Contribution to the characterization and identification of human stability with regard to safety: application to guided transport systems

Vincent Benard[1] Philippe Richard[1] Frédéric Vanderhaegen[2] Patrice Caulier[2]

[1]IFSTTAR -ESTAS 20, Rue Elisée Reclus, 59666 Villeneuve d'Ascq, France

{vincent.benard,philippe.richard}@ifsttar.fr

[2]UVHC -LAMIH UMR CNRS 8201 Le Mont Houy, 59313 Valenciennes, France

{frederic.vanderhaegen,patrice.caulier}@univ-valenciennes.fr

Abstract

This paper presents an original contribution based on the concept of human stability by identifying the associated risks as part of the safety system assessment. The difficulties to take into account human factors in safety studies are first highlighted and definitions of new ways for the integration of human factors based on the existing concepts of stability and resilience are proposed. Although the stability concept is usually defined around a sustainable equilibrium point that induces a feeling of safety control during normal operation, it appears that the stable behaviour of a human operator can lead to risk in certain situations or contexts such as hypo-vigilance, inattention and so on. The core of this paper lays the foundation of human stability for risks assessment. Here, Human stability is defined as the ability of the operator to stay in a stable operating state under specified conditions. This concept is formalized and 3 modes of stability are developed (time, frequency and sequential modes) in order to identify states and change of states of the human stability. The concept of human stability is then applied in the framework of ERTMS/ETCS and shows that sequences of Human stability states and changes of Human stability states may be precursors of risk. Finally, some perspectives highlight the interest of human stability for the definition of risk indicators to assess system safety, by considering the Human operator as a safety/security multi-criteria sensor for the supervision of human-machine systems.

Keywords: Human stability, resilience, safety, transportation application.

1 Introduction

With an opening-up of borders, markets and exchange spaces, people and goods transportation is now a major economical and ecological problem for a large majority of countries. Through various research projects related to transportation systems, this issue is reflected by integrating new technologies, optimizing performances of these systems, but also by improving comfort and safety of passengers and goods. Although there have been technological innovations, the occurrences of incidents/accidents are significant. Statistically, it is highlighted that 30% of these occurrences of incidents/accidents are technical failures, while 70% of them are attributable to human factors (Amalberti, 2001). From this observation, this article aims to present a new concept of safety assessment focused on human operator: the human stability. This new concept is applied to guided transport systems.

This article is divided into 4 parts. The first part of the paper outlines briefly the main methods and tools usually used in dependability to assess guided systems safety and the interest to focus to other concepts like the resilience or stability systems. The second part of the paper justifies the orientation of the research works concerning the new concept of human stability and it proposes a formalization of this parameter. The third part of the paper is an application of this notion of human stability to an ETCS platform within the framework of rail driving. The final part of the paper explains how human stability could be a detector of human errors and risks to the system and presents some perspectives.

2 Safety of guided transport systems: emergence of new issues

The safety in guided transport is integrated throughout the system lifecycle, not only for the regulatory and normative aspects during the design and operation phases, but also for the decommissioning phase. With safety comes the development of operating, supervision and maintenance procedures.

2.1 Safety assessment

To meet the requirements of safety standards, guided systems key players can use a range of methods and tools from hazard assessment (see table 1) that are applicable a priori. Much of these methods and tools focus only on technical aspects of systems and infrastructures without really taking into account the human factors. These traditional tools and methods have shown their limits for the quantitative risk assessment with a growing complexity of systems, some experts suggest to explore new concepts like system resilience without focusing on existing hazard analysis tools (Ligeron, 2006).

In the following subsection, the concept of resilience is presented in more details.

2.2 Safety assessment

By studying resilience in several application fields, it appears that there is no formal definition of this concept and each application domain provides a definition focused on their problems (Goussé, 2005; Martin, 2005; Hollnagel and Woods, 2006; Poupon and Arnoult, 2006; Foussion and Linkowski, 2007; Zieba and Al., 2007; Morel and Al, 2009; Riana and Terje, 2011). It also appears that its terminology is shared with that of the stability concept. Some authors express however differences between the concepts of resilience and stability, although they are close. According to (Lundberg, 2006), the stability is the ability of the system to respond to regular disturbances or events while resilience focuses on unprecedented disturbances or events. Regular events are defined as well-known events (failure machine for example); irregular events are events that it is possible to imagine but, which are normally rare (earthquake for example); lastly, unprecedented events are so rare that normally no organized mechanisms for coping with them exist (flooding of New Orleans for example).

Based on these findings, a definition of the resilience is proposed in (Richard, 2012). Thus, the resilience could be the ability of a system to maintain or return to its original state or to an optimal area of stability. The resilience is able to manage the occurrences of disturbances (see figure 1) by responding:

- in a proactive way : the resilience aims to identify weak signals that may be causing an alarming situation and to correct this situation that might become catastrophic,
- in a reactive way : in this case, the unexpected event happened; the system or/and the operator must react in order to compensate this occurrence,
- or in a curative way : the incident or accident can not be avoided, but the system or the operator is able to limit the consequences of the event.

The system can absorb a disturbance, either by returning to its original equilibrium point after the event occurrence, or by determining a new equilibrium point and by reaching it after an unstable period.

2.3 Taking into account of human aspects

Whatever its level of sophistication and automation, a complex system, such as a guided transport system cannot produce optimum performances and avoid the risk of disastrous events, without the assistance of a human operator who is responsible for the system supervision. To understand the operator as a safety element of the system and not only as a disruptive element, it seems necessary to control the variables that characterize the human behaviour during a dynamic situation (Duquesne, 2005). Although the operator has various faculties and cognitive strategies for problem solving, his behaviour may cause unintended errors in certain circumstances such as deviation of his workload or the manifestation of a dissonance, if he does not evolve in another state. The concept of human stability as defined in the third part of this article aims to highlight and to understand this behavioural duality of the operator, which allows him to be both the weak element and an important element of the system.

Method name	Focused on	comments
PHA (Preliminary Hazard Analysis)	Technical aspects	Can consider human/machine interfaces
FMECA (Failure mode, effects and criticality analysis)	Technical aspects	Can consider suspicious human task that can lead to a system failure
BowTie Method	Technical aspects	To identify causes and effects of an undesired event
Markov states graph	Technical aspects	Quantitative model
RBD (Reliability Block Diagram)	Technical aspects	
HRC (Human Cognitive Reliability)	Human aspects	To assess the operator reliability for a specified task
THERP (Technique for Human Error Rate Prediction)	Human aspects	To assess the human error probabilty during a task in progress
ACIH	Human aspects	To analyse the effects of human unreliablity
HAZOP (Hazardous Operability)	Human and technical aspects	To identify the potentially dangerous drifts

Table 1

3 Human stability formalization and identification

This section allows the characterization of human stability in relation to criteria relating to the human

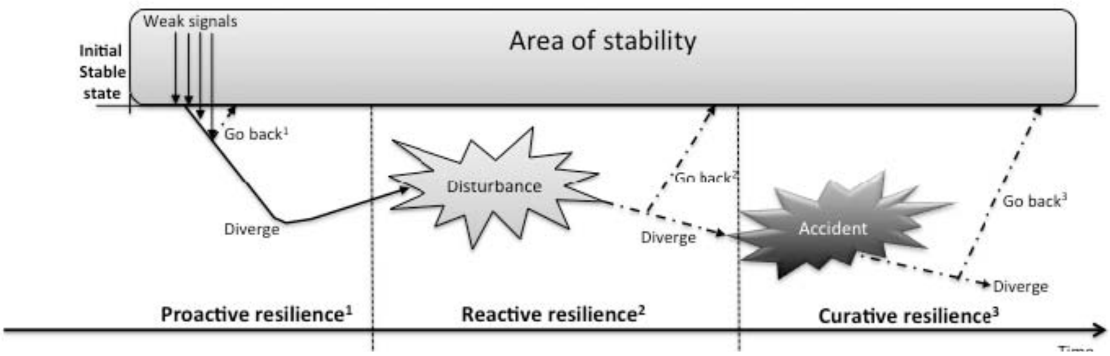

Figure 1 Resilience interpretation

operator. Afterwards, disturbances are regular events, as defined in the subsection 2.2; this hypothesis explains the choice of human stability terminology rather than human resilience terminology.

3.1 Definition

Human stability is defined as the ability of a human operator to be and to stay in a stable state in given conditions (environmental, organisational conditions) for one or more criteria (workload, task achievement, etc; see §3.4). This ability presents the state or the transitions between various states for the operator. These different states are described in the subsection 3.2.

3.2 Formalization

Based on some of the definitions of stability in automation, the human stability refers to a set of states and transitions between these different states (see figure 2):

- Stable state: for the studied criterion, the operator is in a stable state if and only if the value of the criterion is contained between two limit values (Bounded Input, Bounded Ouput principle). These limit values are subject to change depending on conditions in which the operator is (environmental, organizational). A state is considered as stable for the studied criterion if $x(t) < x(t_b)- \alpha_1 x(t_b)$ or $x(t) > x(t_b)+ \alpha_2 x(t_b)$
- Unstable state: one (or more) criterion of human stability diverges. A state is considered as unstable for the studied criterion if $x(t) < x(t_b)- \gamma_1 x(t_b)$ or $x(t) > x(t_b)+ \gamma_2 x(t_b)$
- Indeterminate state: It is an unspecified state.

The transitions between these different states can be formalized by:

- Leaps: these transitions represent the sudden and rapid transition from a stable state to another one.
- Breaks: these transitions represent the sudden and rapid transition from a stable state to an unstable state and vice versa.
- Indeterminate transition: the state of destination is indeterminate.

A transition can be identified if: $x(t) < x(t_b)- \beta_1 x(t_b)$ or $x(t) > x(t_b)+ \beta_2 x(t_b)$ with $\beta_1 > \gamma_1, \beta_2 > \gamma_2$

With $x(t)$, value of the studied criterion at time t; $x(t_b)$, value of the studied criterion at time t_b, α_1 and α_2 lower and upper limit of the stability state (these values are empirically determinate), β_1 and β_2 are the switching amplitude; γ_1 and γ_2 are the divergence amplitude.

Although the process of identifying states and change of state is classic in style for the Human stability, it is different in substance. In contrast to technical systems, « to stay in a stable state for a long time » for a human operator might be dangerous with regard to safety (for example, in monotonous context, the Operator may loose its vigilance).

3.3 Identification

The human stability parameter being formalized, the target is to monitor it during disturbed situation. The identification and detection of states and state changes are determined by the AT (time-dependent algorithm) and AS (sequential algorithm) algorithms.

3.3.1 AT algorithm

The text In order to detect a stable state, the AT algorithm checks at each sampling step that the value of the studied criterion remains around the first measured value. In order to detect an unstable state, the algorithm controls

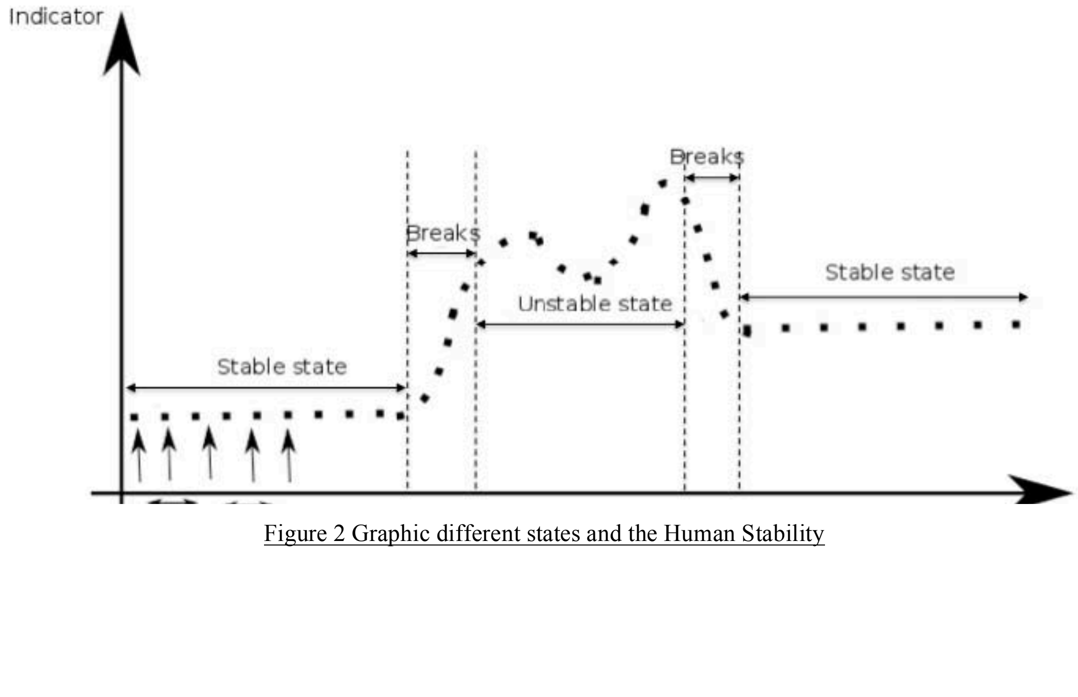

Figure 2 Graphic different states and the Human Stability

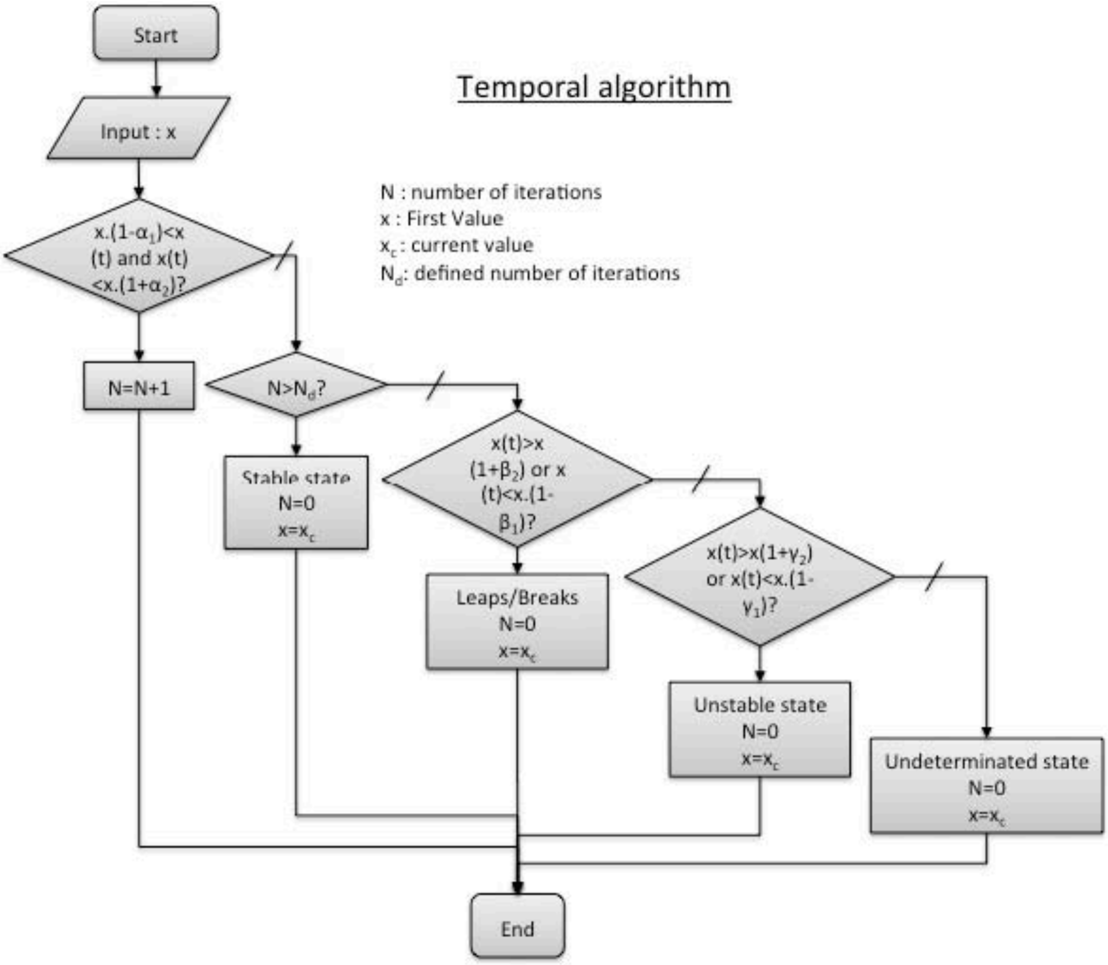

Figure 3 AT algorithm diagram

the divergence of the criterion in relation with the measured values previously. For a transition between states, it aims to determine a brief divergence with high amplitude compared to the previous measurement. This AT algorithm is presented by figure 3.

3.3.2 AS algorithm

The AS algorithm aims to identify sequences of stability states or stability state changes concerning the operator. The goal is to highlight specific sequences (signatures) that can be correlated with disturbances affecting the system or its environment. The figure 4 illustrates the AS algorithm.

The identification algorithms of human stability are evaluated on different criteria from the operator behaviour. The representative criteria are described in the subsection 3.4. For example, when approaching an element of railway infrastructure (level crossing, tunnel or station), the train has to slow down. In a normal situation of this type, we can expect that the driver switches from one stable state to another stable state for the task "speed control" (i.e. to obtain a sequence prescribed or recommended by the designer such as stable/leap /stable).

During a usual situation, driving or supervision tasks are monotonous and repetitive. In the scope of guided transport systems, it is interesting to know both the behaviour and the performances of the operator as well as his intrinsic state for which these kind of tasks can lead to negative effects such as hypo vigilance, fatigue, inattention, etc.). It is mentioned in (Edkins, 2007) that a majority of accidents related to human error in rail transport are linked to attention criteria. In (Richard and Al, 2010), the criteria are classified according to three categories (see figure 5).

The category "state" is intrinsic and is not easily observable. It can assess for example the workload of operator (Sperandio, 1980). It is divided into three aspects:

- "Cognitive" aspect. : cognitive indicators represent the "degree of knowledge monopolized by the Human operator in his/her activity, which are the skill levels, rules and deep knowledge identified by (Rasmussen, 1980).
- “Psychological” aspect. Psychological indicators represent the human operator’s feelings: stress,

dissatisfaction, frustration, inhibition or even guilt.

- "Physiological" aspect. Physiological indicators give indirectly information to the mental work of Human operator: ocular activity (eye movements, gaze direction, blinks), facial recognition, heart rate and speech. Other categories are extrinsic and more easily measurable.
- The "behaviour" category focuses on the system parameters that are directly controlled by the human operator: speed and inter-distance for a guided transport system, for example.
- The "performance measures" category focuses on the compliance of the human operator with driving rules and safety standards and the quality of the product or service. Among the technical evaluation of these indicators include the sense of obligation indicators, the technical characteristics of the added task, or the analysis of the changes in operating behaviour (Spérandio, 1980).

Finally, if this three-dimensional structure of human stability indicators shown in Figure 5 seems generic, the formulation of indicators can be answerable due to the nature of the system.

4 Application to ETCS system

This part presents the application of the human stability to the guided systems field. This application uses the new ETCS rail control system. Two scenarios have been defined: the first, in order to familiarize the 10 selected students, who had no knowledge in railway domain, with the ERTMS platform (http://www.inrets.fr/linstitut/unites-de-recherche-unites-de-service/estas/equipements-scientifiques/simulateurertms.html) during an ordinary driving operation. The second scenario proposes the same course in a disturbed-driving situation with the same selected students.

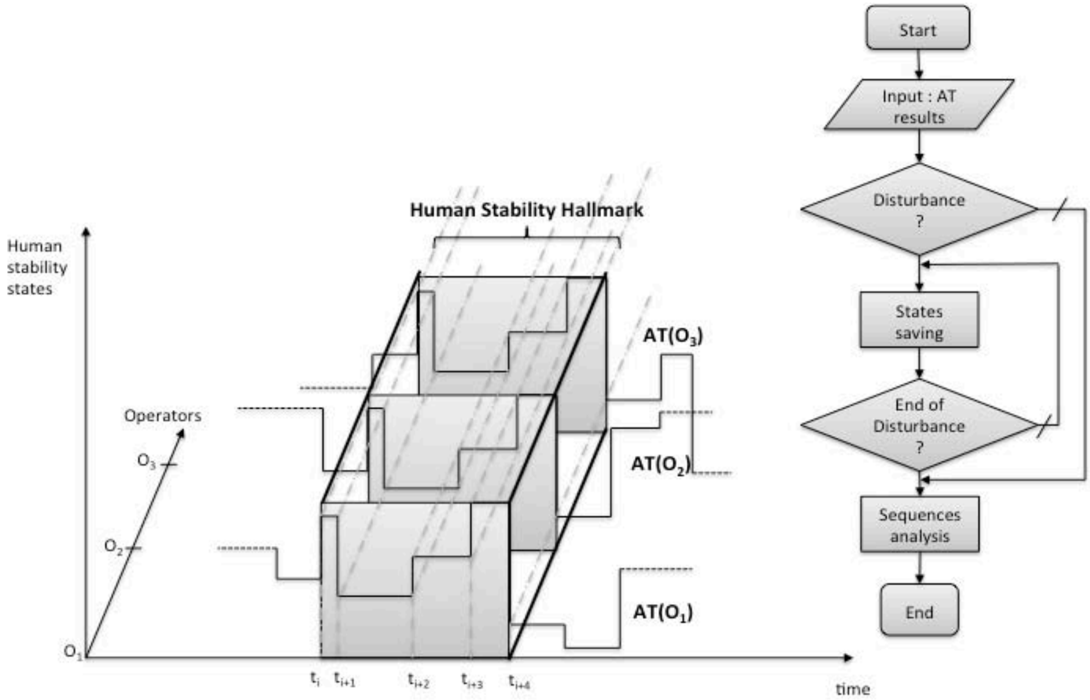

Figure 4 AS algorithm diagram

4.1 The ERTMS platform

The ERTMS platform is made up of various modules (traffic management module, driving module, 3D module to reproduce the driving environment) and is compliant with SRS 2.3.0d (European Railway Agency, 2008). The objectives of this platform are mainly to optimize the traffic management, to certify real railway components and software in a virtual environment, to test driving situation for different rolling stock configurations and to train drivers and maintain their knowledge.

4.2 Experimentations

The simulation put 10 students in a driving situation with a high-speed train and a 60 kilometres long track made up of various infrastructure elements (bridge, stations, tunnels, level crossing, etc). The traffic on this track is light and 6 events are positioned in the track at different milestones in order to disturb the drivers with a work area, a cognitive dissonance (contradictory data between on-board signals and external signalling: Authorization

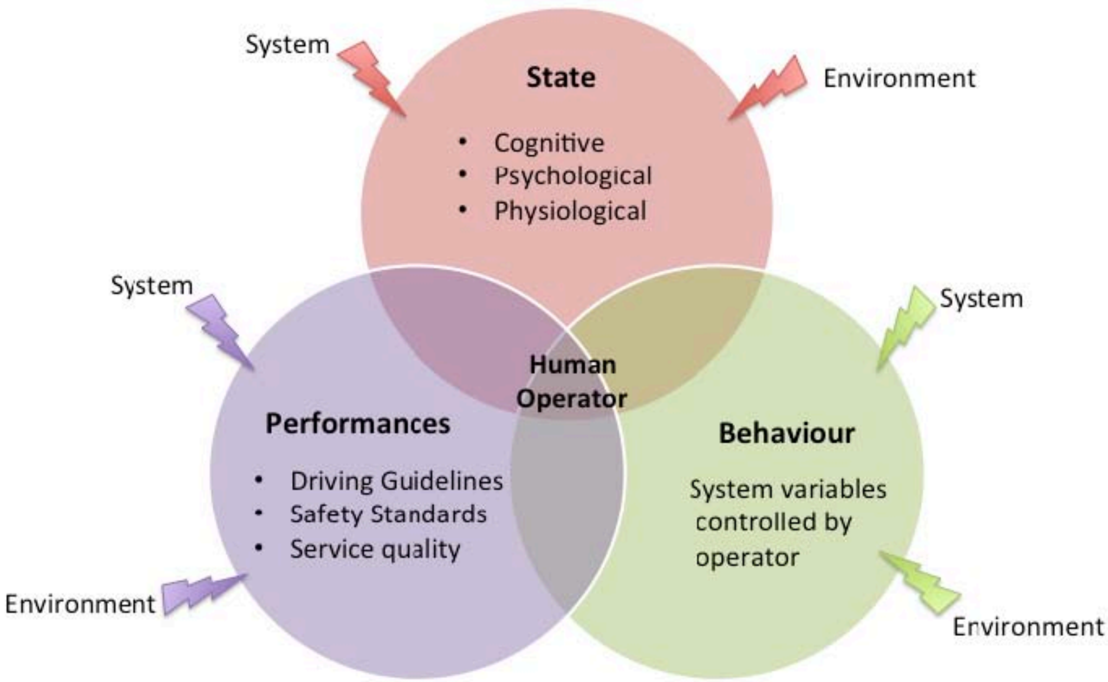

Figure 5 Criteria related to the human operator

by DMI to pass a red light signal), 3 changes of ERTMS level, and a change of driving mode (transition from full supervision to on-sight mode). The experimentation aims to identify via the AT algorithm the states and change of states natures for 3 studied criteria (to respect the speed instructions, to respect the train timetable, to ensure the passengers comfort) in order to extract via the AS algorithm some operator behavioural signatures during the scenario and in particular when disturbance occurs (see figure 6). These criteria are derived from the "behaviour" category defined in 3.4.

state of the operator when a disturbance occurs. It determines what state of stability or transition between states was concerned at the occurrence of the disturbance for the studied criterion. Once the states and transitions identified by criterion for each student, then, AS algorithm allows detecting the signature associated to each disturbance. For the case study here, a signature is considered as the sequence of 3 states or transitions (before the occurrence of the disturbance, at the time of the occurrence of disturbance, after the occurrence of the disturbance).

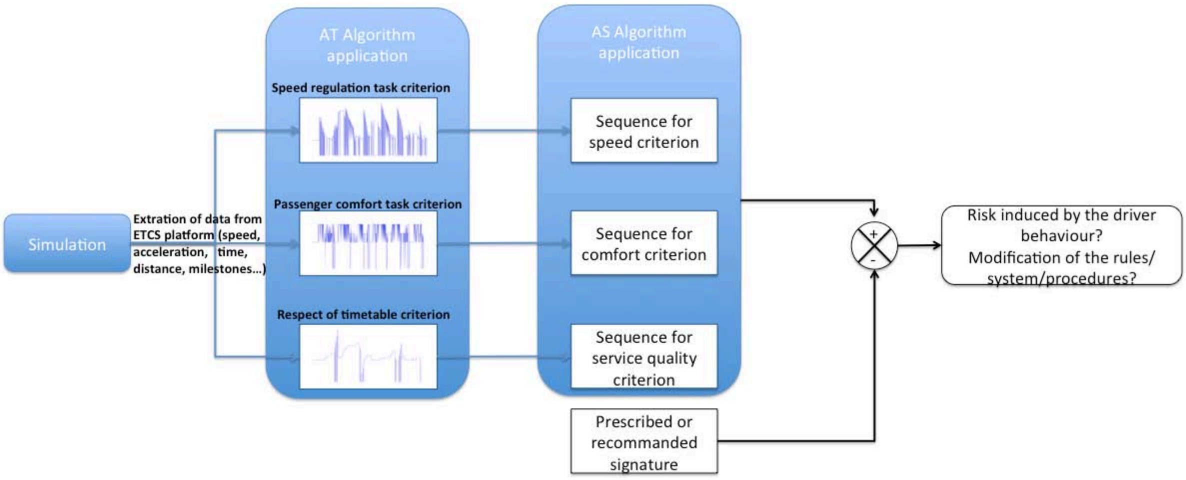

Figure 6 Protocol of the experiment

4.3 Results

The first results obtained by the AT algorithm show the

These signatures are sequences of stability states and can be recommended or risky. Recommended signatures entail a success in the disturbance management while risky signatures entail a failure in the disturbance management. Figure 7 illustrates results for some disturbances examples processed by AT and AS algorithms. It shows the occupation rate of the operator in the different states and transitions for each disturbance and the advised signature when the cognitive dissonance occurs.

State	Criterion	Acquittal fault	To pass the red light	Train collision
Stable state	Speed instruction criterion	42,66%	0%	0%
	Passengers comfort criterion	71,43%	0%	100%
	Service quality criterion	85,71%	100%	100%
Unstable state	Speed instruction criterion	42,86%	0%	100%
	Passengers comfort criterion	0%	0%	0%
	Service quality criterion	0%	0%	0%
Leaps	Speed instruction criterion	0%	0%	0%
	Passenge's confort criterion	21,57%	100%	0%
	Service quqlity criterion	0%	0%	0%
Breaks	Speed instruction criterion	14,29%	100%	0%
	Passengers comfort criterion	0%	0%	0%
	Service quality criterion	14,29%	0%	0%

Data obtained by AT algorithm

Criteria	Detected signature for the cognitive dissonance disturbance (to pass the red light)
Speed instructions	Stable/stable/stable
Passengers comfort	Stable/stable/stable
Service quality	Stable/stable/stable

Data obtained after AS algorithm use

Figure 7 Some results extracted from the AT and AS algorithms for some examples of disturbances

5 Conclusions and perspectives

This paper proposes the study of a new concept for evaluating the behaviour of a Human operator in the man-machine systems: the human stability. The experimentations discussed in this article assess different criteria independently. It allows identifying the nature of states and state changes linked to the parameter "human stability " and highlights recommended and risky signatures for each disturbance. Nevertheless, the mono-criterion study does not seem sufficient. The operator activity cannot be reduced to a single criterion, but may be influenced by a set of criteria from different categories (workload, stress, personal or physical problem...) together influencing the system. The multi-criteria study will require a proposal of a new formalism to improve the study of human stability. It will also be necessary to weight the criteria, i.e. to provide a level of importance for each or a set of criteria. This work suggests a diagnosis of human stability too: why the state of the driver behaviour was unstable at the time of the disturbance occurrence, was this a special case? It would be interesting to investigate deeper with more students in order to assess with an acceptable degree of certainty the signature at the time of the disturbance. It seems interesting too, to develop this concept with an objective of prediction. In (Richard et al., 2009), it is proposed a study on Human operator modelling by the dynamic hybrid system community. This type of model can take into account continuous and discrete components of the Human operator. Another perspective of these works is to extend the study of human stability to the others categories evoked in paragraph 3.4, in particular for the facial recognition (Luong, 2006). In this context, determination of states and transitions via facial recognition application can be done on line. Lately, by combining several criteria of different categories, this kind of application could be implemented in the driver cab in order to identify on line the human stability and to alert the operator when his/her behaviour seems risky.

6 References

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems., Safety science, pp 109-126.

Duquesne,L. (2005). Jugement multicritère d'acceptation individuelle de la signalisation routière variable.

Mémoire de Master Recherche AISIH, UVHC, Valenciennes, France, Juillet 2005.

European Railway Agency (2008), System Requirements Specification (subset 026), technical report.

Fossion, P. and Linkowski, P. (2007), La pertinence du concept de résilience en psychiatrie. In Rev Med Brux. Goussé, V. (2005), Apport de la génétique dans les études sur la résilience : l'exemple de l'autisme. Annales Médico-Psychologiques.

Hollnagel, E. and Woods, D. (2006) Resilience Engineering: concepts and precepts Epilogue: Resilience Engineering Precepts. Ashgate Publishing Ltd.

Ligeron, J.C. (2006), Le cercle des fiabilistes disparus ou critique de la raison fiabiliste. Editions préventique, 2006.

Lundberg, J. and Johanson, B. (2006). Resilience, stability and requisite interpretation in accident investigations. Proceedings of the Second Resilience Engineering Symposium, (Eds.) Hollnagel, E. & Rigaud, E., 191-198

Luong H.V. (2006), Reconnaissance multimodale de gestes de communication non verbale.

Martin S. (2005), la résilience dans les modèles de systèmes écologiques et sociaux, PhD thesis, ENS Cachan.

Morel G. et al. (2009). How good micro/macro ergonomics may improve resilience, but not necessarily safety . Safety Science, 47, 285–294.

Poupon, Y. and Arnould, P. (2006). Mécanique : choix d'un matériau polymère injectable.

Rasmussen, J. (1980). What can be learned from human error reports? In K. Duncan, M. Gruneberg & D. Wallis (Eds.), Changes in working life. Wiley: London.

Riana S., Terje A. (2011). A risk perspective suitable for resilience engineering. Safety Science, 49, 292–297.

Richard, P., Vanderhaegen, F., Benard, V. and Caulier, P. (2009). Proposition of establishment of a model of human-machine system by hybrid dynamic system model applied to guided transport. Proceedings of 8. Berliner Werkstatt Mensch-Maschine-Systeme, Berlin, Germany, 7-9 October 2009.

Richard, P., Benard, V., Vanderhaegen, F., Caulier, P. (2010), Towards the Human stability in transportation systems : concepts and objectives. In IFAC HMS, 2010.

Richard, P. (2012). Contribution à la formalisation et à l'identification de la stabilité humaine au regard de la sécurité : application aux transports guidés. Phd Thesis. Université de Valenciennes et du Hainaut-Cambrésis.

Zieba, S. et al. (2007), Resilience and affordances: perspectives for human-robot cooperation? European Annual Conference on Human Decision Making and Manual Control