Advances in Human Factors of Transportation, Vol. 148, 2024, 364–371
https://doi.org/10.54941/ahfe1005227

Framework for the Evaluation of a Holistic Fitness-to-Drive System for Commercial Drivers in the PANACEA Project

Anna Sjörs Dahlman[1,2], Anna Anund[1,3], Katerina Touliou[4], and Iosu Erauskin[5]

[1]Swedish National Road and Transport Research Institute (VTI), 581 95, Linköping, Sweden
[2]Chalmers University of Technology, Department of Electrical Engineering and SAFER Vehicle and Traffic Safety Centre, 412 96, Gothenburg, Sweden
[3]Rehabilitation Medicine, Linköping University, 581 83, Linköping, Sweden
[4]Hellenic Institute of Transport (HIT), Centre for Research and Technology Hellas (CERTH), 6th Km Charilaou-Thermis, 57001, Thessaloniki, Greece
[5]Datik Información Inteligente, 20009 Donostia, San Sebastian, Spain

ABSTRACT

To support drivers in the future to be fit to drive a system that assess the physical, cognitive, and psychological Fitness-to-Drive of commercial drivers have been developed. In cases of impairment, a complementary cloud-based countermeasures and coaching tool deploy appropriate solutions targeting drivers, operators, and enforcement. The work to identify the use cases and relevant functions were done in the EU funded project PANACEA (Grant Agreement: 953426). To understand what such a "system of system" brings, an evaluation framework incorporating components from several previous frameworks was developed. The framework sets the plan for data collection, analysis, and reporting and is expected to be used in all studies included in the PANACEA project to achieve effective and systematic data collection to provide a solid evidence base for assessment of progress and impact over time. The new evaluation framework was developed by reviewing and combining components from commonly used frameworks in both the automotive and clinical research field, using the FESTA framework as the foundation.

Keywords: Fitness-to-drive, Driver monitoring, Traffic safety, Professional driver, Evaluation

INTRODUCTION

Commercial drivers have a challenging work environment and often report being fatigued or stressed while driving (Filtness, Anund et al., 2019). Key considerations are the impact of shift work, task related fatigue, and impairment due to alcohol or drug use. The PANACEA project aims to create a holistic pre-, during and roadside driving ability monitoring and assessment system i.e. 'Commercial Health Toolkits' (CHT). The system will assess the physical, cognitive, and psychological Fitness-to-Drive of commercial drivers.

In cases of impairment, a complementary cloud-based countermeasures and coaching tool will deploy appropriate solutions targeting drivers, operators, and enforcement. It is intended that the PANACEA system will detect fitness-to-drive prior to starting work and during the work shift. In addition, it will prepare drivers ahead of their future shifts. To enable a systematic evaluation of the PANACEA system in various commercial driver groups, an evaluation framework was developed.

Commonly used frameworks from different fields were reviewed. Previous reviews of evaluation frameworks have concluded that there is an abundance of frameworks available but no single framework that covers all aspects of evaluation (Fynn, Hardeman et al., 2020, Newman-Askins, Ferreira et al., 2003, Yusof, Kuljis et al., 2008). Several transportation system evaluation frameworks exist, but these mostly focus on evaluation of the societal impact and economic benefits (He, Zeng et al., 2010, Newman-Askins, Ferreira et al., 2003) or environmental impact (Jansuwan, Liu et al., 2021).

The PANACEA project has its starting point in the project objectives, which are a combination of technology development, technology evaluation, knowledge creation, and impact assessment objectives. The development itself has it starting point in identified use cases. The framework needs to be flexible enough to allow for a range of different study designs. Moreover, the development of the PANACEA system follows an iterative development process, where the results of initial data collections are to be fed back to refine the PANACEA solutions. Lastly, the framework needs to cover the evaluation of the final PANACEA system, including the technical performance, usefulness and operability, user experiences, safety, socioeconomic impact etc. The purpose of the PANACEA evaluation framework is to create a common framework to be used in all studies in the project to make sure the data are collected in a way that makes it possible to consolidate the results at the end and to provide what is needed for impact analysis.

METHODS

Six frameworks were selected for a review. This was not intended as a systematic review of frameworks, rather a comparison of a few commonly used frameworks from different fields. The objective was not to identify and adopt the optimal model for integration into PANACEA; rather, it was to derive inspiration from preceding research in order to establish a customized framework tailored specifically for PANACEA.

The frameworks chosen for review were the transportation related frameworks FESTA, Trilateral Impact Assessment Framework and System Dynamic modelling, the more general Rainbow framework, and the healthcare related initiatives STROBE and CONSORT and the Framework for Program Evaluation in public health.

The selection of the six frameworks for review was a decision aimed at incorporating expertise and insights from multiple sectors. This approach guarantees that the PANACEA framework is both thorough and flexible, designed to accommodate the specific characteristics and requirements of

various fields. The framework was then used for the planning of data collections within the PANACEA project as described in project deliverable D6.1: Evaluation framework, plans and material (Sjors Dahlman, Anund et al., 2022).

FESTA

The methodology was developed for Field Operational Tests (FOTs) by the European union funded project FESTA (Field opErational teSt supporT Action). The FESTA project developed a handbook on FOT methodology to improve comparability and significance of results at national and European levels (Barnard, Innamaa et al., 2016). A FOT is here defined as a study undertaken to evaluate a function, or functions, under normal operating conditions in road traffic environments typically encountered by the participants to identify real world effects and benefits. FOTs were introduced as an evaluation method for driver support systems and functions with the aim of proving that such systems can deliver real-world benefits. Although the FESTA methodology and handbook was originally developed for FOTs, its basic steps are applicable for a wide range of field and user tests.

Figure 1 shows the FESTA V-diagram, i.e. the steps that are followed during the evaluation. The blue boxes represent the sequential steps to follow and the grey arrows show how to work through these steps from preparing the study, to using the prepared material during data acquisition and finally analysing the collected data.

The starting point in the FESTA is the function or system that will be evaluated. This can be an Advanced Driver Assistance System (ADAS) that has already passed the basic verification but now will be evaluated in an operational environment. In FESTA it is recommended that the system is compared with a baseline condition (i.e., driving without the system).

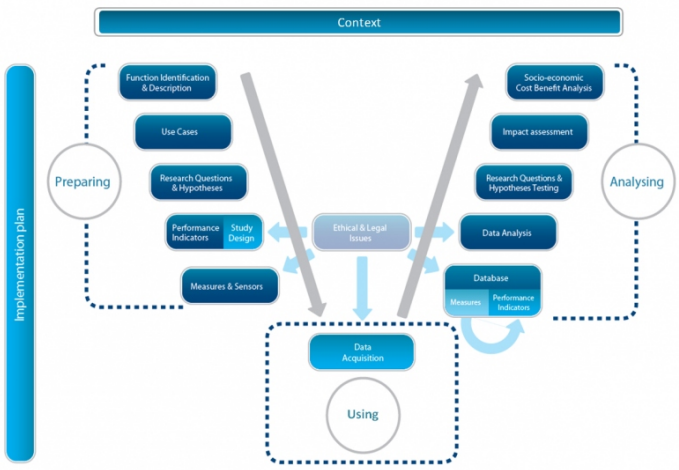

Figure 1: FESTA V-diagram (Barnard, Innamaa et al., 2016).

Trilateral Impact Assessment Framework for Automation in Road Transportation

The trilateral Impact Assessment Framework for Automation in Road Transportation was developed in cooperation between EU, US and Japan (Innamaa, Smith et al., 2018). The purpose was to harmonize the impact assessments performed in the field of automated driving, across the three regions (EC, US and Japan). The framework does not give detailed methodological recommendations, but it aims to facilitate meta-analysis across different studies. Therefore, the focus is on providing recommendations on how to describe the impact assessment study in a way that the user of the results understands what was evaluated and under which conditions. The framework is partly based on the FESTA framework. It is a high-level framework and includes recommendations and advice on; classification of evaluated system/service, common vocabulary, direct and indirect impacts in 12 impact areas, impact mechanisms and paths, recommendations for experimental procedures, recommendations for data sharing, and Key Performance Indicator (KPI) repository. Although, this framework was based on FESTA, it really focuses on impact assessment and automation, hence, only partially fitting the evaluation objectives of the PANACEA project.

System Dynamic Modelling

As a continuation and refinement of the trilateral framework, EU-US-Japan Trilateral Sub-Working Group for Impact Assessment, under the Trilateral Working Group for Automation in Road Transportation, has begun to use system dynamics to gain further insights into potential impacts (Rakoff, Smith et al., 2020). They developed a general framework from which detailed system dynamics models can be created for specific research questions. The work is ongoing, and the goal is to develop a quantitative tool that can help planners and policy-makers understand how highly automated vehicles may fit within the transport system, and to begin to explore consequences of potential actions under various scenarios. The attention is on AVs, but to understand the wider context the framework identifies the major generic roles within the transportation system and considers how they interact within the context of both traditional and new modes. This framework works well in identifying latent variables and complex interrelations, but it requires to be fed with considerable amounts of data, the types usually collected in large scale naturalistic and/ or field tests.

Rainbow Framework

The Rainbow framework developed by BetterEvaluation (betterevaluation .org) describes the evaluation process in 34 different evaluation tasks, grouped by 7 colour-coded clusters. The purpose is to make it easy to choose and use appropriate methods, strategies or processes. It is a general framework that can be used for various types of studies, including Randomized Controlled Trials (RCTs) and Outcome Mapping (OM). The planning

tool can be used to: commission and manage an evaluation; plan an evaluation; check the quality of an ongoing evaluation; embed participation thoughtfully in evaluation; develop evaluation capacity.

The clusters are named manage, define, frame, describe, understand causes, synthesize, and report and support use. Within each cluster, several tasks are listed and for each task a set of options are given. The framework provides many details around the planning, conduction, and reporting of data collections. It also covers general project management aspects that are out of the scope for the PANACEA evaluation framework. Moreover, the tasks in the define cluster and parts of the frame cluster were performed already in the application process and described in the grant agreement.

Framework for Program Evaluation in Public Health

The Framework for Program Evaluation in public health developed by the Centers for Disease Control and Prevention (CDC) is a practical, nonprescriptive tool, designed to summarize and organize essential elements of program evaluation (Milstein and Wetterhall, 1999). The general aim is to improve how program evaluations are conceived and conducted. The framework emphasizes six connected steps that together can be a starting point to tailor an evaluation for a particular effort, at a particular point in time. Because the steps are all interdependent, they might be encountered in a nonlinear sequence; however, an order exists for fulfilling each step and earlier steps provide the foundation for subsequent progress. The framework is purposefully general and thus provides a guide for designing and conducting evaluation projects across many different program areas.

CONSORT and STROBE Statements

In the field of clinical and epidemiological research, there have been several initiatives to standardize the conduction and reporting of studies. Two of them are the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) and the Consolidated Standards of Reporting Trials (CONSORT) initiatives (Altman, Schulz et al., 2001, Von Elm, Altman et al., 2007). Both provide a set of recommendations for the reporting of studies. They offer a standard way for authors to prepare reports of study findings, facilitating their complete and transparent reporting, and aiding their critical appraisal and interpretation. Checklists that focus on reporting how the trial was designed, analysed, and interpreted are available for several types of study designs, e.g., randomized controlled trials, case-control studies, cohort studies, and cross-sectional studies. They emphasize the importance of transparency in the reporting to enable critical judgement of the generalizability and possible bias. These initiatives focus mainly on the reporting of research and do not provide guidelines for the planning and implementation of data collection. However, the checklists can also serve as guidelines of what to consider in the planning of a study.

None of the reviewed frameworks provide a perfect fit for the type of evaluation planned within the PANACEA project. Some of the reviewed

frameworks are more suitable for research projects, driven by research questions, and other frameworks are suitable for innovation projects focusing on evaluating a technical solution. As the PANACEA project is a research and innovation action, it has a combined need and a starting point in specific use cases.

A new evaluation framework was therefore developed by reviewing and combining components from commonly used frameworks in both the automotive and clinical research field. The FESTA methodology (Barnard, Innamaa et al., 2016) was used as the foundation and the various steps in the evaluation process were adapted to suit the purpose of the PANACEA project.

RESULTS

The framework developed within the PANACEA project incorporates components from several of the frameworks reviewed above. The development of the PANACEA system is an iterative process where results from data collections are fed back to refine the system before the final evaluation (Figure 2). Technical validation of the systems used in the data collections are performed before the start of each data collection. The results of the validation are fed back to the relevant activity responsible for the development or integration of the technology. Any issues discovered are resolved before proceeding with the evaluation process. Results from the initial data collections in simulator and roadside pilots are utilized to refine the algorithms for the driver monitoring and countermeasures tools.

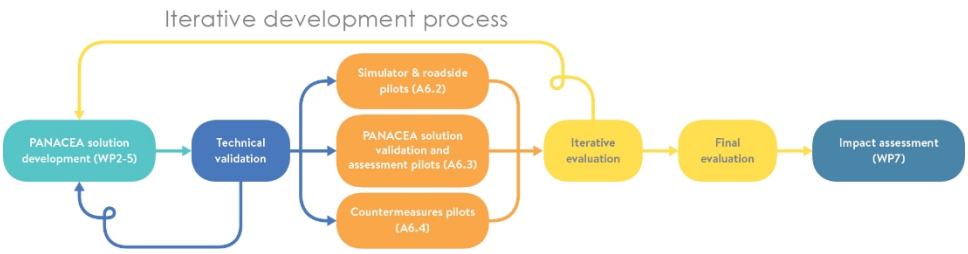

Figure 2: PANACEA iterative development process.

The PANACEA system validation and assessment pilots conduct validation tests to assess the readiness of the CHTs prior to the final evaluation of the PANACEA system at three different pilot sites. Within each pilot site, the aggregation of various technologies and devices connected to the PANACEA system is designated as the Commercial Health Toolkit (CHT). This terminology is chosen to highlight the toolkit's applicability to commercial personnel, such as coach drivers and delivery service riders, within the transportation sector, underscoring its development to improve occupational health and safety for these distinct professional categories. In contrast to the technical validation, this validation focuses on the performance of the full PANACEA system in operation, not the performance of individual sensors or parts. The collected data was used to improve the technologies and their integration to

CHTs and resolve any technology issues. Furthermore, the CHTs' assessment pilots were organised, monitored and executed, to provide data for the final evaluation and impact assessment of the PANACEA system. Final evaluation is about the realisation of the countermeasures' pilots. The evaluation of both the content and the actual online coaching system is performed at the three pilot sites. The data collected was fed back, to further improve the system, see Figure 2.

The various data collections in PANACEA, used for the iterative development and for the final evaluation and impact assessment, will follow the methodology of the framework. The PANACEA final evaluation process is divided into three phases: planning, implementation, and analysis and reporting (Figure 3). Within each phase, there are several steps to follow in the evaluation process. Each box represents a step to follow in the evaluation process. The steps are described as sequential steps in a linear way, where each step provides the necessary input for completion of the next step. The evaluation framework has a mixed methods approach combining qualitative and quantitative methods.

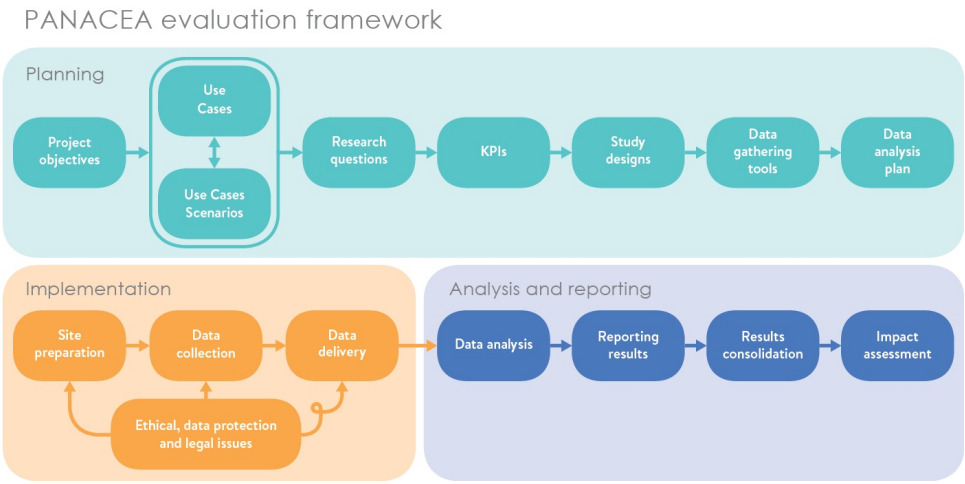

Figure 3: PANACEA evaluation framework.

The steps of the PANACEA framework are used in the experimental plans for the data collections in the project (Sjors Dahlman, Anund et al., 2022). Each step becomes one chapter in the experimental plan, with a general description and an overview of how this will be implemented in the PANACEA project. In this way, various data collections can be harmonized across project activities.

CONCLUSION

An evaluation framework incorporating components from several previous frameworks was developed within the PANACEA project, setting the plan for data collection, analysis, and reporting. The PANACEA framework is used in all studies included in the project to achieve effective and systematic

data collection to provide a solid evidence base for assessment of progress and impact over time. The framework can be useful for similar research and innovation projects.

ACKNOWLEDGMENT

The PANACEA project has received funding by the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 953426.

REFERENCES

Altman, D. G. et al. 2001. The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration. Annals of Internal Medicine 134(8) 663–694.

Barnard, Y. et al. 2016. Methodology for Field Operational Tests of Automated Vehicles. Transportation Research Procedia 14 2188–2196.

Filtness, A. et al. 2019. Bus driver fatigue. London, UK, Transport for London.

Fynn, J. F., Hardeman, W., Milton, K. and Jones, A. P. 2020. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. BMC Public Health 20(1) 1000.

He, J., Zeng, Z. and Li, Z. 2010. Benefit Evaluation Framework of Intelligent Transportation Systems. Journal of Transportation Systems Engineering and Information Technology 10(1) 81–87.

Innamaa, S. et al. 2018. Trilateral impact assessment framework for automation in road transportation. Tech. Rep.

Jansuwan, S., Liu, Z., Song, Z. and Chen, A. 2021. An evaluation framework of automated electric transportation system. Transportation Research Part E: Logistics and Transportation Review 148 102265.

Milstein, B. and Wetterhall, S. F. 1999. Framework for program evaluation in public health.

Newman-Askins, R., Ferreira, L. and Bunker, J. M. 2003. Intelligent transport systems evaluation: From theory to practice. 21st ARRB and 11th REAAA Conference.

Rakoff, H. et al. 2020. Building Feedback into Modelling Impacts of Automated Vehicles: Developing a Consensus Model and Quantitative Tool.

Sjors Dahlman, A. et al. 2022. PANACEA D6.1: Evaluation framework, plans and material. PANACEA project deliverables.

Von Elm, E. et al. 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bulletin of the World Health Organization 85 867–872.

Yusof, M. M., Kuljis, J., Papazafeiropoulou, A. and Stergioulas, L. K. 2008. An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit). International Journal of Medical Informatics 77(6) 386–398.