Transport Research Arena (TRA) Conference

The HADRIAN Novel Human-Machine Interface Prototype for Automated Driving: Safety and Impact Assessment

Marios Sekadakis[a]*, Marianthi Kallidoni[a], Christos Katrakazas[a], Sandra Trösterer[b], Cyril Marx[b], Peter Moertl[b], George Yannis[a]

[a]*National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou str., GR-15773, Athens, Greece*
[b]*Virtual Vehicle Research GmbH, Inffeldgasse 21/A, 8010 Graz, Austria*

Abstract

The current paper was inspired by the HADRIAN project and focuses on exploring the effects of innovative Human-Machine Interface (HMI) prototypes on safety, driving performance, and driver perceptions. Employing driving simulator experiments and questionnaires, this study investigates whether HADRIAN innovative HMI enhances safety and receives positive evaluations from drivers. Specifically, the investigation centers on a driving simulator experiment evaluating novel HMI prototypes designed to improve Automated Driving (AD) at SAE Levels 2 or 3. To facilitate HMI assessment, a tailored safety and impact assessment methodology was developed using unique Key Performance Indicators (KPIs). To benchmark and generate a total score on HADRIAN innovative HMI, a Data Envelopment Analysis (DEA) was deployed based on the aforementioned KPIs. The findings shed light on the influence of HADRIAN HMI innovations on safety and perceived impact when compared to a baseline "state-of-the-art" HMI. Subsequently, a comprehensive discussion is unfolded, highlighting the most significant KPIs that contributed to safety and perceived impact scores. This method and its outcomes can serve as a valuable resource for other HMI stakeholders, enabling them to employ similar human-centered assessment methodologies to assess the safety and perceived impact of potential HMI configurations.

Keywords: Automated Driving; Autonomous Vehicles; Safety & Impact Assessment; Human-Machine Interfaces (HMI); Key Performance Indicators (KPIs); Data Envelopment Analysis (DEA)

1. Introduction

Traffic crashes pose a severe threat to public health, as they account for the 8th most common cause of death worldwide with approximately 1.3 million fatalities annually along with a considerably greater count of non-fatal injuries (20-50 million) (WHO, 2021). The effect of the traffic crashes is also depicted in economic terms, as the economic cost of these accidents is significant in the European Union and worldwide, with estimations ranging from 0.7 to 3.0 million euros for preventing a fatality in 31 European countries in 2005 (Wijnen et al., 2019). Furthermore, European traffic crashes cost 0.4% up to 4.1% of GDP, whereas globally, the World Health Organization (WHO) calculates the total average cost at approximately 3% of GDP, highlighting the substantial impact of traffic crashes on public health, society, and the economy (WHO, 2018; Wijnen et al., 2019; Yannis, Papadimitriou, & Folla, 2014). Analyzing the underlying factors of traffic crashes is crucial for measures to mitigate this public health issue. Studies suggest that human factor is responsible for up to 94% of all traffic crashes (NHTSA, 2015). Nevertheless, the introduction of advanced technologies in the automotive sector such as automated driving (AD) functions promise enhanced safety for all road users. The future introduction of Autonomous Vehicles (AVs) is expected to bring about a significant revolution in transportation systems (Elvik, 2021; Fagnant & Kockelman, 2015). Specifically, high

* Corresponding author. Tel.: +30-210-772-1575;
E-mail address: msekadakis@mail.ntua.gr

market penetration rates of AVs are anticipated to improve traffic safety by eliminating human from driving and thus reducing human error which is the main cause of traffic collisions. However, prior to fully automated vehicles, intermediate SAE automation levels will still require human interactions and interventions with the vehicle. As transitioning to highly automated vehicles, there will still be a need for human interactions with the vehicle. In order to facilitate and realize this interaction between users and vehicles, Human-Machine Interfaces (HMIs) are anticipated to play a major role in this cooperation. This cooperation and interaction of the user and the AV was the main subject of the EU H2020 HADRIAN (**H**olistic **A**pproach for **D**river **R**ole **I**ntegration and **A**utomation Allocation for European Mobility **N**eeds) project (hadrianproject.eu). The HADRIAN project aimed at investigating and providing solutions for seamless and fluid interactions between the driver and the AV.

This study draws its inspiration from the EU H2020 HADRIAN project, with a specific goal of evaluating the impact of innovative Human-Machine Interface (HMI) prototypes on safety, driving performance, and driver perceptions. The current study seeks to address a knowledge gap by developing and providing a comprehensive safety assessment methodology with the total aim of enhancing AV traffic safety. The research was carried out as part of the EU H2020 HADRIAN project. The driver role for automated vehicles is also investigated and assessed using this holistic user-centered assessment approach that evaluates the safety and perceived impact effects with the total aim to achieve high impact and wide-reaching acceptance of automated vehicles.

2. Methodology

This study drew inspiration from the HADRIAN project and assess how innovative HMI prototypes affect safety, driving performance, and drivers' perceptions. This assessment will be based on data gathered from driving simulators and questionnaires. To achieve this, driving simulator experiments were conducted as part of the project in order to test if HADRIAN HMI innovations are safer and beneficial according to drivers e.g., more acceptable, useful, and comfortable. More specifically, This study specifically focuses on a driving simulator experiment that tested various innovative prototypes, including enhancements to Automated Driving (AD) predictability, availability, and continuity, as well as driver monitoring, HMI adaptiveness, and tutoring using technologies such as Heads Up Display (HUD), LED strips, haptic cues on the steering wheel, and an interactive tablet. In the driving simulator experiment, 20 participants tested on a state-of-the-art “baseline” HMI, while 19 participants used the HADRIAN fluid HMI. The simulated route included driving sections with Automated Driving Levels 2 or 3 or manual and in parallel, drivers were instructed to perform a secondary task whenever available. The driving simulator experiment for testing the innovative HMI titled “HADRIAN Integrated fluid HMI” is further unfolded in the following subsection, after the presentation of the KPIs used for the assessment as well as the scoring method follow.

2.1 Driving Simulator: Integrated fluid HMI

It is worth noting that several different HMI configurations were tested within the HADRIAN framework but only one is further exploited in this study since the examined HMI configuration is the most representative since it exploited and combined the most characteristical innovations across all other HMI configurations. For the specific driving simulator experiment, 20 participants drove with a baseline HMI and 19 with a HADRIAN HMI titled “Integrated fluid HMI”. To fully comprehend the assessment outcomes, it is necessary to provide an overview of the study description. The HMIs were tested in the driving simulator of “Virtual Vehicle”. The HMI with all the developed HADRIAN innovations aimed to provide better automated driving predictability, availability, and continuity. The key feature of this HMI is that predefined time was guaranteed for different take over requests i.e., 5 seconds were granted for take over in Automated Driving Level 2 (ADL2) and 15 seconds for take over in Automated Driving Level 3 (ADL3). With driver monitoring (i.e., by tracking eye movements to determine visual distraction, and tracking hands-off-wheel with a camera), it was ensured that unsafe driver states were detected, while tutoring before the drive (tutoring video) and during the drive (through verbal guidance and feedback and the possibility to repeat tutoring video chapters in ADL3) aimed at teaching the driver how to better use the automated driving functions and driver’s responsibilities in different driving modes. This all adds to an adaptive human-computer interface, which is adapted to the driver states and needs by providing tailored information, feedback, and alerts.

Hence, the integrated fluid HMI encompassed the following key functionalities:

- 5 seconds time for take-overs in ADL2, 15 seconds time for take-overs in ADL3: countdown time was displayed to the driver

- Ensured time interval in which ADL3 driving was possible: the duration was displayed to the driver
- Tutoring video before the drive, outlining the driving functions, correct system use, and driver responsibilities
- Verbal guidance and feedback for system use and adaptive tutoring during the drive based on gazing behavior and take-over time during a take-over
- Warnings in case of detected visual distraction or hands-off-wheel during manual or ADL2 driving with times adapted to the driving mode
- Take-over support by providing countdown information, current speed limit, and information about upcoming obstacles, as well as haptic cues at the steering wheel
- Information about vehicle behavior in ADL3 (current speed, detected speed limit, upcoming maneuvers)
- Support of mode awareness, mode changes, and warnings with ambient lighting

In the design of the integrated fluid HMI, these functionalities were addressed by implementing and integrating different interfaces, which serve a specific purpose within the HMI (Fig. 1 provides an overview of the components and the setup in the driving simulator used for the experimental study).

Fig. 1. Components of the integrated fluid HMI of the “Virtual Vehicle” holistic study.

Furthermore, apart from the HADRIAN HMI, the baseline HMI is a critical component of this assessment methodology for comparing the HADRIAN HMI with the baseline which acts as “state-of-the-art” HMI developments. The baseline HMI has been defined as the system against which all HADRIAN fluid-HMIs were evaluated. To establish the baseline HMI, the following setup was carefully considered and rigorously tested by the Virtual Vehicle driving simulator. For a fair comparison, in the baseline condition, it was assumed the same take-over times between the HADRIAN condition and the baseline condition for ADL2 and ADL3. However, in the baseline condition, these times were not indicated by the HMI. Also driving assistance is a common feature in nowadays cars, a hands-off-wheel warning was provided when driving in ADL2. Similar to the HADRIAN fluid HMI, a warning was provided if the driver had hands off the wheel for more than three seconds. Other than that, no adaption to the driver state was provided and no tutoring was available. In Fig. 2, the key features of the baseline HMI are indicated.

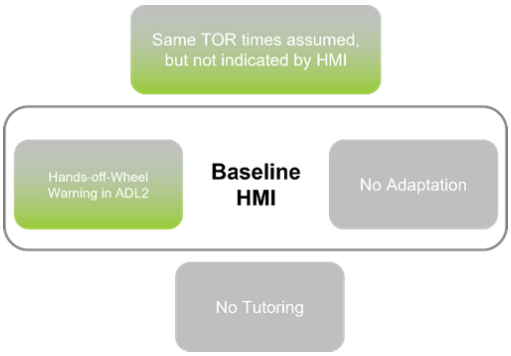

Fig. 2. Key features of the baseline HMI.

The methodology and the key findings from the implementation of the HADRIAN-tailored safety and impact assessment on data from this simulator study are fully reported in the subsequent subsections.

2.2 Assessment KPIs

A safety and impact assessment method designed specifically for the HADRIAN project was developed, incorporating unique Key Performance Indicators (KPIs). These KPIs were calculated by analyzing driving metrics from the simulator and collecting subjective data through questionnaires. This assessment consisted of nine KPIs related to safety and driving performance (e.g., more acceptable, useful, and comfortable), as well as another nine linked to their influence on drivers' perspective. (Fig. 3 & Tab. 1). Descriptive statistics were formulated to gain insights into trends, facilitate comparisons, and provide evidence of the safety and impact advancements achieved by the HADRIAN HMI. Subsequently, a scoring system was devised to benchmark the innovative HADRIAN HMI prototypes. This scoring approach leveraged Data Envelopment Analysis (DEA) applied to the previously mentioned Key Performance Indicators (KPIs). In this research, the participating drivers were regarded as Decision Making Units (DMU), responsible for determining a fixed distance value (output) by adjusting driving attributes, which in turn influenced the safety and perceived impact KPIs (inputs). Through the application of DEA, scores for safety and the perceived impact on the drivers were evaluated.

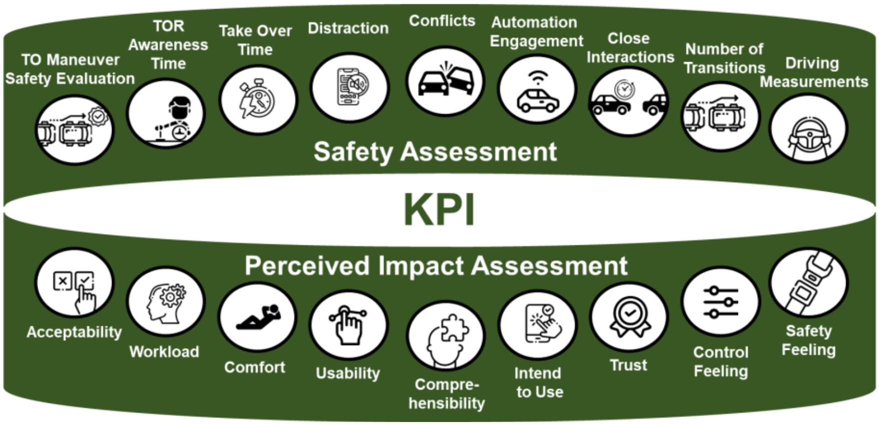

Fig. 3. Set of developed KPIs for safety & impact assessment of HADRIAN HMI innovations.

In Fig. 3, it is depicted the set of the assessed KPIs, and complementary to Fig. 3 the following table (Tab. 1) gives an overview of the KPIs description; the 9 safety KPIs (1.1 to 1.9), which are based on driving and eye-tracking data, as well as the 9 perceived impact KPIs (2.1 to 2.9), which are based on questionnaire responses provided by the driver. This study will not delve into the details of how the Key Performance Indicators (KPIs) are calculated, as its primary focus is on assessing the outcomes. The background, the development as well as the calculation of this assessment methodology, were given to past deliverables of the project (hadrianproject.eu/results/) and previous publications. More specifically, the literature review in light of automated driving acted as the theoretical background of this methodology with the identification of risk factors that were connected with HADRIAN use case drivers i.e., elderly, truck and office workers, was published in a journal paper (Sekadakis, Katrakazas, Santuccio, Mörtl, & Yannis, 2022). The practical considerations and HADRIAN indicators that were used in the safety and impact assessment for seamless interactions through human-machine interfaces were further discussed (Sekadakis, Katrakazas, Clement, Prueggler, & Yannis, 2022). Additionally, the initial extensive list of Key Performance Indicators that were obtained for safe fluid interactions within automated vehicles was presented in this conference publication (Sekadakis, Katrakazas, Santuccio, Mörtl, & Yannis, 2021). Finally, the descriptive insights from the calculated KPIs are fully reported in the latest publication (Sekadakis et al., 2023).

Tab. 1. KPI list for safety & impact assessment of HADRIAN HMI innovations.

KPI ID	***KPI***	***KPI division***	***Description***
Safety KPIs			
KPI 1.1	Take Over Maneuver Safety Evaluation		A scoring metric used to evaluate the safety of take-over maneuvers. This KPI exploits three different driving measurements i.e., speed, longitudinal acceleration, and deceleration in order to estimate the safety score (Ellison, Greaves, & Bliemer, 2015).
KPI 1.2	Take Over Request Awareness Time		Measures the time taken for a driver to be aware of the necessary information needed for the transition from automated driving to manual driving in order to perform a safe TO maneuver.
KPI 1.3	Take Over Time		Measures the time of reaction between automated and manual driving and is calculated by measuring the time it takes for the driver to interact with the vehicle after a request for automated driving level transition.
KPI 1.4	Distraction		Measures the time that a driver is distracted.
KPI 1.5	Conflicts		Measures the total number of conflicts that occur between the driver and all the road users or infrastructure.
KPI 1.6	Automation Engagement	• Level 2 • Level 3	Measures the duration of engagement in specific Automated Driving Levels (ADLs).
KPI 1.7	Close Interactions		A measurement of the number of events where the time to collision (TTC) was less than 3 seconds over a specific trip distance.
KPI 1.8	Number of Transitions	• AD → Manual • Manual → AD	Measures the number of times an autonomous vehicle transitions from automated driving (AD) to manual driving, and vice versa.
KPI 1.9	Driving Measurements	• Speeding Duration • Speed Over the Limit • Harsh Cornerings • Harsh Brakings • Harsh Accelerations	Includes 5 distinct safety metrics and corresponding thresholds.
Perceived Impact KPIs			

KPI 2.1	Acceptability Ratings	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.2	Subjective Workload	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.3	Comfort	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.4	Usability	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.5	Comprehensibility	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.6	Intend to Use	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.7	Trust	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.8	Control Feeling	Calculated using questionnaires distributed to the drivers during the driving experiment.
KPI 2.9	Safety Feeling	Calculated using questionnaires distributed to the drivers during the driving experiment.

2.3 Assessment Scoring

This subsection delves into the methodology used to benchmark and score the safety and perceived impact of HADRIAN HMI developments, serving as the conclusive assessment stage aimed at demonstrating the enhanced safety and positive influence on the driver. Fig. 4 provides an overview of the implementation and application of the safety and impact assessment process, with a specific focus on the final phases of this study. The methodology and the establishment of key performance indicators (KPIs) have been comprehensively documented in prior publications (as mentioned above) and project reports. In this study, the assessment leverages simulator and questionnaire data, which play a pivotal role in the computation of these aforementioned KPIs. Subsequently, the assessment employs Data Envelopment Analysis (DEA), and in the following sections, we delve deeper into the theoretical underpinnings of this analytical technique.

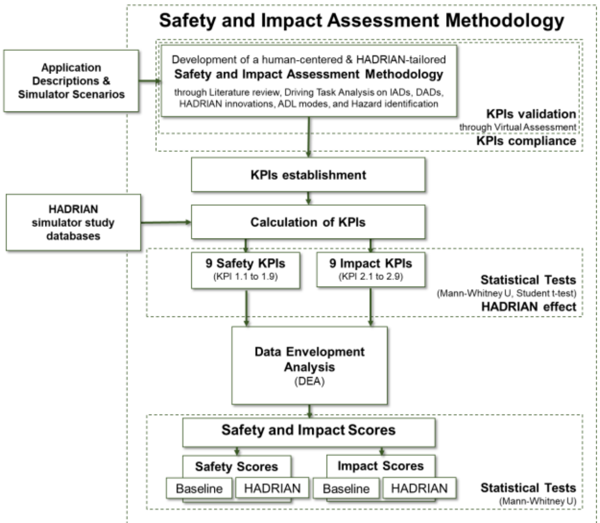

Fig. 4. Flow chart of the implementation and application of safety & impact assessment methodology.

Data Envelopment Analysis (DEA) is a widely recognized method employed in the fields of business, economics, and management to assess the efficiency and productivity of comparable production units. In recent years, DEA has also found expanded application within the transportation sector. The Decision-Making Units (DMU) are technically efficient when the amount of outputs produced is maximized for a given amount of inputs (output-oriented model), or for a given output the amount of inputs used is minimized (input-oriented model) (Farrell, 1957). DEA is a data-oriented technique that determines a DMU's efficiency based on the observed data of the rest of DMUs considering that all DMUs lay on or below the efficiency frontier (Ramanathan, 2003). Several existing studies have examined driving safety by deploying DEA using both driving simulator (Babaee, Bagherikahvarin, Sarrazin, Shen, & Hermans, 2015; Babaee et al., 2014) and naturalistic driving data (Tselentis, Vlahogianni, & Yannis, 2019).

In the present methodology, drivers were identified as Decision Making Units (DMUs) responsible for making various decisions regarding various driving aspects, e.g. distraction, speed, etc. within a specified distance. When drivers took part in both simulation conditions, namely "Baseline" and "HADRIAN," they were individually assessed within each condition. In this evaluation, distance was employed as the output variable because it remains relatively constant across all DMUs during the simulations, while the computed KPIs serve as the input parameters. An input-oriented DEA model was developed aiming to minimize the KPIs (inputs) while maintaining the same distance. This driving efficiency problem was hypothesized to be a constant-returns-to-scale (CRS) problem i.e., outputs are increased proportionally to inputs (Ramanathan, 2003; Tselentis et al., 2019). More specifically, the sum of safety and perceived impact KPIs (inputs) recorded in each simulation trip changes proportionally to the sum of driving distance (output).

Essentially, in this case, as can be seen in Fig. 5, KPIs were divided into 4 sub-groups, i.e., Automated Driving, Unexpected Driving Events, TOR Awareness & Distraction & Speeding, and Questionnaire Responses. This categorization was applied to calculate efficiencies taking into account the diverse nature of KPIs, e.g., events or continuous measurements that have a positive or negative influence on road safety. More specifically, as the input-oriented model aimed to minimize inputs, in the case of the following KPIs sub-groups namely Automated Driving and Questionnaire Responses, the KPIs, which should be maximized in a given distance by the DEA algorithm, were inverted and 1/KPIs were set as inputs in order to invert the influence. It is also important to note that the KPI 1.8, i.e., Number of Takeovers, and Number of Manual to AD transitions, were used as recorded in the given distance and not as normalized (by distance) measures, e.g., Total Number of Transitions/km, to avoid accounting twice the parameter of distance, since it was considered as an output of DEA.

DEA improvement algorithms were performed in R programming language, using the "Benchmarking" package, to calculate the efficiencies of the 4 KPIs sub-groups. As mentioned above, this problem was input-oriented (ORIENTATION="in") and constant-returns-to-scale (RTS="crs"). The total safety score was calculated as the average efficiency of the three objective KPI groups for each DMU, as shown in Fig. 5. The safety and perceived impact scores were finally evaluated if they were statistically significant using the Mann-Whitney U test.

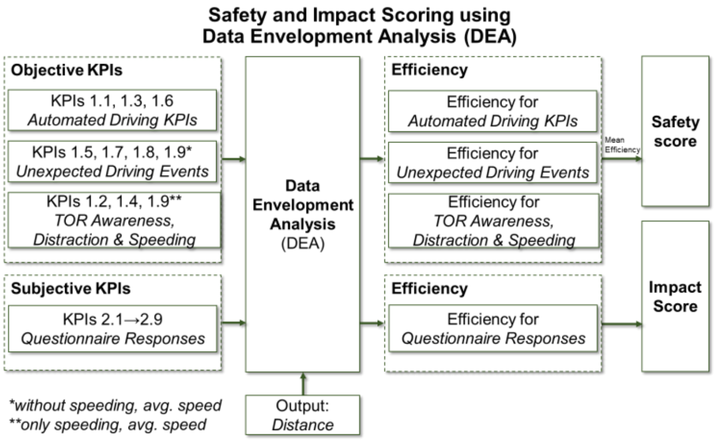

Fig. 5. Architecture of safety and perceived impact scoring using DEA.

Distinct scores/efficiencies were estimated for the "Baseline" and "HADRIAN" HMI conditions, as shown in Fig. 5. The "objective KPIs" and "subjective KPIs" were independently assessed using appropriate data analysis software, followed by the necessary computations to calculate safety and perceived impact scoring. For each safety and perceived impact scoring criterion, the Mann-Whitney U test was conducted. In cases where the calculated p-value fell below the established significance level, the statistical significance of the HADRIAN HMI impact was confirmed.

3. Results and Discussion

This section presents the findings of the assessment. The results reveal the impact of HADRIAN HMI innovations on safety and perceived impact, particularly when compared to a baseline HMI. Concurrently, these effects are analyzed by highlighting the advantageous KPIs and explicating how HADRIAN HMI innovations influenced the safety and perceived impact scores. To begin, descriptive statistics for the safety-related KPIs are presented and discussed. Next, the same analysis are performed for the perceived impact KPIs. Lastly, the examination of the outcomes from the scoring and benchmarking methodology is outlined.

3.1 Safety KPIs

The **driving objective safety KPIs**, namely KPI 1.1 - 1.9, have been included in a table showing the descriptive statistics of each KPI and are presented Tab. 2. The table showcases the participants' performance with descriptive statistics (i.e., mean, standard deviation, minimum and maximum value for each KPI) for the different HMI conditions, i.e., Baseline HMI and HADRIAN HMI. The table overviews the changes in performance on the KPIs, highlighting observed improvements and/or trend that may have arisen as a result of the HADRIAN HMI innovations.

For **KPI 1.1**, the "Takeover Manoeuvre Safety Score," all participants exhibit a similar score for the maneuver, with minimal variability evident in the mean and standard deviation values. However, when examining the trend in a plotted figure (as shown in Fig. 6), a distinct advantage in the Takeover Manoeuvre Safety Score becomes evident for the HADRIAN HMI. This superiority can be attributed to two primary factors: i) the driver's situational awareness before the takeover request and ii) the support provided during the takeover maneuver phase. In the HADRIAN conditions, the driver's situational awareness before the takeover request appears to be higher. Participants were well-informed and adequately prepared for the upcoming maneuver, thanks to the instructional video preceding the simulation experiment. Furthermore, the system informed drivers about the remaining time until the takeover request. Additionally, during the takeover maneuver phase, the HADRIAN system assisted the driver by indicating approaching obstacles, the current speed limit, and providing haptic cues at the steering wheel. The driver's condition and the driving support during the takeover phase contributed to a smoother takeover maneuver characterized by reduced abrupt accelerations and decelerations, as well as fewer instances of exceeding speed limits.

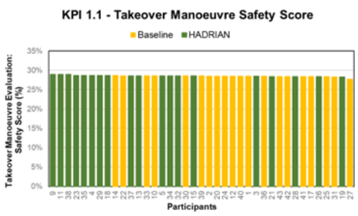

Fig. 6. KPI 1.1 "Takeover Manoeuvre Safety Score" plot.

For **KPI 1.2** "Take Over Request Awareness Time", participants that drove with “Baseline” HMI exhibited an increasing mean value, while HADRIAN participants presented lower "Take Over Request Awareness Time" values. This could be attributed to the countdown information to the driver (5 seconds for ADL2 and 15 seconds for ADL3) accompanied by the tutoring video, the LED indication and the verbal guidance and feedback, allowing drivers to assess all the required insights from the driving environment in less time. As a result of these mentioned enhancements within the same context, in relation to KPI 1.3, which is the "Take Over Time," HADRIAN participants exhibit a noticeable upward trend. This trend can be attributed to the extended period available for reaction and takeover in HADRIAN conditions. Consequently, the HADRIAN driver was not surprised by the takeover request and could effectively utilize most of the designated takeover time, in contrast to the baseline conditions. A longer takeover time proves highly advantageous for the driver, as it allows for a more comprehensive and optimal assessment of the driving environment and conditions before initiating the takeover.

Combining KPIs 1.1, 1.2, and 1.3, the beneficial effect of HADRIAN conditions on **takeover performance** is proven. Not only the driver was more prepared for a takeover request from automated to manual driving with a longer takeover time but, with the indications from the developed system, the driver took less time to scan the necessary information from the driving environment in order to react appropriately prior to the takeover. Furthermore, the HADRIAN supported the drivers to perform a smoother takeover maneuver with reduced values in speed, acceleration, and braking.

For **KPI 1.4** "Distraction Percentage", Baseline participants display significantly longer distraction durations, implying a direct effect of warnings of detected visual distraction or hands-off-wheel in HADRIAN conditions. For **KPI 1.5** "Number of Conflicts", participants demonstrate approximately similar frequencies in both conditions. There are not any significant differences between Baseline and HADRIAN HMIs and this could be attributed since there was no specific system indicating or reducing headways and consequently TTC. For **KPI 1.6** "ADL2 Duration" and KPI 1.6 "ADL3 Duration", similar values are observed with a slightly lower trend for HADRIAN participants, which can be explained due to the fixed durations AD level 2 and 3 scenarios in this study. Also, the values for the **KPI 1.7** "Frequency of TTC <3s" demonstrate increased frequency in baseline HMI conditions, contrary to the KPI 1.5 which the absence of a specific system for reducing headways could produce this result.

For **KPI 1.8** "Number of Takeovers" and "Number of Manual to AD Transitions", there are not many differences in total transitions per km, as there are fixed transitions in both conditions. HADRIAN participants show a slightly higher trend in takeovers and slightly lower in manual to AD transitions, which could be attributed to the design of the HADRIAN simulation. For **KPI 1.9** "Speeding Duration" and "Speed over the Limit", both conditions demonstrate similar mean values with HADRIAN participants tending to present higher values. The information about current speed and detected speed limit may not be effective in affecting driver behavior. For **KPI 1.9** "Harsh Cornerings" and KPI 1.9 "Harsh Accelerations", the total number of these cornering and acceleration events was quite infrequent for all participants and thus the percentage change seems to be extracted great without any essential differences. While for KPI 1.9 "Harsh Brakings", more events are recorded under HADRIAN conditions. HADRIAN drivers may have responded more aggressively in braking due to the indications of speed or provided warnings in the developed HUD.

Tab. 2. Descriptive statistics for driving safety KPIs for Baseline and HADRIAN HMIs.

KPI ID	**KPI**		**Units**	**Baseline HMI**				**HADRIAN HMI**				**Mean Percentage Change**
				Mean	**Standard Deviation**	**Min**	**Max**	**Mean**	**Standard Deviation**	**Min**	**Max**	
KPI 1.1	Take Over Maneuver Safety Evaluation		% Score	0.286	0.002	0.283	0.290	0.285	0.002	0.277	0.290	-0.248%
KPI 1.2	Take Over Request Awareness Time		Seconds	2.475	0.655	1.615	4.871	2.331	0.486	1.809	3.519	-5.814%
KPI 1.3	Take Over Time		Seconds	3.571	0.900	1.651	5.221	4.562	1.299	1.734	6.901	27.738%
KPI 1.4	Distraction		% Distraction Duration	7.995%	7.644%	0.000%	33.391%	0.619%	0.763%	0.000%	2.111%	-92.254%
KPI 1.5	Conflicts		Count	1.450	0.945	0.000	3.000	1.474	0.905	0.000	3.000	1.633%
KPI 1.6	Automation Engagement	• Level 2	% Level Duration	25.874%	0.698%	24.489%	26.791%	25.830%	0.644%	24.342%	27.239%	-0.170%
		• Level 3	% Level Duration	36.691%	0.839%	34.847%	37.896%	36.632%	0.736%	35.527%	38.544%	-0.159%
KPI 1.7	Close Interactions		Count	1.950	0.826	1.000	3.000	1.526	0.772	0.000	3.000	-21.727%
KPI 1.8	Number of Transitions	• AD → Manual	Count	7.700	0.733	6.000	9.000	8.053	1.026	5.000	10.000	4.580%

		• Manual → AD	Count	8.850	1.089	8.000	12.000	8.579	0.769	8.000	10.000	-3.063%
KPI 1.9	Driving Measurements	• Speeding Duration	Seconds	53.704	15.906	32.122	94.644	54.537	14.327	25.118	81.488	1.551%
		• Speed Over the Limit	Km/h	13.596	1.025	11.582	15.270	13.679	1.907	7.339	16.602	0.606%
		• Harsh Cornerings	Count	0.150	0.366	0.000	1.000	0.211	0.419	0.000	1.000	40.351%
		• Harsh Brakings	Count	5.950	3.692	1.000	13.000	7.684	4.191	3.000	19.000	29.146%
		• Harsh Accelerations	Count	1.050	1.099	0.000	4.000	1.263	1.327	0.000	4.000	20.301%

3.2 Perceived impact KPIs

Similarly to the previous subsection, the descriptive statistics of the driver **perceived impact KPIs**, namely KPI 2.1 - 2.9 are presented in the following table (Tab. 3). The table showcases the participants' ratings from the filled questionnaires for the different HMI driving conditions, i.e., Baseline and HADRIAN HMIs.

For **KPI 2.1** "Acceptability Ratings", a higher preference or acceptance towards the baseline HMI seem to be observed. Participants might have rated the baseline conditions more acceptable compared to the HADRIAN HMI, as they are more familiar with simpler and non-interventive vehicles. Regarding **KPI 2.2** "Subjective Workload", HADRIAN participants seem to rate lower the perceived effort required to operate the driving tasks compared to baseline participants. This can be attributed to the fact that the applied HADRIAN innovations provide many insights about the driving environment, reducing mental or cognitive effort. As for **KPI 2.3** "Comfort" and **KPI 2.4** "Usability", HADRIAN participants are observed to rank slightly higher for the HADRIAN innovations, denoting that the HADRIAN conditions are more convenient and usable to the driving experience.

For **KPI 2.5** "Comprehensibility" and **KPI 2.6** "Intend to Use", HADRIAN conditions seem to be ranked lower compared to the baseline ones. The applied innovations and the amount of information provided may increase the overall complexity and understanding of HADRIAN systems, resulting also in less likelihood of further utilization. For **KPI 2.7** "Trust" and **KPI 2.8** "Control Feeling", HADRIAN participants provided moderately higher ratings to the automated systems compared to the baseline participants. These results reveal the high efficiency of the applied HADRIAN systems in generating a sense of confidence and reliability. For **KPI 2.9** "Safety Feeling", there are not many variations between the two conditions, interestingly safety feeling values were reported slightly higher with the HADRIAN HMI.

Tab. 3. Descriptive statistics for driver perceived impact KPIs for Baseline and HADRIAN HMIs.

KPI ID	**KPI**	**Baseline HMI**	**HADRIAN HMI**	
		Median	**Median**	**Average Percentage Change**
KPI 2.1	Acceptability Ratings	5.5	5.0	-7.36%
KPI 2.2	Subjective Workload	35.8	26.7	-9.05%
KPI 2.3	Comfort	1.6	1.8	7.44%
KPI 2.4	Usability	5.4	5.4	1.35%
KPI 2.5	Comprehensibility	2.3	2.3	-2.45%
KPI 2.6	Intend to Use	1.6	1.7	-3.05%
KPI 2.7	Trust	4.2	4.0	1.99%
KPI 2.8	Control Feeling	5.0	5.3	6.13%

KPI 2.9	Safety Feeling	5.2	5.0	0.66%

3.3 DEA Scoring Results

Data Envelopment Analysis (DEA) is the benchmarking method used to evaluate the relative efficiency of driver based on their inputs and outputs, as mentioned in section 2.3. In this study, KPIs were used as inputs, and distance was used as an output to produce efficiency scores and for assessment purposes, efficiency percentages were translated into safety or perceived impact scores that can be translated into a scoring ranging from 0 to 100%. The efficiency scores are illustrated in a boxplot, Fig. 7, to compare the scoring of two conditions i.e., HADRIAN compared to baseline HMI. Safety score was extracted by performing DEA on safety KPIs 1.1 to 1.9. The findings indicate that the HADRIAN HMI exhibited a statistically significant increase in safety score compared to the baseline conditions, as established through a Mann-Whitney U test ($p<0.05$). Specifically, the safety score for the HADRIAN conditions demonstrated a higher average, median, and range of scores, signifying a substantial enhancement in safety scoring and, consequently, in the level of safety for autonomous driving (AD). The baseline HMI scored at 91.9%, and with the integration of HADRIAN innovations, it improved to 95%, reflecting a notable improvement of 3.37%.

The improvement in safety can be linked to the Key Performance Indicators (KPIs) that demonstrated positive effects on safety, as extracted from the descriptive statistics of these KPIs. The KPIs that had a beneficial influence on safety were the increased safety score of take over maneuver (KPI 1.1), decreased takeover request awareness time (KPI 1.2), increased take over time (KPI 1.3), limited distraction (KPI 1.4), and less frequent number of close interaction events (KPI 1.7), may have produced this safety improvement, as presented previously. Upon analyzing with DEA the detailed set of KPI values for each participant an overall safety improvement was revealed for the HADRIAN conditions, despite some KPIs having a negative effect on safety level, probably their general negative trend was not significant enough to negate the overall safety improvement observed in the HADRIAN condition. If the rest of the KPIs were more beneficial and positive to safety scoring, a higher safety improvement would have been caused.

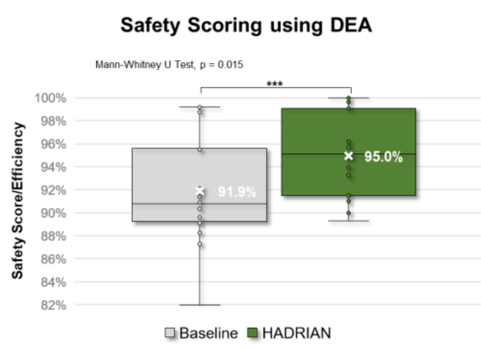

Fig. 7. Safety Scoring using DEA for the different conditions.

In addition, with DEA, the perceived impact score was also extracted by evaluating the impact of the HADRIAN conditions on the driver subjective perspectives. Similarly to safety scoring, the perceived impact score was extracted by performing DEA on subjective KPIs 2.1 to 2.9 exploiting questionnaire responses. As shown in Fig. 8, the perceived impact score for the HADRIAN condition had a lower mean, and median, indicating a slight decrease in perceived impact scoring. It can be affirmed that the baseline HMI score stood at 94.2%, and with the incorporation of HADRIAN innovations, it decreased to 87.1%, reflecting a decline of 7.54%. Interestingly, the majority (5 out of 9) of driver perceived impact KPIs showed a positive effect of the HADRIAN system and only 3 out of 9 showed a negative effect, however this was not sufficient to be drawn in the perceived impact score. It is possible that the driver in the driving simulator study felt unfamiliar with this type of more complex and interventive driving setup, despite the fact that the general effect reported a decrease in subjective workload, an increase in comfort, usability, trust, and control feeling. Nevertheless, it is important to emphasize that this perceived impact score is entirely subjective, indicating that the driver might not have accurately assessed the safety enhancements provided by the HADRIAN innovations, as elaborated in the previous paragraph.

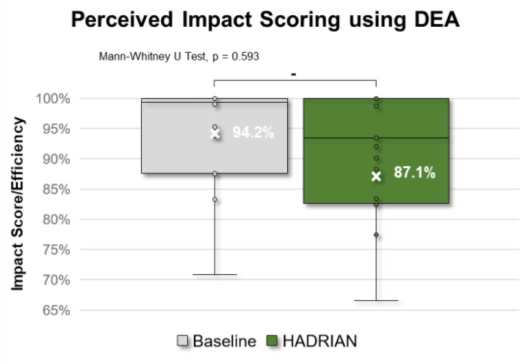

Fig. 8. Perceived Impact Scoring using DEA for the different conditions.

4. Conclusions

As shown in Fig. 7, the results revealed that the studied HADRIAN HMI had a statistically significant higher safety scoring than the "state-of-the-art" or baseline HMI, as determined by a Mann Whitney U test ($p<.05$). As stated previously, the safety enhancement can be attributed to KPIs that were beneficial for safety namely, increased safety score of take over maneuver for the majority of the observations ($p=.306$, Student's t-test), decreased takeover request awareness time for the majority of the observations ($p=.373$, Mann-Whitney U test), significantly increased take over time ($p=.009$, Mann-Whitney U test), significantly limited distraction ($p=.000$, Mann-Whitney U test), observed with the HADRIAN HMI. The rest of the KPIs had either a negative or a neutral effect on safety.

Incorporating the various safety KPIs, it can be demonstrated the beneficial effect of HADRIAN HMI on takeover performance. Not only the driver was more prepared for a takeover request from automated to manual driving with a longer takeover time but, with the indications from the HADRIAN HMI, the driver took less time to scan the necessary information from the driving environment to react appropriately prior to take-over. Furthermore, with the HADRIAN supported the drivers perform a smoother takeover maneuver with lower values of speed, harsh accelerations, and harsh braking. These outcomes should have also benefited from the reduced distraction since the driver was urged by the HMI to stay in the loop at ADL2 with the driving task, as extracted by KPI showing the driver's distraction.

As depicted in Fig. 8, in contrast to the safety scoring, the perceived impact scoring reveals lower values for the HADRIAN HMI. The perceived impact score for the HADRIAN condition exhibited a reduced mean and median, indicating a decline in perceived impact scoring. However, it is noteworthy that among the 9 driver-perceived impact KPIs, a majority of 5 indicated a positive impact of the HADRIAN system, while only 3 showed a negative effect. Nonetheless, this was not substantial enough to impact the overall perceived impact score. It is possible that participants in the driving simulator study may have felt somewhat unfamiliar with the more complex and intervention-oriented driving setup, despite the overall positive effects reported, including decreased subjective workload, increased comfort, usability, trust, and a greater sense of control. It is important to emphasize that this perceived impact score is entirely subjective, indicating that drivers may not have accurately assessed the safety enhancements offered by the HADRIAN innovations.

The methodology and findings of this study hold significance as they can be utilized by any other HMI stakeholder seeking to employ similar human-centered assessment approaches to evaluate the safety implications of AD and how humans interact with potential HMI configurations. This broad applicability ensures that the insights from this study can be beneficial to a wide range of individuals and organizations involved in HMI or AD research and development.

List of abbreviations

Abbreviations	Definitions
KPI	Key Performance Indicators
HMI	Human Machine Interface
AD	Automated Driving
HUD	Heads Up Display
HADRIAN	Holistic Approach for Driver Role Integration and Automation Allocation for European Mobility Needs
ADL	Automated Driving Level
TO	Take Over
TOR	Take Over Request

Declarations

Availability of data and material

All data used during the study are of a confidential nature and may only be provided with restrictions (e.g., anonymized data) upon request to the HADRIAN consortium (hadrianproject.eu).

Funding

HADRIAN (hadrianproject.eu) has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 875597. The document reflects only the author’s view, the Agency and the Commission are not responsible for any use that may be made of the information it contains.

Acknowledgements

The authors of this study would like to thank the entire HADRIAN consortium for their dedication and teamwork. This work is a collaborative effort, without which, each individual contribution would not have culminated in its successful completion.

Competing interests

The authors state that they don't have any clear financial conflicts or personal interests that might have seemed to affect the work described in this paper.

Authors' contributions

Marios Sekadakis: Writing – original draft, Software, Methodology, Data curation. Marianthi Kallidoni: Software, Data curation. Christos Katrakazas: Scoring Conceptualization, Methodology. Sandra Trösterer: Writing – review & editing, Data gathering. Cyril Marx: Data gathering and curation. Peter Moertl: Conceptualization and Supervision. George Yannis: Supervision.

References

Babaee, S., Bagherikahvarin, M., Sarrazin, R., Shen, Y., & Hermans, E. (2015). Use of DEA and PROMETHEE II to assess the performance of older drivers. *Transportation Research Procedia*, *10*(July), 798–808. https://doi.org/10.1016/j.trpro.2015.09.033

Babaee, S., Shen, Y., Hermans, E., Wets, G., Brijs, T., & Ariën, C. (2014). Combining Driving Performance Information in an Index Score: A Simulated Curve-Taking Experiment. *Transportation Research Record*, *2434*(1), 44–51. https://doi.org/10.3141/2434-06

Ellison, A. B., Greaves, S. P., & Bliemer, M. C. J. (2015). Driver behaviour profiles for road safety analysis. *Accident Analysis and Prevention*, *76*, 118–132. https://doi.org/10.1016/j.aap.2015.01.009

Elvik, R. (2021). Can the impacts of connected and automated vehicles be predicted? *Danish Journal of Transportation*, *3*, 1–13. Retrieved from https://levitate-project.eu/wp-content/uploads/2021/02/DJTR-predicting-impacts-CAV.pdf

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253–290. https://doi.org/10.2307/2343100

NHTSA. (2015). *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*.

Ramanathan, R. (2003). *An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement* (SAGE, Ed.).

Sekadakis, M., Kallidoni, M., Katrakazas, C., Trösterer, S., Marx, C., Moertl, P., & Yannis, G. (2023). Impact Assessment of a novel Human-Machine Interface Prototype : A descriptive analysis from the HADRIAN project. *11th International Congress on Transportation Research*, *1*, 1–10.

Sekadakis, M., Katrakazas, C., Clement, P., Prueggler, A., & Yannis, G. (2022). Safety and impact assessment for seamless interactions through human-machine interfaces : indicators and practical considerations. *Transport Research Arena 2022, Transportation Research Procedia*, *00*, 1–8.

Sekadakis, M., Katrakazas, C., Santuccio, E., Mörtl, P., & Yannis, G. (2021). Key Performance Indicators for safe fluid interactions within automated vehicles. *10th International Congress on Transportation Research*.

Sekadakis, M., Katrakazas, C., Santuccio, E., Mörtl, P., & Yannis, G. (2022). Risk factors linked with elderly , truck and office worker drivers : a literature review in light of automated driving. *Advances in Transportation Studies*, *3*(Special Issue), 95–108.

Tselentis, D. I., Vlahogianni, E. I., & Yannis, G. (2019). Driving safety efficiency benchmarking using smartphone data. *Transportation Research Part C: Emerging Technologies*, *109*(June 2018), 343–357. https://doi.org/10.1016/j.trc.2019.11.006

WHO. (2018). *GLOBAL STATUS REPORT ON ROAD SAFETY 2018*. ISBN 978-92-4-156568-4.

WHO. (2021). Road traffic injuries. Retrieved December 27, 2021, from https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Wijnen, W., Weijermars, W., Schoeters, A., van den Berghe, W., Bauer, R., Carnis, L., … Martensen, H. (2019). An analysis of official road crash cost estimates in European countries. *Safety Science*, *113*(December 2018), 318–327. https://doi.org/10.1016/j.ssci.2018.12.004

Yannis, G., Papadimitriou, E., & Folla, K. (2014). Effect of GDP changes on road traffic fatalities. *Safety Science*, *63*, 42–49. https://doi.org/10.1016/j.ssci.2013.10.017