Validation of Driver Monitoring Systems

The future relies on a multi-level and multidisciplinary approach

Authors: Clémentine François and Jérôme Wertz

Publication date: 24 February 2021

Phasya s.a. | Rue du Bois Saint-Jean, 29, 4102 Seraing, Belgium | **www.phasya.com**

Abstract

There is a growing demand for driver monitoring systems (DMS) for various use cases related to safety, comfort, and user experience. And regulators are pushing for DMS and their validation protocols with a view to improving road safety. DMS is becoming more complex, and there are more and more players getting involved in the field. The automotive industry and those with a stake in it are thus facing several challenges in the development and validation of DMS, including as regards efforts to keep them affordable while improving road safety and encouraging user adoption.

This document presents the automotive industry, and those with a stake in it, with:

- an advanced validation methodology based on a multi-level and multidisciplinary approach
- recommendations for enabling a more agile and affordable development of future DMS

It also aims to give them a better understanding of the challenges related to the validation and development of DMS.

<u>Disclaimer</u>

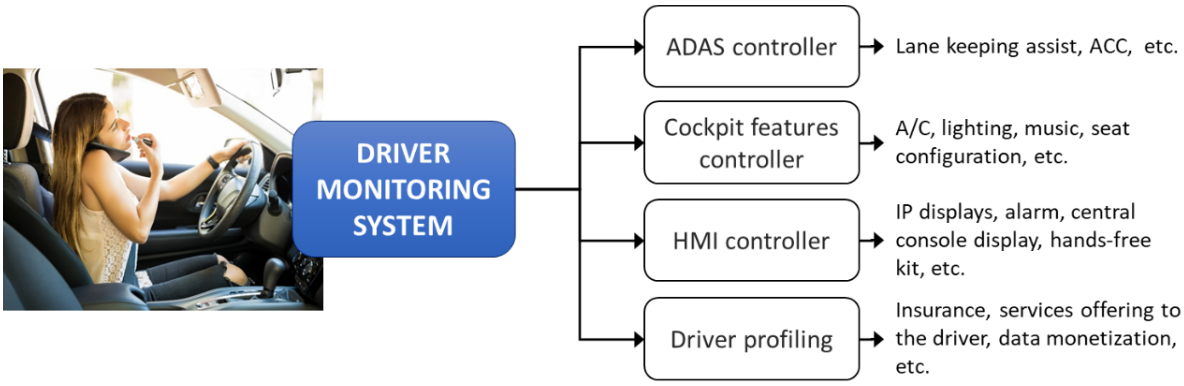
In what follows, Phasya shares its vision about the future of DMS development and validation methodologies. The information in this document is not exhaustive, and does not cover all DMS issues or validation techniques. Its aim is to inform stakeholders in the automotive industry and to spark debate with a view to helping improve DMS, in particular as regards safety.

Table of contents

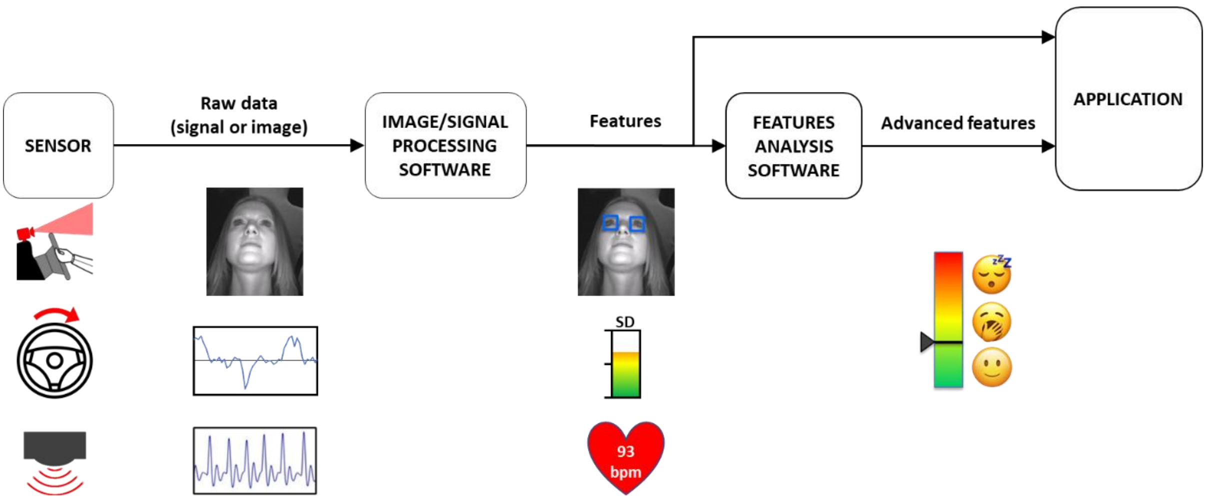
1. Introduction 3
2. Challenges that stakeholders in the automotive industry face 4
3. DMS validation: challenges and requirements 5
4. Multi-level and multidisciplinary validation approach 9
5. Conclusion 11
Abbreviations 12
About Phasya and the authors 13

1. Introduction

A driver monitoring system (DMS) aims to provide information about the driver and their behavior. DMS is thus a key enabler of a range of applications related to safety, comfort, and user experience.



While there are several approaches to driver monitoring, the high-level architecture is similar in each case. One or more sensors collect raw data – a signal or an image – which is processed by a first software layer to extract "features", such as the detection of objects, the opening of the eyes, and the variability of steering angle. In some cases, the features are further analyzed by a second software layer, which provides additional features such as drowsiness detection, eye-gaze heatmap. The features provided by the DMS are then key enablers of several applications and use cases.



The performance of DMS is evaluated through test protocols and metrics that differ according to the sensing modality, the scope of the application, and the supplier. And some stakeholders in the automotive industry want to establish requirements and evaluation protocols for certain safety applications.

In what follows, Phasya shares its vision about the future of DMS development and validation methodologies. It highlights some key factors and challenges for automotive stakeholders as well as paths towards one and another solution.

2. Challenges that stakeholders in the automotive industry face

The development and validation of DMS lead to several challenges for the automotive industry. These can be captured through the following questions.

How can we encourage user adoption?

User adoption relies on the relevance of the information that a DMS provides to the driver. False positives, false negatives, and poorly suited interfaces will discourage user adoption. And the driver needs to feel comfortable about being monitored: there's a risk that a DMS could be seen as Big Brother.

How can we evaluate the performance of a DMS and communicate it efficiently?

While there are various methods of evaluation in different use cases, most still depend on who the individual stakeholder is. Benchmarking some DMS features can thus be a real headache. As DMS safety features become mandatory, a start is made on standardizing evaluation guidelines and protocols. This standardization process, which will take place over a number of years, will be iterative and be subject to limitations because of the lack of objective ground truths for certain features.

How can we improve our development and validation process in order to reduce costs and cut lead time?

The complexity of DMS is increasing as the number of features and applications grows. And the DMS ecosystem is becoming increasingly heterogeneous as more and more specialized players join it. The automotive industry is thus facing a number of challenges in developing future DMS in agile and efficient ways while keeping development and validation processes affordable. Interoperability is a key factor for the next generations of DMS and the automotive industry still has to make that a reality.

How can we make automated and collaborative driving happen?

DMS is a key enabler of L3-L4 of driving automation and the development of advanced driver assistance systems (ADAS). This evolution of the role of DMS creates new challenges. In fact, as the driver becomes a passenger in autonomous driving mode, with their eyes off the road and their hands off the wheel, the vehicle has to check their ability to take back control—and it cannot fully rely on current DMS approaches to do this. In fact, in autonomous mode a DMS based on the analysis of driving behavior such as lane departure and steering movements is useless, while a DMS based on a single camera could suffer from significant data losses caused by activities the driver engages in such as reading a book that blocks the camera's field of view. Additionally, it might be possible to facilitate driver adoption of ADASs through automated ADAS settings that depend on the driver's states.

3. DMS validation: challenges and requirements

3.1. Introduction

The validation of a DMS must take into account quite a large set of technical and human factors. While technical issues can be evaluated rather objectively and straightforwardly, evaluation of human factors is more complex. Moreover, the diversity of factors that impact DMS performance and validation could lead to an extensive landscape of test scenarios.

Taken together, the items listed below give an overview of factors to be taken into account. Though the list is not exhaustive, it aims to provide a better understanding of the challenges and requirements related to the validation of DMS.

3.2. Challenges and requirements

The driver-capabilities pyramid

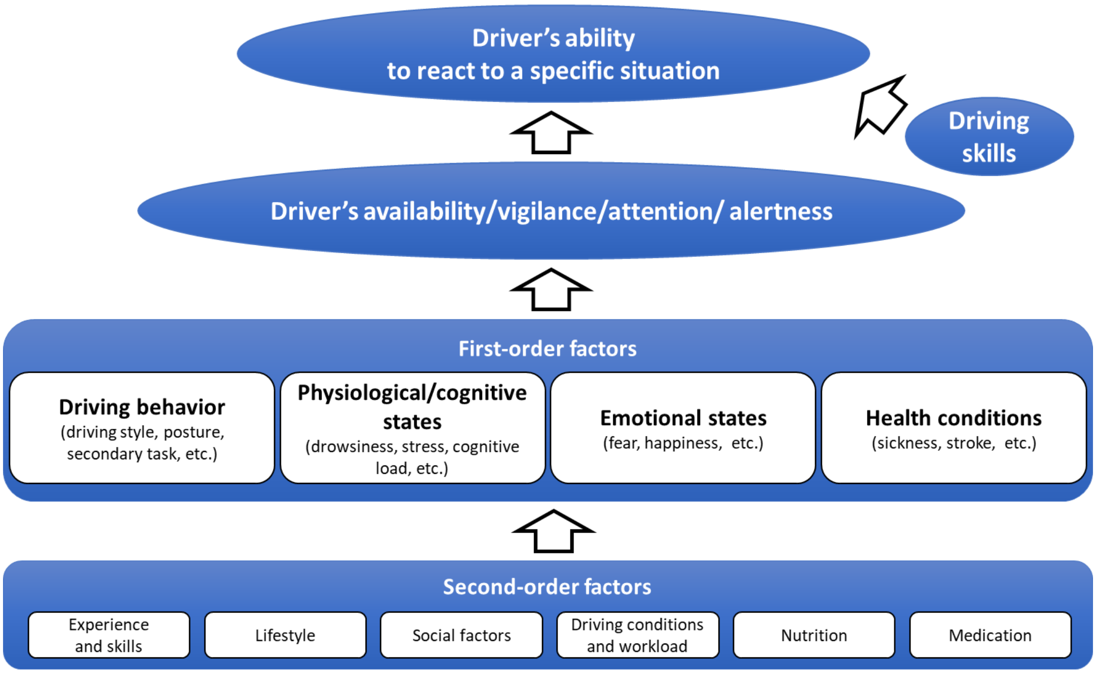
A driver's ability to handle a specific driving situation depends both on their driving skills and their availability/alertness. The monitoring of the driver's availability/alertness is quite complex for two reasons:

- the driver's availability/alertness depends on several first-order factors – see the graph hereunder – such as their behavior and cognitive states
- these factors depend in turn on other, second-order factors (see below) such as lifestyle, road conditions, and experience.

In order to support applications that enhance safety, comfort, and user experience, the DMS has to provide information about the first-order factors. Indeed, by providing these applications with information on the underlying states that impact the driver's availability and their alertness, they are able to work in a relevant way.

From a validation standpoint, the evaluation protocols and the related ground truths must make it possible to distinguish among the first-order factors. Evaluation protocols also require appropriate control of some of the second-order factors.

The following graph gives an overview of the factors that can impact driver's capabilities.



Physiological and cognitive states

Physiological and cognitive states cannot be simulated in a relevant way by a human being. The evaluation protocol thus has to induce the state in question in a natural way.

Physiological and cognitive states are not binary, but progressive. For instance, before a driver falls asleep, they will be in a state somewhere between fully awake and fully drowsy/falling asleep. It is thus quite challenging to define key performance indicators (e.g., sensitivity and specificity) for the ability to specifically detect such states.

There is no fully objective single ground truth for each state, and there probably never will be. Thus, no single ground truth will ever meet all the requirements for validating a particular state. For instance, the table below compares four ground truths for wakefulness and drowsiness, based on the following requirements:

- Objective data: the ground truth is based on fully objective data. (In some cases, the ground truth data has to be scored in a subjective way, so the data is almost but not fully objective).
- Direct: the ground truth measures the state directly. Behavioral data serve as indirect measurements of the state.
- Continuous: the ground truth is collected continuously.
- Non-disturbing: the collection of the ground truth does not disturb the driver's tasks or behavior.

Ground-truth category	Example	Ground-truth characteristics			
		Objective data	Direct	Continuous	Non-disturbing
Physiological	EEG	Almost	V	V	V
Behavior	SDLP	V	X	Almost	V
Self-evaluation	KSS	X	V	X	X
Video rating	ORD	Almost	X	V	V

Emotional states

The challenges created by emotional states are similar to those created by physiological and cognitive states (on which see the previous section).

Driver diversity

These factors include size, ethnicity, gender, and age; whether the driver is using eyeglasses, has piercings, is wearing make-up, or has a face mask on; and what clothes they are wearing and how their hair is styled. The robustness of a DMS when it comes to driver diversity is a key evaluation criterion for the validation of some DMS features. In these cases, the validation protocol has to ensure a sufficient level of driver diversity in the test scenarios.

Driving conditions

Ideally, a DMS should work in all driving conditions. Driving conditions include factors such as luminosity, the quality of any road markings present, weather conditions, traffic conditions, and how much dust there is in the immediate environment or on the sensor. Because driving conditions may affect the sensing modality – the quality and availability of raw data and features – as well as driver behavior, the evaluation protocol has to ensure the validation of the DMS under all relevant conditions.

Sensors and data

Depending on the application, one or more sensors and data items are used, such as the driver camera, the angle of the steering wheel, and a seat-pressure sensor. In some cases, one sensor can be used for several applications, while a particular application can be supported by various sensing techniques that are used alone or in combination. Moreover, sensors and data specifications may vary from one DMS supplier to another.

Together, these factors lead to the creation of several sensor architectures and data specifications. The evaluation protocol must thus be designed for each sensor architecture and each set of data specifications.

Features and applications

A DMS may support several use cases related to safety, comfort, and user experience. Each application has its own requirements in terms of DMS features such as the direction of the driver's gaze, the detection of objects, the detection of drowsiness, and identification of the driver, as well as the specifications of each feature. The evaluation protocol must thus evaluate each feature required by the applications.

3.3. Conclusion

The combination of all the above challenges and requirements could theoretically lead to extensive data-collection scenarios that would be quite demanding in terms of cost and lead time, in particular if the DMS is to be validated as a whole.

The following example[1] gives insights – simplified for illustrative purposes – into the size of the dataset that could potentially be required in order to validate a camera-based DMS for detecting drowsiness. The dataset combines all the evaluation factors, and includes at least 90 drivers (=3*2*3*5) performing 24 driving scenarios (=3*2*2*2), for a total of 2,160 evaluation scenarios.

Evaluation Factor	Description of Scenario	No. of Evaluation Scenarios
Driver ethnicity	Asian/European/African	3
Driver gender	Male/female	2
Driver size	< 160 cm/160-190 cm/> 190 cm	3
Driver age	16-30/30-45/45-60/60-75/> 75	5
Eyeglasses	Without/prescription glasses/sunglasses (1 driver can perform all scenarios)	3
Make-up (only female) OR Beard (only male)	Without/with (1 driver can perform both scenarios)	2
Luminosity	Night/day	2
Driver sleep deprivation	Without/with	2

In addition, some evaluation factors required for the validation of a DMS depend on the sensor and application. Thus, for the same application or the same sensor system, the validation protocol may vary from one DMS to another.

Where each DMS is validated as a whole for one or more applications, the entire validation process may have to be repeated should there be any changes, even if these are related to just one DMS layer. Such an approach would hamper the agile and affordable development of a DMS.

From the standpoint of human factors, the validation of certain advanced features of a DMS requires to combine several ground truths, because these features cannot be evaluated against one single, fully objective ground truth. Moreover, skills in the study of human factors are required in order to design a suitable validation protocol and analyze the results of the evaluation.

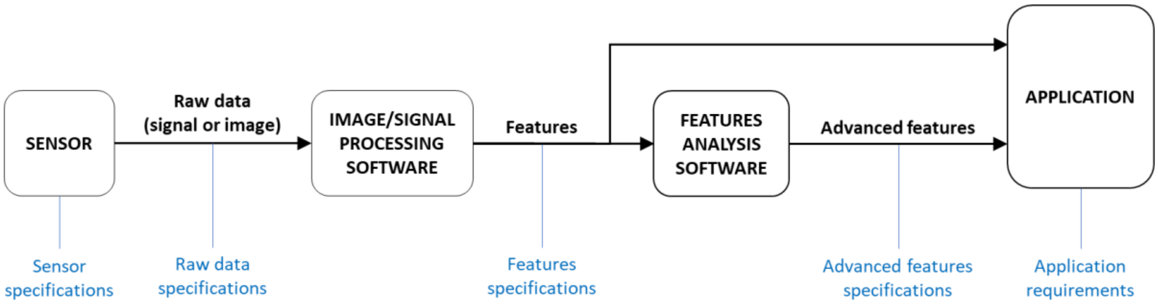
The methodology whereby an entire DMS is validated as a single black box is a thing of the past. New validation techniques that are available for each DMS layer enable the development of more agile approaches to DMS validation and development. The future relies on a multi-level and multidisciplinary approach as described below.

[1] This example does not take into account the potential of the synthetic-image dataset.

4. Multi-level and multidisciplinary validation approach

4.1. Introduction

The next generations of DMS will become more and more "multi-", with multiple sensors, multiple features, and multiple applications. The development of DMS should thus become more and more agile by relying on a modular design approach. The DMS thus requires clear specifications and validation on every layer. Indeed defining, for each layer, specifications and requirements for input data will allow the automotive industry to make changes to a particular layer without having to perform the entire validation process for the whole DMS again.



Certain advanced features of the DMS are dedicated to understanding driver behaviors and states. And they have to allow applications to work in the best ways, including as regards interfaces with the driver. For these reasons, the validation of a DMS requires multidisciplinary skills beyond engineering that include expertise in human factors and medicine.

4.2. Approach to validation

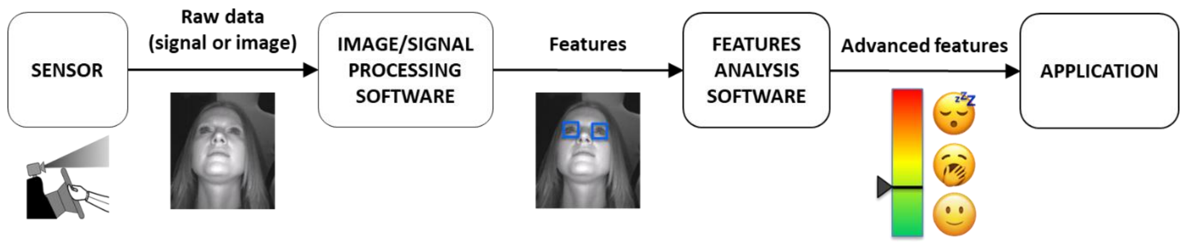
In the approach to validation that is proposed, each DMS layer is validated independently. For each such layer, validation is based on the following questions, which also take into account the specifications for input data of the next layer:

- Which specifications of the layer are to be validated? Which ones are required for the next layer?
- What ground truths are needed?
- What are the validation results?
- What factors impact the output data?
- What skills are required?

Based on these questions, the protocol for evaluating a given layer will be suitable for that layer while ensuring its interoperability/compatibility with the next one. This approach makes it possible to isolate the challenges and requirements related to each layer while complying with the input data requirements of the next layer. And it makes it possible to simplify the validation protocol in terms of required test scenarios – fewer scenarios for one layer than for the whole DMS – while keeping the DMS modular.

4.3. Validation approach – Uses cases

The illustration below shows the validation approach to two DMS layers in the case of drowsiness detection with a camera-based DMS. This example is kept simple for a sake of clarity.



DMS Layer	Image Processing Software	Feature Analysis Software
Input data requirements	Infrared images of the driver's face, including specifications for resolution, frame rate, and field of view.	- Right/left eye detection, including the minimum availability of this data - Right/left eye opening, including the minimum availability of this data and its minimum accuracy
Output data (= data to be validated)	- Right/left eye detection - Right/left eye opening	Level of drowsiness
Ground truth	Manual labelling of images	- KSS - EEG
Validation results	- Specificity and sensitivity of eye detection - Accuracy with which the opening of the eyes is measured.	- Specificity and sensitivity of drowsiness detection based on KSS - Correlation with KSS and EEG
Evaluation data set	Data set collected with a large sample of drivers in several different driving conditions in order to cover most situations encountered in real life, including as regards driver diversity.	Data collected with several drivers in several sleep conditions in order to cover a range of states of wakefulness/drowsiness.
Required skills	Engineering, statistics	Sleep medicine, human factors, engineering, statistics

5. Conclusion

While assistive driving technologies are becoming more and more fully integrated into vehicles, there is a growing demand for driver-centric solutions. The DMS is becoming a key enabler of safety, comfort, and user-experience applications as well as of L3-L4 of driving automation. In fact, the next generations of DMS will be more and more "multi-", with multiple sensors, multiple features, and multiple applications. This evolution is already driving the automotive industry to rethink the development and validation approach in order to keep it affordable and agile.

Moreover, because safety is the primary driver of DMS, and because regulators are pushing for validation guidelines and protocols, the automotive industry must work closely with regulators to develop these. This work should lead to validation guidelines and protocols that are affordable and as simple as possible while improving road safety and ensuring user adoption.

Through a multi-level and multidisciplinary approach to validation, the automotive industry will be able to meet the challenges that future DMS run into in terms of development and validation. Indeed, beyond the validation, this approach also tends to provide guidelines for developing a modular DMS architecture that will enable the affordable and agile development of DMS.

In our opinion, now is the time for the automotive industry to start setting and continually updating standards for the specifications of the interface for each DMS layer, in order to enable interoperability between DMS layers. This interoperability will facilitate the modular design of, and the validation process for, DMS as well as comparisons between similar solutions for the same layer.

We are now at the dawn of driver monitoring – and it's time to make it modular.

Abbreviations

ACC	Adaptive cruise control
ADAS	Advanced driver assistance system
DMS	Driver monitoring system
EEG	Electroencephalogram
HMI	Human machine interface
KSS	Karolinska Sleepiness Scale
ORD	Observer Rating of Drowsiness
SDLP	Standard deviation of lateral position

About Phasya and the authors

Phasya

Phasya s.a. specializes in software for monitoring physiological and cognitive states such as drowsiness, stress, and cognitive load, from the analysis of physiological data such as ocular data and heart rate.

Phasya was founded in 2015, following several years of research and development at the University of Liège. For more than 10 years, the Phasya team has been developing a unique software portfolio for monitoring physiological and cognitive states as well as unique multidisciplinary expertise that combines skills in engineering, human factors, and medicine.

Phasya offers advanced software solutions to the automotive industry for the development of the next generation of driver monitoring systems. Based on its more than 10 years of multidisciplinary expertise, the Phasya team also offers services related to the testing and validation of DMS.

In 2019, Phasya received a CLEPA[2] Innovation Award for its innovations in software for driver monitoring and its validation.

https://www.phasya.com/

Clémentine François

Clémentine Francois is Chief Scientific Officer, co-founder, and board member at Phasya. She is responsible for R&D, innovation, and scientific partnerships at Phasya.

In 2020, she won a Silver Award at AutoSens Young Engineer of the Year Awards.

Clémentine François holds an MSc in biomedical engineering and a PhD in engineering from the University of Liège in Belgium. In 2010, she started her career as a researcher at the University of Liège working on a technology for monitoring drowsiness. In 2014 and 2016, she co-founded and organized SomnoSafe, an international symposium on drowsiness, vigilance, and safety. In 2015, she co-founded Phasya. For more than 10 years, she has been fully committed to developing innovative solutions for the monitoring of human physiological and cognitive states in order to improve safety in both transport and industry.

Jérôme Wertz

Jérôme Wertz is Chief Executive Officer, co-founder, and board member at Phasya. He is also a delegate of the Belgian automotive industry on behalf of Agoria in the CLEPA Working Groups related to driver monitoring for EU regulations and Euro NCAP.

Jérôme Wertz holds an MSc in industrial engineering from the Gramme Institute in Liège, Belgium. He also holds an executive master's in Management. In 2009, he started his career as a researcher and project manager at the University of Liège, working on the development of a technology for monitoring drowsiness. In 2014, he co-founded and organized SomnoSafe, an international symposium on drowsiness, vigilance, and safety. In 2015, he co-founded Phasya. For more than 10 years, he has been fully committed to developing and bringing to market innovative technologies for improving safety in transport and industry.

In 2015, Jérôme Wertz was named an "Innovator under 35" by the MIT Technology Review for his innovations in the monitoring of drowsiness.

---

[2] CLEPA is the European association of automotive suppliers.