MODELING HUMAN-MACHINE INTERACTION FOR THE ASSESSMENT OF HUMAN RELIABILITY

Daniel Schwencke[1], Jan Grippenkoven[1], Karsten Lemmer[1]

[1] *German Aerospace Center*
Institute of Transportation Systems
Lilienthalplatz 7, 38108 Braunschweig
Germany

{daniel.schwencke|jan.grippenkoven|karsten.lemmer}@dlr.de

There has been recent research on the analysis and assessment of human-machine interaction in railways based on barriers. Considering the theory of multiple resources, we take this approach further and develop additional assessment criteria as well as a more elaborate notation. We report on how the results of a simple assessment of some barriers compare to those obtained from a different method.

Introduction

Human error plays a crucial role in the majority of critical railway events. In the past the focus was set mainly on technical developments in order to improve safety. Operators were more and more replaced by technical systems to enhance safety. The corresponding shift from active physical work towards supervision tasks changed the requirements on humans and their working environment needed to enable reliable performance; in particular, cognitive aspects of monitoring became more important.

The methods to analyze and assess human reliability that are used in railway practice today, however, most often do not adequately reflect these changes. Either they consist of oversimplified approaches to human reliability or they are based on approaches and collected data from other industries (e.g. nuclear industry or air traffic) with different characteristics and influencing factors. European endeavors towards a common risk based approach to railway safety increase the need for an adequate and quantitative assessment of human reliability.

These challenges are addressed in the project SMSmod (“System Mensch-Sicherheit modellieren”) funded by the German research foundation (DFG). In this paper we present those parts of the project pertaining to further developments of a method proposed by Hammerl (2011). First, we recall the original method for the analysis of safety-relevant human-machine interaction and the assessment of human reliability. Second, we enlarge the set of assessment criteria of the method where Wickens’ multiple resource theory is taken into account (Wickens & McCarley 2008). Third, we revise the diagrammatic presentation which is part of Hammerl's work and introduce a more explicit representation of the assessment criteria for safety-relevant human-machine interaction. Finally, we report on results of the application of a simple approach to the assessment of human reliability using the method.

Modeling the Human-Machine Interaction of a Barrier

We briefly recall the approach of Hammerl (2011) to the analysis of prevention and protection mechanisms (so-called "barriers") in railways. This analysis mainly results in a diagram displaying the interaction of a human with one or more technical systems (referred to as "the machine"), the interaction implementing the safety function of the barrier. The analysis serves as a basis for the assessment of the barrier's reliability, including in particular the influence of the human on the safety function, as we shall see in the sections below. It also serves to draw conclusions about the impact of human actions on safety. For a more comprehensive introduction we refer to (Talg, Hammerl & Meyer zu Hörste 2012).

Barriers

Sklet (2006) defines barriers as "physical and/or non-physical means planned to prevent, control, or mitigate undesired events or accidents". Hollnagel (2008) distinguishes four types of barriers:

- physical barriers, which prevent an event from taking place or block or mitigate its effects by their mere existence,
- functional barriers, which set up one or more pre-conditions that have to be met before an action can be carried out,
- symbolic barriers like signs, which require an act of interpretation in order to achieve their purpose, and
- immaterial barriers like rules, which lack material form or substance in the situation where they are applied.

Moreover, barrier function and barrier systems are distinguished: the barrier function is the specific way in which the barrier fulfills its purpose. The barrier system is the organizational and/or physical structure, without which the barrier function could not be achieved. One barrier may consist of several barrier systems.

Symbolic and immaterial barriers always depend on human actions, functional ones may depend on human actions, but often in a limited way. Physical barriers are independent of human actions and thus are not considered in the sequel.

Analysis of the Human-Machine Interaction

First of all we observe that for functional barriers the machine fulfills the barrier function, whereas for symbolic and immaterial barriers the human does. For further analysis, we decompose the realization of the barrier function into three steps:

- barrier initiation (BI), i.e. information perception by the entity which performs the barrier processing,
- barrier processing (BP), i.e. deriving the necessary actions from the information, and
- barrier execution (BE), i.e. carrying out the derived action.

Drawing the last two steps in the human-machine system, we obtain three types of interaction as displayed in Figure 1. Types A and B correspond to symbolic and immaterial barriers, type C represents functional barriers. An arrow between human and machine shows an information flow or an action, a loop on the human side describes cognitive information processing. The barrier execution arrow of type B is the only one that crosses the border of the human-machine system which is for example the case if a train driver passes information about an object on the neighboring track to the dispatcher.

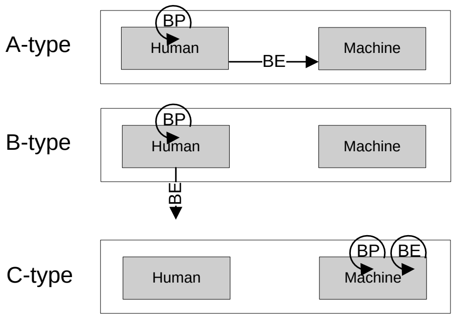

Figure 1: Basic types of human-machine interaction

In order to complete the interaction, different patterns for the barrier initiation may be added to the basic interaction types from Figure 1. Figure 2 shows on the left-hand side the combination of type A with information retrieval from one barrier system with active initiation (a signal that requests the observer to do something).

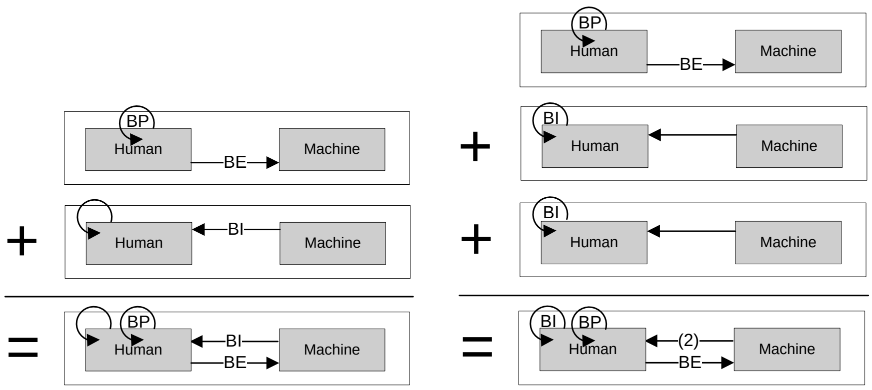

Figure 2: Combination of basic type A with different barrier initiation patterns

On the right-hand side, we have again type A which this time is combined with information retrieval from two barrier systems with passive initiation (the human needs to recognize the problem).

In the same way, the basic type B may be combined with different barrier initiation patterns. For type C, we distinguish two kinds of initiation patterns: either an erroneous human action may initiate a safety function of the machine (type C-I), or a safety function may be deactivated by the human (C-II). For a detailed discussion of possible initiation patterns for the basic types see Chapter 6.2 of (Hammerl 2011).

Safety Implications of Different Types of Barriers

In Table 1 we give an overview of the allocation of the safety function and of the risk for the different basic interaction types. Using barriers of type A or B, a "normal" human error may directly cause a safety risk; using barriers of type C-I or C-II, a technical function is additionally involved.

Table 1: Allocation of safety function and safety risk

Barrier Type	Allocation of Safety Function	Allocation of Risk
A/B	Human performs safety function	Risk in case of a human error
C-I	Reaction of technical system in case of a human error	Risk in case of a human error *and* a technical system failure
C-II	Human may deactivate safety function	Risk in case of a technical system failure *or* an error of commission

Table 2: Consequences of "normal" errors and errors of commission

Barrier Type	"Normal" Human Error	Error of Commission
A/B	Failure of barrier function (safety risk!)	Useless execution of barrier function (possibly loss of performance)
C-I	Error is remedied by technical system (possibly loss of performance)	Performance function not needed
C-II	Useless execution of barrier function (possibly loss of performance)	Barrier deactivation despite needed barrier function (safety risk!)

Consequently, we must distinguish between "normal" human errors and *errors of commission* (i.e. actions where the human operator acts in good faith, which are right on the level of the behavior, but wrong in the particular situation). As shown in Table 2, depending on the type of barrier and the kind of human error the consequences for safety differ. Most important are normal errors in connection with barrier types A/B and errors of commission in connection with type C-II since in those cases a safety risk arises.

Criteria for the Assessment of Human Reliability

It is argued by Hammerl (2011) that the human reliability is not solely dependent on the type of barrier (symbolic, immaterial or functional). The hypothesis which leads us to the assessment of human reliability using the "human-barrier interaction" is that an increasing complexity of the human-machine interaction is equivalent to a more error-prone barrier. Based on the above analysis of the steps (BI, BP and BE) for establishing a safety function, we next set up complexity criteria which may be combined in order to obtain an estimation of the reliability of barriers, see below.

Criteria Given by Hammerl

Table 3 gathers several criteria from (Hammerl 2011) that may be taken into account for the assessment of human reliability.

Further Criteria

As explained below, the theory of multiple resources already was the reason to incorporate the criterion "multiple use of a sensory channel" in Table 3. But it also leads us to additionally consider the following criteria in the SysML notation for the assessment of human reliability:

- Barrier system: availability over time (a sensory channel is needed precisely in the time slot where the barrier system, e.g. a signal, is available)
- Barrier system: number of supporting sensory channels (according to the multiple resource theory reception via different sensory channels is more reliable)
- Whole barrier: further parallel tasks (the multiple use of a sensory channel may not only be caused by different barrier systems of one barrier, but also by other tasks that need to be performed in parallel)

There are still further criteria we propose to include for the assessment of human reliability since they affect the cognitive complexity:

- Barrier system: number of states of the barrier system carrier (less states better)
- Barrier processing: probability to derive the wrong action despite good intention (assessment according to human decision making literature)
- Barrier execution: ergonomics and usability factors (e.g. number of degrees of freedom to fulfill a task, feedback)

Multiple Resources in Human-Barrier Interaction

The incorporation of the concept of multitasking into human barrier interaction described in this section follows basic thoughts of Wickens' theory of multiple resources (2008). One of the main premises of this theory is that the human capacity to process environmental information is limited. A key factor that determines the complexity of a working environment for the human is the temporal order of appearance of relevant information. Tasks can either be dealt with stepwise or they have to be coordinated with other tasks in a parallel fashion.

Table 3: Complexity criteria for the human-machine interaction

Criterion	Effect on Reliability	Explanation
Criteria for the Complexity of the Barrier Initiation		
Barrier initiation	Active better than passive	Active initiation = by a signal/sign (i.e. from the machine side); passive initiation = the human needs to recognize a problem by himself
Number of barrier systems	Small number better	The number of barrier systems the human needs to retrieve information from during the barrier initiation phase
Spatial distance of barrier systems	Short distance better	The distance is short in case there is only one barrier system or several systems close by, e.g. on the driver's desk; it is high if e.g. a driver needs to consider barrier systems from the trackside and inside the cabin
Multiple use of a sensory channel	The less the better	Maximum number of barrier systems that require the use of the same sensory channel
Criteria for the Complexity of a Single Barrier System		
Frequency of interaction	Higher frequency better	Describes whether interaction with the barrier system is an everyday task or rather seldom (e.g. rare signals or signs for the train driver)
Presentation of information	Explicit and unambiguous better	Rates how easily information can be perceived and interpreted correctly subject to its presentation
Criteria for the Complexity of the Barrier Processing		
Time between barrier initiation and execution	Shorter time better	The time period (the necessary delay of the barrier execution) in which the human needs to keep in mind the pending execution of the barrier (and in which he might forget about the remaining actions)
Criteria for the Complexity of the Whole Barrier		
Available time	Longer time better	The time until the barrier function needs to be executed at the latest (which affects pressure/stress)

The role of the factor "number of barrier systems" as described in Table 3 therefore has to be extended by the factor time. In human-barrier interaction the number of barrier systems the human needs to retrieve information from especially gets important when the time to attend all of the important information is limited. If the number of barrier systems that have to be attended at the same time exceeds one during barrier initiation phase, the problem of interference can occur due to the necessity of parallel processing and limited human perceptual resources.

According to the multiple resource theory, when it comes to parallel processing in an early stage of perception, the modality (visual or auditory) addressed by different kinds of information plays a key role. This idea is tackled in the aspect "multiple use of a sensory channel" in Table 3. The negative effect of interference is the strongest if two (or more) subtasks demand processing in the same perceptual modality at the same time. In the railway domain, this could occur e.g. when a train driver approaches a signal besides the track that he should attend and at the same time has to monitor the tachometer and mile posts for controlled speed reduction. Note that in all of these subtasks the visual attention of the driver is required. A probable consequence of the resulting interference could be that one of the three sources of information can not be attended adequately (or is even missed, in case of the signal). The consequence would be disturbed barrier processing, which could either lead to a declined speed adjustment or even to a so called SPAD (signal passed at danger).

As stated earlier, in early information processing the problem of interference only arises when the same modality is required. When multiple information occurs in a way that it touches different perceptual modalities (e.g. regarding a signal while listening to an operator's radio message), parallel processing is usually less challenging. The incorporation of human processing capabilities under multitasking conditions over time is a major advance of the revised notation for human-barrier interaction presented in the following section. Note that there are methods like CPM-GOMS (Gray, John & Atwood 1992) that follow similar ideas but only apply to human-computer interaction which is insufficient for a train driver's work place.

More Detailed Interaction Diagrams

Consider Figure 3 where the human-machine interaction was modeled for a barrier against speeding according to the method described above.

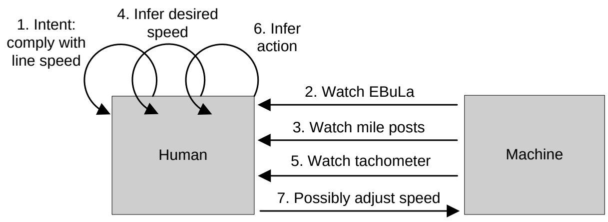

Figure 3: Barrier against speeding – simple diagram

Steps 1-5 belong to the barrier initiation which involves the three barrier systems EBuLa[1], mile posts and tachometer, steps 6 and 7 represent the barrier processing and execution, respectively (so this interaction is of type A). Some of the above

[1] Electronic system on the driver's desk displaying time table and speed restrictions

criteria (e. g. whether a technical system or the human initiates the barrier) are visible in the diagram, others (e. g. the time between barrier initiation and execution) are not. Additionally, this notation offers no possibility to get an overview of interferences in human information processing during the parallel execution of different subtasks.

Using SysML Sequence Message Charts

We present a more elaborate notation for the human-machine interaction based on sequence message charts from the SysML modeling language (Figure 4). This notation additionally allows to model temporal aspects and to highlight how multiple resources are affected during parallel human information processing as well as different technical devices as part of the barrier. Most of the identified criteria for human reliability assessment become explicit in the new notation.

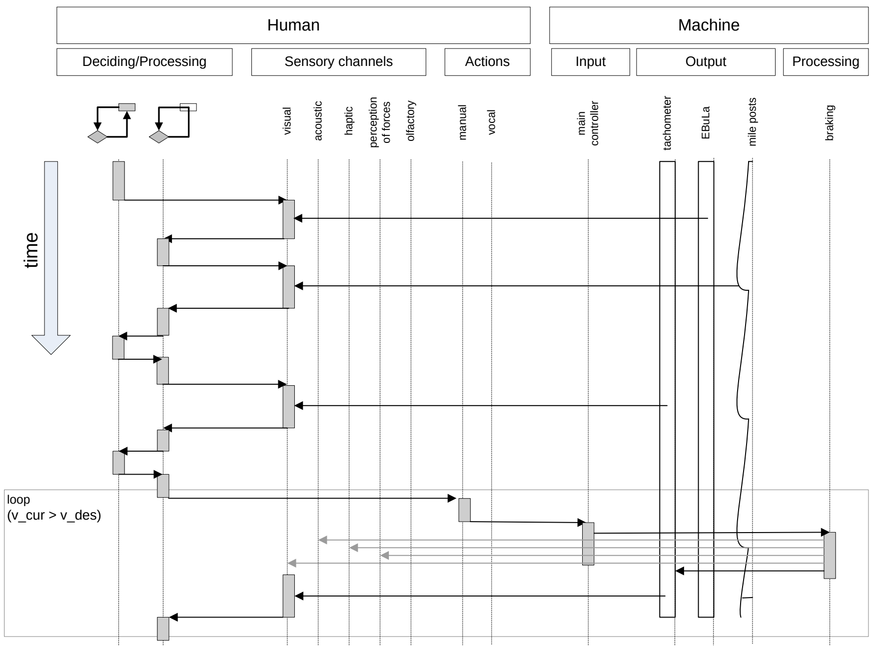

Figure 4: Barrier against speeding – sequence message chart

The barrier against speeding from Figure 3 modeled in a sequence message chart is shown in Figure 4. Human and machine are decomposed into inputs, processing and outputs which are called sensory channels, deciding/processing and actions for the human. The human processing is divided into non-automated and automated (simple and frequently performed) processing. Time flows in downwards direction. The activity of a part of the human-machine system is indicated by the filled gray rectangles which are drawn on its dashed lifeline; for the machine outputs, the white areas indicate the availability to the human. The black arrows show the main information flow, the gray ones show supporting information flow. The area in the bottom bordered by the gray line indicates a closed-loop control.

To meet concerns about limited human resources, the notation as presented in Figure 4 offers the possibility to combine sequences of more than one task over time in one chart, to model parallel processing. The human-barrier interaction in the task “attending a trackside signal” for example could be implemented in the sequence chart of the barrier presented above. Overlapping rectangles on the lifelines e.g. of sensory channels or actions offer a direct visual impression of tasks in which interference could become a problem if they occur together at the same time.

In the theory of multiple resources, besides the dimension of the perceptual modality (visual or auditory), Wickens (2008) further differentiates visual perception (focal and ambient vision), tasks (spatial and verbal codes) and stages of information processing (perception, cognition and responding). An implementation of those dimensions into the presented SysML sequence message charts appears reasonable.

Human Reliability Assessment

Application of Hammerl’s approach

As demonstrated by Hammerl (2011), his criteria may be used to compare different barriers with the same purpose with respect to human reliability. Hammerl also outlines an approach capable of comparing barriers with different purpose. We report on our experience applying this approach, which so far is only based on the following criteria (cf. Table 3): barrier initiation, presentation of information of a barrier system, frequency of interaction with a barrier system, and available time. The second and third criteria are taken together in order to determine whether the human processing of a barrier system is automated or not. According to Table 4, one obtains a score between 1 (highest reliability) and 7 (lowest reliability) for a barrier. For example, in case of the above barrier against speeding (passive initiation, no time pressure, three times non-automated human processing) this yields a score of 5.

Table 4: A possible approach to a barrier’s reliability (Hammerl 2011)

		no time pressure		time pressure	
		active initiation	passive init.	active initiation	passive init.
number of non-automated human processing	0	1	2	3	4
	1	2	3	4	5
	2	3	4	5	6
	>2	4	5	6	7

Comparing to the CAHR method

A central part of the SMSmod project is to analyze some event data provided by Deutsche Bahn using the second generation human reliability assessment method CAHR (“Connectionism Assessment of Human Reliability”) developed by Sträter (2000). This way we obtained basic reliabilities for some human-machine interactions we could compare to our results. Despite the small number of interactions that could be modeled with both methods in a comparable way so far, the comparison indicates that Hammerl's approach is generally valid: an interaction rated with a higher score than another one was accordingly rated less reliable than the other one by CAHR in all cases. The comparison also gave an impression of the range of reliabilities that is covered by the scores. It appeared that the scores are nearly linearly distributed on a certain reliability interval, but might not sufficiently differentiate high reliabilities.

Summary and Future Work

We extended the work on human-barrier interaction by Hammerl (2011) in several directions: further complexity criteria were proposed as well as a more elaborate notation, partly based on Wickens' multiple resource theory. The extensions preserve the original aim to develop a method which is based on cognitive psychological findings but can be applied by railway engineers without a psychological background. The method was used for a simple assessment of the reliability of some barriers which proved to be sound when compared to results of the CAHR method.

One direction for future work might be to investigate parallel tasks via overlay of sequence message charts. Also, the sequence message charts could be completed by some new notation in order to make the few remaining criteria visible. Towards a semi-quantitative assessment of human reliability, the next step would be to integrate further criteria taking interdependencies of the criteria into account. For a better scaling and calibration reliability results for further barriers from other methods are needed; in order to calibrate the impact of important performance shaping factors, one might use again CAHR results complemented by simulation studies.

References

Gray, W. D., John, B. E. and Atwood, M. E. (1992), The précis of Project Ernestine or an overview of a validation of GOMS. In P.Bauersfeld, J.Bennett & G.Lynch (eds.) *CHI Conf. on Human Factors in Comput. Sys.*, 307-312, (ACM Press)

Hammerl, M. 2011, *Analyse der menschlichen Einflussfaktoren und Zuverlässigkeit im Eisenbahnverkehr*, Ph.D. thesis, Technische Universität Braunschweig

Hollnagel, E. 2008, Risk + barriers = safety? *Safety Science*, 46, 221-229

Sklet, S. 2006, Safety barriers: Definition, classification, and performance. *Journal of Loss Prevention in the Process Industries*, 19, 494-506

Sträter, O. 2000, *Evaluation of Human Reliability on the Basis of Operational Experience*. GRS-170, (Gesellschaft für Anlagen- und Reaktorsicherheit, Köln).

Talg, M., Hammerl, M. & Meyer zu Hörste, M. 2012, Integration of human factors to safety assessments by human barrier interaction. In *Railway Safety, Reliability and Security: Technologies and Systems Engineering*, (IGI Global)

Wickens, C. D. & McCarley, J. S. 2008, *Applied attention theory*, (CRC Press)