To cite this version :

José Alejandro BETANCUR, Nicolás GÓMEZ, Mario CASTRO, Frédéric MERIENNE, Daniel SUÁREZ - User experience comparison among touchless, haptic and voice Head-Up Displays interfaces in automobiles - International Journal on Interactive Design and Manufacturing - Vol. 12, n°4, p.1469-1479 - 2018

User experience comparison among touchless, haptic and voice Head-Up Displays interfaces in automobiles

J. Alejandro Betancur[1] · Nicolás Gómez[1] · Mario Castro[1] · Frederic Merienne[2] · Daniel Suárez[1]

Abstract

This paper evaluates driving experiences when using a Head-Up Display (HUD) system by different interaction methods, since HUDs in automotive market do not follow standard interaction methods (Betancur et al. in Int J Interact Des Manuf 12(1):199–214, [3]), and therefore it is difficult to identify what types of these methods are the most appropriated in terms of usability for being implemented in HUDs. This paper focuses on comparing the mental workload that a driver could experience while interacting with a specific HUD visual interface by touchless gestures, haptic or voice methods. Then, test subjects (n = 15) performed conjointly a driving activity and a set of in-vehicle tasks by using these interaction methods throughout a HUD. The experiments show that the haptic method was better accepted than the touchless gestures and voice methods. Nevertheless, the touchless gestures method was explored under some specific usability configurations, in order to demonstrate that it was not significantly different from the haptic method.

Keywords Head-Up Display · Human–machine interface · Vehicle

1 Introduction

Regarding the authors' perspective, the *Interactive approach* concept explores different ways for interacting with a technical system, mainly looking for enhancing the user experience and performance. Consequently, this paper proposes a comparison among different interaction methods applied to a HUD system, in order to identify the best and worst qualities of every one of these from drivers' point of view. According to the above, the interactive development here proposed is focused on evaluating and comparing methodologically different ways for interacting with a technical system, which is useful for solving interaction problems in engineer design processes for industry. Also, the experimentation tests between the proposed HUD system and the drivers generated different interaction models, which can be quantitatively described in terms of mental workload estimators, for then being compared and complemented among these same; in this way, these models can be explored under an engineering analysis, and even discussed in term of scientific hypotheses.

The HUD systems are devices that produce a virtual image in front the driver's field of view [2], as shown in Fig. 1. These systems have been broadly analysed in order to determine its influence on the driver behaviors. As a result, currently there are different studies that confirm the advantages of automotive HUD systems in terms of mental workload [12, 30], driving attention [8, 32], active safety [19], traffic accidents [23], among others. Nonetheless, regarding traffic accidents, Liu and Wen [23] propose that the time of glances from the road to the HUD systems must be less than 2 s, in order to avoid as much as possible hazard situations. Therefore, to avoid overcoming such time limit, an efficient interaction method and visual friendly interface are strongly necessary when using HUD systems, which is the main reason why this

---

✉ J. Alejandro Betancur
j-betancur@javeriana.edu.co

Nicolás Gómez
gomez.nicolas@javeriana.edu.co

Mario Castro
mariocastro@javeriana.edu.co

Frederic Merienne
frederic.merienne@ensam.eu

Daniel Suárez
d-suarez@javeriana.edu.co

[1] Departamento de Ingeniería Industrial, Pontificia Universidad Javeriana, Bogotá 110231, Colombia

[2] Lispen EA7515, Arts et Métiers, UBFC, HeSam, Institut Image, 2 rue T. Dumorey, 71100 Chalon-sur-Saone, France

Fig. 1 HUD projections proposed by enterprise 1 and 2 [18, 28]

research study focuses on comparing some of the main and current interaction methods.

According to the bibliometric classification method proposed by Betancur et al. [3], most of the studies about HUD interfaces are focused on proposing a faster driver acceptation, by using commonly touchless [9], haptic [16] and voice commands [22]. Consequently, in the present study a set of experiments were developed to determine the most safety-oriented methods for interacting with a proposed HUD visual interface; this, with the aim to analyse how the obtained results depend not only on the proposed HUD visual interface, but also on the way it is handled.

Moreover, regarding different studies about driving benefits for HUD systems when using Audio-Visual (AV) feedbacks [5, 15, 17], these last ones were implemented on the HUD visual interface here proposed. Also, this interface was designed according to visual design patterns that have been established by previous studies, mainly for hierarchical and list-based menu structures [24, 25].

Next subsection describes previous well-known research works focused on analysing automotive HUD interfaces, underlying design parameters, experimental considerations, HUD interface proposals, among others. Section 2, specifically Sects. 2.1 and 2.2 present the research objectives and the methodology proposed to fulfil these. Subsections 2.3 and 2.4 present in depth the experimental implementation procedure proposed for the methodological approach; then, results and discussions are detailed in Sect. 3 and 4 respectively. Finally, conclusions and future research steps about the obtained data are discussed in Sect. 5.

1.1 Related works: driver interfaces and driver behaviour

Lif et al. [22] considered in a simulated combat vehicle the impact of four display configurations on the driver reaction time, threat identification, and mental workload. Here, 20 non-military subjects aged between 18 and 26 years old were asked to align and fire, as quickly as possible, the heading of combat vehicle towards any threat displayed, while they were driving according to navigation information. Driving indications were given by a tactile display, and threat information was displayed using four different configurations (using both HUD and Head-Down Display): visual +3D audio, tactile +3D audio, tactile only, and multimodal (visual, tactile and 3D audio). An Analysis of variance (ANOVA) was used to analyse performance data. This study concluded that visual +3D audio displays provided better precision and shorter reaction times, while multimodal configuration had higher rates on perceiving threat position. These types of researches are some of the most outstanding ones, specifically those that explore some auditory stimuli configurations for interacting with HUD functionalities; as proposed by Jakus et al. [17], who explored some auditory stimuli configurations for handling HUD functionalities, finding that audio-visual HUD system configurations allow to keep attention on the road instead of on HUDs, which was really valuable when driving situations required full visual attention; also, these system configurations allowed to use HUD visual interfaces in less time and with less physical effort.

Luzheng et al. [4] proposed a P300 Brain Computer Interface (BCI) in the context of HUDs, mainly for predicting the user's driving destination. In this way, three healthy males between 20 and 30 years old were asked to respond to the P300 stimuli, at 9 possible locations (a 3*3 matrix). The data collection was composed by two sets: the first one was responsible for recording the EEG potentials of all test subjects, while the second set consisted of characters that were displayed on the windshield by a HUD system, and the timing in which they were flashed to the test subjects. Luzheng et al. [4] used the Principal Component Analysis and the Linear Discriminant Analysis for determining the driving destinations desired by the users. This study demonstrated that the P300 BCI is feasible since the accuracy of user's destination was 75%.

Hosseini et al. [16] introduce a new generation of automotive augmented reality systems for driving assistance, including a passive stereo night vision system that captures heat energy from objects, in order to detect 3D positions of potential collision partners; but also, it detects the driver's gaze direction, so that warnings are displayed at the exact position according to the viewing direction on a full-windshield. The fusion of stereo vision and thermal cameras allow a better background objects identification, based on their temperature as well as their 3D positioning. Hosseini et al., describe algorithms and real time information processing requirements for identifying driver's viewing direction and for recognizing, classifying and calculating distances to obstacles, as well as full-windshield projectors selection. Functionalities for dynamical visualization of collision warnings were demonstrated on a prototype.

Simon et al. [29] proposed improvements to an Advanced Driver Assistance System. This approach pursues an increase on road safety by displaying on a HUD not noticeable enough road signs. In this study, 30 observers were asked to look at road images randomly flashed for 5 s and count for the "no entry" signs. The positions and durations of subject's fixations at images were collected by a remote eye-tracker. Then, a second phase required participants to rate (Between 0-10) the saliency of the above mentioned signs. Results suggest that the saliency is related to the size of a sign and its Intrinsic Saliency (the saliency of road signs computed through a learning and classification function). This lead to define a Computational Size-Dependent saliency estimator, which allows to decide whether the sign is salient enough or not. This information can be used by intelligent vehicles to display only the poorly noticeable road signs.

Götze et al. [13] summarized the qualitative and quantitative requirements about automotive Human Machine Interfaces (HMI) for urban areas, taking into account its main components: a HUD, an instrument cluster, and an acceleration force feedback pedal. The study included different sources and almost 150 confirmed requirements; for instance, qualitative requirements applied to a HUD system include: minimization of textual content, as well as avoid driving scene overlapping, using colours red and blue at night, using animations to catch attention unless for warning display, and overlapping of content, among others. On the other hand, quantitative requirements for a HUD systems indicate for instance the minimum and ideal character height at a 3 m distance (10.47 mm = 12” and 17.45 mm = 20” respectively), display duration of road signs alarms (≥ 5 s), maximum time of glances from the road to a HUD (< 2 s), maximum numbers of colours (4), and suggested luminance of a HUD (≥ 17.000 cd/m^2).

2 Methods

2.1 Problem understanding

Automotive HUD systems differ radically from each other, mainly because these do not follow driver-based interaction design patterns [3], and therefore among all possible options it is not clear which interaction method is the most advisable (from a safety-oriented point of view) for a specific HUD visual interface proposal. According to this, the following three interaction methods were tested and compared for a specific HUD visual interface:

- Haptic buttons (H)
- Voice commands (V)
- Touchless gestures (T)

Consequently, this study is focused on estimating drivers’ mental workload when using a specific HUD visualization throughout the H, V or T interaction methods, identifying which of these methods induce a lower mental workload on the driving activities. This estimation is proposed by solving the following questions:

- Which interaction method require less time to activate the proposed in-vehicle tasks?
- Which interaction method implies a lower blink rate?
- Which interaction method is better accepted by the test subjects?

2.2 Problem development

Firstly, in subsection 2.3 a characterization to describe the test subjects in terms of visual performance is detailed, mainly regarding the Visual Acuity (VA) test using the Landolt Rings symbols[1] (LR). Then, all test subjects were prompted to drive several times a specific car in a track (20 min or 10 laps), in order to compare their lap time and in this way evidencing significant differences among their driving performance; at the same time, during these tests their respective average blink rates were measured and compared for various driving situations, which according to different authors it is an indicator of mental workload demand [10, 20, 34]. For performing all the above, a basic driving instrumental setup was developed, which allows to recreate (in terms of visual effort and mental workload) a real driving situation when the proposed HUD system is being used.

Afterwards, a HUD interface was proposed in Sect. 2.4, specifically in Sect. 2.4.1, regarding the design parameters previously mentioned in Sect. 1.1; also, ensuring by the results obtained in Sect. 2.3 that every test subject would have the needed VA for identifying all graphic details included in the proposed HUD visual interface. The test subjects were also asked for driving while using this interface for 20 min.

After all test subjects became accustomed to driving while using a HUD system, Sect.2.4.2 indicates how to evaluate the proposed HUD visual interface by using H, V or T as interaction method, in order to determine which of these contribute to a lower mental workload demand. Thereby, this paper proposes to measure and compare by every of these methods: the drivers’ acceptance (qualitatively), the activation time of a proposed set of in-vehicle tasks, and also the blinks rates during the development of these tasks.

Additionally, for this last one measure, a comparison between driving with and without the proposed HUD visual interface by H, V or T was carried out, identifying which of these methods have a higher impact on the driver’s’ mental workload demand. Section 2.4.3 detail and justify the experimental conditions applied for obtaining the above-mentioned measures. In Sect. 3, the results obtained from Sect. 2 were presented, detailing usability relations, best performances,

---

[1] LR symbol: it is an optotype proposed by Edmund Landolt, which is focused on the visual acuity evaluation.

significant differences, and also exploring the T interaction method as a reliable option for improving the visual understanding of the HUDs while driving. Section 4 present a discussion not by interaction method, but by in-vehicle task; also, this section suggests some implications about using the proposed HUD visual interface, the blink rate differences when using or not a HUD, and finally proposing a mental workload indicator regarding the previous results. Conclusions about the previous research steps and future works are enunciated in Sect. 5.

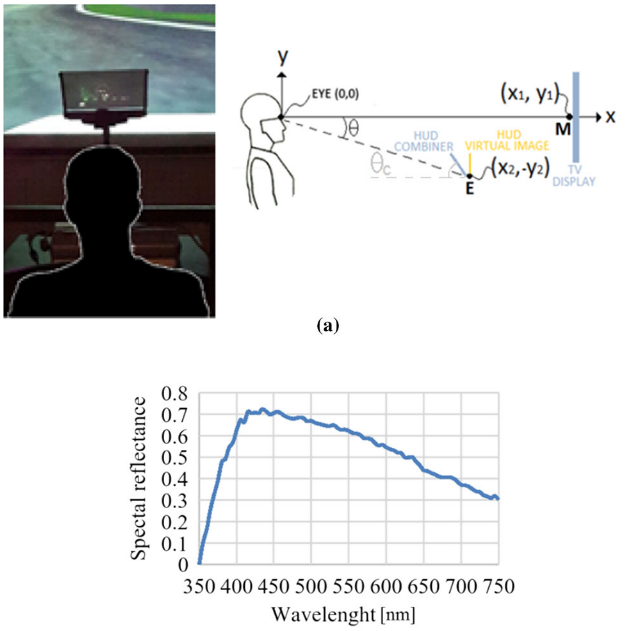
2.3 Test subjects description and driving instrumental setup

Primarily, a VA test was implemented for all test subjects in order to compare their respective visual qualities [1], looking for selecting a homogenous testing sample. As a result, 15 test subjects (8 males and 7 females) from 18 to 34 years old (M = 21.6, SD = 4.7) were selected. In this way, all test subjects have not only normal or corrected-to-normal sight and hearing, but also similar visual acuity (Right eye: M = 1.079, SD = 0.403, Left Eye: M = 1.055, SD = 0.424).

Regarding that some studies support differences in terms of reaction times between right and left handed subjects [27], all selected test subjects were right hand. Finally, every test subject provided a written consent to participate in all the experiments involved by this research, and also declared not having history of psychiatric, neurological or dramatic ophthalmological illness.

The proposed driving instrumental setup is exposed in Fig. 2a), which is configured in order to simulate the same visual effort the test subjects would experiment while driving. In this way, regarding an emmetropic adult human looking at the optical infinity (6 m) [12], and the human depth of focus as 0.43D (2.3 m) [7], in this setup the roadway screen and the HUD virtual image were placed at both of these distances respectively ($x_1 = 6.0$ m, $y_1 = 0$ m; $x_2 = 2.3$ m, $y_2 = -0.4$ m). Moreover, the HUD virtual image has an extension of 7.92 cm x 15.12 cm (vertical × horizontal); finally, the HUD combiner was made of a material I (thickness 2.15 ± 0.05 mm) with a reflectance depicted in Fig. 2.b).

In relation to the driver´s horizontal gaze line, the HUD virtual image has 10° as maximum eccentricity value, and the HUD combiner has 45° of inclination angle, as shown in the Fig. 2a by θ and θc respectively. This configuration allows moderately to avoid the cognitive capture phenomena, which occurs when driving responses are strongly affected due to the processing information from a HUD image [12]. In this way, the drivers are not prompted to move their heads to view the proposed HUD visual interface, but just their sight and using their peripheral field of view.



(a)

(b)

Fig. 2 **a** Test subject - instrumental setup configuration. **b** Reflectance spectrum of the proposed HUD combiner

2.4 Proposed HUD interface proposal and comparisons

2.4.1 HUD visual interface

An operative structure was proposed regardless the HUD visual interface, which is composed by 2 main group of options: music and air conditioning; each of these categories gathers other group of options, and so on, until a maximum of 4 ramifications, which means 3 levels in depth as indicated in the following function tree:

- 1. Music
  - 1.1 AM
    - 1.1.1 Pre-set stations
      - 1.1.1.1 Station 1
      - 1.1.1.2 Station 2
      - 1.1.1.3 Station 3
      - 1.1.1.4 Station 4
      - 1.1.1.5 Move backward
    - 1.1.2 Move backward
  - 1.2 FM
    - 1.2.1 Pre-set stations
      - 1.2.1.1 Station 1
      - 1.2.1.2 Station 2
      - 1.2.1.3 Station 3
      - 1.2.1.4 Station 4
      - 1.2.1.5 Move backward

    - 1.2.2 Move backward
  - 1.3 Volume
    - 1.3.1 Loud
    - 1.3.2 Move backward
  - 1.4 Devices
    - 1.4.2 CD player
      - 1.4.2.1 Next
      - 1.4.2.2 Move backward
  - 1.5 Back to the main menu

- 2. Air conditioning

  - 2.1 Air flow direction
    - 2.1.1 Frontward and downward
    - 2.1.2 Move backward
  - 2.2 Air flow
    - 2.2.1 Frontal windshield
    - 2.2.2 Back windshield
    - 2.2.3 Recycle A: it recycles the air flow using just the air inside the car
    - 2.2.4 Recycle B: it recycles the air flow using the air outer the car
    - 2.2.5 Move backward
  - 2.3 Move backward

The green, blue, yellow and orange colours were applied respectively for each level in depth, avoiding any possible interaction effects among them; then, a visual interface for the proposed HUD systems was designed regarding the above-mentioned functional structure and the design parameters depicted in the Sect. 1.1. Also, as an auditory feedback every displacement and selection option produce the same specific sound.

On the proposed instrumental setup, all related auditory feedbacks (engine, brakes, HUD outputs, etc.) and also the in-vehicle tasks reading (pre-recorded throughout the TTSreader sound clips generator [33]) were implemented by using two computer speakers (Conexant smartaudio HD) placed close to the test subjects. The above, looking for emulating the in-vehicle car audio speakers, which are close to drivers' seat.

2.4.2 HUD interaction method

Then, the H, V, and T interaction methods for the proposed visual interface were stated. The HUD visual interface was handled by pressing 3 buttons on the steering wheel (backward option, forward option, and select option), which correspond to 3 touchless gestures, and also to 3 voice commands. In addition, the T interaction method was designed to be used by hand gestures without moving the hand away from the steering wheel. The H, V and T interaction methods were implemented by configuring a Genius Speed Steering wheel 3MT (without force pedal feedback), a Microsoft Kinect 360 and a 5DT Data Glove respectively.

In this way, all proposed interaction methods allow for the proposed HUD visual interface, going up in depth by selecting an in-vehicle function, and also going down in depth by selecting the *Move backward* option. In any case, all in-vehicle functions of any of the proposed levels in depth are always visible.

2.4.3 Mental workload measures

In the cognitive context, the mental workload concept has been considered as "the portion of an individual's limited mental capacity that is actually required by task demands" [26], but also as the interaction between the requirements of a driving task, the circumstances under which it is performed, the skills, the behaviours, and the perceptions [14]. Then, the evaluation of mental workload could help to understand the effort that a test subject suffers while driving, specifically when analysing visual attention tests.

According to the above, and as proposed by Stern et al. [31] and Faure et al. [11], the first suggested experiment was focused on going as fast as possible around a specific car-track[2] configuration without using HUD systems. Then, for the proposed HUD visual interface, a set of in-vehicle tasks were developed while driving, using the H, V and T interaction methods, measuring for every one of these methods and tasks:

- Blink rate: usually while the test subject become more concentrated, there is a reduction on his/her blinking frequency [6]. In this sense, it is valuable to consider this reduction as an indicator of mental workload demand; for instance, in some studies blink rates and blink durations have been considered as mental workload indicators [10, 20]. This measure was taken while the test subjects drove with and without a HUD system.

---

[2] Driving a car in a specific test track, both detailed as follows: Sport utility vehicle (automatic transmission): mass = 1000 kg, motor revolution per minute (min–max) = 3.000–10.000, differential ratio = 3.67, front brake torque = 6000 Nm, back brake torque = 5500 Nm, wheel mass (4 wheels) = 23 Kg, wheel radius = 0.4 m, dynamic friction (wheel-ground) = 0.025, static friction (wheel-ground) = 0.9.

Track (without traffic): 3 lane track (11.6 m wide) arranged as follows: straight line of 1119.63 m, curve of radius 162.23 m, straight line of 378.67 m, curve of radius 104.78 m, straight line of 1811.40 m, 3 consecutive curves of radius 108.90, 209.63, 209.63 m respectively, straight line of 198.58 m, 4 consecutive curves all of radius 211.60 m, straight line of 176.09 m, 2 consecutive curves of radius 213.68 and 207.66 m respectively (downhill), straight line of 176.09 m, 3 consecutive curves of radius 206.58, 206.58, 190.24 m respectively (uphill). Good visibility conditions without traffic.

Table 1 Proposed in-vehicle tasks

Task	Description
1	Play the following CD track
2	Set the airflow direction downward and toward the frontal windshield
3	Turn the volume up (loud)
4	Press AM and select pre-set station 3
5	Press FM and select pre-set station 4

- Activation time: the activation time is the time required for a test subject to develop a requested action from the instant in which this action was fully described until the instant the requested action is fully completed. According to the above, the test subjects were prompted to develop the in-vehicle tasks indicated in Table 1.
- Driver acceptance: a quantitative result about the User Experience (UX) was obtained; this, by asking the test subjects for ranking from 1 to 10 the difficulty level tolerated after developing the proposed interaction tests.

3 Results and analysis

For estimating the mental workload of each test subject different type of data were collected; then, due to the non-normality that some data groups presented, these were analysed by using a the Kruskal–Wallis test in SPSS Statistics 22® and PAST 2.17c®, for which homogeneity of variance assumption was verified.

3.1 Driving performance and attention

First, after being introduced to the proposed study, 10 lap times were recorded per test subject and compared among these (without using the proposed HUD); also, control limits were established to identify atypical data cases and then applying the Kruskal–Wallis test. Overall comparisons among the test subjects were separated by gender, in order to validate homogeneity of variance. The results show significant differences between male and female subjects ($H = 46.71$, 1 d.f., p value $= 8.24\text{E}-12 < 0.05$), where men tend to have lower lap times than women (mean rank: male = 46.61 s, female = 92.58 s).

Within the male group, the Kruskal–Wallis test suggests there are significant differences between subjects ($H = 20.03$, 7 d.f., p value $= 0.0055 < 0.05$); but, after applying Mann–Whitney pairwise comparisons corrected by Bonferroni, results indicate that there are not significant differences between any pair of the male test subjects (p values > 0.05, for all pairs).

On the other hand, the female group showed very different frequency distributions among test subjects, which did not allow to validate the homogeneity of variance assumption; the above could be explained by the wide range of driving experience of women selected for this study.

The same procedure was applied to the blink rates measured while driving. The Kruskal–Wallis test suggest that there are significant differences between subjects within the male group ($H = 18.7$, 6 d.f., p value $= 0.005 < 0.05$), but once the Bonferroni correction was applied to the Mann–Whitney test, it confirmed that there are not significant differences between any pair of subjects. On the other hand, the results of the Kruskal–Wallis test for female group do not show significant differences between subjects ($H = 5.6$, 2 d.f., p value $= 0.061 > 0.05$).

3.2 Mental workload by interaction method

As mentioned above, each of the test subjects was requested to activate the in-vehicle functions depicted in Table 1, while measuring different workload estimators: blink rates, activation time, and driver acceptance values.

3.2.1 Blink rate

This data set was also analysed through the Kruskal–Wallis test, finding not enough evidence to confirm that there were significant differences among the blink rates, when using the HUD visual interface throughout the proposed H, V and T interaction methods ($H = 2.934$, 2 d.f., p value $= 0.231 > 0.05$). The Table 2 shows the mean ranks for the blink rates.

3.2.2 Activation time

The activation time data did not meet normality assumptions. Consequently, as it was necessary to validate homogeneity of variances for running the Kruskal–Wallis test, control limits were established for identifying atypical cases, which could be related to unexpected distractions suffered by the test subjects; then, 5 atypical cases were eliminated from the 225 observations in the original data base. This procedure allowed to verify homogeneity of variances and using the Kruskal–Wallis test.

Table 2 Mean rank for blink rate results

Variable	Interaction method	N	Mean rank (Blinks/min)
Blinks	H	44	59.39
	V	45	70.43
	T	45	72.50

Table 3 Mean rank for activation time results

Variable	Interaction method	N	Mean rank (ms)
Activation time	H	74	61.53
	V	73	130.41
	T	73	140.23

Table 4 Mean rank for the driver acceptance metric

Variable	Interaction method	N	Mean rank (NA)
Driver acceptance	H	15	14.57
	V	15	26.13
	T	15	28.30

Results showed that the proposed interaction methods produce a significant effect on the activation time of the overall in-vehicle functions ($H = 66.875$, 2 d.f., p value $= 0.000 < 0.05$). Furthermore, to identify which of the above mentioned interaction methods showed significant different activation times, the Mann–Whitney pairwise comparison was used, as well as a Bonferroni correction, to make results more accurate.

Specifically, H has significantly different activation times than both, V and T, but the last two do not show differences between them. Results suggest that H produce significantly lower activation times than V and T, as seen in the mean ranks showed in Table 3.

3.2.3 Driver acceptance

A quantitative exploration about the user opinion was done by answering from 1 to 10 (easiest-hardest): how much mental demand the H, V and T interaction methods are requesting for developing the proposed in-vehicle tasks? The Kruskal–Wallis test show that difficulty perceived by drivers is significantly different, at least between one pair of these methods ($H = 9.657$, 2 d.f., p value $= 0.0079 < 0.05$). The Table 4 shows mean ranks by interaction method for the proposed driver acceptance metric.

The above could be corroborated with the Mann–Whitney post hoc test, suggesting that H produce a significantly higher acceptance than V and T (for H-V: p value $= 5.6\text{E}-11 < 0.05$, and for H-T: p value $= 7.4\text{E}-13 < 0.05$), but there is not significant differences between the two latter (p value $= 0.81 > 0.05$ for V–T). Also, this analysis was consistent when using the Bonferroni correction.

3.3 Touchless exploration

Regarding previous studies about the future implementation trends of the T interaction methods in HUDs [3], a deeper exploration about this was proposed; specifically, it was tested a T interaction method (T') that allow to access the visible options of the proposed visual interface directly by doing a hand gesture, avoiding to navigate inside this visual interface by using the backward, forward, and select commands. Also, this T' interaction method was designed to be used without moving the hand away from the steering wheel.

The same 3 test variables were considered for this analysis. Therefore, for the blink rate analysis no significant differences were found among the proposed H, V, T and T' interaction methods (Kruskal–Wallis test, p value $= 0.307 > 0.05$).

Then, the activation time data of all four interaction methods were transformed using Log10, in order to validate homogeneity of variances. The results of the Kruskal–Wallis test indicate that there were significant differences between the H, V, T and T' ($H = 122.7$, 3 df., p value $= 0.000 < 0.05$). Then, according to the Mann–Whitney test, there is no significant differences between H and T', but significant differences were found between V and T' (p value $= 8.1\text{E}-16 < 0.05$) as well as T and T' (p value $= 1.5\text{E}-17 < 0.05$). The Table 5 shows means and medians for all H, V, T and T'.

These results suggest that the H and T' methods have lower activation times than the other two interaction methods. Also, there is not enough evidence to confirm that V and T are significantly different from each other. Figure 3, show graphically this behaviour.

The driving acceptance might be related to the activation time analysis, as Kruskal–Wallis test indicates the presence of significant differences between the 4 interaction methods, and Mann–Whitney pairwise comparisons show that H is different from V and T (p value $= 0.0077 < 0.05$ and p value $= 0.009 < 0.05$ respectively for H-V and H-T). On the other hand, there are not significant differences between H and T' interaction methods (p value $= 0.38 > 0.05$). Nevertheless, the analysis does not show significant differences between

Table 5 Mean and median activation time for all the in-vehicle functions and interaction method

Variable	Interaction method	Mean (ms)	Median (ms)
Activation time for all the in-vehicle functions	H	28583.4	24,369.5
	V	48581.3	46,830.0
	T	52698.0	51,617.0
	T'	24683.4	23,948.0

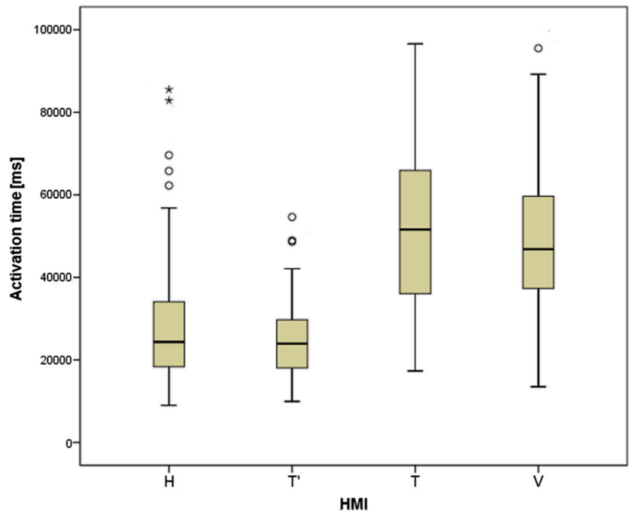


Fig. 3 Boxplot for activation time by interaction method (circles: atypical values, asterisks: extreme values)

the method T’ and V nor T (p value = 0.137 > 0.05 and p value = 0.135 > 0.05 respectively for T’–V and T’–T). This could suggest that not only the activation time is a more accurate indicator for mental workload, but also that these results could be attributed to less familiarity with touchless interaction methods.

4 Discussion

4.1 In-vehicle tasks

It is possible to analyse the above results, not regarding the proposed interaction methods but the in-vehicle tasks, indicating the best suited tasks for these methods; however, due to possible advantages or problems that a HUD visual interface handled by different input types could produce on the UX [17] (mixed input types of HMI may confuse driver), this analysis should be taken into account for future research; making special emphasis on evaluating mental workload of using different input methods for different tasks in the same visual interface. However, according to the above a basic analysis in term of activation time was developed, in order to show, in the most specific and reliable way, the relations between the proposed in-vehicle tasks, and the H, V, T and T’ interaction methods.

Results are consistent along all 5 tasks, as it can be illustrated by Fig. 4, where mean activation times by task are higher for T and V in comparison to H and T’. Furthermore, although there is not significant differences between T’ and H, the former has a lower mean activation time than the latter, in 4 of the 5 tasks.

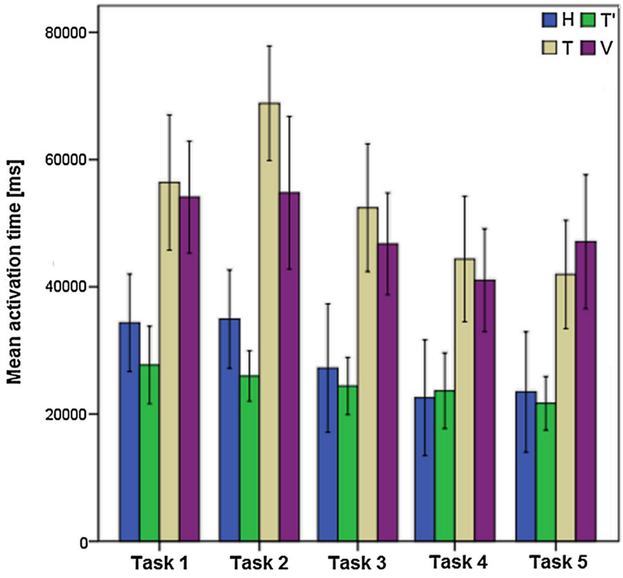


Fig. 4 Mean activation time by interaction method for all tasks (Error bars: 95% confidence level)

4.2 Visual Interface complexity

Recalling the previous subsection, regardless the type of interaction method in terms of activation time, the proposed in-vehicle functions are a significant factor according to Kruskal–Wallis test (p value = 0.00 < 0.05). Consequently, the task 1 produces significant higher activation times than the tasks 4 and 5, according to Mann–Whitney pairwise comparisons. In the same way, task 2 presents significantly higher values than tasks 3, 4 and 5. This could be explained as the first task requires the test subject to navigate from the Main menu through the Music menu while reading the options displayed, and then making decisions for finding and selecting the appropriate options. Conversely, while performing tasks 4 and 5, the test subjects just have to go back one or two levels inside the Music menu for completing the requests; and in this way, they performed these tasks without having to go back to the main menu.

On the other hand, the task 2 has the largest quantity of instructions (for the H, V, T and T’) to be fulfilled, which could explain the higher activation times observed for this task in comparison to task 3, 4 and 5. Furthermore, no significant differences were found between tasks 1 and 2, nor tasks 4 and 5, as shown in Fig. 5.

4.3 Blink rate implications

An additional analysis was performed in terms of blink rates per minute, aiming to identify whether or not significant differences are found when driving with and without the HUD system.

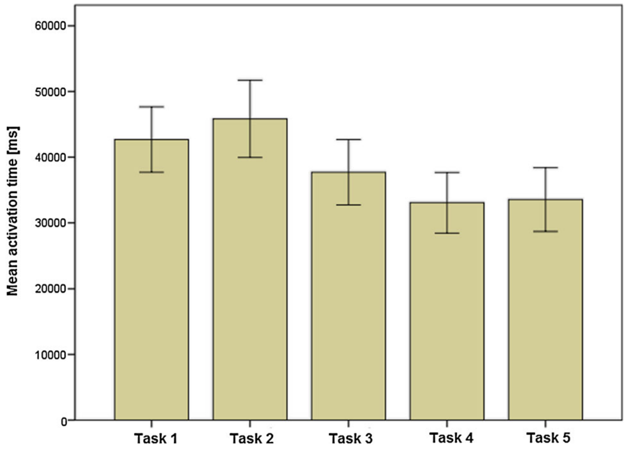


Fig. 5 Mean activation time by in- vehicle task (Error bars: 95% confidence level)

According to the obtained results from the Kruskal–Wallis test, and under the experimental configuration here proposed, there are not significant differences within all groups analysed (i.e. H, V, T, T' and absence of the HUD visual interface). This might be explained by the short number of test subjects selected for this study since other similar studies with above 20 participants report blink rates as a reliable mental workload metric [21]. These results are displayed in Fig. 6.

Therefore, in this case the blink rate analysis must be complemented by other workload indicators, such as activation times related to tasks, or lap times recorded by test subjects while driving, with and without, HUD and HMI configurations.

4.4 Proposed indicator

In order to make a comparison between the different measures studied above (activation times, blinks rate and driver acceptance), obtained results for each of these were classified in 10 groups, where the group 10 would be the biggest in terms of mental workload. In that sense, for example the highest activation time would be assigned to group 10, as it may be related to a greater workload for users. Regarding the blinks rate, the relationships are reported to be inversely proportional, which means lower blink rates are associated to higher workload levels.

According to this, a single indicator was calculated for each interaction method; so, firstly, the mean of the classification groups was computed by method, and then the mean of the three measures was calculated. As explained before, blink rates did not have significant different values from one method to other, which is why it was not taken into account for this indicator. In this sense, the Table 6 show the final indicator for all interaction methods.

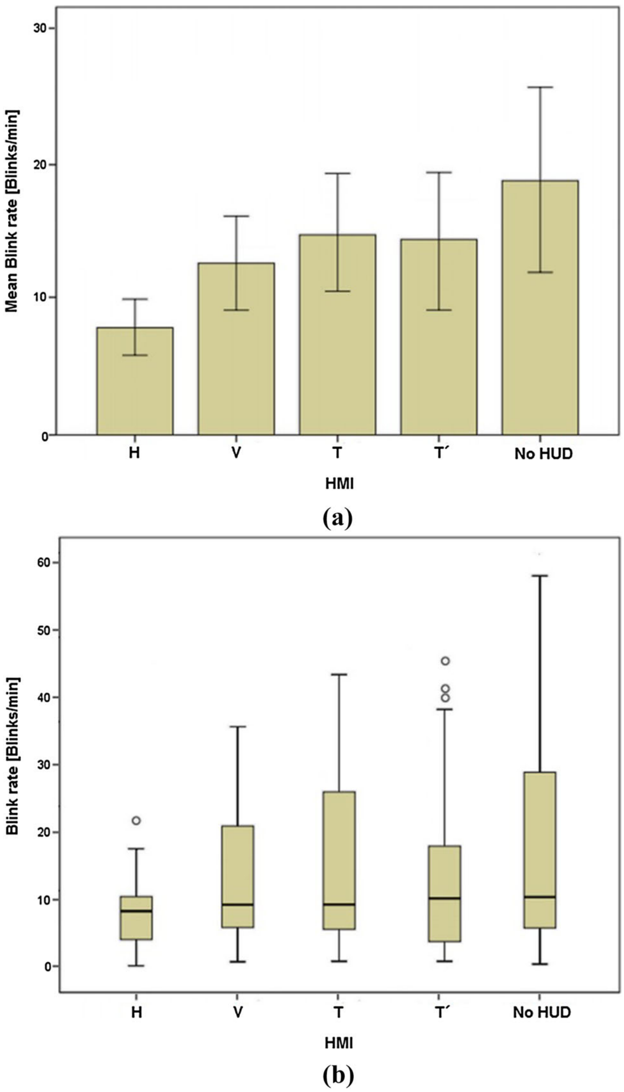


Fig. 6 (Error bars: 95% confidence level). **a** Mean blinks rate by interaction method. **b** Boxplot for all blink rate groups

Table 6 Consolidated indicator values

Variable	Interaction method	Mean rank (NA)
Consolidated indicator	H	2.99
	V	5.12
	T	5.56
	T'	3.31

However, this analysis just explores whether or not the proposed indicator, for the interaction methods here proposed, could be related to the statistical results already obtained; but in order to corroborate the above, a correlation analysis among the measured variables is strongly suggested for future works.

5 Conclusions

For the H, V and T interaction methods, there is a common trend on the results obtained from the activation time and the driver acceptance measures, being H better than V and T; and also, finding that there are not significant differences between the two latter. Therefore, under the experimental configurations here proposed, the results suggest that the activation time and the driver acceptance are similar indicators for estimating the mental workload on HUD systems, although further analysis in terms of correlation between these two metrics is strongly suggested. However, the activation time could seem to be the most accurate indicator in this case. The above is not applicable to the blink rate analysis, which indicate that there is not significant differences among the proposed interaction methods.

Regarding that H has not only a tactile feedback when pressing the steering wheel buttons, but also a visual and auditory stimuli from the HUD visual interface; it is possible that the longer activation times for the in-vehicle tasks while using the V and T methods, respond to the absence of the haptic feedback. In this sense, while using T the test subjects must corroborate the changes in the HUD through visual and auditory stimuli; the above may coincide with the results of Lif et al. [22], who found that a configuration of several stimuli (visual, auditory and tactile) allows to recognize information with greater precision. However, a better user-oriented T' interaction method could overcome the lack of tactile feedback, showing no significant differences with respect to H, as indicated in subsection 3.3.

The above-mentioned conclusions are corroborated by analysing separately the proposed in-vehicle tasks results. Additionally, regarding the Fig. 4, T' has a lower standard deviation (9055.9 ms) than H (16294.9 ms), which could make T' a more desirable option for being implemented in a HUD visual interface. Additionally, in terms of blink rates no significant differences were found between using and not a HUD system while driving under the H, V, T and T' interaction methods.

Nevertheless, it must be considered that the usability results of this study were obtained from a hierarchical and list-based HUD visual interface; and therefore other types of visual interfaces could change the results here obtained. For instance, regarding the proposed HUD visual interface, it seems that showing just one type of options (Music or Air Conditioning) with the less possible levels in depth, could enhance the interaction performance regardless the interaction method.

Finally, as future work more test subjects and in-vehicle functions to develop this experimental procedure are strongly suggested, looking for wider and more reliable conclusions. Moreover, applying the HUD usability results here proposed, to other commercial HUD images is an alternative for validating its advantages. Also, an auditory and tactile feedback exploration must be applied to T', in order to identify whether or not, significant better results than H are possible.

Funding This study has been financially supported by the Pontificia Universidad Javeriana (Bogota, Colombia). Project: "Análisis del comportamiento del conductor cuando se utilizan sistemas automotrices Head-Up Display", ID 07228.

References

1. Bach.: Manual of the Freiburg Vision Test 'FrACT', Version 3.9.8 (2017). http://www.michaelbach.de/fract/media/FrACT3_Manual.pdf. Accessed 10 March 2017
2. Betancur, A.: Physical Variable Analysis Involved in Head-Up Display Systems Applied to Automobiles, Augmented Reality - Some Emerging Application Areas (2011)
3. Betancur, J.A., Villa-Espinal, J., Osorio-Gómez, G., Cuellar, S., Suárez, D.: Research topics and implementation trends on automotive head-up display systems. Int. J. Interact. Des. Manuf. **12**(1), 199–214 (2016)
4. Bi, L., Luo, N., Xin'an F.: A brain-computer interface in the context of a head up display system. In: 2012 ICME International Conference on Complex Medical Engineering, pp. 241–244, Kobe Japan, Oral Presentation (2012)
5. Biever, W.J.: Auditory-based supplemental information processing demand effects on driving performance. Master's Thesis. Virginia Tech (2002)
6. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. **44**, 58–75 (2014)
7. Campbell, F.W.: The depth of field of the human eye. Ootica Acro **4**, 157–164 (1957)
8. Charissis, V., Papanastasiou, S., Vlachos, G.: Comparative study of prototype automotive HUD vs. HDD: collision avoidance simulation and results. In: Proceedings of the Society of Automotive Engineers World Congress (2008)
9. Doshi, A., Cheng, S.Y., Trivedi, M.M.: A novel, active heads-up display for driver assistance. IEEE Trans. Syst. Man Cybern. Part B **39**, 85–93 (2009)
10. Eggemeier, F.T., Biers, D.W., Wickens, C.D., Andre, A.D., Vreuls, D., Billman, E.R., Schueren, J.: Performance assessment and workload evaluation systems: analysis of candidate measures. Technical Report No. HSD-TR-90-023. Brooks Air Force Base, Armstrong Aerospace Medical Research Laboratory, TX (1990)
11. Faure, V., Lobjois, R., Benguigui, N.: The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. Transp. Res. Part F: Traffic Psychol. Behav. **40**, 78–90 (2016)
12. Gish, K., Staplin, L.: Human Factors Aspects of Using Head UP Displays in Automobiles: A Review of the Literature. Office of Crash Avoidance Research National Highway Traffic Safety Administration US Department of Transportation Washington, DC (1995)
13. Götze, M., Bißbort, F., Petermann-Stock, I., Bengler, K.: A careful driver is one who looks in both directions when he passes a red light-increased demands in urban traffic. In: International Conference on Human Interface and the Management of Information, pp. 229–240 (2014)
14. Hart, S.G., Staveland, L.E.: Development of a multi-dimensional workload rating scale: Results of empirical and theoretical

research. In: Hancock, P.A., Meshkati, N. (eds) Human Mental Workload, pp. 139–183 (1988)

15. Horrey, W.J., Wickens, C.D.: Driving and side task performance: the effects of display clutter, separation, and modality. Hum. Fact. J. Hum. Fact. Ergon. Soc. **46**(4), 611–624 (2004)
16. Hosseini, A., Bacara, D., Lienkamp, M.: A system design for automotive augmented reality using stereo night vision. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings, pp. 127–133. IEEE (2014)
17. Jakus, G., Dicke, C., Sodnik, J.: A user study of auditory, head-up and multi-modal displays in vehicles. Appl. Ergon. (Part A) **46**, 184–192 (2004). https://doi.org/10.1016/j.apergo.2014.08.008
18. Kee, E.: BMW heads-up display now in technicolor. http://www.ubergizmo.com (2011). Accessed 15 June
19. Kosaka, H., Uematsu, K., Kurosaki, A., Noda, M. et al.: Evaluation of a new in-vehicle HMI system composed of steering wheel switch and head-up display when a driver encounters sudden danger. SAE Technical Paper 2007-01-0439 (2007)
20. Kramer, A.F.: Physiological metrics of mental workload: a review of recent progress. In: Damos, D.L. (ed.) Multiple Task Performance. Taylor & Francis, London (1991)
21. Li, R., Chen, Y.V., Sha, C., Lu, Z.: Effects of interface layout on the usability of in-vehicle information systems and driving safety. Displays **49**, 124–132 (2017). https://doi.org/10.1016/j.displa.2017.07.008
22. Lif, P., Oskarsson, P.A., Lindahl, B., Hedström, J., Svensson, J.: Multimodal threat cueing in simulated combat vehicle with tactile information switching between threat and waypoint indication. In: Symposium on Human Interface, pp. 454–461 (2011)
23. Liu, Y.-C., Wen, M.-H.: Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. Int. J. Hum Comput Stud. **61**(5), 679–697 (2004)
24. Merenda, C., Smith, M., Gabbard, J., Burnett, G., Large, D.: Effects of real-world backgrounds on user interface color naming and matching in automotive AR HUDs. Paper presented at the 2016 IEEE VR 2016 Workshop on Perceptual and Cognitive Issues in AR (2016)
25. Niu, T., Hsieh, Y., Liao, S., Kuo, H.: A head-up display illuminator design and virtual-image estimation method. In: Digest of Technical Papers—SID International Symposium, vol. 45, no. 1 (2014)
26. O'Donnel, R.D., Eggemeier, F.T.: Workload assessment methodology. In: Kaufman, B.K., Wiley, T.J. (eds.) Handbook of Perception and Human Performance, vol. II, pp. 41–42. Cognitive Processes and Performance, New York (1986)
27. Peters, M., Ivanoff, J.: Performance asymmetries in computer mouse control of right-handers, and left handers with left- and right-handed mouse experience. J. Mot. Behav. **31**(1), 86–94 (1999)
28. Renault Mégane.: Head-up Display—Interior Front Seats Wallpaper (2016). http://www.caricos.com/cars/r/renault/2016_renault_megane/images/31.html. Accessed June 15
29. Simon, L., Tarel, J.P., Brémond, R. (2009) Alerting the drivers about road signs with poor visual saliency. In: Intelligent Vehicles Symposium (2009)
30. Soomro, S.R., Urey, H.: Design, fabrication and characterization of transparent retro-reflective screen. Opt. Express **24**(21), 24232–24241 (2016)
31. Stern, J.A., Boyer, D., Schroeder, D.: Blink rate: a possible measure of fatigue. Hum. Factors **36**(2), 285–297 (1994)
32. Tretten, P., Gårling, A., Nilsson, R., Larsson, T.C.: An on-road study on head-up display preferred location and acceptance levels. In: Proceeding of the Human Factors and Ergonomics Society, vol. 55, no. 1, pp. 1914–1918 (2011)
33. TTSreader.: Text-to-speech reader Demo. https://ttsreader.com/ (2017). Accessed 15 June
34. Wilson, G.F., Fisher, F.: The use of cardiac and eyeblink measures to determine flight segments in F4 crews. Aviat. Space Environ. Med. **62**, 959–962 (1991)