Refining a User Behaviour Model Based on the Observation of Emotional States

[1]Yuska P. C. Aguiar, [1,2]Maria F. Q. Vieira
[1]LIHM, Electrical Engineering Department, UFCG
Campina Grande, Brazil
[2]CESIP, Dept of Electronic & Electrical Engineering, University of Strathclyde
Scotland, UK
yuska.aguiar@lihm.ufcg.edu.br, fatima@dee.ufcg.edu.br

Edith Galy
Université de Provence (Aix-Marseille I), Département de Psychologie Cognitive et Expérimentale
Centre de Recherche PsyCLE
Aix-en-Provence, France
edith.marie@univ-provence.fr

Jean-Marc Mercantini, Charles Santoni
Université Paul Cézanne (Aix-Marseille III), Département Génie Industriel et Informatique, LSIS
Marseille, France
jean-marc.mercantini@lsis.org, charles.santoni@polytech.univ-mrs.fr

***Abstract*— User behaviour models are important tools to study human error in the industrial context. With a programmable user model it is possible to simulate user activity, analyse the influence of context on user behaviour and impact of user behaviour on task outcomes. This paper proposes a procedure to refine user behaviour models. The procedure has been developed to support the analysis of accident and incident reports in the operation of electrical power systems. The procedure involves observing the user interacting with a system simulator that replicates situations described in accident and incident reports. This paper focuses on the emotional components of behaviour observed during the interaction.**

Keywords-User behaviour model; data gathering on human emotions; interaction observation; human error studies.

I. INTRODUCTION

The analysis of accidents and incidents is essential for the study of human error and central to strategies for preventing these: Human Computer Interaction (HCI) adaptation and improvement in training and task adaptation. Report analysis is the traditional path followed by many authors [3][15][23].This approach was adopted by the authors when investigating human error in the operation of an electrical power systems company in Brazil. The studies performed were based on a corpus of 31 reports of accidents and incidents that occurred over a ten-year period. Among the results, these studies resulted in a prototype of a system's operator behaviour model [1][17].

The operator model simulates the dynamic behaviour of a system operator performing tasks during situations and contexts that lead to error. This model has proved to be an important tool for studying, and understanding, human error in that it allows situation contexts to be simulated through the parameterization of behavioural variables cited in the error reports. The original intension was to replicate scenarios in which human error occurred and alter user behaviour to experiment with new scenarios. The parameterization of behavioural variables allowed the simulation of a range of external and internal aspects influences on operator behaviour.

The analysis of the corpus of reports revealed that the error reports focus on the technical aspects of the scenario in which the error occurred but rarely address the operator's state while performing the task, often omitting relevant information relating to emotion and behaviour. In its initial stage the model complexity was kept low, with a small number of variables representing operator state. Tiredness, stress, inattention and confusion, were the causes most frequently mentioned in the reports. To refine the original user behaviour model, and in order to better understand the error context, more information about the status of the operator during the performance of a task is required. The proposed refinement consists of adding new characteristics to the model, i.e., extending the set of variables that represent the operator's state. This requires immersing the operator in the work context and replicating scenarios described in the accident and incident reports in order to observe behaviour. The scenarios must account for the wider environment (e.g., the occurrence of lighting, noise, etc.) as well as the immediate environment. This will be achieved with a simulator that replicates the working environment, with all the objects needed to perform the task, as comprehensively as possible [24]. In order to analyse interactions and highlight elements that contribute to the occurrence of error, the observation should be informed by methods and tools found in psychology, such as task analysis and the observation of emotional components.

This paper proposes the application of Scherer's Components Model of Emotion (CME) [16] to the observation, recording and analysis of the emotional components of operator behaviour. The CME model considers emotion to be an episode of interrelated, synchronized, state changes of subsystems, that are a result of the evaluation of an external or internal stimulus event. The model components are: cognitive appraisal, physiological reactions, behaviour tendencies, motor expression, and subjective feeling (emotional experience). CME was used to structure a range of relevant emotion-measuring methods that can be used to find relevant emotions during the observation of user-system interaction (or with its representation such as the simulator used in this research). The emotions identified, and their relationships, are incorporated into the user behaviour model. To guide the observation an experimental protocol is needed that focuses, among other things, on the user behaviour components. The protocol consists of a set of procedures, activities and

 ISBN: 978-1-61208-155-7

documents that support the observers when planning, conducting and reporting an experiment.

This paper describes how emotions can be measured in accordance with CME and how this knowledge can be integrated into an experimental protocol devised to observe user interactions with systems. The text is organized as follows: Section 2 presents the CME and tools employed to collect emotion data, Section 3 presents the experimental protocol in its original formulation, and Section 4 presents an adaptation of the protocol with CME components. In the final section, the authors consider the future direction and developments this work which is still in progress.

II. COMPONENT MODEL OF EMOTION (CME)

With the advance, and popularization, of interactive technologies, the users' emotional state has become a valuable source of information with the potential to improve the interaction mechanisms offered by a system [20]. Studying human reactions to emotional episodes allows improved understanding of human behaviour.

Mahlke [9][10] identifies how usability and emotional reactions can determine a user's overall appraisal of a system and thus influence his future decisions and behaviour. This work used CME to structure a range of relevant emotion-measuring methods and was adopted as the starting point for the work described here. What follows is an explanation of CME, a list of emotions that can be measured and the tools necessary to make these measurements. It is proposed that these emotions will be integrated into the user model. Although quantitative measurement of the behavioural tendency component of CME has been referred in previous studies is the least explored in the literature and this component is not addressed here.

A. *Cognitive appraisals*

Cognitive appraisals are defined as a quick evaluation of a situation that can direct emotional responses (positive or negative). Based on a review of the literature, Demir et al. [4] propose the following set of appraisal components: consistency of motives, intrinsic pleasure, expectation confirmation, standard conformance, agency, coping potential, and certainty. The tool Geneva Appraisal Questionnaire (GAQ) [16] assesses the result of an individual's appraisal process in the case of a specific emotional episode. GAQ aims to measure: intrinsic pleasantness, novelty, goal/need conduciveness, coping potential and norm/self-compatibility. The emotions measured by GAC are: anxiety, irritation, contentment, joy, sadness, disgust, fear, anger and surprise.

B. *Physiological reactions*

These can be expressed in cardiovascular, electrodermal and respiratory measures. Kreibig [7] presents a review of investigations of different emotions using a range of emotional induction paradigms. The review argues that the elements most often investigated are distributed in three categories: (i) cardiovascular measurements, i.e., heart rate (HR), systolic and diastolic blood pressure (SBP) and heart rate variability (HRV); (ii) respiration rate (RR); and (iii) skin conductance level (SCL). These are also measurements of anxiety, contentment, joy and fear.

C. *Motor expressions*

These are postural, vocal and facial expressions. This work addresses only facial expressions. To use facial expressions it is necessary to classify and correlate them with the appropriated emotion (or set of emotions). The most widely reported work in this context is that of Ekman & Friesen [5], known as the Facial Action Coding System (FACS).Its adoption, however, requires a highly skilled professional. This work adopts a simplified system, FaceReader [22], for automatic real time analysis of facial expressions. FaceReader allows the measurement of the following emotions: sadness, disgust, fear, anger and surprise.

D. *Subjective feelings*

Subjective feelings refer to the unique mental and bodily experience during a particular event [19]. Scherer claims that no objective method for measuring the subjective experience exists. To access it one must ask the individual to report on his/ her experience. The Self-Assessment Manikin (SAM) [8] - is a non-verbal scale, using schematic manikins to represent the different feelings (anxiety, hope, boredom, relaxation, irritation and contentment). SAM manipulates the valence, the arousal, and the dominance dimensions. An alternative, the Activation-Deactivation Adjective Check List (AD-ACL) [21], is a multidimensional tool to test various transitory arousal states (interest, irritation, contentment, joy and fear). The tool considers four sub-scales to measure the relation between energetic and tense arousal: energy, tiredness, tension, and calmness. The Geneva Emotion Wheel (QEW) [19] - is a verbal self-reporting instrument in which the participant is asked to indicate the emotional intensity for a single emotion (or a blend of several emotions) on 20 distinct emotion families (including interest, irritation, contentment, joy, sadness, disgust, fear, anger and surprise). Five degrees of intensity are represented by circles of different sizes.

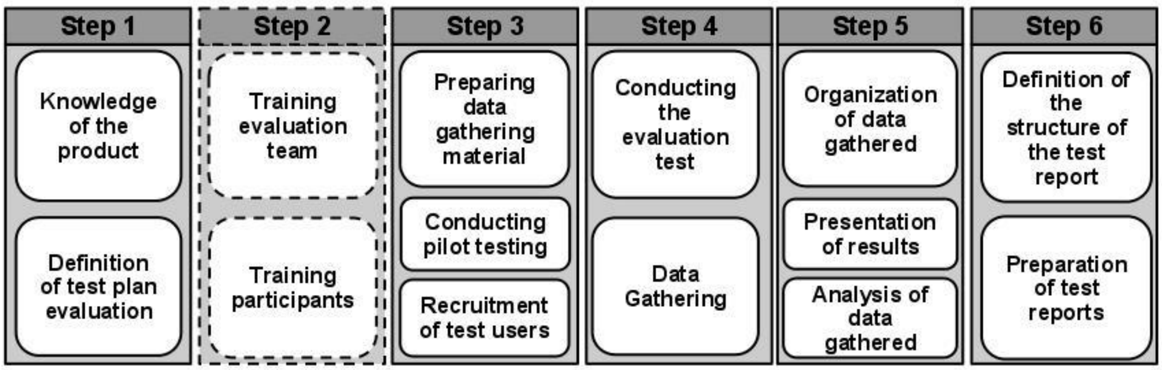
E. *Other tools*

In addition to the tools listed above, there are two others that are relevant to this study. The Objective and Cognitive Profile of the User (POCUS) [18] structures a system user profile using the categories: personal, professional, contextual, physical, psychological and clinical. The NASA Task Load Index (NASA-TLX) [6], employed to measure mental workload, employs three dimensions: behaviour (effort and performance), task (physical, mental and temporal demands), and subjective (frustration).

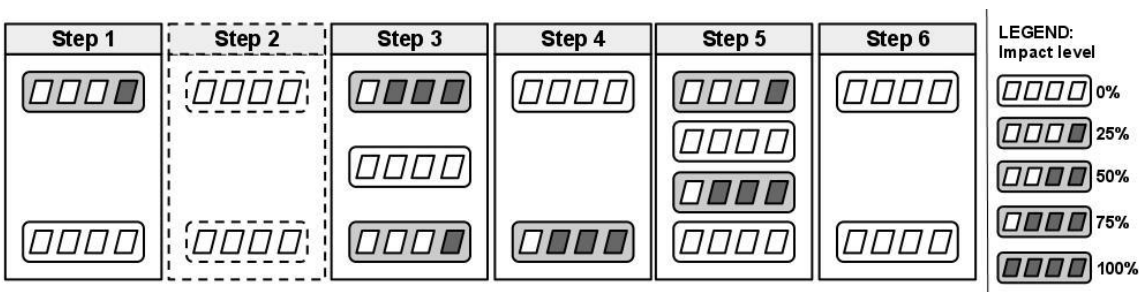
III. PROTOCOL FOR EXPERIMENTAL OBSERVATION OF THE INTERACTION (PEOI)

The Protocol for Experimental Observation of Interaction (PEOI) [2] structures the usability recommendations found in the literature [11][13][14]. PEOI was conceived to support the observation of a system-user interaction focusing on usability. It is adaptable to a range of usability testing contexts (in the laboratory, in the field and in situ) and to different product complexities. It was employed in the usability evaluation process of the electric system simulator [24].

The protocol is organized in six steps each of which consists of a process defined by a set of activities. The steps and respective processes are illustrated in Figure 1. Step 1: Planning the Test characterizes the product, its context and its users. Step 2: Training (when needed) prepares the evaluation team and/or the test participants with the product's context of use, tools and methods. Step 3: Preparation and Validation of the Test: structures the test, develops the necessary supporting materials (preparation) and performs a pilot test (validation) with a recruited participant. Step 4: Conducting the Test and Data gathering: executes the experiment resulting in a sample of data. Step 5: Data Tabulation and Analysis structures and organizes the gathered data for analysis and results in a diagnostic for the product-user interaction. Finally, Step 6 - Presentation of the Results specifies the form, content and media to report the experiment and its results.



1 Figure 1 Steps and Processes in PEOI



2 Figure 2 Distribution of proposed changes in PEOI

PEOI associates a set of methods employed for data gathering (observation, interview, questionnaire, document analysis) to four categories of data. General data is gathered through interview, and aims to clarify the test objectives. All four methods are employed to gather data on: task, product and context of use. Pre-interaction gathers data on the user profile (personal, professional and contextual using all methods except observation). Interaction gathers data on subjective indicators using observation. It also gathers objective indicators about user activity using all methods except document analysis. Post-interaction gathers data on user satisfaction level with the product or system under test using all methods except document analysis.

Given that the original purpose of the experimental protocol is to support usability evaluation, the target data and the methods used to gather data have proved adequate. Given the interest in extending the protocol to support the understanding of user behaviour, however, it needs to be adapted to gather data about the user emotional state. This additional data collection is associated with the interaction and post-interaction data categories. The adaptation of PEOI is described in the following section and was based on the CME approach.

IV. PEOI'S ADOPTION OF CME'S TOOLS

In spite of extending the protocol to support user behaviour observation, no changes are required in its general structure, thus its steps and processes remain the same. The changes required are: (i) extension of the range of data to be gathered; (ii) adding new methods for data collection; and (iii) including new activities to be performed by the evaluation team during the experiment.

Given the new aspects of interest in the pre-interaction step, it is proposed that POCUS be adopted when gathering data on the user profile. The extended profile identifies personality and temperament traits. During the interaction and post-interaction steps, the focus becomes that of gathering data about the operator´s emotional state. In these two steps the elements of interest are the following thirteen emotions (according to Geneva Affect Label Coder (GALC´s dictionary) [19]): anger, anxiety, boredom, contentment, disgust, fear, hope, interest, irritation, joy, relaxation, sadness and surprise.

With respect to data gathering methods four new groups are proposed: a) physiological measurement to gather physiological reactions, b) face recognition to gather motor expressions (FaceReader), and c) self-reporting to gather subjective feelings (SAM/AD-ACL/GEW).

The overall changes to team activities during the experiment concern the measurements to be performed relating to behaviour observation. Figure 2 illustrates, at an abstract level, how these changes map onto the protocol. In Figure 2, the processes associated to each step are represented as rectangles, and within each process the corresponding activities are represented as small parallelograms. The figure highlights where changes are being proposed and the overall level of change proposed within each step/process. The level of change represents the volume of new activities introduced in the process (five levels are represented: 0%, 25%, 50%, 75% and 100%).

The processes Knowledge of the Product, Recruitment of Test Users and Scheduling of Test Sessions and Data Gathering can be supported by POCUS. The highest impact of protocol adaptation is on these processes which relate to the test planning activity, test plan execution and data gathering.

In the process Preparing Data Gathering Material four activities were modified: (1) defining which data to gather in order to include variables related to the operator´s emotional state, (2) including tools for data gathering comprising cognitive appraisal (GAQ), physiological measures, motor expressions (FaceReader), subjective feelings (SAM/AD-ACL/GEW), user profile (POCUS) and workload (NASA-TLX), (3) specifying the tools and resources required to collect physiological reactions (HR, HRV, SBP, RR, SCL), (4) preparing the artefacts required to perform the experiment, e.g., questionnaire, forms/cards and self-reporting.

In the process Data Gathering three activities were modified: (5) pre-test activities in which POCUS is applied while measuring physiological variables which will be used as a reference for later comparison with the values collected during task activity; (6) conduct the observation in which physiological variables (HR/HRV/SBP/RR/SCL), motor expressions (FaceReader) and subjective feelings (SAM/AD-ACL) are measured; (7) conduct post-test activities in which cognitive appraisal (GAQ), subjective feelings (GEW), and workload (NASA-TLX) are measured.

In Step 5, the process Analysis of Data Gathered reflects all the changes introduced in the previous steps. The data gathered in Step 4 will impact the analyses process because it requires the correlation analysis between subjective and objective indicators.

Not all the tools and data types in the extended protocol must be adopted in every experiment of course. The choice of data types and corresponding data gathering tools depends on the specific aim of the observation. It is likely, therefore, that a specific experiment will only encompass a subset of the human behaviour related variables (i.e., subset of emotions) to be observed.

V. Proof of Concept

An instantiation of the modified protocol is underway. It consists of an experiment in which a product developed for use in a critical situation is being used in order to validate the protocol and support the selection of tools used in the analysis of user behaviour during the task.

The experiment was conducted in the Research Center of Psychology of Cognition, Language, and Emotion of at the *Université de Provence*, in France. It consisted of eight test sessions (including the pilot), during which users were observed with the aid of the protocol. The evaluation team was comprised by one psychologist, two usability experts and two usability trainee students, with varying levels of knowledge about the protocol.

The product under observation was the *Generateur de Plans d'Intervention* (GENEPI) [12], which consists of a decision support system, to assist in preparing contingency plans to deal with maritime accidents. The usage scenario was the communication of an accident followed by preliminary data, which then had to be complemented by the user through the communication with various agencies such as a weather office. The user's goal is to propose an action plan, consistent with the situation of the accident, and the additional data obtained, in the shortest time possible. The whole process of supplying information about the accident (when requested), along with frequent calls demanding for an initial plan, was simulated by the research team during the test sessions.

The data collected from the test sessions is currently being analyzed, and will be used to evaluate the new version of the protocol in terms of ease of use, effectiveness and efficiency. More specifically the data will be used to analyze the impact of the changes made in the protocol, on its artefacts, on the actors and the interaction between them. It is also intended to select those tools that are most appropriate for the acquisition of knowledge about the emotions and behaviour of users. Finally it is intended to assess the adaptability of the protocol to a different context from that for which it was conceived (i.e., the observation of electrical systems operation).

From this experiment, it should be possible to identify any requirement for protocol refinement, before it is applied in the context of electrical systems and used to improve the user behaviour model.

VI. Final Considerations

The human error study based on document analysis (in particular error reports) will benefit from a user behaviour model extracted from the observation of the operator when confronted with the work context. To be a useful tool, the programmable user model must be refined with data gathered using tools and methods from the domain of psychology. This paper asserts that experiments to gather data on user behaviour must be supported by an experimental protocol. It describes such a protocol in the context of a product usability evaluation adapted to human behaviour observation. The observations supported by the protocol focus on aspects related to the user behaviour and emotional

 ISBN: 978-1-61208-155-7

state. The proposed protocol supports: (i) identification of the emotions to be measured; (ii) identification of tools to be used when observing the user-system interaction; and (iii) structuring of the data gathering process to be employed during the interaction observation.

One innovative aspect of the work reported is the simultaneous application of different tools to collect data about the operator's emotion and behaviour, during their work activity, in order to evaluate their effectiveness. Another is the proposal of the experimental protocol itself. Although the existing literature cites both practices on usability evaluation of products and the observation of user behaviour in psychology, methods and procedures have not yet been made available in the form of a systematized protocol to support the reporting, interpretation and replication of experiments.

A limitation of the proposed protocol is that it has been designed to observe the behaviour of individuals under stress interacting with critical systems. Under these conditions human reactions are amplified and thus more easily measured. Furthermore, the observation is conducted in a simulated environment. This imposes two corresponding challenges for the future extension of this work: (i) to evaluate the applicability, efficiency and effectiveness of the protocol outside the domain of critical systems, and (ii) to investigate the degree to which the data collected by the protocol reflects that which would be collected in a real working environment.

In the current version of the protocol, data collection is focused on the following aspects: workload, attention, emotion and behaviour; complemented by data on the profile and personality of the operators. However, it is intended to extend this to include other aspects of user behaviour. This depends on identifying additional tools to collect human behaviour data. The observation of such supplementary factors will only be useful, of course, if the ranges of the metrics describing them in which human errors occur can be identified.

Improved understanding of human behaviour during situations leading to error will complement traditional information (such as accident and incident reports) and lead to reduction in the incidence of human error. Information on user behaviour projected onto a user programmable model will allow the relationship between user behaviour, work context and error occurrences in specific working scenarios to be investigated experimentally.

REFERENCES

[1] Netto, A. V. S., Aguiar, Y. P. C, Scherer, D., and Vieira, M. F. Q. Context analysis during task execution: an operator model. In: IADIS Multi Conference On Computer Science and Information Systems, 2009, Algarve, pp. 113-120.

[2] Aguiar, Y. P. C. and Vieira, M. F. Q. Proposal of a protocol to support product usability evaluation. Fourth IASTED International Conference Human-Computer Interaction, 2009, St. Thomas, pp. 282-289.

[3] Cellier, J. M. L'erreur humane dans le travail. In : Les facteurs humains de la fiabilité dans les systèmes complexes, by J. Leplat and G. De Terssac. 1990, Marseille.

[4] Demir, E., Desmet P.M.A., and Hekkert, P. Appraisal Patterns of Emotions in Human-Product Interaction. Int. J. of Design, 2009, pp.41-51.

[5] Ekman, P. and Friesen, V. W. Facial action coding system: A technique for the measurement of facial movement. Palo Alto, Ca.: Consulting Psychologists Press, 1978.

[6] Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press, 1988, pp. 139-183.

[7] Kreibig, S. D. Autonomic nervous system activity in emotion: a review, Biological Psychology 84, 2010, pp. 394–421.

[8] Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm, A. O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. Psychophysiology, 30 (1993), pp. 261-273.

[9] Mahlke, S. and Lindgaard, G. Emotional Experiences and Quality Perceptions of Interactive Products. In J. Jacko (Ed.) Human-Computer Interaction, Part I, HCII 2007, LNCS 4550 pp.164-173. Berlin: Springer.

[10] Mahlke, S. and Thüring, M. Studying Antecedents of Emotional Experiences in Interactive Contexts.CHI 2007 Proceedings, pp. 915-918. New York: ACM Press.

[11] Mayhew, D. J. The usability engineering lifecycle: a practitioner's handbook for user interface design, San Francisco, CA: Morgan Kaufmann Publishers Inc., 1999

[12] Mercantini J.-M., Guerreiro C. V. S, and de Freitas D. D., Designing a Software Tool to Plan Fight Action against Marine Pollutions. The 7th International Mediterranean and Latin American Modelling Multiconference, October 13-15 2010, pp. 93-99, Fes, Morocco.

[13] Nielsen, J. Usability Engineering. Morgan Kaufmann, 1994

[14] Preece, J., Rogers, Y., and Sharp, H. Interaction Design: Beyond Human-Computer Interaction. 2nd Edition, 2007

[15] Rasmussen, J., O. M. Pedersen, A. Carmino, M. Griffon, and P. Gagnolet. Classification system for reporting events involving human malfunctions. RISO National Laboratory, Dennmark: RISO National Laboratory, 1981.

[16] Scherer K. R. Appraisal Considered as a Process of Multi-Level Sequential Checking; in K. R. Scherer, A. Schorr and T. Johnstone (eds.) Appraisal Processes in Emotion: Theory, Methods, Research, pp. 92-120. New York and Oxford: Oxford University Press, 2001.

[17] Scherer, D., Netto, A. V. S., Vieira, M. F. Q. and Aguiar, Y. P. C. Programming a user model with data gathered from a user profile. In: ADIS Multi Conference on Computer Science and Information Systems, 2010, Freiburg. , 2010. pp. 139-146.

[18] Scherer, D. and Vieira, M. F. Q. Accounting for the Human Error when Building the User Profile. In: Third IASTED International Conference Human-Computer Interaction, 2008, Innsbruck, 2008. v. 1. pp. 132-137.

[19] Scherer, K. R., What are emotions? And how can they be measured? Social Science Information 44, 2005, pp. 695–729.

[20] Tayari, I. and Le Thanh, N. "Modélisation des états émotionnels par un espace vectoriel multidimensionnel" Research Report 2009.

[21] Thayer, R. E. Toward a psychological theory of multidimensional activation (arousal). Motivation and Emotion, 2, 1978, pp. 1-34.

[22] den Uyl, M., Kuilenburg, H. V., and Lebert, E. FaceReader: an online facial expression recognition system. Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research, 2005, Wageningen, The Netherlands, pp. 589-590.

[23] van Eekhout, J. M., and W. B. Rouse (1981). Human Errors in Detection, Diagnosis, and Compensations for Failures in the Engine Control Room of a Supertanker. IEEE Transactions on System, Man, and Cybernetics, 1981, 12 ed., pp. 813-816.

[24] Vieira, M. F. Q., Nascimento, J. A. N., Scaico, A., Santoni, C., and Mercantini, J.-M. A Model Based Operator Training Simulator to support Human Behavior Studies. Transactions of the Society for Computer Simulation, v. 86, pp. 41-51, 2010.

 ISBN: 978-1-61208-155-7