Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson

doi: 10.3850/978-981-18-5183-4_S06-09-254-cd

Identification of human driver critical behaviors and related reliability evaluation in real time

Chao He

Chair of Dynamics and Control, University of Duisburg-Essen, Germany. E-mail: chao.he@uni-due.de

Dirk Söffker

Chair of Dynamics and Control, University of Duisburg-Essen, Germany. E-mail: soeffker@uni-due.de

The role played by humans is becoming more and more important as the proportion of human-related accidents is increasing in industry and traffic. Human error taxonomies and their applications in driving context improve the understanding of human error mechanisms in situated driving context. In previous works, the authors provide a human performance reliability score (HPRS) which can be applied to driving data using the modified fuzzy-based CREAM (cognitive reliability and error analysis method) approach. The data clustering approach FN-DBSCAN (fuzzy neighborhood density-based spatial clustering of application with noise) with genetic algorithm is applied to automatically generate membership functions characterizing the driving behaviors individually. The driving behaviors and the mechanism of human error to the corresponding HPRS numbers are not analyzed in previous works. In this contribution, the classification of human driver error and its application in driving context is reviewed. The driving behaviors and the different human errors with continuously calculated values are analyzed to investigate what really happens. Human driver reliability is evaluated especially in situated context, this means dynamically changing situations (on a second-timescale). The newly developed approach provides a dynamic measure and therefore allows to dynamically identify critical situations during operation in real time. As example the supervision of an interacting human driver is shown.

Keywords: Human error taxonomy, human reliability analysis, dynamic context, FN-DBSCAN algorithm, modified CREAM, driving behavior.

1. Introduction

The development of automation is shifting the role of humans from active controlling to passive monitoring McDonnell et al. (2021). Human operators should maintain situation awareness and manually take control when automation is incapable of dealing with the problem. In human-machine systems, the role played by humans is becoming more and more important as the proportion of human-related accidents is increasing. In traffic context the US national highway traffic safety administration (NHTSA) stated that 94 % of traffic accidents are related to human factors Singh (2015). Many advanced driver assistance systems (ADAS) are developed to assist in some human driver operations and monitor human vigilance to avoid accidents Moujahid et al. (2018). The society of automotive engineering (SAE) defines six levels of automation regarding driving from level 0 of no automation to level 5 of full automation Cárdenas et al. (2020). From the definitions, it is detected that the human driver is not able to be decoupled with driving activities even with full automated vehicle as the driver still needs to monitor the driving situations and possibly to takeover the vehicle. Therefore, the analyzing of human driver errors is still vital for driving safety.

Human reliability analysis (HRA) provides structured methods to evaluate human reliability in different application. The fundamental step in HRA is the identification of human errors. To successfully identify human errors, human error classification is necessary Philippart (2018). Human error classification is used to analyze errors occurring in accidents and anticipate potential error that may happen. Human error classification improves the understanding of human error mechanisms and suggests the hints to avoid human errors Baysari et al. (2009). With the development, many HRA methods are developed and classified into ”three generations”. The so called ”first generation” considered the human as a mechanical component who fails to execute tasks. In the so called ”second generation”, context is the most important factor affecting human reliability and more cognitive models are considered. Further improvements related to pre-existing methods are driven by the limitations and deficiencies of the ”second generation” methods. The dynamic progression of human behavior is considered and the studies have focused on defining the database of HRA to overcome the shortages of empirical data for the development and validation of HRA models. These data-based and dynamic-considered methods are classified as ”third generation” HRA, these methods are still in development Di Pasquale et al. (2013). In HRA methods, cognitive reliability and error analysis method (CREAM) is wildly applied to conduct a retrospective analysis of event and prospective analysis for the design of high-risk systems or process.

The widely used ”first generation” and ”second generation” of HRA methods are static as time is less involved Zio (2009). When situated context and dynamic cognitive and actions are considered, these methods are not suitable, an adaption should be generated to integrate dynamic features. A HRA method considered as dynamic should account for the evolution of performance shaping factors (PSFs) and their consequences to the outcome of events. Moreover, dynamic human reliability focuses on a detailed step-by-step breakdown of human actions and intentions with time Boring and Rasmussen (2016). For continuously changing driving context, static HRA is not suitable, so a definition for dynamic HRA is required.

In the previous work He et al. (2021), the authors provide a human performance reliability score (HPRS) which is applied to driving data collected from driving simulator with the modified fuzzy-based CREAM approach to evaluate human driver reliability in dynamic context. The applied data clustering approach is fuzzy neighborhood density-based spatial clustering of application with noise (FN-DBSCAN). Here, the used gradient of the membership functions of the clustered data is not modeled optimally (too steep), which leads to abrupt HPRS changes. This leads to misinterpretations, so that the fuzzy-based filter must be improved accordingly. The challenge is to simultaneously model the filter sensitive for dynamic driving situations.

The following sections make up this contribution: In section 2, human error taxonomies are briefly reviewed and human errors in driving context are discussed. In section 3, the modified fuzzy-based CREAM approach is presented to explain procedures of automatic generation of membership functions based on driving data and the steps to obtain HPRS. Driving behaviors and the mechanism of human error with the continuously calculated values are analyzed to investigate what really happens in section 4. The summary and outlook are provided in section 5.

2. Human driver error taxonomies

The application of human error classification schemes is common in complex safety critical systems. Different human error taxonomies have been developed to understand human error mechanisms.

2.1. *Human error definitions*

Two views are distinguished between ”old view” and ”new view” of human error Dekker (2017). In the ”old view”, human error is the cause of trouble and it is a simple problem, when all systems are working well, people just need to pay attention and comply to avoid human errors. People can, and must, achieve zero errors, zero injuries, and zero accidents. In the ”new view”, human error is a symptom of deeper trouble and the complexity of generating human error is depending on the complexity of the organization and environment. People can, and must, enhance the resilience of the people and organization.

In the process of understanding human error, different definitions and related glossary of terms have been proposed. Swan and Guttman defined human error as an error that is simply an action which is out of tolerance, where the limits of the tolerance is defined by the system Swain and Guttmann (1983). From Rasmussen's point of view Rasmussen (1983), human error can only be described with reference to human objectives or expectations, it depends on the explicit situation. From Reason Reason (1990), it is obtained that human error is taken as a universal term to comprise all the occasions which a planned sequence of mental or physical activities fails to generate the intended outcome, and these failures cannot be associated to the intervention of some chance agency. Hollnagel defined human error as an erroneous action which fails to generate the expected result and/or which produces an unwanted consequence Hollnagel (1998). In Dhillon's definition Dhillon (2017), human error is the failure to execute a stated task that could result in interruption of scheduled operations or damage to property and equipment.

2.2. *Human error taxonomies*

Various human error taxonomies have been proposed. Three dominated taxonomies are reviewed in this contribution, which are Rasmussen's skill, rule, and knowledge error Rasmussen (1986), Reason's slips, lapses, mistakes and violations Reason (1990), and Hollnagel's phenotypes and genotypes Hollnagel (1998).

In Rasmussen's skill-,rule- and knowledge-based (SRK) behavior model, errors are affected by skills, experience and familiarity with the situation encountered. The generic error-modeling system (GEMS) is applied to classify these errors Reason (1990).

- Skill-based behavior is developed without conscious control as smooth, automated, and highly integrated patterns. Skill-based error is typical detected in routine repetitive work.
- In rule-based behavior, the actions are often controlled by a memory-based stored rule or procedure.
- The performance which is goal-controlled during unfamiliar situations, which no rules for control are available is knowledge-based behavior.

Reason classified human errors into slips, lapses, mistakes and violations. When combining with Rasmussen's SRK model, skill-based errors correspond to slips and lapses, rule-based and knowledge-based errors are related to mistakes.

- Slips are errors which result from some failures in the execution of an action sequence. Slips can be seen as externalized actions not conducting as planned.
- Lapses are errors which result from failures in the storage stage of an action sequence. Lapses are generally used for more covert error forms, including failures of memory.
- Mistakes are failures in the inferential and/or judgemental processes in the selection of an objective. Mistakes are more subtle than slips and harder to detect.
- Violations relate to actions habitual or isolated departure from rules and regulations.

In Hollnagel's CREAM approach, it is stated that human actions/errors are all to some extent cognitive, indicating that they are not able to be properly described without consideration of human cognition. Human error can be identified as phenotypes and genotypes.

- Phenotype concerns with the manifestation of an erroneous action. It can be divided into action at wrong time, action of wrong type, action at wrong object and action in wrong place/sequence.
- Genotype refers to the possible causes such as the functional characteristics of the human cognitive system that are assumed to contribute to an erroneous action. Human related genotypes can be further divided into observation, planning, interpretation, temporary person related causes and permanent person related causes.

2.3. *Driving errors*

Identification and classification of driving errors contribute to the understanding of human error mechanisms and the development of assisted driving systems. In Stanton and Salmon (2009), the existing driving error taxonomies are reviewed and a generic driving error taxonomy with underlying psychological mechanisms is proposed including action errors, cognitive and decision-making errors, observation errors, information retrieval errors and violations. From Khattak et al. (2021), driving errors can be classified into recognition errors, decision errors, performance errors, physical condition related errors, experience/exposure errors and violation. With the development of assisted driving systems, the role of human driver is gradually switching from active maneuvering to passive monitoring, so the related driving errors are also changed. Some new driver errors occur in assisted driving because new driving tasks may lead to misunderstanding and/or inappropriate reactions on drivers Noy et al. (2018).

3. Modified fuzzy-based CREAM approach

The modified fuzzy-based CREAM approach established in He et al. (2021) is applied for automatic generation of membership functions and calculation of human performance reliability score (HPRS) to realize the individualized human reliability evaluation in real time.

3.1. *CREAM*

The CREAM approach as a so called "second generation" of HRA approach is applied for retrospective analysis of historic events and a prospective analysis for the design of high-risk systems or processes. It provides the human cognition model to illustrate the information processing which is denoted as contextual control mode (COCOM). It assumes that the degree of human operator's control on context is the most significant index for human performance reliability estimation. Four control modes in CREAM are defined, which are scrambled control, opportunistic control, tactical control, and strategic control. Strategic control is related to the highest reliability and scrambled control has the lowest reliability.

Common performance conditions (CPCs) represents the most vital factors in operation context, which are similar with the concept of performance shaping factors. There are nine CPCs defined in CREAM. Each CPC includes several levels and related expected effects on performance reliability which are improved, not significant and reduced. The CPC score could be calculated as [$\sum$ reduced, $\sum$ improved]. In this case, human performance reliability is determined with control mode map Hollnagel (1998).

To apply CREAM into driving context, a new CPC list has to be defined which includes the number of surrounding vehicles, time to collision (TTC), ego-vehicle speed, longitudinal acceleration, lateral acceleration, traffic density, and general visibility.

3.2. *Automatic generation of membership functions*

(i) Fuzzy logic: Fuzzy logic is used for modeling the imprecise modes of reasoning that play an essential role in human decision ability in an environment of uncertainty and imprecision Zadeh (1988). It considers the degree of truth of statements continuously between true (1) and false (0). To define the related membership function, the core and support points and membership function shape should be known. In this contribution, trapezoidal shape is selected.

(ii) FN-DBSCAN algorithm: To define the core and support points in membership functions, the fuzzy density clustering method Ulutagay and Nasibov (2008) fuzzy neighborhood density-based spatial clustering of application with noise (FN-DBSCAN) is applied.

(iii) Genetic algorithm: In FN-DBSCAN, the parameter of fuzzy cardinality threshold needs to be predefined, therefore, genetic algorithm is applied for the optimization of parameter.

To generate cores and support values of membership functions, FN-DBSCAN algorithm is applied. To evaluate the fitness of chromosomes, train and test data are fuzzified and through a KNN algorithm, fitness is calculated.

3.3. *Human performance reliability score (HPRS)*

The CPC levels are divided by data clustering, When membership functions of CPCs are generated, they will be assigned to different levels with corresponding expected effects on performance reliability which are improved, not significant, and reduced. In this case, each CPC score is calculated and human performance reliability score is generated with the sum of each CPC score.

In general, the steps to calculate HPRS are following: i) Execute genetic algorithm to obtain optimal value of fuzzy cardinality threshold ii) Apply the FN-DBSCAN to calculate cores and supports of membership functions of CPCs iii) Assign CPC levels and related effects on reliability ot membership functions to calculate CPC scores iv) Add up all CPC scores to get the final HPRS.

4. Results and analysis

4.1. *Data generation platform*

A driving simulator (SCANeR^{TM} studio, Fig. 1) is applied to collect driving data. Data with ego-vehicle dynamics (speed, steering angles, etc.) and surrounding vehicle status (TTC, lateral shift, etc.) relative to ego-vehicle are collected to evaluate driving behavior and human driver reliability.

Fig. 1. Driving simulator laboratory, Chair of Dynamics and Control, U DuE

The scenario in this work is a two-way high way with three lanes in each way. Normal daytime weather condition is implemented in this scenario. Other interacting vehicles are introduced during ego vehicle driving, so the ego vehicle driver is required to complete the driving tasks such as acceleration, deceleration, maintaining the speed, and changing lanes following the traffic rules. Human driver reliability is dynamically changing with driving maneuvering in different situations.

4.2. *Experimental results*

In this contribution, an example data set is contributed by a human driver with a valid driving license for eight years with approximately 250 kilometers per weekly driving and experience with driving simulator. The driving data between 400 s and 520 s are selected to generate the membership functions and the HPRS.

Four CPC data including ego-vehicle speed, TTC, longitudinal acceleration, and lateral acceleration are clustered and membership functions are generated for each CPCs. The CPC scores of traffic density and general visibility are defaulted to 1 as the scenario is simple with normal daytime weather condition and the lane are relatively empty. When no vehicle nearby, the effect of CPC of number of surrounding vehicles on performance reliability is improved, when there are no more than 2 vehicles around, the effect is not significant, when it is more than 2 surrounding vehicles, the effect is reduced.

The membership functions of the clustered CPC data are shown in Fig. 2. For the CPC of speed, three membership functions are generated. Therefore, the first membership function (green) could be assigned to improved effects, the second membership function (blue) could be assigned to not significant effects, and the last one (red) is assigned to reduced effects. On the contrary, for the CPC of TTC, the assignment should be the opposite, where the first membership function (green) is assigned to reduced effects and the third membership function (red) is assigned to improved effects as lower TTC indicates higher time pressure for human driver to recognize the situations and take actions. There is only one membership function in CPCs of longitudinal acceleration and lateral acceleration. In this case, the membership function with membership degree of 1 is assigned to not significant effects and the other part is reduced effects.

When the effects are assigned, the CPC score is calculated as presented in Fig. 3. It can be observed that CPC score of lateral acceleration is fluctuating frequently.

With the calculated CPC scores, the final HPRS values is calculated as Fig. 4. In He et al. (2021) the control mode levels in HPRS have been determined where strategic level has the value of 2 in HPRS and tactical level is -1. In this example, all

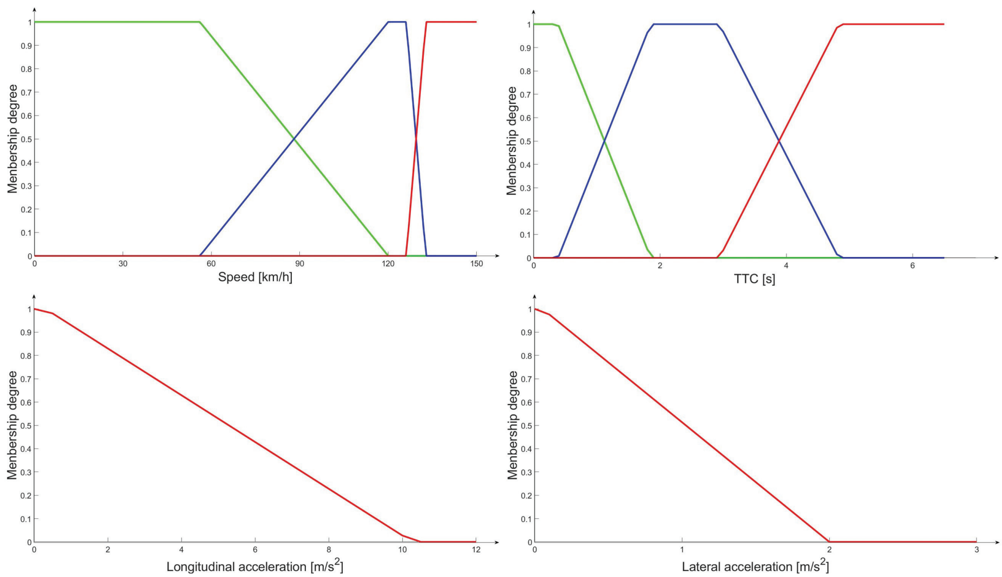

Fig. 2. Membership functions of CPCs

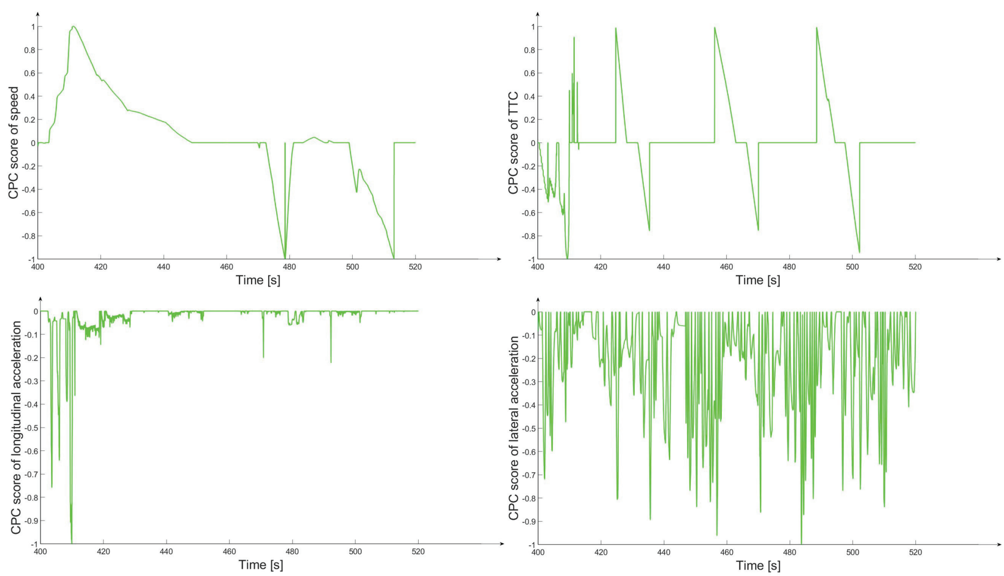

Fig. 3. Four CPC scores

values of HPRS is larger than -1, so tactical level is not shown. It can be detected that most of the HPRS values are above strategic level indicating that the driving performance is reliable. It is still meaning to investigate human driver performance in the peaks and valleys of HPRS, especially the dynamic changes of human performance during the valleys time. With the continuously calculated HPRS, human driver performance can be supervised in real time.

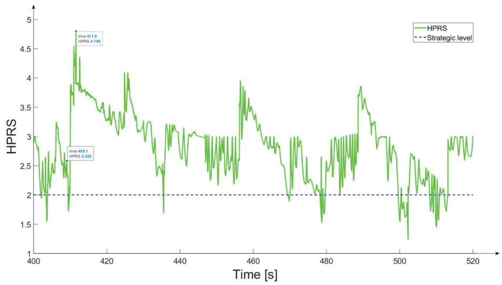

Fig. 4. HPRS results

4.3. *Analysis of human driver critical behaviors*

In this work, the time period between 409.1 s and 411.6 s in HPRS is studied to examine details of the driving process (as example). In this period of time, HPRS dramatically fluctuates with firstly decreasing to the bottom then increasing to the top. The situation could be described as follow:

- At the time of 409.1 s, the speed, TTC, longitudinal and lateral acceleration are 82.6 km/h, 0.46 s, 0.95 m/s^2, and 0.05 m/s^2, respectively. It should be noticed that before time of 409.1 s, the driver already pressed the brake pedal to slow down the speed, but the TTC is still decreasing and finally reached the minimum of 0.39 s.
- At time of 409.3 s, the TTC decreased to 0.39 s, the driver suddenly pressed the brake pedal hardly inducing the longitudinal acceleration increasing to -10.2 m/s^2 with the speed of 63.5 km/h and the related TTC of 1.42 s at time of 409.9 s.
- After the sudden and hard braking, at time of 411.6 s, the ego-vehicle speed decreased to 56.3 km/h with the related TTC of 4.6 s and longitudinal acceleration of 0.36 m/s^2.
- During maneuvering, the lateral acceleration has two relatively large fluctuations, one is at time of 409.3 s and the other is at time of 410.6 s with both of the absolute value of 0.73 m/s^2.

From the description of the critical situations, it can be detected that the human driver has perceived the situation (e.g. short time of TTC) and has pressed the brake pedal to slow down the ego-vehicle speed. However, the human driver did not identify that the TTC is still decreasing to critical level so the TTC between the ego-vehicle and front vehicle is misjudged. When the driver finally identified that the distance is too close, hard braking to slow down the speed is implemented. The values of lateral acceleration waving indicate the stress of driver when the critical situation is suddenly identified.

To classify the human driver critical behaviors, the taxonomy established in Stanton and Salmon (2009) is applied. It can be concluded that the human driver misjudges the TTC, so the error mode is classified into misjudgment which is related to situation assessment of cognitive and decision-making errors in the underlying psychological mechanism.

5. Summary and outlook

In this contribution, human driver critical behaviors are identified and evaluated in real time with a new human performance evaluation approach based on modified fuzzy-based CREAM approach. The driving data collected from driving simulator are clustered with FN-DBSCAN to generate individualized membership functions of different CPCs. The critical situation in driving process is analyzed, and the critical driving behavior is identified as misjudgment of TTC between ego-vehicle and front vehicle. This work contributes to generate a new measurement-based understanding of human error-related critical situations in situated context, here applied to a traffic example.

As a next step, some other data clustering approaches, such as genetic-based membership function parameter-estimation (GMFPE), and probabilistic-based data clustering approach could be applied to more driving data generated from different human drivers. Meanwhile, the control mode levels can be compared with SRK framework to define the value of new levels for better estimation of critical situations.

Acknowledgement

The research in this paper is partly supported by China Scholarship Council through scholarship received by the first author for his Ph.D. study at the Chair of Dynamics and Control, UDE, Germany.

References

Baysari, M. T., C. Caponecchia, A. S. McIntosh, and J. R. Wilson (2009). Classification of errors contributing to rail incidents and accidents: A comparison of two human error identification techniques. *Safety Science 47*(7), 948–957.

Boring, R. and M. Rasmussen (2016). GOMS-HRA: A method for treating subtasks in dynamic human reliability analysis. In *Proceedings of the 2016 European Safety and Reliability Conference*, pp. 956–963.

Cárdenas, J. F. S., J. G. Shin, and S. H. Kim (2020). A few critical human factors for developing sustainable autonomous driving technology. *Sustainability 12*(7), 3030.

Dekker, S. (2017). *The field guide to understanding 'human error'*. CRC press.

Dhillon, B. S. (2017). *Safety, reliability, human factors, and human error in nuclear power plants*. CRC Press Florida.

Di Pasquale, V., R. Iannone, S. Miranda, and S. Riemma (2013). An overview of human reliability analysis techniques in manufacturing operations. *Operations management*, 221–240.

He, C., Y. Y. Lum, K. Y. Lee, and D. Söffker (2021). Human reliability estimation based on fuzzy logic-modified CREAM approach. In *2021 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, pp. 45–50. IEEE.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method (CREAM)*. Elsevier.

Khattak, A. J., N. Ahmad, B. Wali, and E. Dumbaugh (2021). A taxonomy of driving errors and violations: Evidence from the naturalistic driving study. *Accident Analysis & Prevention 151*, 105873.

McDonnell, A. S., T. G. Simmons, G. G. Erickson, M. Lohani, J. M. Cooper, and D. L. Strayer (2021). This is your brain on autopilot: Neural indices of driver workload and engagement during partial vehicle automation. *Human factors*, 00187208211039091.

Moujahid, A., M. E. Tantaoui, M. D. Hina, A. Soukane, A. Ortalda, A. ElKhadimi, and A. Ramdane-Cherif (2018). Machine learning techniques in ADAS: A review. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pp. 235–242. IEEE.

Noy, I. Y., D. Shinar, and W. J. Horrey (2018). Automated driving: Safety blind spots. *Safety science 102*, 68–78.

Philippart, M. (2018). Human reliability analysis methods and tools. In *Space Safety and Human Performance*, pp. 501–568. Elsevier.

Rasmussen, J. (1983). Human errors process control. In *Contribution to NATO Workshop on Origin of Human Error, September, Bellagio, Italy*.

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

Reason, J. (1990). *Human error*. Cambridge university press.

Singh, S. (2015). Critical reasons for crashes investigated in the national motor vehicle crash causation survey. Technical report, National Highway Traffic Safety Administration, Washington, D.C., USA, DOT HS 812 115.

Stanton, N. A. and P. M. Salmon (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science 47*(2), 227–237.

Swain, A. D. and H. E. Guttmann (1983). Handbook of human-reliability analysis with emphasis on nuclear power plant applications. final report. Technical report, Sandia National Labs., Albuquerque, NM (USA).

Ulutagay, G. and E. Nasibov (2008). FN-DBSCAN: A novel density-based clustering method with fuzzy neighborhood relations. In *8th International Conference on Application of Fuzzy Systems and Soft Computing (ICAFS-2008)*, pp. 101–110.

Zadeh, L. A. (1988). Fuzzy logic. *Computer 21*(4), 83–93.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability engineering & system safety 94*(2), 125–141.