshore Drillers

Ruby ROBERTS[a], Rhona FLIN[a] and Jennifer CLELAND[a]
[a]*University of Aberdeen*

ABSTRACT

Introduction: A critical human aspect in maintaining safety in industrial settings is the ability of workers to maintain Situation Awareness (SA). This cognitive skill is thought to influence subsequent decision making and performance. A domain in which maintenance of accurate SA is vital is that of the Oil and Gas industry, particularly drillers who work in a high-risk, fast paced work environment. This paper presents preliminary findings from a new study which aims to identify and map the cognitive components associated with drillers' SA during well control. **Methods:** The study uses cognitive task analysis techniques of observation within a state of the art drilling simulator and critical incident interviews with experienced drillers to investigate drillers' SA The goal is to create a Hierarchical Task Analyis of drillers' tasks that require situation awareness whilst conducting routine and hazardous wells operations.

Results: This is a new study and preliminary results will be discussed at the doctoral colloquium.

KEYWORDS

Situation Awareness; Drilling; Task Analysis.

INTRODUCTION

Definition & Theory

A plethora of definitions have been proposed to define situation awareness (SA). However, the most dominant and widely cited definition of SA was proposed by Endsley (1988): SA is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (p.79). In addition, Endsley's (1995) three level information processing model of SA is the dominant model in the field, being the most investigated and applied in industrial domains. It describes SA as an internally-held, cognitive product of the three hierarchical levels, perception, comprehension and prediction. It is widely accepted that SA is a safety critical factor, as inaccurate SA can lead to poor decision making and unnecessary risk taking, increasing the likelihood of an accident (Stanton, Chambers & Piggott, 2001). For the purposes of this project, Endsley's three level model of SA will be utilized.

Offshore Drilling

A domain in which possession and maintenance of accurate SA is vital is that of the Oil and Gas industry. For example, a survey of offshore installation managers indicated that they felt that loss of care and attention was one of the main causes of accidents on offshore production platforms and drilling rigs (O'Dea & Flin, 1998). Failures of SA have not just been identified within accidents in the industry but also in large scale disasters. The investigation into the Deep Water Horizon drilling rig disaster in the Gulf of Mexico in which 11 lives were lost and caused the worst oil spill in US history has highlighted failures of SA and risk perception as root causes of the disaster (Report to the President, 2011). In fact, it would be difficult to investigate any large scale oil and gas disaster and not find human failures associated with inaccurate SA (Sneddon, Mearns & Flin, 2006). Therefore, it would appear that SA research in the offshore oil and gas industry would have obvious applications.

A particular sample of offshore workers whose SA is vital is that of the drilling crews. Drilling personnel are involved in very high risk activities on the drilling rig or production platform to build or maintaining the underground wells. This high-risk, interactive, fast paced work environment requires high quality SA to constantly monitor and comprehend the state of the well and the drill floor operations so as to be able to make the best decisions and keep accident risk to a minimum.

The literature surrounding SA in the oil and gas industry is sparse. Sneddon et al. (2006) investigated the role of SA in an offshore environment within drilling crews. It was found that accident analysis supported SA as a key factor in errors and subsequent accidents with isolation from events at home, fatigue and stress being perceived to be most significant contributory factors to reduced awareness. More recently, Sneddon, Mearns and Flin (in press) showed that stress and fatigue could be detrimental influential factors on drillers' SA, using their Work Situation Awareness rating scale. Despite this increased understanding, the fundamental cognitive components

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

required for drillers' SA have not been identified. Therefore, the current research project aims to investigate the cognitive components associated with drillers' SA.

Current Study

This project aims to indentify and map the cognitive components associated with drillers' SA during well control. The research will be conducted by identifying the cognitive skills and associated behaviours in relation to SA for the driller whilst he or she is conducting well control tasks (studies 1 and 2). The information collected will then be used to produce a Hierarchical Task Analysis (HTA) for drillers whilst conducting one or more hazardous well control tasks to identify the steps and cognitive components involved in completing these steps (study 3).

METHODS

Overall, human factors methods commonly used in other industrial domains for data collection and examination (e.g. Stanton, 2005) will be transferred into the new domain of the oil and gas industry Research in other high-risk, high-reliability industries such as aviation and healthcare has sought to understand the cognitive components of SA by producing taxonomies of skills and associated behaviours (Endsley & Garland, 2000; Phipps, Meakin, Beatty, Nsoedo & Parker, 2008). This approach would use two frequently used human factors data collection methods of simulator observations and critical incident interviews for cognitive task analysis (Stanton, 2005; Seamster, Redding & Kaempf, 1997). The HTA method identifies a goal or task and aims to break down the steps to complete this goal into small steps. Cognitive task analysis is then conducted by inferring the cognitive skills and components associated with each of the broken down steps (Stanton, 2005). The observation and interviews provide information to give the most accurate inference of cognitive skills. HTA is appropriate as it has previously been applied to the drilling domain (Stanton & Wilson, 2005). Overall, this method is also appropriate as it is very similar to the frequently used SA requirement analysis which identifies exactly what the operator's SA consists of and what information is required for the operator to complete their tasks or goals (Endsley, 1993). This method involves unstructured interviews with subject matter experts, goal directed task analysis and questionnaires in order to identify the relevant SA requirements. As this project in its infancy with only an initial set of observations and interviews, only preliminary results will be discussed at the doctoral colloqium. However, the expected results will be noted along with the methods and analysis used.

Study 1 – Observation

The aim of study 1 is to identify the cognitive skills and associated behaviours of drillers' SA whilst conducting their most hazardous tasks in the simulator. This is being conducted through observation of approximately 20 drillers' behaviour whilst performing scenarios of hazardous tasks in a drilling simulator. Observation is an ideal method for initial data gathering as it provides the naive researcher with the opportunity to both learn about the new setting and collect initial data (Stanton, 2005). Both live and video recording observation techniques are used such as taking field notes and check lists of behaviours. The data collected focuses on visible behaviours and skills associated with SA, such as where their information sources are and awareness of what the rest of the team are doing, as well as cognitive aspects such as cue recognition and prospective memory. The data will be analysed by collating the information from the observations into recurring themes associated with SA models to produce a catalogue of behaviours and skills associated with drillers' SA. These data can then be later used in study 3's HTA to make inferences about the cognitive components associated with the behaviours and skills.

Study 2 – Critical Incident Interviews

The aim of study 2 is to elicit knowledge from a sample of drillers about their own SA and the involvement of SA in well control events that they have experienced. Critical incident interviews were selected as a data collection method as they both elicit knowledge about tasks and the environment, as well as provide a gauge of SA role in incidents (Crandall, Klien & Hoffman, 2006). The interview schedule includes asking about the individuals' job and tasks, aspects of their SA, as well as talking through a challenging day. The challenging day focuses on a difficult well control task or coping with a well control event e.g. a kick. During this discussion, the interviewee is asked about their own SA, if and how it was involved in the causation of the well control event and the recovery from the well control event.

A sample of 10 drillers plus 10 assistant drillers will be interviewed either via the telephone or during visits to the simulator training days. Once the interviews are transcribed, they will be coded to identify recurring aspects of the drillers' SA. The coding method will include key aspects of SA as highlighted in the literature focusing on Endsley's (1995) three level model. These data will provide the foundation for the hierarchical task analysis by identifying the cognitive components and requirements for accurate SA.

Hierarchical and Cognitive Task Analysis

HTA has previously been used successfully to identify the steps involved in complex tasks allowing for the inference of associated cognitive components in other high-risk, high-reliability industries such as health care (Phipps et al., 2008), as well as in the oil and gas industry (Stanton & Wilson, 2005). The aim of study 3 is to produce a HTA for drillers whilst conducting the most hazardous task(s) to identify the steps and cognitive components involved in completing these task(s). The data from studies 1 and 2 will be used to inform this process, with the aid of a drilling expert to ensure that the task(s) are accurately described. The task(s) to be broken down will be chosen in collaboration with the collaborative company and data from studies 1 and 2, with the expectation that they will cover hazardous aspects of well control. The actual process of HTA will use software to produce the model such as Task Architect. Cognitive analysis will then be conducted on the individual steps to identify the cognitive components associated with drillers' SA during well control tasks.

CONTRIBUTION TO THE FIELD

As a project of this nature has not previously been conducted in the drilling domain, it is expected to produce a data set of behaviours, skills and cognitive components of drillers' SA that has not been collected before. It is hoped that these cognitive components and skills will be able to be generalised to other high risk, high reliability domains as well as inform, theoretical concerns related to SA. It is also expected that this greater understanding can be used to aid industry in indentifying and recording SA within their own competence assurance processes, incident records, as well as being used to inform the design of interventions to improve safety, such as training against identified types of errors. More generally, it is likely that novel measurement and data collection methods will be developed in response to issues that arise during the course of the project.

ACKNOWLEDGMENTS

Acknowledgement: This doctoral research project is sponsored by Maersk Drilling.

REFERENCES

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds. A practitioner's guide to cognitive task analysis*. Cambridge, Mass: MIT Press.

Endsley, M. (1988). Situation Awareness Global Assessment Technique (SAGAT). *In: Proceedings of the National Aerospace and Electronics Conference (NAECON), IEEE, New York,* pp. 789-795.

Endsley, M.R. (1993). A survey of situation awareness requirements in Air-to-Air Combat fighters. The *International Journal of Aviation Psychology,* 3, 157-168.

Endsley, M.R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors,* 37, 32-64.

Endsley, M.R., & Garland, D.J. (2000). *Situation Awareness Analysis and Measurement.* Mahwah, NJ: Lawrence Erlbaum Associates.

O'Dea, A., & Flin, R. (1998). Site management and safety leadership in the offshore oil and gas industry. *Safety Science,* 37, 39–57.

Phipps, D., Meakin, G. H., Beatty, P. C. W., Nsoedo, C., & Parker, D. (2008). Human factors in anaesthetic practice: insights from a task analysis. *British Journal of Anaesthesia,* 100, 333-343.

Report to the President (2011). National Commission on the Deepwater Horizon Oil Spill and Offshore Drilling. Washington.

Seamster, T. L., Redding, R. E., & Kaempf, G. L. (1997). *Applied Cognitive Task Analysis in Aviation.* London: Ashgate.

Sneddon, A., Mearns, K., & Flin, R. (2006). Safety and situation awareness in offshore crews. *Cognition, Technology and Work,* 8, 255–267.

Sneddon, A., Mearns, K., & Flin, R. (in press). Stress, fatigue, situation awareness and safety in offshore drilling crews. *Safety Science.* DOI: http://dx.doi.org/10.1016/j.ssci.2012.05.027.

Stanton, N. A. (2005). *Human Factors Methods: A Practical Guide for Engineering and Design.* Farnham: Ashgate Publishing.

Stanton, N.A., Chambers, P.R.G., & Piggott, J. (2001). Situational Awareness and safety. *Safety Science,* 39, 189-204.

Stanton, N. A., & Wilson, J. (2005). *Safety and performance enhancement in drilling operations by human factors intervention (SPEDOHFI). HSE Report 264.* London : HSE Books.

How sounds and soundscapes can help to maintain situational awareness

Jukka SEPPÄNEN[a]
[a]*Kymenlaakso University of Applied Sciences, Finland*

ABSTRACT
Introduction: Main research questions of this doctoral paper are how sound perceptions are connected to the understanding of emergent situations and how often they are in an important role in sense-making during critical and crisis situations. **Method**: All completed maritime investigation reports (N=99) made by the Safety Investigation Authority in Finland during the last ten years were analysed. **Results and discussion:** Sound perceptions or the lack of them were mentioned in 35 reports (35%) and in eleven reports, which is 11 % of all reports, the correct understanding of emergent situations was mainly based on the observations of natural sounds.

KEYWORDS
Sound; soundscape; situational awareness.

INTRODUCTION

This doctoral paper is a part of the author's dissertation in the field of psychology of music with the title of "Sounds of fear and hope: How to enhance mental performance during emergency situations" for University of Jyväskylä, Finland. The dissertation is multidisciplinary because of the tight connection to the human factors science. The first supervisor is Ph.D. Suvi Saarikallio, Academy of Finland Research Fellow, while another supervisor is adjunct professor Ph.D. Veli-Pekka Nurmi, executive director of Safety Investigation Authority, Finland.

The original idea in the dissertation is to find out what the role and significance of sounds, soundscapes (including silence) and music is during demanding situations in sociotechnical systems. The first research question is what kind of information sounds can carry, i.e. how soundscapes can maintain situational awareness. The second research question will be how soundscapes affect feelings, i.e. whether sounds can awake or startle feelings or even cause fear. The third research question will be how sounds and music can serve to enhance human performance in addition to physical and psychological methods such as relaxation and mental images. This doctoral paper is meant to answer to the first question with the following sub-questions:

1. how sound perceptions are connected to the understanding of emergent situations, and
2. how often they are in an important role in sense-making during critical and crisis situations

Scientific background

Younger generations listen to music mainly to regulate their emotions and mood (Saarikallio, 2007) and to create their own sonic environments, also when walking in the heavy city traffic. The absence of sound signals has been proved to increase risks in this kind of environments. (Lichenstein et.al., 2012) Other findings also indicate that both the frequency and severity of driving violations are higher when drivers listen to aggressive music than during trips when no music is played. (Brodsky & Slor, 2012)

When sounds in scuba-diving were studied by the author, it was found that sounds are very significant signals for divers. Divers use sounds as an information source and many kinds of underwater sounds and also the silence cause strong emotions. The underwater silence was merely considered as positive and nice, and the technogenic noise as noxious and frightful. (Seppänen & Nieminen, 2004)

The term soundscape is the designation of any human-audible sounding environment. Soundscapes are usually divided into three basic elements: *keynote sounds* outline the character of the acoustic environment, *sound signals* are consciously listened foreground sounds and *soundmarks* are sounds which are unique to a particular environment. (Schafer, 1994) The airborne soundscape in natural environment at sea is dominated by keynote sounds of the wind and sea. (Lee & Lu, 2009) There are also sound signals, i.e. sounds by, for example, other vessels, sea birds, fog signal devices and occasionally some specific sound marks, e.g. sounds of the packing ice, sounds of jumping humpback whales.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

When examining the acoustic environment onboard from the OOW's[8] point of view, monotonous keynote sounds including the tiring isothermal air conditioning and some sound signals (e.g. alarms, radio calls) are typical of a modern closed bridge. However, a proper look-out by sight and hearing shall be maintained at all times. (COLREG, 5§) If a sound signal from the operational environment, e.g. from another vessel's foghorn, is not observed, or if there are disturbances (e.g. wrong filtering, unsuccessful interpretation, long reaction time) it results in a shortage of situation awareness which may be critical regarding the required decision-making. (Factors contributing to fatigue)

According to Endsley (1995), situational awareness is based on the perceptions in the current situation, the comprehension of the current situation and the projection of the future status. As a part of situational awareness, auditory inputs allow to assess and monitor the socio-technical environment of the craft. It has been stated, for example, that all crew members know the sounds of a normal-functioning craft, while on the other hand, unexpected sounds or the lack of sounds may alert them to possible malfunctions, failures, or hazards. (Hearing and noise)

METHODS

Soundscape is also a method and an idea to study the interrelation among sound, environment and human. It is the human that feels, thinks and remembers the sound and environment and then response to the sound and environment. (Lee & Lu, 2009). The priority in this article is "to employ natural sounds" (Schafer, 1994) i.e. to find primary sounds as naturalistic elements to maintain situational awareness; neither sound alarms nor voices of the communication inside the bridge or via radio are taken into account.

To find out what has been the importance of sounds in noticing emergencies, all maritime accident and incident investigation reports made by the Safety Investigation Authority in Finland (SIAF) during the last ten years, 2003-2012, with the total number of 99 completed (by 18 March 2013) reports have been analysed.

RESULTS

Number of observations

Altogether 47 separate sound observations or instances of lack of observation have been mentioned in 35 different safety investigation reports (35 % of all reports), and altogether 39 different words were found to be used describing the nature of the sounds or noise. Words used more than once were collision (in Appendix numbers 3, 4, 27, 28, 30), loud (9, 36, 37, 42), cry (32, 33, 35), vibration (1, 26, 31), slam (7, 8, 10), crash (2, 6), short (11, 17) and start (21, 47). Allegories have been used in two cases (1, 20).

Sources of the sounds mentioned in the reports and descriptions of the nature of the sounds are collected in Appendix. Descriptions are taken literally from the reports by the SIAF and their English translations. When necessary, Finnish descriptions have been translated by the author.

Roles of the sounds

The role of the sounds is estimated based on the actions taken after them. In eleven reports, sounds heard by operators have been in a major role i.e. the correct understanding of emergent situations was based on the sound observations, and due to them at least someone has taken some immediate steps to control the emergency. Actions are listed below, starting from very severe cases:

- they heard a rattling noise and the sound of objects sliding on the deck above => they rushed up and out from the superstructure (44)
- when heard five short blasts given by Silja Serenade => he realized the situation and started the turn immediately (17)
- he heard the other seaman's cry => when he turned his head towards the sound he saw the seaman's legs disappear over the rail (33)
- a deck hand called out, "stop, stop, help" => the skipper switched off the engine and turned to look towards the aft (34)
- he heard an unfamiliar sound coming from the water => immediately he caught hold of a torchlight and detected a seaman in the water (38)
- after hearing a noise => the AB rushed to the site and saw that part of the crane had been derailed and collapsed (42)
- after stopping their own craft, the patrol boat crew heard crying for help => they moved ahead towards the sounds and saved lives of both men (35)
- the trainee heard a sound of something falling => a moment later he noticed that the Electrician had fallen about three meters (40)
- he heard an exceptionally sudden and strong raise in the engine noise => the first engineer narrowly escaped from the danger zone before the engine raced and broke down (20)

8 Officer on watch

- on hearing the noise of the collision => she immediately ran down to see what had happened (4)
- he heard the crash => and returned to the bridge (6)

In all other 24 reports sound observations have been in a minor role, i.e. also some other information sources were mentioned or the sounds were considered to have no significant role.

Conclusions

From the perspective of the soundscapes studies, only some keynote sounds are mentioned: a sound of the storm, a rumble noise of the tugboat (both in 46) and a loud hit of the wave (9). Also normal vibrations as such are characteristic to the ship environment. Blasts given or not given by vessels (11-18) are signal sounds and other sounds mentioned refer to temporary soundmarks, very unique to the particular ongoing situation onboard.
Based on the findings and previous studies, the role of sounds and their connections to the understanding of emergent situations is described in figure 1. As the figure shows, sounds carry information about onboard environments of different kinds (natural, operational, socio-technical) and are essential tools in individual and social processes.

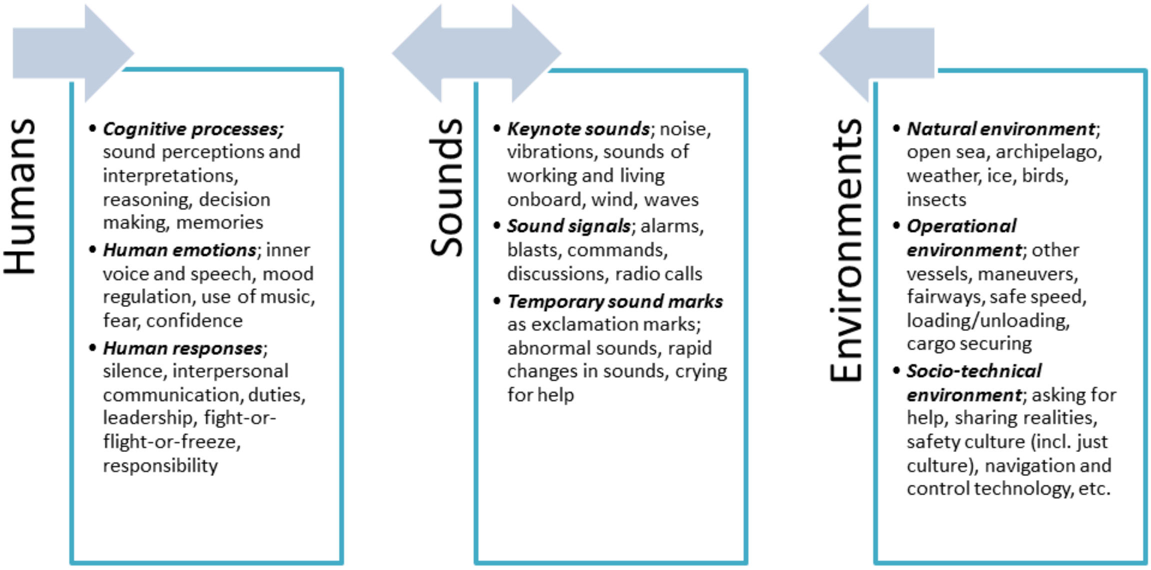

Figure 1. The role of soundscapes in maritime. After semicolons, there are only some examples. (Modified from Truax, 1984 and Wrightson, 2000).

When sounds have been mentioned, in almost every third cases they have been in a very significant role. Of the 99 analysed reports, in eleven reports, which is 11 % of all reports, the correct understanding and sense-making of emergent situations was mainly based on the sound observations. Further studies are needed to determine the role of sounds as tools in individual processes: how and how often sounds, e.g. music, musical memories or inner voices are used to regulate mood and enhance mental performance.

DISCUSSION

In the generalisation used in soundscape studies, post-industrial soundscapes are defined as "lo-fi" soundscapes: meaningful sounds can be masked due to the fact that ambient noise and sonic information can mutate easily into anti-information. Pre-industrial soundscapes are considered to form exactly opposite environments: in "hi-fi" soundscapes all sounds can be heard clearly, and the so called acoustic horizon may reach several miles. (Wrightson, 2000)
The findings of this study – e.g. in three cases blasts given by a nearby vessel were not heard (11, 15, 16) and in two cases blasts were not even given (14, 18) – indicate that acoustic environments inside modern ships correspond unquestionably to the "lo-fi" soundscapes. Particular sounds are not always heard and individual's aural space (see Wrightson, 2000) is sometimes so reduced that sounds are underrated. On the other hand, when the look-out by hearing or other kind of proper alertness has been maintained, people have been able to hear blasts and cries for help (especially 17 and 35, also 33, 34, 38).
The listening of sounds in onboard environments is not a historical relic from the times of "hi-fi" soundscapes in tall ships and steamers: sounds and soundscapes will always be very essential factors in maintaining situational awareness in maritime operations and in demanding situations.

REFERENCES

Brodsky, W. & Slor, Z. (2012) Background Music As A Risk Factor For Distraction Among Young Drivers: An IVDR (in-vehicle data recorder) Study. Available at http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/137_Proc.pdf

COLREG. The International Regulations for Preventing Collisions at Sea 1972.

Endsley, M.R. (1995) Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64.

Factors contributing to fatigue and its frequency in bridge work. Investigation report S3/2004M. Safety Investigation Authority, Finland. Available at http://www.turvallisuustutkinta.fi/en/Etusivu/Tutkintaselostukset/Vesiliikenne/VesiliikenneTeematutkinnat/1266334150682

Hearing and noise in aviation. FAA Hearing Brochure. Available at http://www.faa.gov/pilots/safety/pilotsafetybrochures/media/hearing_brochure.pdf

Lee, Y-C. & Lu, L-S. (2009) Resonance rather than Solo: Shaping a Regional Image with Soundscape. National Yunlin University of Science and Technology. Available at http://www.designresearchsociety.org/docs-procs/DRS2010/PDF/076.pdf

Lichenstein, R. & Smith, D. & Ambrose, J. & Moody, L. (2012) Headphone use and pedestrian injury and death in the United States: 2004-2011. *Injury Prevention*, 18(5), 287-90.

SIAF, Safety Investigation Authority in Finland. Home page at http://www.turvallisuustutkinta.fi/en/Etusivu

Saarikallio, S. (2007) Music as mood regulation in adolescence. Diss. University of Jyväskylä, Finland.

Seppänen, J. & Nieminen, M. (2004) Measurements and Descriptions of Underwater Noise in Finland. Geophysica (2004), 40(1-2), 23-38. Available at http://www.geophysica.fi/pdf/geophysica_2004_40_1-2_023_seppanen.pdf

Schafer, R.M. (1994) The soundscape: Our sonic environment and the tuning of the world. Rochester, Vt. Destiny Books. United States.

Truax, B. (1984) Acoustic Communication. Norwood, NJ: Ablex Publishing. United States.

Wrightson, K., 2000. An Introduction to Acoustic Ecology. Soundscape, Vol. 1, 10-13. Available at http://www.glennbach.com/wrightson_intro_AE.pdf

APPENDIX

The source and the nature of the sounds mentioned in safety investigation reports. All reports can be found at http://www.turvallisuustutkinta.fi/en/Etusivu using the report number mentioned.

Hull or similar

1. vibrations and metallic sound, like pushing trolley along rugged surface, D3/2011M
2. sound of the crash, C4/2010M
3. grinding sound, collision was smooth and sounds were faint, C3/2010M
4. noise of the collision, C3/2009M
5. scraping sounds from the bottom, C3/2008M
6. crash, C4+C5/2006M
7. slam was heard, everyone was silent for a while, C2/2005M
8. slamming sounds, C4/2003M
9. loud hit of the wave under the stern, B1/2003M
10. slam in the bow when hitting ice, C1/2003M

Fog signal device

11. short blast given, not heard, D2/2011M
12. NoD. (= no description), blasts were given as long as there were electricity onboard, D18/2009M, 1/2010M
13. NoD., blasts given by ships in the area, used to help in search, D11/2009M, 1/2010M
14. NoD., no sound or light signals were made by either vessel, C5/2008M
15. NoD., fog signal device was at use, but there was no audio lookout onboard another vessel, C1/2007M
16. long blasts were given, not heard, C7/2005M
17. five short blasts were given, C4/2005M
18. NoD., no sound signals were made by either vessel, C3/2004M

Engine

19. engine pounding forcefully, N5/2012M
20. exceptionally sudden and strong raise, like a whining chainsaw, C1/2009M
21. start of the emergency generator, C7/2006M
22. rumble noise, C9/2004M
23. chinking sound from the valve, B1/2003M
24. abnormal sound from the bearing area, B1/2003M
25. strange engine sound, C5/2003M osa II

VDR[9]

26. sounds of the vibration, C4/2009M
27. sound of the collision, MS BIRKA EXPRESS & HANSE VISION, BSU Ref.: 20/09
28. collision noise, items dropping from tables, C5/2008M
29. impact to the pier, MS FINNLADY, BSU Ref.: 211/08
30. collision sounds, C6/2006M
31. vibrations, B1/2004M

Crying for help

32. men cried help but it was not heard, D17/2009M, 1/2010M
33. seaman's cry, B1/2009M
34. deck man called out stop stop help, B2/2008M
35. patrol boat crew heard crying for help, C5/2003M osa I

Underwater sound

36. loud commotion, C3/2009M
37. loud and dull sound, B2/2004M

Other

38. unfamiliar sound from the water, MV CARISMA DMAIB 201204183
39. attention drawing sound from tracks of the (floatable) digger, D4/2011M
40. sound of the falling, D10/2010M
41. sound "tscha" and a sound of water in pipes, MV SEA WIND SHK Rapport RS 2011:01 Dnr S-211/08
42. AB heard huge noise, master heard loud noise, B1/2006M
43. hissing noise of escaping compressed air, C7/2006M
44. rattling noise and objects sliding on the deck above, sliding and rattling sounds, C7/2006M
45. cracking sounds were heard before, while heavy sea, B2/2004M
46. sounds of the storm, C9/2004M
47. sound of the automatically starting bilge pump could help to notice leak, C5/2003M osa I

[9] VDR (Voyage data recorder) is a data recording system designed for all vessels, sometimes also called Black box for ship, which inter alia include audio information from the bridge and bridge wings. VDR data is usually utilized e.g. in accident and incident investigation.

The role of argumentation in the resolution of ambiguous situations: an exploratory study in the field of supervision in the nuclear industry

Johanna MERAND[a&b], Françoise DARSES[a] and Cecilia DE LA GARZA[b]
[a]*ACSO (Action and Cognition in Operational Situations department), IRBA (French Armed Biomedical Research Institute)*
[b]*Human and Organisational Factors in Socio-technical Systems - Management of Industrial Hazards, EDF R&D*

ABSTRACT
This exploratory study is part of the evaluation, using a full-scale simulator, of a new organisation for a nuclear plant operating team. The purpose of the study is to test the possibility of there being a link between the characteristics of the operating activity and the argumentative activity of the teams when managing ambiguous situations. A comparative analysis of the operating activity, the "decision pathway" and the argumentative discourse of the three future operating teams was carried out during the resolution of an ambiguous situation. Results show that, according to the team, the "decision pathway" differs in terms of the aims set or prioritised by the team (alongside those required by the procedure). According to the types of aims, we will highlight the different argumentative patterns.

KEYWORDS
Decision-making; cooperation; argumentative process; supervision; ambiguous situations.

INTRODUCTION

This study is part of a process to design new operating methods for a next generation nuclear power plant control room. By "operating methods", we mean the procedures, human-machine interfaces and organisational requirements that operators must observe in order to supervise and manage the process. The new team organisation features four operators: an action operator (AOP), a strategic operator (SOP), an operations shift manager (OSM) who is in charge of the operating team and a safety engineer (SE). The goal is to ensure that "human-machine" and, more broadly speaking, "human-organisation-machine-interfaces-procedures" relations work properly and promote effective and safe emergency operation.

This study was carried out within the context of an evaluation process that has been implemented for over ten years, in relation with the designers of a new nuclear reactor (Labarthe, & De la Garza, 2011; De La Garza, Labarthe & Graglia, 2012). The specific feature of this study is that this *Human Factors* evaluation campaign was carried out by a multi-disciplinary EDF team (ergonomics, reliability and safety experts), using a full-scale simulator of the future control room. This study looks at one of the aspects of this evaluation, which concerns the processes of cooperation deployed within the operating team (Burke et al., 2000; Crichton et Flin, 2004; Hoc, 2001) in situations that appear to be "poorly defined" or "ambiguous" (Klein, 1997). In the case analysed, given the fact that we are in a design situation, the "poorly defined" or "ambiguous" situations are mainly linked to operating methods that have not been completely finalised and the teams not having completed their training programme which might, for example, result in a lack of understanding when faced with major technological changes. Feedback from plants in operation suggests that we can find similar situations that were not foreseen in the design phase and for which the operating methods are not optimised. The operating teams can therefore, in natural situations, be faced with "ambiguous" situations.

PURPOSE

The purpose of this exploratory study is to describe the underlying processes of cooperation involved in the resolution of ambiguous situations, in cognitive terms. To cooperate, team members needs to interact. It is even more necessary when team members have different roles and thus possess information appropriate to their own role. In emergency operation situation, team members are often engaged in an argumentative process in order to share their mental representation or to confront their point of view in order to make an effective and appropriate decision.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Numerous studies were interested in these argumentative processes in the nuclear industry (Carvalho et al., 2005; Chung et al., 2009; De La Garza et Le Bot, 2008; Kim et al., 2010a,b; Lee et al., 2012; Park et al., 2012; Stachowski et al., 2009) but also in other domains such as the medical domain (Pelayo, 2007), the military aviation (Bourgeon, 2011) and manufacturing (Darses, 2006). The speech act coding schemes developed in these studies are based on various models of team cooperation (Hoc, 2001; Hélie et Loiselet, 2000; Hollnagel et Woods, 2005; Endsley, 1995), the choice of which depends partially on the object of study: to explore the link between interaction patterns and team performance (Stachowski et al., 2009; Kim et al., 2010b), to study the impact of a new technology (Chung et al., 2009) or a standard communication protocol on team processes and team performance (Kim et al., 2010a)...

This study seeks to determine whether the characteristics of the argumentative process developed by the operators can be correlated to the specific characteristics of operations and the decision pathway[10] involved when managing ambiguous situations. By looking at the involvement of the various roles in the argumentative process of the "decision pathway", this study seeks to establish the extent to which the new team organisation influences the generation of argumentative processes promoting safe and efficient operating activity or not. By analysing the argumentative process, we should, in addition to understanding the individual and collective aspects of argumentation, be able to identify areas where the design of the operating methods and the organisation can be improved, so as to make the system more robust when dealing with situations that were not anticipated or foreseen in the design phase.

METHODOLOGY

Population

Three teams recruited for the operation of the future nuclear reactor took part in this evaluation campaign. All those involved have previous operating experience and the training program for future reactor is in progress. The teams were trained on the processes, HMI, imaging and the specific skills required for each position in the operating team for the purposes of the evaluation.

Technical configuration

The simulator features an exact replica of the future control room (spatial layout of the stations, HMI, procedures, etc.) and the dynamics of the future technical system are almost totally replicated. This enabled the operators to play out various ecological use scenarios defined by a team of experts for the requirements of the evaluation. The design of the entire socio-technical system (procedures, HMI…) is not yet finished.

An instructor workstation next to the control room was used to view the progress of the scenarios and the on-screen display on each of the four computer workstations in the control room. Headsets also enabled the various conversations between members of the team to be followed in real-time. The instructors simulated telephone calls with the various on-site specialist sectors. Each simulation, lasting around three hours, was followed by a group debriefing with the operating team, led by the evaluation team.

Collection of data

The study is focused on an emergency operation situation, carried out on a full scale simulator that should be handled in the right way in the future by the teams. The emergency operation procedures cover "nominal" emergency situations, where everything goes as expected. However, when a failure occurs (covered or not by operation procedures), situations could require operators to perform a considerable activity of diagnosis and even problem-solving. This may or may not be facilitated by the HMI (accessing information in the HMI could be easy or not), the procedures themselves (the action may be found in the sequence of the procedure or not) or by the knowledge of the operators. This study looks at this type of situation, in this case, the failure of a process controller that has not been clearly understood by the team.

The data collected are:

- Notes taken from the observations made by the evaluation team;
- Verbal exchanges and behaviour observed in video and audio recordings;
- Copies of on-screen displays and overviews of the control room.

Selection of the sequences analysed

Our study focussed on managing the cooling system following a partial cooling system process controller failure. In the scenario in question, when certain threshold parameters were reached, the partial failure of the cooling system process controller was quickly compensated by the automatic start-up of a diversified controller fulfilling the same function. While dealing with this failure, the teams had to continue to apply the procedures geared towards ensuring the complete cooling of the facility until such time as the required thresholds were reached.

[10] The decision pathway means all the discussions, consultations and deliberations, as well as the aims set and prioritised by the team, related to the resolution of a specific problem.

Processing of the data

For each team, the sequences featuring the topic of cooling were extracted from the verbal protocols. These sequences could be very short: very basic information (e.g. "it's 2°C cooler") or a longer discussion, designed to solve a problem. In this case, problem-solving related to: (i) A misunderstanding about the changes to the facility, (ii) the perception that what was proposed by the procedure was unsuitable for the changes to the facility, and (iii) the need to prioritise an action given the changes to the facility and/or available resources (human and technical).

Coding of the sequences

A coding system was defined with the aim of modelling the cognitive and collective processes involved in the decision pathway. The graphic representation of the coding includes a time line, where T_0 corresponds to the switch to emergency operation triggering initiation of a specific emergency procedure adapted to the degraded state of plant. This coding, an example of which is shown in Figure 1, enables the pathway of the decision-making process to be described exhaustively, specifying:

- The key events
- The required steps, and the steps actually taken
- The aims pursued, whether they are required by the procedure (in black) or set/prioritised by the team (in red in the coding)
- The human resources assigned to achieve each required aim or each aim set/prioritised by the team and any shift in the scope of the roles
- Communication activities (information sharing, approval requests and decision-making)
- Individual cognitive activities (anticipation of a step, detection of an event, etc.)

In this article, we will focus exclusively on analysing exchanges between operators which correspond to:

- *Arguments aimed at evaluating and understanding the situation* (answering the question, "what?") which are symbolised by a circle in the graphic coding. Here is an example of this type of argumentbetween the Operations Shift Manager (OSM) and the Strategic Operator (SOP): OSM - "It doesn't look like the partial cooling is running... SOP - "Yes, hang on, I'm just getting a message, Auxiliary Steam Distribution System (ASDS) restore request. "
- *arguments about procedures* that the operators are going to implement to solve the problem (answering the question "how? ") which are symbolised by a triangle. Here is an example of a argument about procedures: SOP - "Now he's working on isolating it fully from the Steam Generator (SG). I'll let him finish that and before he (NB: the other operator, the AOP) moves onto something else, I'll get him to change module." OSM - "But the SG level won't go up as long we haven't lowered the pressure, will it... unless you do it." SOP - "If you like." … OSM - "There's no point interrupting him, there are two things to do: the isolation and the partial cooling." … SOP - "OK, that's fine. ...".

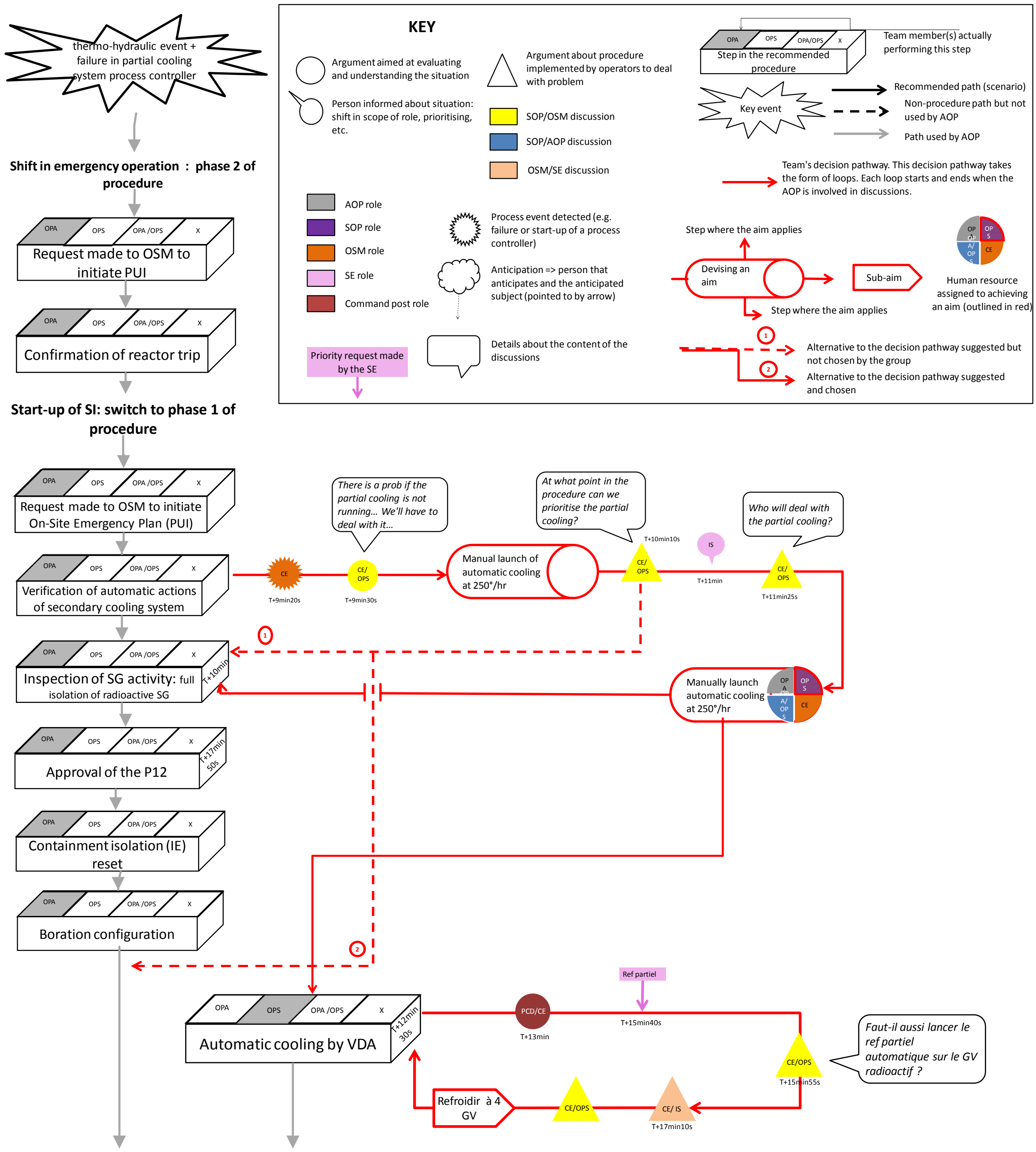

Figure 13. extract from the coding of a team A sequence

Indicators of the teams' operating activity

The operating activity of the operators was characterised according to the following indicators:

- *Correct evaluation of states E1* (the facility is not cooling) *and E2* (the facility is cooling)
- *Correct diagnoses D1* (partial cooling system process controller failure) *and D2* (start-up of the diversified cooling system process controller), corresponding to the states E1 and E2 respectively.
- *Time (in minutes and seconds) elapsed between E1and D1, and time elapsed between E2 and D2*
- *Number of attempts required to achieve the aim*
- *Time (in minutes and seconds) elapsed between T0 and the first attempt to achieve the aim*
- *Efficiency of the aim (yes/partial/no): effective and optimum achievement of the aim* (e.g. the facility cools correctly and sufficiently quickly until it reaches the required threshold)
- *Relevance of the aim (yes/no): given the changes in the state of the facility, the aim set remains relevant* (e.g. cooling is no longer relevant because it was implemented too late);
- *Safety of the aim (yes/no): the aim is achieved without there being any risks for personnel, the local population and the environment.*

RESULTS

Characteristics of the teams' operating activity

The three teams correctly and quickly diagnosed (D1) the facility as being in state E1: the malfunction of the process controller (partial cooling). In the studied scenario, they didn't immediately have the means to diagnose the starting up of the diversified automatism (D2), even if they had identified the E2 installation state: a cooling in spite of the automatism failure. From the operation point of view, impacts on the general goal achievement are different, i.e. in the cooling of the installation. The cognitive analysis of the argumentation and decision-making process is thus relevant to be able to understand some of the difficulties met by the teams.

Role of argumentation in the activity of the teams

In this section, we will describe the *argumentative loops* developed by the three teams as they tried to resolve the problem. We will analyse the differences and similarities by team.

Regarding the first argumentative loop

As regards the first *argumentative loop*, which follows detection of the malfunction (partial cooling system process controller failure) and is shown in figure 2 for each team, we can make the following points:

- There are no *arguments about procedures* for teams B and C
- In contrast, team A has several *arguments about procedures*. This can be explained by the fact that, given that the facility was identified as being in state E1 (not cooling), this team quickly set itself an aim (compensate for the malfunction by manually triggering the cooling system process controller), at the same time as the actions of the AOP. From this point on, the arguments of team A, in the first loop (and the next one), were more arguments about procedures than arguments aimed at evaluating and understanding the situation and its consequences (two of the three discussions are arguments about procedures).

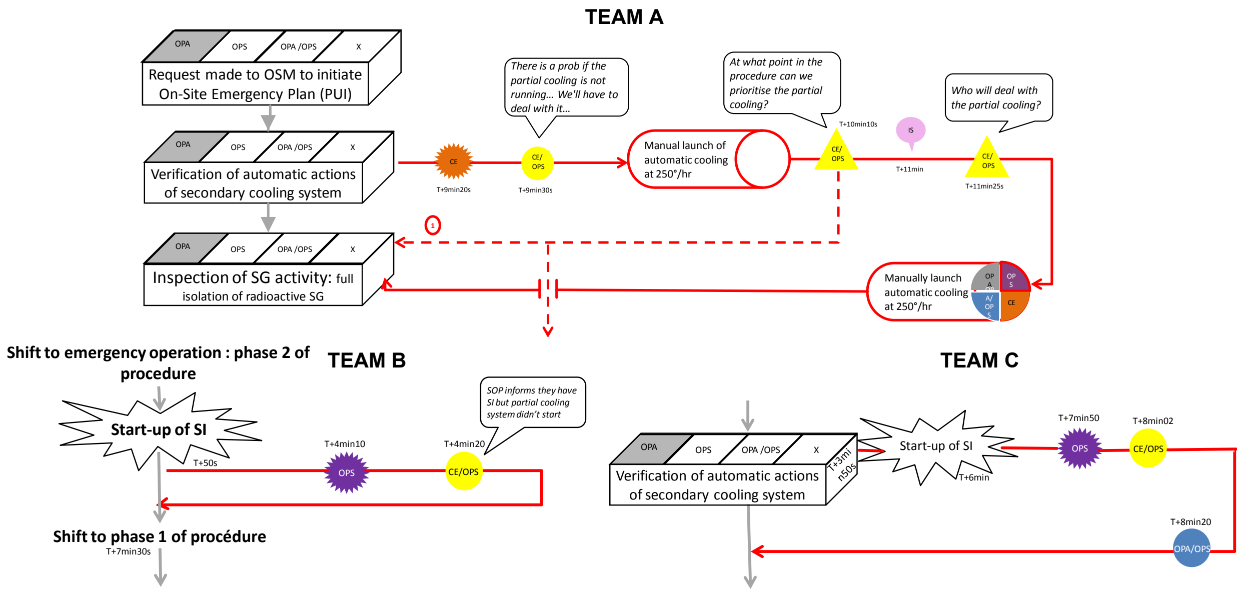

Figure 2. Comparing the teams' first argumentative loop

Regarding all of the argumentative loops developed during the scenario

Analysis of the argumentative process in each team — Team A held twice as many arguments (there are 32 in total) to deal with the cooling system malfunction than team B (18 arguments) and team C (16 arguments). Each argumentative loop developed by team A includes the formulation of one or more aims. Team B, which in the simulated situation achieved the aim too late to be efficient, only set itself one aim at the same time as applying the procedure (delegation of the monitoring of an action to the AOP). Team C devised or prioritised two aims, only one of which was implemented.

Number of arguments and prioritising of the teams — For all three teams, arguments about cooling were principally held between the SOP and the OSM. There were nonetheless ten times fewer SOP/OSM discussions in team B than in the other teams. Team B did not prioritise achieving this aim at any point. During the debriefings, the SOP and OSM confirmed their intention "not to focus on that and slow down operations, given that we knew the instructions indicated that we would have to do it later (cooling to compensate for the malfunction of the process controller)". Achieving this aim (cooling) was therefore not their main concern. Indeed, at the same time as managing the malfunction, we can see that team B held several discussions about resolving another problem: the rising water level in the radioactive steam generator. Solving this second problem was given priority over the rest of the operations by assigning this task to the SOP.

Roles of team members involved in arguments about aims set or prioritised by the team — During the scenarios, all of the teams devised or prioritised aims, at the same time as addressing the aims required by the procedure. Depending on the situation, this consisted of:

- A totally separate aim or sub-aim (for example, the aim of implementing cooling or the sub-aim of cooling with the four SGs). This was the case for teams A and C;

Or evaluating whether the targeted state had been properly attained, especially when the facility's response time was longer. This was the case for teams B and C. Typically, this involved monitoring the outcome of actions such as cooling or even depressurisation. This evaluation sometimes involved performing adjustments when the state of the facility did not totally match the targeted state.

The aims consisting of evaluating whether the targeted state had been properly attained systematically followed argument between the AOP and the SOP, not involving the OSM, whereas devising a totally separate aim or sub-aim followed argument between the SOP and the OSM. This pattern of behaviour, which may be described as defining a totally separate aim or sub-aim following a SOP-OSM discussion and deciding whether the targeted state has been properly attained following a AOP-SOP discussion, was also seen in other cases of problem-solving not covered in this study (rapid rise in the pressuriser level and rapid rise in the water level of the radioactive steam generator).

DISCUSSION AND CONCLUSION

The description of the "decision pathway" relating to a specific malfunction, in terms of the operating activity and the argumentative loops developed by the teams, can be used to shed light on the way in which a team takes decisions when faced with an ambiguous or poorly defined situation.

This "decision pathway" differs from one team to another. For example, when faced with an unexpected event, in this case the malfunction of the partial cooling system, a team may decide to act quickly by devising new aims, by reassigning its resources to implement these aims. The members of team A therefore hold arguments more frequently and set themselves aims to deal with the malfunction quickly. On the other hand, in terms of safety, the first attempt to achieve the aim does not subsequently appear to be the most suitable even though it resulted from a discussion between the SOP, OSM and SE.

Team B, however, preferred to wait and see if time and/or the procedure would resolve the problem before acting. For this team, the attempt to achieve the aim ended up being no longer relevant (not useful) given the state of the facility. We can draw an analogy here with the conclusions of Amalberti (2001) regarding the behaviour of operators faced with errors in dynamic and complex situations, even though in our case, we are not concerned with errors but rather with the specific state of the facility: many detected errors are not dealt with by operators because they can resolve themselves over time, as the situation changes. It it therefore necessary to weigh up the risks incurred by the specific state of the facility, possible changes to the situation and the resources involved (physical and cognitive) if action is to be taken. Team B did not weigh up the situation correctly.

More specific analysis of the roles taking part in discussion of the aims set or prioritised by the team also enables us to formulate hypotheses about the degree of abstraction and distance from the action that the various roles are likely to have, in ambiguous situations. Therefore, when devising aims close to operational tasks, in this case delegating monitoring of the outcome of an action that the AOP has just performed, it would appear that cooperation between the involved parties occurs more through horizontal controls (De La Garza & Weill-Fassina, 2000) between the AOP and the SOP and through "task-oriented cooperation" as described in Hoc's model (2001). The hypothesis that a degree of "task-oriented cooperation" seems to make sense given the nature of the functions of the AOP (completion of operating tasks) and the SOP (checking that operating tasks have been completed properly). As far as devising the aim is concerned, vertical controls (De La Garza & Weill-Fassina, op.cit.) between the SOP and the OSM would enable them to coordinate themselves with a greater degree of abstraction, similar to the level of "planning-oriented cooperation" (Hoc, 2001). In the sequences analysed, the AOP was not involved at this second level of cooperation at any point. Meanwhile, the SOP was actively involved at both levels of cooperation.

In all of the scenarios, we noted that discussions predominantly involved two member. Just one discussion involved three member (each operator exchanging information) and there were no discussions involving four operators. The AOP appeared to be unable to be a resource involved in decision-making discussions because he seemed to busy with his operating activities. However, studies in the nuclear industry, as in other sectors, have shown how collective decision-making can make a socio-technical system more robust. In the rest of the study it will be necessary to specify the organisational characteristics that promote or handicap collective discussion, and which therefore lead to more safe and efficient operation.

The "decision pathway" is dependent on the individual mental representation of the situation and on what their colleagues know and do. It is therefore essential to take an in-depth look at the types of arguments put forward by the various roles. Analysis of these arguments will enable the dynamics of the roles in this new team organisation to be described more effectively in the process of cooperation, and in situations where requirements and technical support systems are not always enough to guide operators (in "unconsidered situations"). However, taking into account the fact that some of the operating methods are still in the process of being designed, the situations identified as "ambiguous" or poorly "defined" will be reviewed beforehand in order to identify any points that may be improved and therefore contribute to the reliability of the socio-technical system.

GLOSSARY

AOP: Action Operator
ASDS: Auxiliary Steam Distribution System
HMI: Human-Machine Interface
IE: containment isolation
OSM: Operations Shift Manager
P12: permissive 12
PAF: switch to cold shutdown
PCD: command post
PUI: On-Site Emergency Plan
SE: Safety Engineer
SG: Steam Generator
SI: Safety Injection
SOP: Strategic Operator
TBS: Turbine Bypass System
VDA: atmospheric steam dump system

REFERENCES

Amalberti, R. (2001). La maîtrise des situations dynamiques. *Psychologie Française,* 46 (2), 105-117.

Bourgeon, L. (2011). *Mécanismes cognitifs et rôle du collectif dans la persévération : gestion des évènements imprévus dans l'activité de pilotage militaire*. Thèse de Psychologie Ergonomie, Université Toulouse 2 Le Mirail, Toulouse.

Burke, C.S., Cannon-Bowers, J.A., & Salas, E. (2000). Teamwork: emerging principles, *International Journal of Management Review,* 2(4), 339-356.

Carvalho, P.V.R., dos Santos, I.L., & Vidal, M.C.R. (2005). Nuclear power plant shift supervisor's decision making during microincidents. *International journal of industrial ergonomics, 35, 619-644.*

Chung, Y.H., Yoon, W.C., Min, D. (2009). A model-based framework for the analysis of team communication in nuclear power plants. *Reliability Engineering and System Safety*, 94, 1030-1040.

Crichton, M.T., & Flin, R. (2004). Identifying and training non–technical skills of nuclear emergency response teams. *Annals of Nuclear Energy*, 31, 1317-1330.

Darses, F. (2006). Analyse du processus d'argumentation dans une situation de reconception collective d'outillages*. Le Travail Humain*, 69 (4), 317-347.

De la Garza, C., Labarthe, J-P., & Graglia, L. (2012). The contribution of ergonomics to risk analysis in the design process: the case of a future control room. *Work: A Journal of Prevention, Assessment and Rehabilitation,* 41, 730-736.

De la Garza C., Le Bot P. (2008). *Emergency operation in nuclear power plants : proposition of analysis protocol to highlight collective team performance in simulation situations*. International conference on Probabilistic Safety Assessment and Management, Proceedings of the PSAM9, Hong Kong, China.

De la Garza, C., Weill-Fassina, A. (2000). Régulations horizontales et verticales du risque. In T.H. Benchekroun, & A. Weill-Fassina (Eds.), *Le travail collectif. Perspectives actuelles en ergonomie* (pp. 217-234). Toulouse, France : Octarès.

Endsley, M.R. (1995b). Toward a theory of situation awareness in dynamic systems. Human Factors 37(1), 32–64.

Hélie, P. & Loiselet, A. (2000). *Human-human cooperation in airborne combat system: Designing CSCW assistances*. Paper presented at *6th S.E. Ergonomics Society Conference and 4th Asia Pacific Conference on Computer Human Interaction*. Singapore, Nov.

Hoc, J-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *Int. J. Human-Computer Studies,* 54, 509-540.

Hollnagel, E., & Woods, D.D. (2005). Joint cognitive systems: Foundations of cognitive systems engineering. Broken Sound Parkway : CRC press Taylor and Francis Group.

Kim, M.C., Park, J., Jung, W., Kim, H., & Kim, Y.J. (2010a). Development of a standard communication protocol for an emergency situation management in nuclear power plants. *Annals of Nuclear Energy, 37, 888-893.*

Kim, S., Park, J., Han, S., & Kim, H. (2010b). Development of an extended speech act coding scheme to observe communication characteristics of human operators of nuclear power plants under abnormal conditions. *Journal of loss prevention in the process industries, 23, 539-548.*

Klein, G. & Zsambok, C.E. (1997). Naturalistic Decision Making. New Jersey: Lawrence Erlbaum Associates.

Labarthe, J-P., & De La Garza, C. (2011). The human factors evaluation program of a control room: The French EPR approach. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 21, 331-349.

Lee, S. W., Park, J., Kim, A. ryum, & Seong, P. H. (2012). Measuring situation awareness of operation teams in NPPs using a verbal protocol analysis. *Annals of Nuclear Energy*, *43*, 167–175.

Park, J., Jung, W., & Yang, J-E. (2012). Investigating the effect of communication characteristics on crew performance under the simulated emergency condition of nuclear power plants. *Reliability Engineering and System Safety, 101, 1-13.*

Pelayo, S. (2007). *D'une coopération verticale à une planification coopérative des actions : le cas de la gestion des prescriptions thérapeutiques hospitalières*. Thèse de Psychologie, Lille : Université de Lille 2.

Rasmussen, J., Pejtersen, A.M., & Schmidt, K. (1990). Taxonomy for cognitive work analysis. Roskilde, Denmark: Risø National Laboratory. (Risø report M-2871).

Stachowski, A.A., Kaplan, S.A., & Waller, M.J. (2009). The benefits of flexible team interaction during crises. *Journal of Psychology, 94 (6), 1536-1543.*

Subjective workload and performance of young drivers faced to unexpected pedestrian crossings

Julie PAXION[a, b], Chloé FREYDIER[a], Edith GALY[b] and Catherine BERTHELON[a]
[a] *French Institute of Science and Technology for Transport (IFSTTAR)*
[b]*Aix-Marseille University (AMU)*

ABSTRACT

Introduction: The aim of the present study is to identify which of subjective workload dimensions are influenced by driving experience and situation complexity, and which ones influence driving performance. **Method**: Fifty-seven young drivers (15 traditionally trained novices, 12 early-trained novices, 15 with three years of experience and 15 with at least five years of experience) were randomly assigned to three situations (simple, moderately complex and very complex) in a driving simulator. Self-reported levels of workload during unexpected pedestrian crossings were collected by a questionnaire (NASA-TLX) between each situation. **Results and discussion:** Three workload subscales decreased with experience and increased with situation complexity, and one subscale was correlated to driving performance.

KEYWORDS

Subjective workload; driving experience; situation complexity.

INTRODUCTION

Epidemiological studies show that young novice drivers have a risk of accident two to four times higher than young experienced drivers (Triggs, 2004). Driving impairments come especially for novices who often have a lack of routine automation (De Craen, Twisk, Hagenzieker, Elffers, & Brookhuis, 2008; Fuller, 2002). In the complex system of traffic (environment – road user – vehicle), the origin of failures mainly come from interactions between the three elements of this system (Van Elslande, 2003). Interactions between environment and road users are particularly interesting inasmuch 90 % of road accidents are due to human errors (Amditis, Andreone, Pagle, Markkula, Deregibus, Rue, Bellotti, Engelsberg, Brouwer, Peters, & De Gloria, 2010). Indeed, subjective workload, i.e. cost perceived depending on the interaction between tasks required and drivers' state (Hockey, 2003), can damage performance if its level is either too high (overload) or too low (underload) (Meister, 1976 ; in De Waard, 1996). Situation complexity provokes thus more or less workload which is perceived differently depending on individuals. Organizing knowledge into schemas with learning provokes a task automation leading to a decrease of workload (Tricot, 1998). It can also lead to a cognitive readiness by anticipating and scheduling situations already known, which is necessary to take an efficient decision, especially in complex situations (Cegarra, & van Wezel, 2012). Conversely, if the activity is not entirely automatized, performing the task implies producing a higher effort. For complex tasks, despite subjective workload enhancement, driving performance can be maintained because compensatory mechanisms are gradually set up with practice (Amalberti, 1996; Cegarra & Hoc, 2006). Nevertheless, required effort can be too high for individual's abilities and can therefore generate overload characterized by an increase of workload leading to performance impairments. Thus, for the same driving situation, the activity can be controlled, or automatized throughout experience enhancement, which leads to a higher effort for novices than for experienced drivers (Patten, Kircher, Östlund, Nilsson, & Svenson, 2006). Therefore, the threshold from which drivers perceive an overload not only depends on situation complexity but also on skills acquired during the driving.

Moreover, some aspects of subjective workload could be more or less influenced by driving experience, and could more or less influence driving performance. Indeed, Hart and Staveland (1988) identified six combinations of relevant factors which characterize subjective workload: mental demands (amount of mental and perceptual activity required), physical demands (amount of physical activity required), temporal demands (amount of pressure felt due to the rate at which the task elements occurred), own performance (successful assessment in doing the task required and satisfaction assessment in accomplishing it), effort (difficulty assessment in having to work mentally and physically to accomplish the level of performance) and frustration (assessment of feeling insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent during the task). According to Gaillard (1993; in Collet, Averty, & Dittmar, 2009), effort and stress are two components that determine workload, with stress enhancement which could increase effort depending on task demands and individuals. Therefore, our hypothesis is that both frustration (associated to stress) and effort

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

should increase with situation complexity and decrease with driving experience. Moreover, task difficulty (mental and physical demands) is associated to task demands and should also depend on situation complexity and driving experience. Conversely, temporal demands should be influenced by these two factors through subjective anxiety enhancement, as time pressure imposed by the situation first activates the emotion and have thus an indirect effect on workload (Monod & Kapitaniak, 1999; in Galy, Cariou, & Mélan, 2012). Finally, performance estimation should be positively correlated to objective driving performance.

METHOD

Participants

Fifty-seven young drivers (33 males and 24 females) were divided up into four groups depending on their driving experience. Two groups were composed of novice drivers who obtained their driving license within the last two months, with 15 Traditionally Trained Drivers (TTD)[11] aged between 18-20 (M = 19, SD = 0.84) and 12 Early-Trained Drivers (AAC – Apprentissage Anticipé de la Conduite[12] aged 18. The two other groups consisted in 15 drivers aged 21 arriving at the End of the three-year Probationary Period (EPP)[13], and 15 Experienced Drivers (ED) aged between 23-30 (M = 26.87, SD = 2.97) with at least five years of experience.

Experimental setup

The experiment was carried out on the SIM²-IFSTTAR fixed-base driving simulator equipped with an ARCHISIM object database (Espié, Gauriat & Duraz, 2005). The driving station comprised one quarter of a vehicle (see Figure 1). The image projection (30 Hz) surface filled an angular opening that spanned 150° horizontally and 40° vertically. The vehicle had an automatic gearbox and was not equipped with rearview mirrors.

Figure 1. Driving simulator

Procedure and statistical analyses

Participants drove on three different rural driving situations (22.5 kms each) in a counterbalanced order. The simple situation consisted in a straight national road with two ways, without any traffic. The moderately complex situation included right and left curves (length: 600 m, radius: 300 m). The most complex situation had double and sharper curves (length: 300 m, radius: 120 m), with oncoming traffic. Whatever the situation was, three scenarios implying a pedestrian were included. The pedestrians, hidden by a billboard, a bus stop or a tree in a random order, crossed the road around 2.7 seconds before the participant arrived at his level. The instructions were to drive at 90 km/h.

The NASA-TLX questionnaire (Hart & Staveland, 1988) assessed the subjective level of workload after each circuit. It comprised 6 subscales: Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration. For each subscale, participants estimated their workload during the last circuit on a 20 points scale ranged from 0 = ‘very low’ to 20 = ‘very high’. For Own Performance dimension, the scale was ranged from 0 = ‘success’ to 20 = ‘failure’. The questionnaire has been modified in order to separate subjective workload associated with the different portions of the three circuits. For each subscale of the questionnaire, the question was attributed to the pedestrian crossings.

Repeated measures ANOVA were carried out in order to test the effects of driving experience and situation complexity on each subjective workload dimension attributed to pedestrian crossings. Bonferroni post-hoc tests were subsequently used for pairwise comparisons. Linear regressions analyses then tested the effect of each subjective workload dimension (except from Own Performance dimension) on the number of collisions with pedestrians (performance). The link between subjective and objective performance was tested via a Bravais-

[11] TTD: 20 hours of driving lessons with an instructor.

[12] AAC: 20 hours of driving lessons with an instructor and an additional driving practice with an adult during 3,000 km., driving learning permitted to start at the age of 16.

[13] EPP: from the driving license exam, partial license during three years with restrictions as speed limitation and only 6 points instead of 12 points.

Pearson's correlation. For all analyses, statistical significance was set at p≤.05. To clarify the reading, non significant effects are not presented.

RESULTS

Effects of driving experience on each subjective workload subscale

Driving experience significantly decreased Own Performance (F(3,53)=6.54; *p<.001*), Effort (F(3,53)=2.91; *p<.05*) and Frustration (F(3,53)=4.62; *p<.01*). Own Performance was higher for Early-Trained Drivers (AAC) than for drivers at the End of the Probationary Period (EPP) and Experienced Drivers (ED). Effort and frustration were higher for AAC then for ED (See Table 1 below).

Table 1. Each subjective workload subscale for each group

Subscales	TTD, Mean (SD)	ETD, Mean (SD)	EPP, Mean (SD)	ED, Mean (SD)
Mental demands	13.42 (4.83)	15.92 (3.94)	13.47 (4.46)	12.51 (4.64)
Physical demands	12.42 (4.97)	14.06 (5.15)	11.29 (5.44)	9.47 (4.58)
Temporal demands	13.33 (5.13)	15.53 (5.22)	14.38 (4.24)	12.71 (4.56)
Own performance	8.73 (5.73)	12.42 (6.02)	7.27 (4.90)	6.64 (3.41)
Effort	12.91 (4.63)	15.50 (3.88)	13.13 (4.56)	11.38 (4.67)
Frustration	11.20 (5.71)	13.19 (5.67)	9.00 (4.98)	7.80 (4.41)

Effects of situation complexity on each subjective workload subscale

Situation complexity had a significant effect on Mental Demands (F(2,106)=3.91; *p<.05*), Physical Demands (F(2,106)=6.05; *p<.01*) and Effort (F(2,106)=3.77; *p<.05*). Indeed, each of these three dimensions was lower in simple situation than in very complex situation. Physical Demands were also lower in simple situation than in moderately complex situation (see Table 2 below).

Table 2. Each subjective workload subscale for each situation

Subscales	Simple, Mean (SD)	Moderately complex, Mean (SD)	Very complex, Mean (SD)
Mental demands	12.68 (5.02)	14.02 (4.36)	14.46 (4.37)
Physical demands	10.61 (5.17)	12.02 (5.44)	12.44 (5.07)
Temporal demands	13.23 (5.22)	13.86 (4.77)	14.63 (4.53)
Own performance	7.49 (4.97)	9.00 (5.09)	9.23 (6.16)
Effort	12.16 (4.71)	13.49 (4.46)	13.68 (4.72)
Frustration	9.89 (5.39)	10.32 (5.57)	10.23 (5.71)

No interaction effect between driving experience and situation complexity were observed on subjective workload dimensions.

Subjective workload subscales and collisions' number relationships

In this model, none of the subjective workload dimensions tested influenced the number of collisions. However, the correlation analysis showed that performance was more assessed as a failure when the number of collisions increased (r = .61 ; p < .001) (see Figure 2 below).

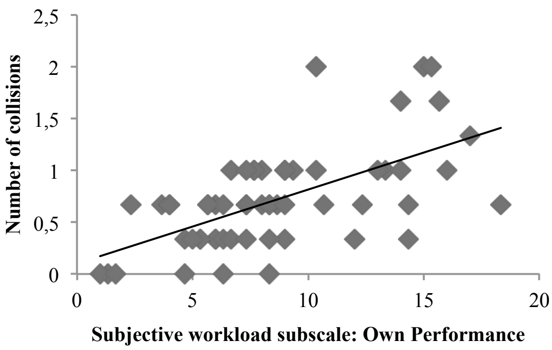

Figure 2. Correlation between Own Performance and collisions' number

DISCUSSION

Three subjective workload subscales (Own Performance, Effort and Frustration) decreased with driving experience. Early-Trained Drivers assessed thus more their own performance as a failure than more Experienced Drivers, probably because they had more collisions with pedestrians than more Experienced Drivers who have more knowledge to take an efficient decision (Cegarra, & van Wezel, 2012). Traditionally Trained Drivers didn't consider they failed more than the other groups. They probably overvalued their abilities and undervalued their accident risk (optimism bias, see McKenna, 1993) compared to more experienced drivers who met more hazardous events, helping them to develop situation awareness (Underwood, Chapman, Bowden, & Crundall, 2002). Early-Trained Drivers assessed that they made more effort in accomplishing the task required than Experienced Drivers. It thus reveals that from five years of experience, drivers have sufficiently automatized the driving to have more available resources to deal with unexpected situations as a sudden pedestrian crossing. Moreover, Early-Trained drivers also assessed they had more frustration than Experienced Drivers. This result indicates that as drivers think they don't make a great effort, they feel quite relaxed during the task. More data about their subjective anxiety should reveal that Experienced Drivers feel less anxious and therefore less frustrated than novice drivers. Traditionally Trained Drivers should make more effort and should be thus more frustrated than Experienced Drivers inasmuch novices haven't automatized all driving tasks (De Craen et al., 2008; Fuller, 2002). A comparison between objective and subjective workload could show whether the absence of difference between these groups is due to a false self-assessment from novices or not. However, mental and physical demands didn't vary as a function of driving experience, revealing a high variability between the groups.

As expected, the three subscales Mental Demands, Physical Demands and Effort increased with situation complexity. Nevertheless, situation complexity didn't provoke a variation of frustration, which is probably due to the same pedestrian crossing scenario for the three situations.

According to our hypothesis, own performance was positively correlated to driving performance.

Training in simulator with complex and/or unexpected situations could therefore help novice drivers to construct a mental representation of these situations and thus automatize driving tasks in order to decrease their subjective workload and improve their performance.

AKNOWLEDGEMENTS

We thank the simulation team at IFSTTAR (LEPSIS: Laboratoire Exploitation, Perception, Simulateurs et simulations), notably Isabelle Aillerie for designing the displays.

REFERENCES

Amalberti, R., (1996). *La conduite des systèmes à risques*. Paris: Presse Universitaires de France (PUF).

Amditis, A., Andreone, L., Pagle, K., Markkula, G., Deregibus, E., Rue, M.R., Bellotti, F., Engelsberg, A., Brouwer, R., Peters, B., De Gloria, A., (2010). Towards the Automotive HMI of the Future: Overview of the AIDE-Integrated Project Results. *IEEE Transactions on Intelligent Transportation Systems, 11 (3)*, 567–578.

Cegarra, J. & Hoc, J-M. (2006). Cognitive styles as an explanation of experts' individual differences: A case study in computer-assisted troubleshooting diagnosis. *International Journal of Human-Computer Studies, 64(2),* 123-136.

Cegarra, J., & van Wezel, W. (2012). Revisiting decision support systems for cognitive readiness: A contribution to unstructured and complex scheduling situations. *Journal of Cognitive Engineering and Decision Making*, *6(3)*, 299-324.

Collet, C., Averty, P., & Dittmar, A. (2009). Automatic nervous system and subjectives ratings of strain in air traffic controllers. *Applied Ergonomics, 40*, 23-32.

De Craen, S., Twisk, D.A.M., Hagenzieker, M.P., Elffers, H., & Brookhuis, K.A. (2008). The development of a method to measure speed adaptation to traffic complexity: Identifying novice, unsafe, and overconfident drivers. Accident Analysis and Prevention, 40 (4), 1524-1530.

De Waard, D. (1996). The measurement of drivers' mental workload. Traffic Research Center, Thesis, 127 p.

Espié, S., Gauriat, P., & Duraz, M. (2005, September). Driving simulators validation: The issue of transferability of results acquired on simulator. Paper presented at the Driving Simulation Conference North-America (DSC-NA 2005), Orlondo, FL.

Fuller, R. (2002). Human factors and driving. Human Factors for Highway Engineers, 77-97.

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiologie*, 83, 269-275.

Hart, S. G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Emperical and Theoritical Research. Advances in Psychology, 52, 139-183.

Hockey, G. R. J. (2003). Operator Functional State as a Framework for the Assessment of Performance Degradation. In: G.R.J. Hockey, A.W.K. Gaillard & O. Burov (Eds.), Operator Functional State: The Assessment and Prediction of Human Performance Degradation in Complex Tasks. Amsterdam: IOS Press, pp. 8-23.

McKenna, F. P. (1993). It won't happen to me: unrealistic optimism or illusion of control? *British Journal of Psychology, 84,* 39-50.

Patten, C. J. D., Kircher, A., Östlund, J., Nilsson, L. & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. Accident Analysis & Prevention, 38(5), 887-894.

Tricot, A. (1998). Charge cognitive et apprentissage. Une presentation des travaux de John Sweller. *Revue de Psychologie de l'Education, 1,* 37-64.

Triggs, T.J. (2004). Simulation evaluation of driver performance changes during the early years of driving. Proceedings of the driving simulation conference, Paris, France, 421-430.

Underwood, G., Chapman, P., Bowden, K. & Crundall, D. (2002). Visual search while driving: Skill and awareness during inspection of the scene. *Transportation Research Part F*, *5(2)*, 87-97.

Van Elslande, P. (2003). Erreurs de conduite et besoins d'aide : une approche accidentologique en ergonomie. *Le travail humain, 66*, 197-226.

Improvisation in Safety Critical Situations: An Analysis of Improvisation Incidents in Led Outdoor Activities

Margaret TROTTER[a], Paul SALMON[b] and Michael LENNÉ[a]
[a]Monash Injury Research Centre, Monash University, Melbourne, VIC 3800, Australia
[b]University of the Sunshine Coast Accident Research (USCAR), Faculty of Arts and Business, University of the Sunshine Coast, Maroochydore, QLD 4558, Australia

ABSTRACT

Introduction: Appropriate, effective improvisation can enhance organisations' resilience in safety critical situations; however, conditions across work systems that influence improvisation are not well understood. This PhD aims to identify the factors influencing improvisation from a systems perspective. **Method**: 20 Critical Decision Method interviews were conducted with led outdoor activity leaders and the transcripts analysed using Leximancer™ software. **Results and discussion:** 35 concepts across 12 major themes were identified. The themes confirmed influencing factors identified in the literature (e.g. decision making, knowledge, situation awareness) as well as several important new themes (e.g. time and students). The concept map analysis demonstrates the important links between concepts and themes that will provide guidance for the construction of a systems model of improvisation.

KEYWORDS

Decision-making; outdoor recreation, improvisation, systems safety, critical decision method.

INTRODUCTION

Improvisation – the spontaneous real-time conception and execution of a novel response to a situation that is beyond the boundaries for which a system has prepared (Trotter, Salmon & Lenné, 2012) – provides a means by which organisations can adapt in response to safety critical situations (Grøtan, Størseth, Rø & Skjerve, 2008). Improvisation can therefore enhance organisational resilience; however, how to support appropriate and effective improvisation remains ambiguous. While improvisation can mitigate catastrophic systems failure, such as in the ditching of US Airways flight 1549 (National Transportation Safety Board, 2010), it can also produced less favourable outcomes, such as in the Mangatepopo Gorge incident in which seven people died (Brookes, Smith & Corkill, 2009). Determining how organisations can enhance appropriate, effective improvisation while preventing inappropriate, ineffective improvisation requires an understanding of the factors that influence improvisation in safety critical situations; however, as yet there is no comprehensive model of these factors and their interactions (Trotter et al., 2012).

The existing improvisation research examines influencing factors independently rather than from a systems-based perspective as would be in line with contemporary ergonomics approaches (e.g. Rasmussen, 1997; Leveson, 2011). As a result, the conditions across the system that influence improvisation are not well understood. This represents a significant research gap, which currently prevents the concept from being realised within safety critical systems. To address this I devised a five stage research agenda aimed at identifying and modelling the factors and their interactions influencing improvisation from a systems perspective.

- **Stage 1** critiques existing knowledge of improvisation in safety critical situations and the extent to which it has been investigated from a systems perspective (Trotter et al., 2012).
- **Stage 2** examines the appropriateness of a particular systems-based framework for the study of improvisation using a case study methodology to examine two improvisation incidents (Trotter, Salmon & Lenné, submitted 2012).
- **Stage 3** uses a survey methodology to determine the relevance of the factors identified from the literature to a particular complex sociotechnical system, the led outdoor activity (LOA) domain (Trotter, Salmon & Lenné, submitted 2013).
- **Stage 4** identifies the factors influencing improvisation across the LOA system using in-depth analysis of Critical Decision Method interviews (CDM; Klein, Calderwood & MacGregor, 1989) focused on LOA improvisation incidents in order to produce a systems-based model improvisation in this domain.
- **Stage 5** establishes how LOA organisations can best support appropriate, effective improvisation.

This paper outlines a preliminary analysis of the CDM data from Stage 4 of the research program.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

METHOD

Participants

Participants were eleven male and nine female employees of a major Australian LOA organisation with a mean age of 29 years (SD 4.6). Participants were outdoor leaders currently working in the field with on average 5.7 years of LOA experience (SD 3.7).

Procedure

The CDM method (Klein et al., 1989) was adapted to study improvisation by having the interviewees focus specifically on instances of improvisation and including a series of improvisation-specific probes. Interviews lasted approximately 2 hours and were conducted individually by one interviewer. Each interview consisted of four stages:

- **1. Identification:** Interviewer determined that participants' incidents fit the improvisation definition.
- **2. Summary**: Participants summarised their improvisation;
- **3. Timeline**: Interviewer created a timeline of the improvisation;
- **4. Probes:** Participants responded to a series of probes about the improvisation;

All interviews were recorded and then transcribed using Microsoft Word.

Data Analysis

A content analysis of the transcripts was performed using the text analysis software tool Leximancer™. Leximancer™ identifies key concepts and relationships, and then creates semantic networks from raw data. Concept names and themes are derived based on co-occurrence in the texts. A statistical algorithm is used to create a two dimensional concept map. The initial Leximancer™ analyses were exploratory so that concepts were extracted from the transcripts automatically. Following this, similar concepts were combined manually and the analysis run again based on these revised concept lists. Themes were automatically produced and mapped by Leximancer™. Analyses were produced for each individual case as well as for the aggregate of all cases. This paper focuses on the aggregate analysis.

RESULTS

Incidents ranged in potential severity from minor (e.g. crossing a ford safely without appropriate equipment) to life threatening (e.g. facing a hurricane during a kayaking trip). All improvisation incidents recounted by participants resulted in positive safety outcomes for participants. The aggregate analysis yielded 35 emergent concepts (Table 1) with 'Knowledge' and 'Time' the most relevant. These concepts were also relevant in 18 and 17 individual cases respectively. Leximancer™ identified 12 dominant themes within the aggregate text, labelling them according to the central concept within each (Table 2). 'Decision' was the most highly connected theme, followed by 'Situation' and 'Knowledge'. These themes also appeared in seven, five, and nine of the case-specific concept maps respectively. The connections between themes and concepts are illustrated in the concept map (Figure 2).

Table 1. Concepts identified from the aggregate analysis

Concept	Count	Relevance*	Concept	Count	Relevance
Knowledge	282	100%	Use	106	38%
Time	272	96%	Option	94	33%
Students	210	74%	Someone	90	32%
Decision	183	65%	Day	90	32%
People	175	62%	Person	86	30%
Situation	168	60%	Risk	76	27%
Ability	164	58%	Level	68	24%
Work	159	56%	Area	68	24%
Experience	150	53%	Rescue	64	23%
Group	141	50%	Thinking	58	21%
Training	140	50%	Trip	56	20%
Happen	132	47%	Comfortable	52	18%
Organisation	124	44%	Plan	52	18%
Water	123	44%	Place	51	18%
Different	116	41%	Better	46	16%
Felt	114	40%	Everything	38	13%
Take	112	40%	School	37	13%
Look	111	39%			

* The relevance percentage represents the percentage frequency of text segments coded with that concept, relative to the most frequent.

Table 2. Themes and their component concepts from the aggregate analysis

Theme	Connectivity	Concepts
Decision	100%	Decision, happen, group, felt, comfortable, thinking
Situation	99%	Situation, work, different, person, place
Knowledge	84%	Knowledge, people, someone, better
Training	63%	Training, water, rescue, level
Experience	62%	Experience, ability, use
Time	59%	Time, take, everything
Students	25%	Students
Organisation	23%	Organisation, risk
Option	22%	Option, area
Look	20%	Look, plan
Trip	16%	Trip, school
Day	20%	Day

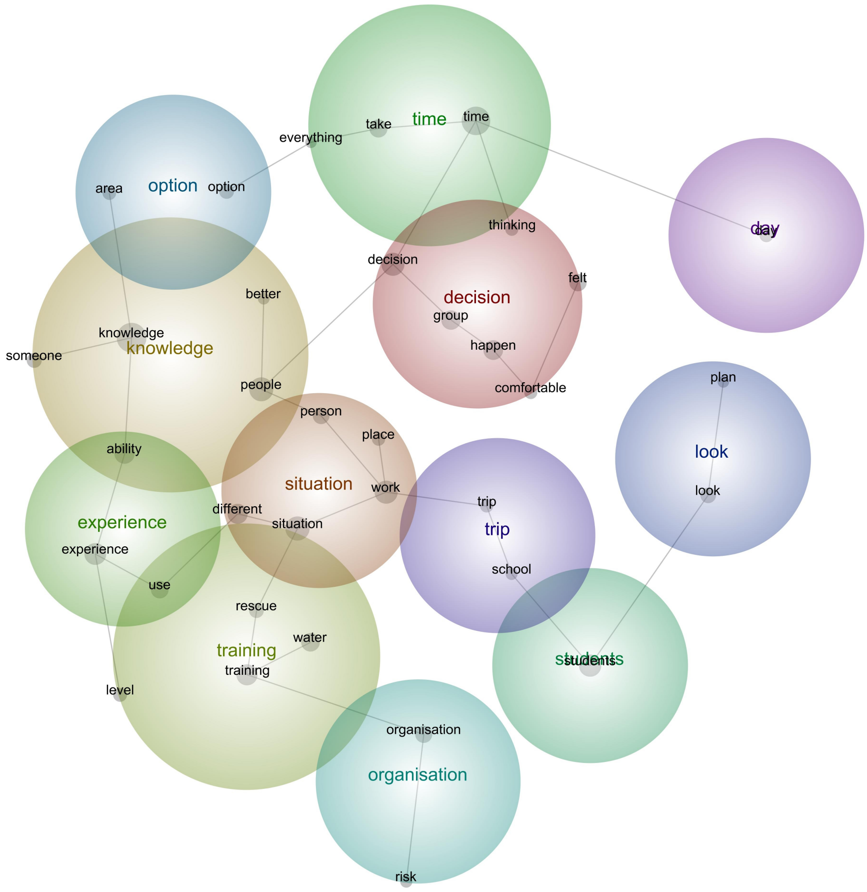

Figure 2. Aggregate analysis concept map (theme names in large font). Note: The size of each node represents the centrality of the concept, while the pathways denote the most likely connections between concepts.

Key Themes

Decision Theme

The 'Decision' theme includes concepts associated with participants' decision-making processes. The prominence of this theme is unsurprising given than decision-making is key component of improvisation (Mendonça, 2007). The concepts 'Thinking' and 'Felt' are included in this theme. Many participants reported thinking through the consequences of their improvised solutions; however, many also referred to unconscious processes such as intuition, particularly in relation to determining when a situation required improvisation, for example:

> *"My gut feeling, intuition, which was saying this situation isn't right, it isn't good, I don't feel that these students are safe…"*

Interestingly, the concepts 'Decision' and 'Thinking' are directly connected to the 'Time' concept' in the network whereas 'Felt' is not. This indicates that the time available affects improvisers' ability to exhaustively think through options and consequences in order to make a decision is more than it effects intuitive processes. In very time constrained situations, participants may rely more on intuitive decision-making.

Situation Theme

The 'Situation' theme reflects the 'Person', 'Place' and 'Work' factors apparent in a situation; for example, *"It was a combination of elements that made the situation different: the weather the slipperiness, the particular person and the location that meant you couldn't see."*

These concepts resemble the three internal factors identified in models of outdoor decision-making (e.g. Boyes & O'Hare, 2003). A key differentiating concept within this theme is 'Different', which relates to aspects of the situation that are different to what participants have previously experienced. The 'Situation' theme overlaps with, 'Knowledge', and Training', in particular, with knowledge of leaders' and students' abilities, and rescue training.

Knowledge Theme

The 'Knowledge' theme reflects knowledge of the people involved in the situation, those available to provide assistance, the location, different activities and techniques. It also represents knowledge of the reasons behind the rules, for example:

> *"Having that good sound knowledge base allows you to know the rules and where you have got room to move."*

This theme overlaps with the 'Experience' theme, particularly through the concept of ability. Many participants reported that their knowledge of student abilities develops over the course of programs and that their knowledge of fellow leaders' abilities came from working with them previously. Also overlapping is the 'Options' theme, particularly around knowledge of the area and the options that it affords, for example, the location of access points.

DISCUSSION

This research forms part of a program aimed at identifying and modelling the factors influencing improvisation in safety critical situations from a systems perspective. This research represents preliminary analysis of CDM data from Stage 4 of this program. The Leximancer™ analysis identified 35 concepts and grouped these into 12 major themes, the most connected of which were 'Decision', 'Situation' and 'Knowledge'. These themes are consistent with those identified in the literature as influencing improvisation (Trotter et al., 2012). The concept map enhances our understanding of improvisation by identifying the dominant connections between the concepts and themes, a vital element of any systems-based model.

Another important characteristic of a systems-based model is that it reflects factors and interactions from across different systems levels. While this analysis identifies the higher-level themes 'Organisation' and 'School', nothing at the regulatory or government levels of the system is identified. This is most likely the consequence of all CDM participants being at the leader level and therefore potentially unaware of how external factors affect their organisation. To address this, semi-structured interviews with all management level staff at the same organisation were conducted and will be included in the subsequent in-depth interview analysis.

This analysis also identified a number of factors that not identified in previous research, the most dominant of these being 'Time' and 'Students'. Examination of the 'Time' theme revealed variations in the timescales over which improvised responses were executed. While some improvisation incidents required only one decision, then execution of this response over a number of minutes (e.g. rescuing a capsized kayaker), others involved multiple decisions and executions over a longer timescale (e.g. protecting kayaking students from a hurricane). This analysis indicates that time is not just a defining characteristic of improvisation but an influencing factor. This has significant implications for the definition of improvisation and will be explored further through in-depth analysis of the incident timelines.

The 'Student' theme was also identified in the Stage 3 survey and may be specific to LOA. Boyes and O'Hare (2003) include student ability in their model of outdoor adventure decision-making; however, in a more generalised model of improvisation, this factor may be subsumed by the concepts 'People' or 'Ability'.

Leximancer™ analysis is based on word frequency and co-occurrence within raw data. It provides a useful interpretive guide for an exploratory examination of influencing factors, indicating potential factors and interactions. Meaning, however, can lie not just within specific words, but also in the spaces between. In the next phase of analysis manual coding will be conducted to identify and understand factors in sufficient detail to allow the development of a comprehensive model of improvisation from which recommendations can be devised. The structure of this context map will guide model construction.

ACKNOWLEDGMENTS

Ms. Trotter's contribution to this research was supported by an APA Postgraduate Research Scholarship. Dr. Salmon's contribution was funded through the Australian NHMRC Post doctoral training fellowship scheme. The authors would like to acknowledge Dr Natassia Goode for her advice on the draft manuscript.

REFERENCES

Boyes, M. A. & O'Hare, D. (2003) Between Safety and Risk: A Model for Outdoor Adventure Decision-making, *Journal of Adventure Education and Outdoor Learning*, 3, 1, 63-75.

Brookes, A., Smith, M. & Corkill, B. (2009) Report to the Trustees of the Sir Edmund Hillary Outdoor Pursuit Centre of New Zealand, Mangatepopo Gorge Incident, 15 April 2008.

Grøtan, T., Størseth, F., Rø, M. H. & Skjerve, A. B. (2008) Resilience, Adaptation and Improvisation – Increasing Resilience by Organising for Successful Improvisation, Proceedings of the 3rd Symposium on Resilience Engineering. Antibes, Juan-Les-Pins, France. Available from: http://www.resilience-engineering.org/RE3/papers/Grotan_Storseth_Ro_Skjerve_text.pdf (21 August 2011).

Klein, G., Calderwood, R. & MacGregor, D. (1989) Critical Decision Method for Eliciting Knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 3, 462–472.

Leveson, N. G. (2011) Applying Systems Thinking to Analyze and Learn from Events, *Safety Science*, 49, 1, 55–64.

Mendonça, D. (2007). Decision support for improvisation in response to extreme events: learning from the response to the 2001 World Trade Centre attack, *Decision Support Systems,* 43*,* 952–967.

National Transportation Safety Board (2010) *Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549, Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009*. Aircraft Accident Report NTSB/AAR-10/03. Washington, DC.

Rasmussen, J. (1997) Risk Management in a Dynamic Society: A Modelling Problem, *Safety Science*, 27, 2/3, 183–213.

Trotter, M. J., Salmon, P. M. & Lenné, M. G. (2012) Improvisation: Theory, Measures and Known Influencing Factors, *Theoretical Issues in Ergonomics Science*, Online, 1-24.

Trotter, M. J., Salmon, P. M. & Lenné, M. G. (Submitted Nov 2012) Irnpromaps: Applying Rasmussen's Risk Management Framework to Improvisation Incidents, *Safety Science.*

Trotter, M. J., Salmon, P. M. & Lenné, M. G. (Submitted Mar 2013) Improvisation in Led Outdoor Activities: An investigation into its occurrence and influencing factors. *Australian Journal of Outdoor Education.*

Using the notion of mental models in design to encourage optimal behaviour in home heating use

Kirsten M A REVELL, Neville A STANTON
University of Southampton, UK

ABSTRACT

Introduction: Understanding how to influence householder's energy consuming behaviour, could inform far reaching strategies to combat climate change. A Mental Model (MM) approach to design, to encourage optimal behaviour was explored. Challenges exist in accessing, describing and analysing user MMs and associated behaviour. **Method**: A method that considered bias in interpretation was developed, involving a structured interview, concept maps and graphical self-reported behaviour. Using this method, 6 householders in matched accommodation, over winter 2011/2012, participated in a home heating case study. Thermostat set point data was also collected from participant's households. A home heating expert was interviewed using the same method, for comparison. **Results and discussion:** Key variations in MMs of home heating were found. The differences in user MMs from each other, and an expert, were insightful in explaining non-optimal home heating operation. These suggest design solutions that could promote or compensate for user mental models to influence energy consumption.

KEYWORDS

Mental models; Design; Behaviour change; Domestic energy consumption; Home heating

INTRODUCTION

User behaviour with home heating systems, contributes to climate change

The U.K. has legislated to cut carbon emissions by 80% by 2050 (*Climate Change Act 2008*) with 25% of total UK carbon emissions from domestic customers. Lutzenhiser and Bender (2008) report that variations in domestic energy use are due to the behavioural differences of householders. Home heating accounts for over 25% of domestic energy use in the UK (Department of Energy and Climate Change, 2012). This research focuses on how the concept of mental models (MMs) can be applied in design to elicit behaviour change to reduce domestic heating use.

Mental models research could inform behaviour change strategies

Mental models are thought to be representations of the physical world (Johnson-Laird, 1983; Rasmussen, 198, Veldhuyzen & Stassen, 1976), constructs that can explain human behaviour (Kempton, 1986; Wickens, 1984) and internal mechanisms allowing users to understand, explain, operate and predict the states of systems (Craik, 1943; Gentner & Stevens, 1983; Hanisch et al. 1991; Kieras and Bovair, 1984; Rouse and Morris, 1986) The notion of mental models has been used in design for the development of interfaces (Carroll & Olson, 1987; Norman, 2002; Jenkins et al., 2010; Williges, 1987), to promote usability (Mack and Sharples, 2009; Norman 2002) and in the human factors domain, to enhance performance (Bourbousson et al., 2011; Grote et al., 2010; Stanton & Baber, 2008; Stanton & Young, 2005) and reduce error (Moray 1990a, Stanton & Baber, 2008).
Kempton (1986) proposed that different patterns of behaviour when operating a home heating thermostat result from the user holding different mental models of how the heating system works. This association is yet to be proven, and may no longer be relevant. Kempton (1986) identified two typical types of mental models of home heating. These represented common elements found in his participants' individual 'mental models' (this is distinct from concepts such as 'team' or 'shared' mental models) . The 'valve' model considered the thermostat like a gas valve. The 'feedback' model recognized the thermostat as an automatic switch based on temperature, without considering the thermodynamics of the dwelling. Kempton (1986) proposed specific behaviour characteristic evident in householders with a 'valve theory', may result in lower consumption than those with a 'feedback theory'. Further 'typical' models of the home heating thermostat have been offered by Norman (1988), and Peffer et al. (2011) who describe 'timer' and 'on/off switch' models respectively. Understanding the cause of user mental models of home heating, and their effect on behaviour, offers a novel approach to influencing domestic energy consuming behaviour, to help combat climate change.

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

Contributions to mental model methods and specific insights into domestic heating domain

A comprehensive literature review identified that consideration of bias in methods of access, description and analysis of mental models, is essential for practical application of the notion (Bainbridge, 1992; Revell & Stanton, 2012; Rouse & Morris, 1986; Wilson & Rutherford, 1989). A generic framework for considering bias in the research of knowledge structures was developed. This framework informed the development of methods to capture and analyse user's mental models and behaviour with home heating systems. Insights into the specific context of domestic home heating behaviour are expected by exploring the link to mental models, heating system design and domestic energy consumption. Application of these insights in design strategies will also contribute to the field.

METHODOLOGY & RESEARCH UNDERTAKEN

Development of the Quick Association Check

A quick, inexpensive, method for exploring association between users mental models of home heating systems, and their behaviour, was sought. The Quick Association Check (QuACk) is a structured interview method which includes activities and templates to produces verified outputs ready for analysis. Examples of the key outputs from QuACk, describing a user mental model, a self-report of user behaviour, are shown in figures 1 and 2.

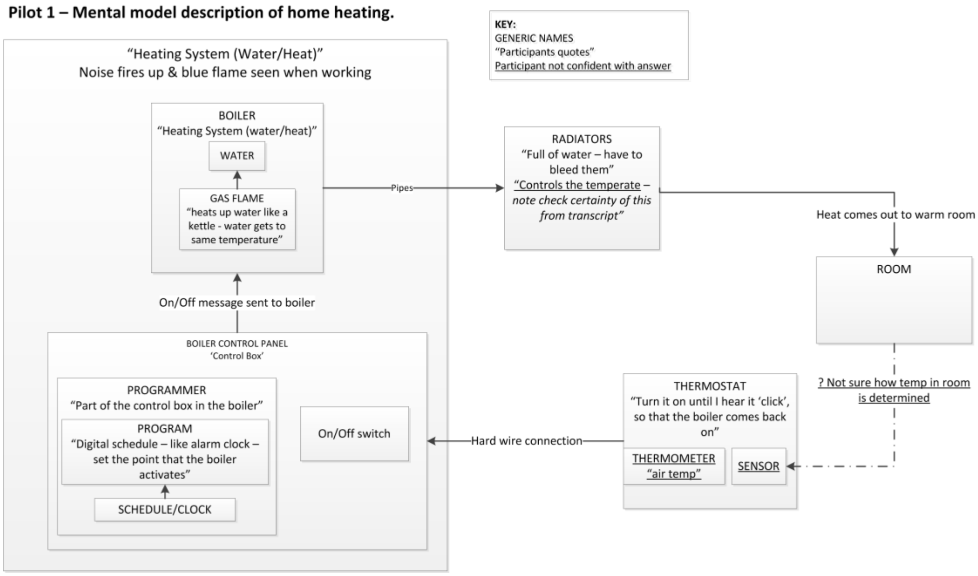

Figure 14. Mental model descripton of home heating from QuACK

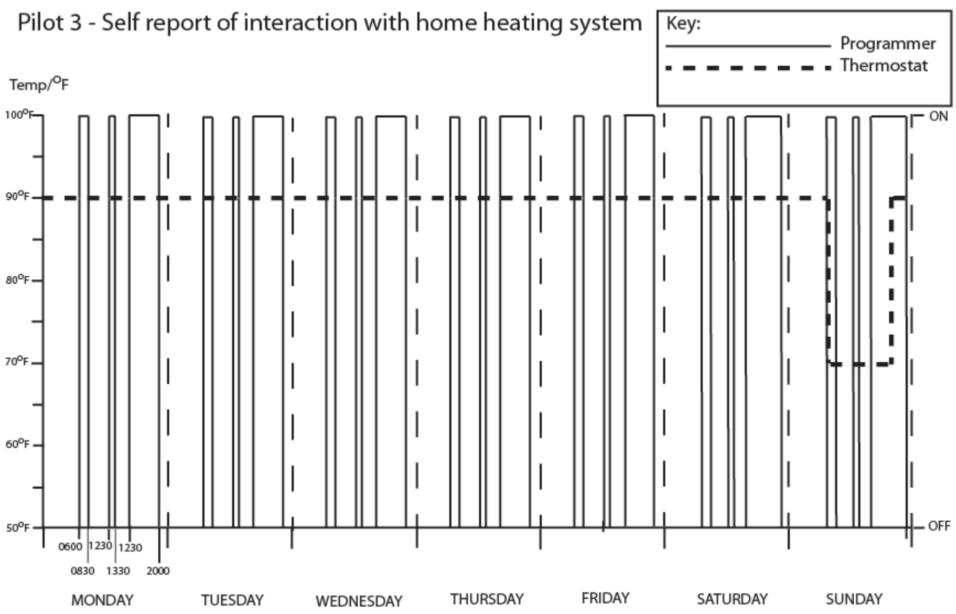

Figure 15. Self report of home heating interaction from QuACK

The development of the QuACk was undertaken systematically (see figure 3). This approachcould benefit researchers exploring the association between mental model and behaviour in other contexts and domains.

Hancock and Szalma (2004) emphasised the need to embrace and integrate qualitative methods in ergonomics research. QuACk was therefore developed using case studies and participant observation to provide rich feedback, that drove iterative developments to the prototype.

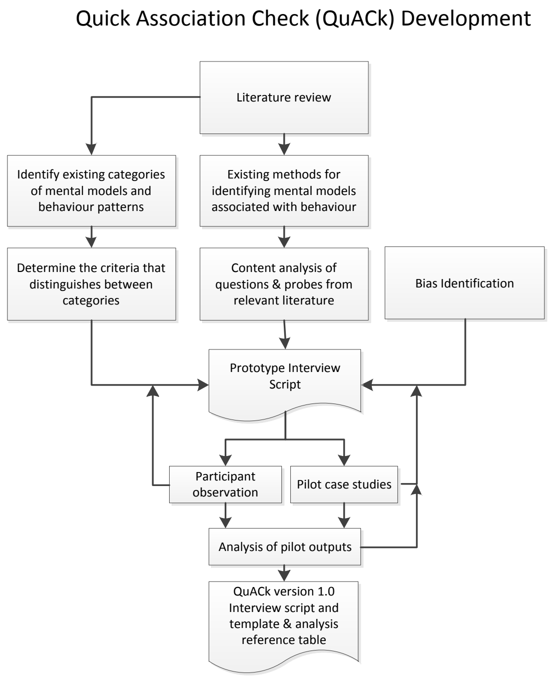

Figure 16. The stages of development of the Quick Association Check (QuACk)

Conducting a naturalistic case study of home heating

The QuACk method was applied to a case study comprised of non-randomly selected, postgraduate students with families, new to the UK and residing in semi-detached university owned accommodation. Accommodation, heating devices and insulation levels were matched, so variations in mental model descriptions could be attributed to characteristics of the participant, rather than the environment. Interviews with 6 participants from 5 households were undertaken. The impetus for this case study, was to seek evidence of the 4 typical mental models described in the literature, and explore association with user behaviour patterns.

Capturing an expert mental model of home heating, to inform design specifications

Norman (1986) described how problems users have interacting with devices was caused by the 'gulf of execution and evaluation'. He attributes this gulf to differences between a device's 'design model' and the 'user's model' of the device. To capture a representation of the 'design model' of home heating, an expert from the heating control manufacturer who provided the devices for the naturalistic study was interviewed using QuACK. This provided verified outputs in the same format as the case study data, allowing us to gain insights into the differences between the expected and actual way householders think and interact with their heating system. These insights are key to inform the development of design specification intended to promote appropriate behaviour through better user mental models.

RESULTS & DISCUSSION

Mental models research should view domestic settings as complex systems

Different types of mental models were identified from the naturalistic case study, showing that people view the home heating system in quite different ways. Switch and feedback mental model types were useful for categorizing the thermostat, and all types identified in the literature helped to categorize other heating controls. A range of different strategies for using the system were reported, as well as differences in users' goals, the number of people who operated the heating system, and the control devices favoured. The interplay of the thermodynamics of participants' houses and responsiveness of the heating system installed, also coloured householders' mental models and behaviour choices. The impact of these variables on user behaviour emphasised the need to consider the household as a complex, sociotechnical, dynamic setting.

Key differences in expected and actual use of home heating systems

Comparisons of the 'design model' derived from an expert in home heating, and 'user's models', concluded that householders need assistance to bridge Norman's (1986) gulf of execution and evaluation. The role of control devices and the way they interact varied in user mental models. Some participants with an appropriate mental model at the system level, but inappropriate mental model at the control level reported non-optimal behaviour. The usability of control devices (e.g. programmer), also affected the choice of control when adjusting comfort levels. These findings provide direction to design specifications targeted at promoting appropriate mental models at both the system and device level.

CONCLUSION

The 1986 study by Kempton, inspired the focus of this research. Our research has built on this by; a) devising a systematic generic method for exploring association between mental models and behaviour, b) gaining an understanding of the way domestic home heating behaviour is linked to mental models of the heating system, and, c) comparing user mental models to an expert model to identify the gap in expected and reported thinking and behaviour. Further work will be investigating home heating control as a complex system, involving goals, strategies and considering device usability. The insights gained from this research will help inform approaches to design that aim to reduce energy consuming behaviour in the home, as a novel approach to help combat climate change.

ACKNOWLEDGMENTS

This research reported in this paper was supported by the Intelligent Agents for Home Energy Management project with funding from EPSRC (EP/I000143/1).

REFERENCES

Bainbridge, L. (1992). Mental Models in Cognitive Skill: The Example of Industrial Process Operation. In Y. Rogers, A. Rutherford & P. A. Bibby (Eds.), Models in the mind : theory, perspective and applications (pp. 111). London: Academic Press.

Bourbousson, J., Poizat, G., Saury, J., & Seve, C. (2011). Description of dynamic shared knowledge: an exploratory study during a competitive team sports interaction. Ergonomics, 54(2), 120-138.

Carroll, J. M., & Olson, J. (Eds.). (1987). Mental models in human-computer interaction: Research issues about what the user of software knows. Washington, DC: National Academy Press.

Climate Change Act 2008. Retrieved 10th November, 2011, from http://www.decc.gov.uk/en/content/cms/legislation/cc_act_08/cc_act_08.aspx

Craik, K. J. W. (1943). The nature of explanation. Cambridge, England: Cambridge University Press.

Grote, G., Kolbe, M., Zala-Mezö, E., Bienefeld-Seall, N., & Künzle, B. (2010). Adaptive coordination and heedfulness make better cockpit crews. Ergonomics, 53(2), 211-228.

Hancock, P. A., & Szalma, J. L. (2004). On the relevance of qualitative methods for ergonomics. Theoretical Issues in Ergonomic Science, 5(6), 499 - 506.

Hanisch, K. A., Kramer, A. F., & Hulin, C. L. (1991). Cognitive representations, control, and understanding of complex systems: a field study focusing on components of users' mental models and expert/novice differences. Ergonomics, 34(8), 1129 - 1145.

Jenkins, D. P., Salmon, P. M., Stanton, N. A., & Walker, G. H. (2010). A new approach for designing cognitive artefacts to support disaster management. Ergonomics, 53(5), 617-635.

Johnson-Laird, P. N. (1983). Mental models: towards a cognitive science of language, inference and consciousness Cambridge, UK: Cambridge University Press.

Kempton, W. (1986). Two Theories of Home Heat Control. Cognitive Science, 10, 75-90.

Kieras, D. E., & Bovair, S. (1984). The Role of a Mental Model in Learning to Operate a Device. Cognitive Science, 8(3), 255-273.

Lutzenhiser, L. a. S., & Bender, S. (2008). The average American unmasked: Social structure and difference in household energy use and carbon emissions.: ACEEE Summer Study on Energy Efficiency in Buildings.

Mack, Z., & Sharples, S. (2009). The importance of usability in product choice: A mobile phone case study. Ergonomics, 52(12), 1514-1528.

Mental models in human computer, i., Caroll, J. M., & Olson, J. R. (1987). Mental models in human computer interaction : research issues about what the user of software knows : Workshop on software human factors : users' mental models : Papers: [np].

Moray, N. (1990). Designing for transportation safety in the light of perception, attention, and mental models. Ergonomics, 33(10), 1201-1213.

Norman, D. A. (2002). The Design of Everyday Things. New York: Basic Books.

Peffer, T., Pritoni, M., Meier, A., Aragon, C., & Perry, D. (2011). How people use thermostats in homes: A review. Building and Environment, 46, 2529 - 2541.

Rasmussen, J. (1983). Skill rules and knowledge; signals signs and symbols; and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, 13(3), 257-266.

Rasmussen, J., & Rouse, W. B. (Eds.). (1981). Human detection and diagnosis of system failures. New Yourk: Plenum Press.

Revell, K. M. A., & Stanton, N. A. (2012). Models of models: filtering and bias rings in depiction of knowledge structures and their implications for design. Ergonomics, 55(9), 1073-1092.

Rouse, W. B., & Morris, N. M. (1986). On Looking Into the Black Box: Prospects and Limits in the Search for Mental Models. Psychological Bulletin, 100(3), 349-365.

Stanton, N. A., & Baber, C. (2008). Modelling of human alarm handling response times: a case study of the Ladbroke Grove rail accident in the UK. Ergonomics, 51(4), 423-440.

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. Safety Science, 27(2-3), 149-159.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. Ergonomics, 48(10), 1294-1313.

Veldhuyzen, W., & Stassen, H. G. (Eds.). (1976). The internal model: What does it mean in human control? New York: Plenum.

Wickens, C. D. (1984). Engineering psychology and human performance. London: Merrill.

Williges, R. C. (1987). The Society's Lecture 1987 The Use of Models in Human-Computer Interface Design. Ergonomics, 30(3), 491-502.

Author Index

9 791092 329001