

Article

Gaze-Head Input: Examining Potential Interaction with Immediate Experience Sampling in an Autonomous Vehicle

Aya Ataya, Won Kim, Ahmed Elsharkawy and SeungJun Kim * 

Gwangju Institute of Science and Technology, School of Integrated Technology, Buk-gu, Gwangju 61005, Korea; aya.ataya@gist.ac.kr (A.A.); kimwon30@gm.gist.ac.kr (W.K.); elsharkawy@gist.ac.kr (A.E.)

* Correspondence: seungjun@gist.ac.kr

Received: 1 December 2020; Accepted: 15 December 2020; Published: 17 December 2020



Abstract: Autonomous vehicles (AV) increasingly allow drivers to engage in secondary tasks such as eating or working on a laptop and thus require easy and reliable interaction inputs to facilitate communication between the driver and the vehicle. However, drivers report feeling less in control when driving is no longer the primary task, which suggests that novel approaches for assessing satisfaction regarding AV decision-making are needed. Therefore, we propose an immediate experience sampling method (IESM) that learns driver preferences for AV actions. We also suggest gaze-head input (G-HI) as a novel input in an AV. G-HI provides a hands-free, remote, and intuitive input modality that allows drivers to interact with the AV while continuing to engage in non-driving related tasks. We compare G-HI with voice and touch inputs via IESM for two simulated driving scenarios. Our results report the differences among the three inputs in terms of system usability, reaction time, and perceived workload. It also reveals that G-HI is a promising candidate for AV input interaction, which could replace voice or touch inputs where those inputs could not be utilized. Variation in driver satisfaction and expectations for AV actions confirms the effectiveness of using IESM to increase drivers' sense of control.

Keywords: gaze-head input interaction; experience sampling method; automated driving; non-driving related task

1. Introduction

As automated driving technology develops, as defined by the Society of Automotive Engineers (SAE level 3 and above [1]), drivers will increasingly relinquish the primary driving-related task to the autonomous vehicle (AV) [2,3]. This paradigm shift will allow drivers to engage more freely in non-driving related tasks (NDRTs) such as working on a laptop, eating/drinking, chatting on a smartphone, social networking, playing games, tending to a child, relaxing, or sleeping [2,4–8]. These NDRTs may be accompanied by a change in position inside the vehicle, for example, reclining the seat to use a laptop would position the driver farther from the dashboard than during normal driving [9] which may influence the driver's user experience (UX). Of course, convenient and supportive vehicle operation during NDRTs requires that the driver and AV be able to communicate effectively. Driver–vehicle interaction can be initiated by either the driver or vehicle: drivers may need to control the vehicle's settings (e.g., turn down the radio) or assume control (e.g., change lanes or stop) [10–12], whereas the AV may need to ask the driver to make a decision (e.g., about an upcoming road hazard) instead of asking the driver to assume manual control [13,14].

Typically, drivers interact with AVs via voice input (VI) or a touch panel, standard input interactions in today's vehicles [15]. However, when drivers are engaged in NDRTs, usually one or both hands

may be occupied [3] or they may have reclined the seat to have more space [9], which could impede interactions, particularly via the touch panel interface. Considering what drivers need to comfortably perform NDRTs requires exploring a variety of different input interactions as a critical element in AV design [10,12,16]. Consequently, even though drivers are familiar with touch and voice controls in today's vehicles, driver-vehicle interactions will become increasingly complex as AV capabilities mature. This suggests that a variety of input interactions are needed to identify the input interaction that best supports NDRTs in AVs.

Previous studies have shown that drivers often feel out of control and have low acceptance of this during automated driving [17,18]. For example, AVs that do not support both single-lane and multiple-lane crossovers may not be accepted by drivers who want to cross multiple lanes of traffic at once. Consequently, a user-centered orientation and an understanding of the factors that influence a driver's satisfaction and expectation about AV actions in many scenarios will increase driver acceptance for AV performance.

One way to understand driver satisfaction about vehicle actions is the experience sampling method (ESM), a reliable measure of subjective experience, where feedback on vehicle actions is collected during or after the trip [19]. However, implementing ESM in non AVs requires special precautions regarding safety, and data should be collected only when the vehicle is stopped [20]. This lag time between the vehicle's action and the driver's evaluation can diminish data quality when more than one action needs to be evaluated [21]. However, AV drivers could respond to ESM questions immediately following the vehicle's action of interest, without risks to safety. Therefore, we propose an immediate experience sampling method (IESM) to collect drivers' satisfaction and expectations about AV actions in two simulated driving scenarios. In addition to more precise data collection, IESM could lead to a quicker and automated learning cycle for the AV to understand the driver's preferences. Implementing IESM in a novel context requires intuitive interaction channels to help drivers convey their answers smoothly, even while engaged in NDRTs.

In this study, we employ the IESM via an in situ survey, which allows drivers to report more accurate UX close to ground truth in two simulated driving scenarios. We suggest gaze-head input interaction (G-HI) in AVs as a remote, hands-free, and intuitive input interaction and compare its performance with touch and voice input interactions. These three input interactions could be used by drivers while performing an NDRT, as IESM compels participants to interact with the AV, thus evaluating vehicle actions using one input interaction at a time. Therefore, we tested the performance of three input interactions while collecting UX using IESM. Our findings will contribute to the interaction design of future AVs to better support NDRTs. Also, implementing IESM, as a novel context in AVs, may present a crowdsourcing for collecting UX that can contribute to the AV's motion planning and decision-making.

The goal of this study is to compare gaze-head input interaction as a novel remote input in AVs with voice and touch input (TI) to assess its viability. Touch input was selected as a baseline of all interactions because it has been the most common input channel in cars for decades [15], whereas voice input is the fastest growing remote interaction modality in AV studies [12,14]. This comparison considers the participants performing an NDRT (using a smartphone to read news article) and then asks them to evaluate the vehicle's actions in two different scenarios via IESM. Drivers' perceived workload, system usability, and reaction time are measured to assess the performance of using the three input interactions. Moreover, IESM is qualitatively evaluated by collecting participants' satisfaction and expectation levels for the AV actions in the two simulated driving scenarios.

This study has been designed to answer the following research questions:

- How does performance of G-HI compare to touch and voice input while drivers engage in an NDRT in AV?
- How does IESM enhance the driver's perception of the AV's general performance, enhance the accuracy of the collected data, automate data collection, and accelerate the AV learning cycle?

2. Related Works

2.1. Input Interactions

Previous AV studies have investigated the use of touch panels for drivers to select a suggested proposition (e.g., pass the car ahead or hand over driving control) when the automated system has reached its limit [14,22], or to select AV maneuvers (e.g., change lanes) via touch input implemented on the steering wheel [23]. However, these approaches require the driver's hands to be free to interact with the AV—a less-than-ideal condition when drivers engage in NDRTs—and close enough to the steering wheel to reach the touch panel. Some participants also disliked interacting via touch input which required they divert their gaze from the road to the touch interface while interacting [14].

In addition to touch interaction, hand gesture has been applied for maneuver-based intervention [12] and to control in-car lateral and longitudinal AV motions [24]. While hand-gesture has shown good usability as a remote input interaction, problems with overlapping commands, rates of misrecognition, and arm fatigue diminish its usefulness, and drivers struggle to remember each gesture [12,24]. Furthermore, hand gesture input may require the driver to stop performing an NDRT to interact with the AV.

Voice and gaze inputs can supplement the drawbacks of touch or hand gesture inputs by supporting remote and hands-free input interactions. Voice input control, one of the most common input interactions in AVs, has been implemented to support driver-vehicle cooperation and to select vehicle maneuvers [12,14]. In previous studies, drivers preferred voice over touch input and hand gestures, as voice commands do not require physical movement and allows them to keep their eyes on the road during hands-free interaction. However, voice input may not work well in a noisy environment (e.g., a conversation with other people in the vehicle), and drivers do not completely trust speech recognition or may be confused about the appropriate command to initiate a desired action [12,14].

Gaze input interaction has been shown to be faster and more precise than hand gestures [25], and tracking eye movement has been shown to accurately indicate user attention [26,27]. However, to confirm or complete a selection via gaze, users must dwell for a few seconds on an object [28,29]. This increases cognitive load and causes eye fatigue [30], obvious problems during driving and limitations for gaze input as a single modality.

To eliminate the drawbacks of using gaze input as a single input modality, a gaze combined multimodal interaction has been suggested [26], with gaze determining the object and the combined modality to trigger the selection. This study [15], showed that gaze-touch and gaze-gesture [31] inputs can be used to select objects on a head-up display (HUD) while interacting with in-vehicle infotainment systems. Similarly, researchers in [32] presented a gaze-voice input, where gaze was used to activate the voice input. However, these combined inputs still suffer from the limitations of the voice, touch and gesture single modality inputs mentioned above.

Combining gaze with head movement may address most of the aforementioned limitations of a single modality [28,33,34]. In gaze-head input interaction, gaze is used to select an object and a head nod is used to trigger the selection [28,35], where head movement can be defined by orientation angles of the head (roll, pitch, and yaw) [36]. G-HI is natural and intuitive, as it reflects human-to-human interaction and is viable using state-of-the-art eye trackers [25]. In previous studies in augmented reality [37], virtual reality [25,38], or even collaborative gaming [39], G-HI has achieved good usability and high accuracy, as it supports hands-free input and does not require users to reposition themselves with high accuracy [40]. It has also been shown to be a fast, natural, convenient and accurate way to interact with mobile screens and desktop computing [35,41]. Based on these findings, we believe that using this interaction modality combination with AVs via the HUD might show good performance. Unfortunately, due to safety issues and limited demand, G-HI has not been discussed in automotive vehicle research to date. However, it could be considered as an input interaction for the AV interactive environment.

2.2. User Experience Data

ESM has been used in some non-autonomous vehicle studies to collect UX data after the trip [20] and emotional levels during traffic congestion [42]. These studies have relied on paper post-hoc questionnaires or smartphones, which requires participants to stop the vehicle and thus interrupt or otherwise distract them from the primary task; either option could pose safety risks [19,20]. While in AV, researchers collected UX data about AV actions in many scenarios via post-hoc surveys [43]. However, to best learn about drivers' acceptance of AV actions, we employ IESM via an in situ survey immediately following the AV action. IESM in AV allows researchers to query drivers at any moment during the driving scenario without compromising safety to identify their experiences, i.e., preferences or needs. Therefore, our IESM survey was triggered by a real-time questionnaire to reduce biases and reliance on participants' ability to recall earlier experiences accurately [44].

3. Experiment Design

3.1. Procedure

We evaluated three input interactions in four sessions (two scenarios with two AV actions) and one NDRT for (N = 32) participants within-subject in a driving simulator. Our goal was to evaluate the use of gaze-head input compared to two input interactions (touch and voice inputs) while drivers perform an NDRT in the simulated AV as shown in Figure 1a. Participants range in age from 26–35 years (*Age: M* = 29.1; *SD* = 3.96; *Gender: Male* = 19, *Female* = 13) and all had a driver's license, which guaranteed that they were aware of our tested driving scenarios. Participants were recruited from our university and have a similar education level. They were recruited with a counterbalancing design (3×4) between the three input interactions and the four automated driving actions considering the learning effects. They were informed about the experiment, completed a consent form, then answered preliminary questions regarding personal information, including to self-report mobility problems related to the experiment design; no participants reported problems in moving their eyes, nodding head, talking, or touching. Calibration for eye-tracking was performed prior to a practice session to familiarize them with all three input interactions. After that, the main experiment was conducted and interim-questionnaires were completed after each session. After completing the final interim-questionnaire of the last session, a post-questionnaire was provided to collect participant's preferences of input interactions and how their preferences effected by the engagement of the NDRT combined with open-ended opinions. Participants were compensated with approximately \$10 after completing the entire experiment.

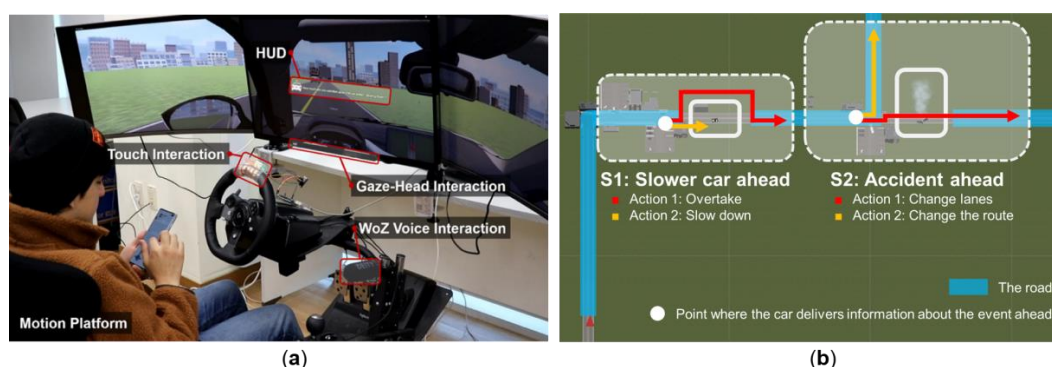


Figure 1. Experiment setup (a) Simulation environment (motion platform, full windshield head-up display (HUD); gaze-head input interaction (eye-tracker); touch input interaction (touch panel); voice input interaction (microphone); participant reading from a smartphone (non-driving related task) while sitting in the autonomous vehicle (AV) simulator motion platform); (b) destination and 2D user interface map implemented in Unity that shows two driving scenarios and four AV actions (two actions for each scenario).

3.2. Driving Scenarios

The AV first drove on an urban road at 30 km/h for approximately 1 min, as shown in Figure 2a, while participants were instructed to read a CNN news article of their choice with their own smartphones to moderately engage their attention in an NDRT [45] as illustrated in Figure 1a. Task design was based on a previous study that identified the most desired activities in an AV is using a personal digital assistant [2,7]. We allowed participants to choose their own news article to read, as we assumed they would choose the article they found most interesting and thus engaging. Our experiment scenarios were designed following previous studies that focused on scenarios where drivers assist the AV in decision-making [13,14]. Participants were asked to rate their preferences about each AV actions in two scenarios, Slower car ahead and Accident ahead, via IESM displayed on the HUD where it was introduced as the most preferred display in AVs and provides a higher situation awareness [46], Figure 2d. In the Slower car ahead scenario, the AV would either (a) slow down or (b) overtake the car ahead. In the Accident ahead scenario, the AV would either (a) change lanes or (b) change the route, as shown in Figure 1b. In each scenario, the AV notified drivers about which action it was performing and why as shown in Figure 2b,c.

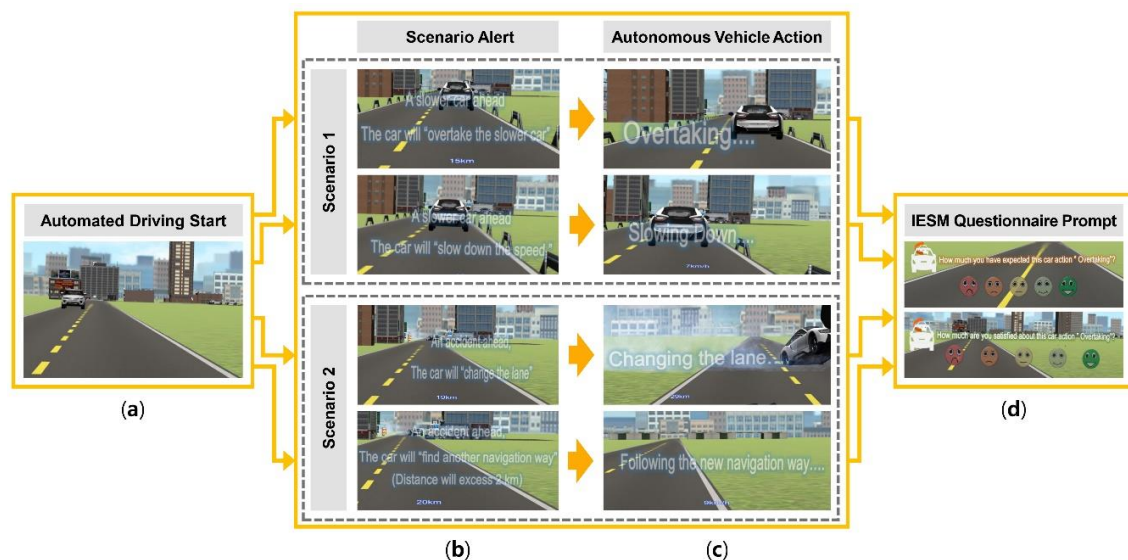


Figure 2. Illustration of our experiment scenarios sequence. (a) The main road view where the AV starts moving; (b) scenario alert (notification message shown on HUD to explain the current scenario to the participant before it starts performing the action); (c) AV scenario action (a notification message shown on HUD while the action occurs); (d) two immediate experience sampling questions were shown on HUD immediately after the completion of each action.

To transfer the driver's attention away from a smartphone and toward the studied scenario, a notification sound was generated, followed by the appearance of a HUD message to illustrate information about the current scenario and the AV's actions, as in Figure 2b,c, and then the IESM questionnaire was conducted after the AV completed the action as in Figure 2d. For example, after the notification, the AV would first display the message "A slower car ahead" to explain the driving situation, followed by the message "Overtake the car ahead" to notify drivers about its action. After the messages were displayed, the IESM questionnaire was shown. Subjects were required to respond to the IESM after each AV action using the randomly assigned input interaction considering the counterbalancing. The detailed experiment flow is illustrated in Figure 3.

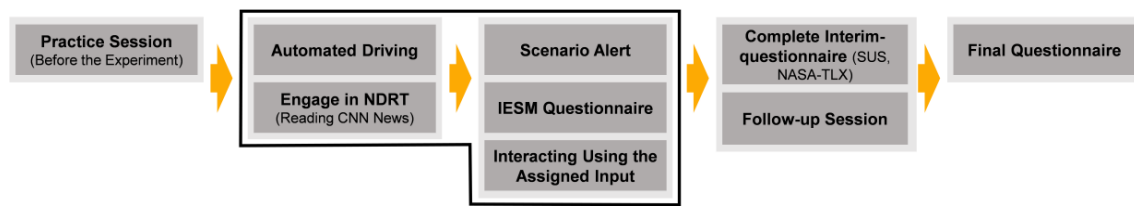


Figure 3. Experiment flowchart.

3.3. Apparatus

3.3.1. Touch Input (TI)

We developed our own TI interface with a 5-strips capacity touch panel placed on the steering wheel following the implementation in [15,47]. TI was implemented using Arduino Uno serial communication connected to a host PC; the IESM interface was implemented on the Unity platform. To promote friendly and intuitive interaction between the driver and the IESM, the TI interface implemented a 5-point Likert scale that uses face emoticons that correspond to statements describing positive and negative situations (1 = very sad face, comparable to “Strongly Disagree”; 5 = very happy face, comparable to “Strongly Agree”), which were explained to participants before the experiment, as shown in Figure 4a.



Figure 4. (a) Touch input interaction implemented on the steering wheel with 5-Likert scale presents using five emoticon faces that describing positive and negative situations; (b) A Wizard of Oz (WoZ) voice input was presented using microphone attached on the simulation platform to mimic the standard voice input for the participants; (c) a Tobii 4C eye tracker was implemented on the center monitor for gaze-head input interaction, drivers gaze at an icon of interest, then nodding by (16°) to confirm the selection.

3.3.2. Voice Input (VI)

As current voice input technology is under development due to unfavourable recognition rates compared to other input interactions [48], a Wizard of Oz (WoZ) approach is applied to fully support perfect recognition in parallel to this study [48]. To best perform the role of WoZ with no mistakes, a microphone was implemented on the right side to simulate the voice input and all choices were interconnected with keyboard shortcuts and the participant’s selection was performed by the experimenter immediately after the participant uttered a choice, as illustrated in Figure 4b.

3.3.3. Gaze-Head Input (G-HI)

G-HI was implemented with Tobii Eye Tracker 4C using Software Development Kit (SDK) to detect the driver’s gaze movement and head position. G-HI is defined as using eye gaze as a cursor to select an object of interest then nodding the head forward approximately less than 16 degrees which requires a small movement to confirm the selection, following this study [36]. By tracking the head position, the degree of nodding can be calculated and thus used to select the answer, as shown in Figure 4c.

3.3.4. Driving Simulation

Our experiment was conducted in an indoor driving simulator setup (Atomic A3 motion platform, Logitech G920 steering wheel and pedal), a high-fidelity driving simulator that transfers the motion force of the virtual vehicle to the driver's seat. The simulation was created with Unity, and the interface was connected with the driving simulation after we implemented the input device for each input interaction. The driving scene was displayed on three side-by-side 32" monitors (5760×1080). Three input interaction devices were attached to our test bed to collect IESM data and test our input interactions: an eye-tracker attached to the bottom of the center monitor for G-HI, touch panels on the steering wheel for TI, and a microphone on the right side of the driver's seat for VI, Figure 1a.

3.3.5. Immediate Experience Sampling Method (IESM)

In our study we used an IESM questionnaire which appeared on the HUD immediately upon completion of each AV action. The questionnaire asked about the subject's expectations about the vehicle's action (IESM1) (e.g., "How much did you expect the car to slow down?") and subject's Satisfaction about the vehicle's action (IESM2) (e.g., "How satisfied are you with the vehicle's decision to change the route?"). Expectations followed from how subjects typically perform this action in the same scenario in their manual driving, where satisfaction reflected a subject's satisfaction level about the AV's action in each scenario. These two questions told us whether driver satisfaction about AV actions was affected by daily driving experience. Also, the benefit of using a HUD is that the drivers' eyes will be on the windshield and ready to interact and report their experiences about the AV's actions immediately after the vehicle provides the notification sound. Subjects responded to IESM questions as shown in Figure 2d using three different input interactions (G-HI, TI, and VI), one input interaction at a time, while engaged in an NDRT.

4. Results

For each session, we collected recorded video and interview responses and analyzed data on workload, subjective preference, task performance. All results were subjected to paired sample *t*-test and one-way repeated measures analysis of variance (ANOVA) tests for post-hoc tests at a 5% confidence level. Equivalence tests were performed to analyze pairwise comparisons (Bonferroni corrected or Games–Howell corrected as post-hoc tests after checking the homogeneity of variances).

4.1. Input Interactions

4.1.1. Reaction Time

Participants' reaction time (the interval between the appearance of the IESM questions to participants' selection) was calculated using the stored live video to better understand the performance of each input interaction, as faster input helps to keep the driver involved in NDRT.

We measured the average reaction time for each input across all scenarios and two IESM. The one-way repeated measure ANOVA test showed a statistical significance, $F(2, 62) = 3.81, p < 0.027$. As Figure 5 shows, G-HI had the fastest reaction time ($M = 3.76, SD = 1.10$), followed by TI ($M = 4.14, SD = 1.23$) and VI ($M = 4.60, SD = 1.54$). Reaction time for VI was about 1 s greater than the G-HI $p < 0.029$.

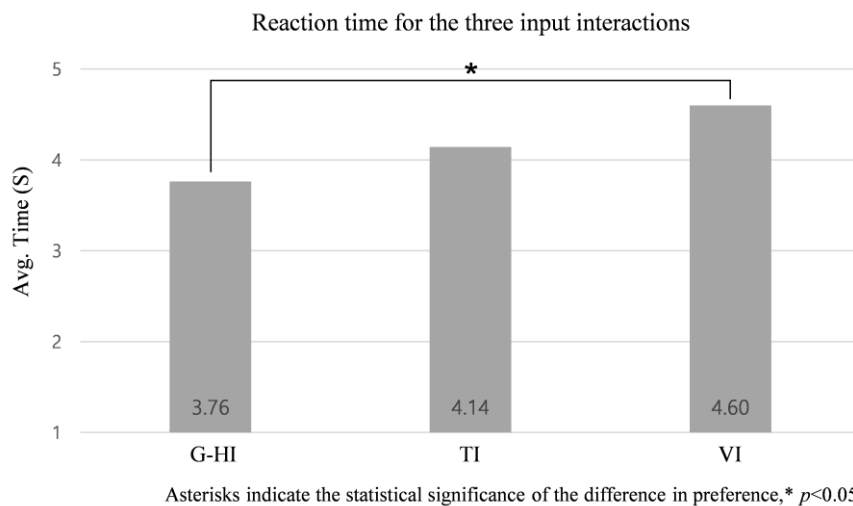


Figure 5. Average of reaction time for each input interaction (gaze-head input, touch input, voice input).

To understand which scenario most influenced reaction time, we analysed the reaction time in each scenario for each input. The ANOVA test shows a significant difference in the changing lanes action $F(2, 62) = 7.39, p < 0.001$ but no significant differences in the other actions as shown in Figure 6. When changing lanes, reaction time using VI ($M = 4.54, SD = 1.47$) was about 1.4 s greater than G-HI ($M = 3.19, SD = 1.43, p < 0.002$), where reaction time in G-HI was 1.16 s faster than TI ($M = 4.35, SD = 1.64, p < 0.01$).

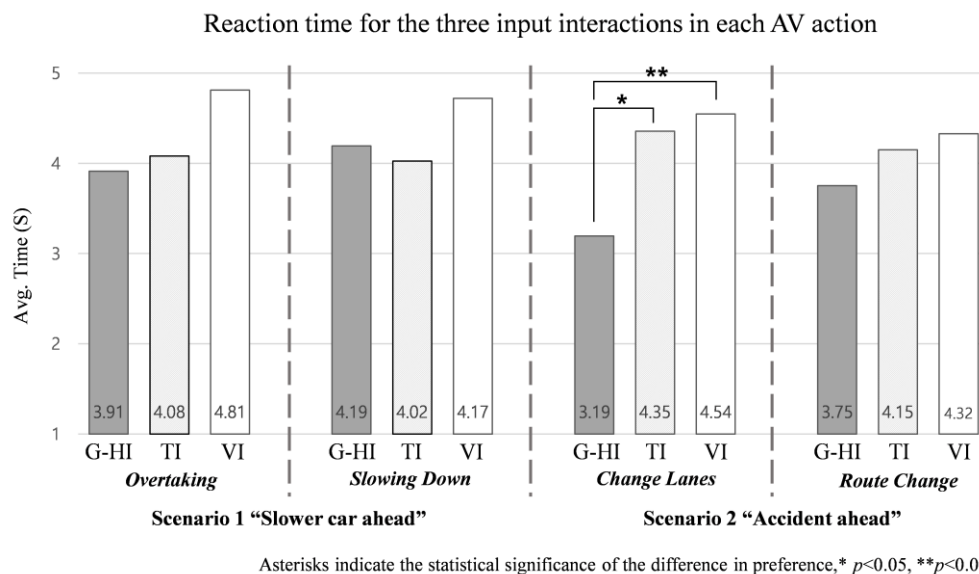


Figure 6. Reaction times for three input interactions considering each AV action in two driving scenarios.

4.1.2. System Usability

We identified and compared participants' subjective usability, on how easily and efficiently they used each input interaction through the system usability scale (SUS), evaluated on a scale of 0 to 100 in increments of 10. SUS scores of more than 71 points indicate that the system is acceptable with five steps anchored with "Strongly Disagree" and "Strongly Agree" [49].

The one-way repeated measures ANOVA tests shows significant differences $F(2, 62) = 5.69, p < 0.005$ in SUS. TI scored highest ($M = 75.31, SD = 16.97$), followed by VI ($M = 71.87, SD = 15.83$) and G-HI ($M = 63.35, SD = 22.12$), as shown in Figure 7. The SUS score of TI was significantly higher than in G-HI.

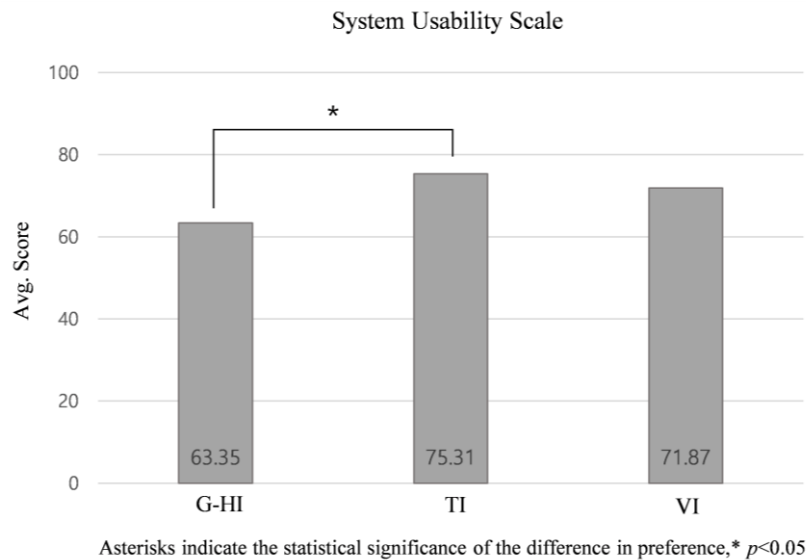


Figure 7. Average system usability scale (SUS) scores for the three input interactions.

To determine the most influential item in the SUS result, we analysed the 10 increment items of SUS for each input (Q1. I would like to use this system frequently, Q2. This system unnecessarily complex, Q3. This system was easy to use, Q4. I would need the support of a technical person to be able to use this system, Q5. The various functions in this system were well integrated, Q6. There was too much inconsistency in this system, Q7. Most people would learn to use this system very quickly, Q8. This system was very cumbersome to use, Q9. I was very confident using the system, Q10. I needed to learn a lot of things before I could get going with this system.), as shown in Figure 8. The ANOVA test found a significant difference in the SUS item Q4 (“I think that I would need the support of a technical person to be able to use this system”), $F(1.62, 50.34) = 6.32, p < 0.006$, where TI ($M = 80.46, SD = 26.74$) scored significantly higher than G-HI ($M = 65.6, SD = 29.61, p < 0.004$) and VI ($M = 72.65, SD = 27.20, p < 0.04$). Also, SUS item Q6 (“I thought there was too much inconsistency in this system”) showed a significant difference, where $F(2, 62) = 9.07, p < 0.000$, where G-HI ($M = 51.56, SD = 31.7$) scored significantly lower than TI ($M = 75.0, SD = 21.99, p < 0.001$) and VI ($M = 69.53, SD = 25.97, p < 0.013$). The least influential item in SUS is item Q9 (“I felt very confident using the system”), $F(1.89, 58.7) = 6.39, p < 0.004$, where TI ($M = 78.1, SD = 23.54$) was highly confident compared to G-HI ($M = 57.03, SD = 33.1, p < 0.010$), as shown in Table 1.

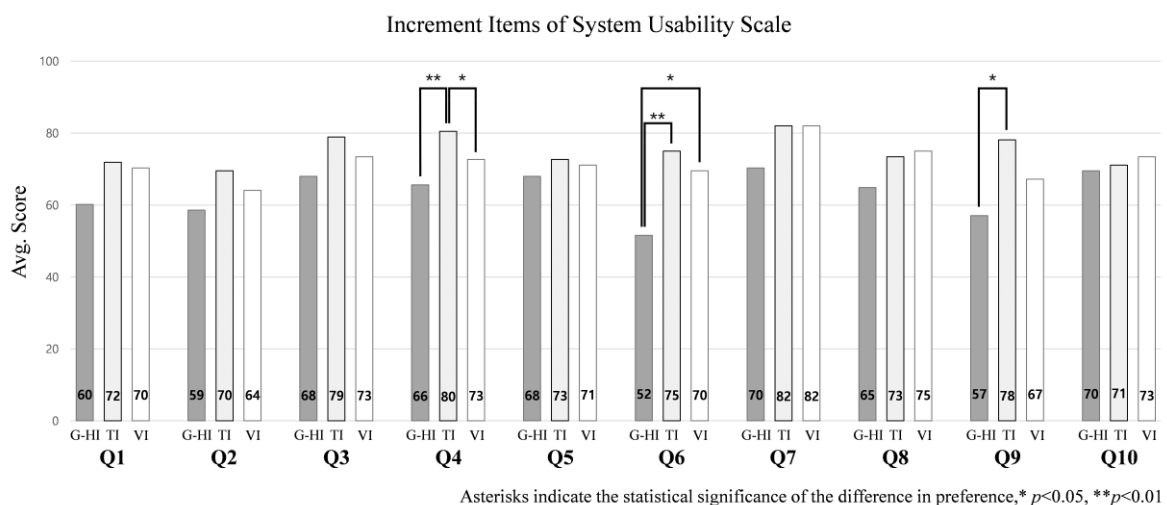


Figure 8. Ten increment items of system usability scale scores for the three input interactions.

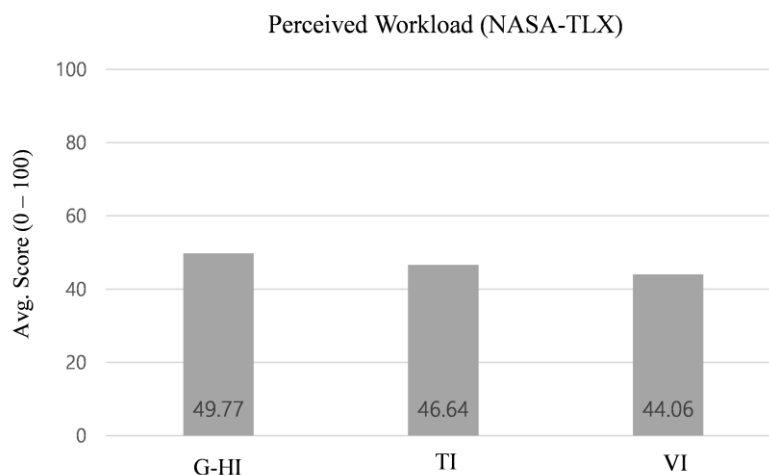
Table 1. The mean (M) and standard deviation (SD) values for the ten increment items of system usability scale scores for the three input interactions.

Input Interaction		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Gaze-Head	Mean	60	59	68	66	68	52	70	65	57	70
	SD	(31.7)	(30.8)	(31.3)	(29.6)	(23.1)	(31.7)	(28)	(30.4)	(33.1)	(27.5)
Touch	Mean	72	70	79	80	73	75	82	73	78	71
	SD	(22.7)	(30.9)	(23.8)	(26.7)	(20.4)	(22)	(19.3)	(26.9)	(23.5)	(33.7)
Voice	Mean	70	64	73	73	71	70	82	75	67	73
	SD	(28)	(29.1)	(24.5)	(27.2)	(21.2)	(26)	(20.3)	(22)	(28.7)	(28.3)

4.1.3. Perceived Workload

Subjective perceived workload was measured by the National Aeronautics and Space Administration Task Load Index (NASA-TLX), which measures the level of six dimensions (mental demand, physical demand, temporal demand, effort, performance, and frustration) to determine an overall workload rating. Each dimension is evaluated on a scale of 0 to 100 [50].

The ANOVA test shows no significant differences in overall workload rating in the NASA-TLX, $F(2, 62) = 1.17$, $p < 0.315$ between the three input interactions, G-HI ($M = 49.77$, $SD = 17.83$), TI ($M = 46.64$, $SD = 16.47$), and VI ($M = 44.06$, $SD = 13.71$), as shown in Figure 9.

**Figure 9.** National Aeronautics and Space Administration Task Load Index (NASA-TLX) scores for the three input interactions.

We also analysed the six dimensions of the NASA-TLX to determine the most influenced NASA-TLX factor of workload rating for each input. As shown in Figure 10, The ANOVA test shows a significant difference in mental demand $F(1.57, 48.96) = 5.48$, $p < 0.012$, with significant differences shown between G-HI ($M = 10.68$, $SD = 7.83$) and TI ($M = 6.47$, $SD = 6.16$, $p < 0.02$), and no significant differences between VI ($M = 7.43$, $SD = 5.91$) and TI or G-HI. Also our results show a significant difference in frustration, $F(2, 62) = 4.29$, $p < 0.018$. VI ($M = 5.89$, $SD = 8.23$) shows higher frustration than TI ($M = 2.95$, $SD = 4.88$, $p < 0.03$) and also G-HI ($M = 7.89$, $SD = 10.51$, $p < 0.05$) shows higher frustration than TI. On the other hand, the ANOVA test also showed statistical significance in physical demand $F(1.89, 58.74) = 3.98$, $p < 0.02$. where G-HI ($M = 5.208$, $SD = 5.97$) causes a lower physical load than TI ($M = 9.10$, $SD = 8.47$, $p < 0.03$) and VI ($M = 5.207$, $SD = 5.94$, $p < 0.028$) shows a lower workload than TI, as illustrated in Table 2.

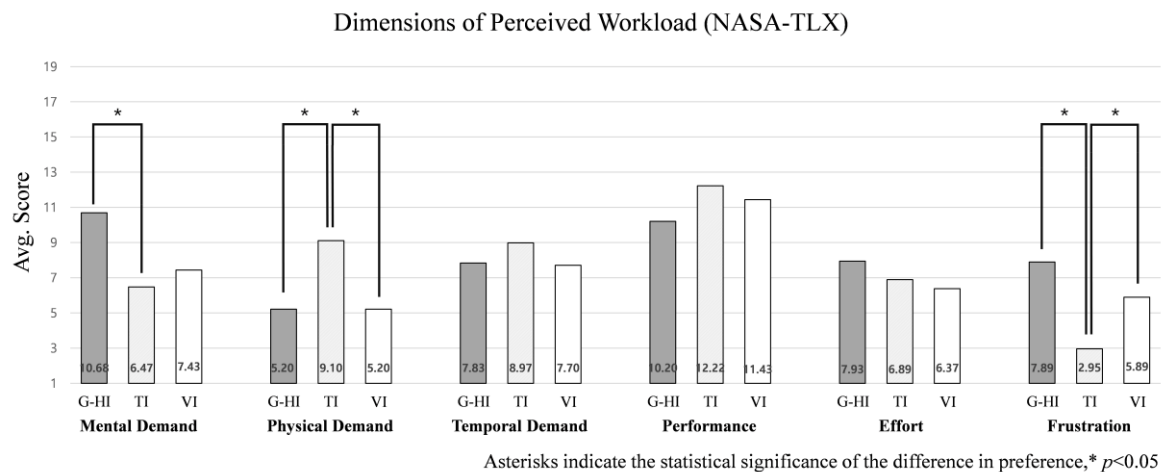


Figure 10. The six dimensions of NASA-TLX for the three input interactions.

Table 2. The mean (M) and standard deviation (SD) values for the six dimensions of NASA-TLX for the three input interactions.

Input Interaction		Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
Gaze-Head	Mean	10.68	5.2	7.83	10.2	7.93	7.89
	SD	(7.83)	(5.98)	(6.46)	(9.85)	(6.59)	(10.51)
Touch	Mean	6.47	9.1	8.97	12.22	6.89	2.95
	SD	(6.16)	(8.47)	(7.5)	(12.37)	(6.58)	(4.89)
Voice	Mean	7.43	5.2	7.7	11.43	6.37	5.89
	SD	(5.91)	(5.95)	(5.77)	(10.69)	(4.59)	(8.23)

4.2. IESM Responses

We designed an immediate questionnaire based on the IESM to rate participants' satisfaction and expectations about the AV's actions in two common scenarios: (i) slower car ahead, and (ii) accident ahead. Via two questions for each scenario, a participant's satisfaction and expectation were collected through 5-point Likert scale (1 = "Strongly Disagree" to 5 = "Strongly Agree") in sequence, as shown in Figure 2d.

In response to the two IESM questions for each scenario, a paired sample *t*-test shows a significant difference in (IESM2) "Satisfied" in all AV actions with an average ($M = 3.47$, $SD = 0.60$), $t(31) = -2.41$, $p < 0.02$, than in (IESM1) "Expected" the AV actions with an average ($M = 3.31$, $SD = 0.52$). This implies that drivers' satisfaction about AV action is not influenced by their expectations based on their own manual driving.

4.2.1. Scenario 1

The collected data were statistically analysed by two scenarios with a paired sample *t*-test. In the Slower car ahead scenario there were no significant differences between the actions overtaking the car ahead ($M = 3.66$, $SD = 1.43$) and slowing down to follow the car ahead ($M = 3.11$, $SD = 1.41$), $t(31) = 1.17$, $p < 0.25$, even though overtaking the car ahead scored higher than slowing down. However, driver satisfaction (IESM2) ($M = 3.48$, $SD = 0.61$), $t(31) = -2.49$, $p < 0.018$ was significantly higher than driver expectation (IESM1) ($M = 3.30$, $SD = 0.51$) using a paired sample *t*-test, which means that satisfaction about AV action is not influenced by drivers' expectations in a Slower car ahead scenario. They prefer the AV to overtake the slower car ahead even if they do not usually perform that action in their own driving.

4.2.2. Scenario 2

In the accident ahead scenario even *changing the route* scored higher than changing lanes, though there was no significant difference between changing lanes ($M = 3.24$, $SD = 1.49$) and changing the route ($M = 3.53$, $SD = 1.34$), $t(31) = -0.650$, $p < 0.52$ using a paired sample t -test. However, a paired sample t -test showed that there were no significant differences between driver expectation (IESM1) ($M = 3.32$, $SD = 0.66$), $t(31) = -1.54$, $p < 0.113$ and driver satisfaction (IESM2) ($M = 3.45$, $SD = 0.69$) in the accident ahead scenario. This implies that drivers' satisfaction about AV action is influenced by their expectations in an accident ahead scenario.

5. Discussion

In this study, we proposed gaze-head input as a hands-free, remote, intuitive input interaction to be implemented in autonomous vehicles. We tested reaction time, workload, system usability, and subjective preference of our proposed gaze-head input interaction against voice and touch input. An immediate experience sampling method was also used as a platform to assess each interaction and to enhance driver's perception of AV performance.

5.1. Reaction Time

Even though TI and VI are commonly used in our daily life, G-HI was the fastest input interaction and significantly faster than VI, as shown in Figure 5. To use G-HI, drivers stare at the windshield and use a quick eye movement for selection and a head nod to trigger the selection. On the other hand, when using TI, drivers must divert their eyes from the windshield to a touch panel and stretch or adjust their body posture to reach the panel; that physical effort takes time. When using VI, participants used sentences of varying length to trigger their selection: some said "the first one" or "select the first object". The time the user takes to decide which words correctly represent the desired selection likewise affect total reaction time.

The automated driving platform responded in two ways for each of our two scenarios, as shown in Figure 1b. In the slow car ahead scenario, the vehicle overtakes the slow car or slows down behind it. In the accident ahead scenario, the vehicle changes lanes to pass the accident on the right side or changes the trip route to avoid passing near the accident. We found that using G-HI enables faster reaction time for the three-input interactions in all vehicle actions (overtaking, slowing down, changing lanes, and changing route). However, for lane changes, reaction time using G-HI is significantly faster than using TI or VI, as shown in Figure 6. When changing lanes, the driver will pass the accident in the nearest lane, mimicking human nature where drivers often continue to look at an accident. Therefore, using VI and TI are not optimal input interactions in such cases, as both input interactions create high cognitive and physical loads, as shown in Figure 6. Moreover, the usage of G-HI shows a consistency even during the motion of the driving platform, users were able to reach and trigger their selection comfortably.

5.2. System Usability

G-HI scores the lowest for system usability, as shown in Figure 7, while TI scores significantly higher. There is no significant difference in usability between G-HI and VI. However, VI is familiar to most participants because of technologies like Google Assistant and Siri. The SUS assesses each input's usability subjectively. Most participants have no experience using G-HI, unlike voice and touch inputs, which we believe contributed to its preference and usability rating; as drivers become more accustomed to it, we expect G-HI preference and usability ratings to improve. Moreover, participants noted the advantage of G-HI in open-ended questionnaires: "Gaze-head is fast and promising, but it can be improved (e.g., showing a cursor for eye movement)" Participant 16; "During a call or [when] my hands are busy, gaze-head will be better" Participant 1; "Gaze-head interaction was easy to use" Participant 11.

5.3. Perceived Workload Measurement (National Aeronautics and Space Administration Task Load Index, NASA-TLX)

Workload measurement reflects no significant differences in using the three input interactions, however using G-HI scores the highest, as shown in Figure 9. Looking closer at the six factors in the NASA-TLX calculation as shown in Figure 10, G-HI creates a significantly higher cognitive load and frustration, but its physical load is significantly lower than the other input interactions. For example, using TI interrupts user engagement in the NDRT “using smart phone” and thus causes an extra physical burden. This indicates that hands-free input interactions (G-HI and VI) are more effective while engaged in NDRTs, which confirms findings that drivers are likely to use both hands to engage in NDRTs in AVs [3]. Participant behaviour when performing G-HI shows clear confusion because of the lack of familiarity compared to VI and TI. When we asked participants to nod to trigger the selection, some participants rolled their heads instead of pitching them even though we trained them how to perform the action. This may be due to cultural differences, as in some societies rolling the head rather than nodding is an expression of assent. Consequently, G-HI needs more time before it can be a universal input interaction.

5.4. Influence of Non-Driving Related Task (NDRT) in Input Interaction

After participants finished the experiment, they reported their overall preference for input interaction: VI (56.3%), TI (31.3%), and G-HI (12.5%). Then, they reported their overall preference when engaged in an NDRT: VI (56.3%), TI (6.3%), and G-HI (37.5%). Participants consistently preferred voice input, whether engaged in an NDRT or not, whereas preference for touch input decreased by 25% if the user was engaged in an NDRT. Conversely, preference for G-HI jumps by a corresponding percentage when the user is engaged in an NDRT. Therefore, we believe G-HI could assist drivers in interacting with the vehicle, especially when they are engaged in an NDRT.

5.5. Immediate Experience Sampling Method (IESM) Performance

IESM was used by participants to evaluate each input interaction immediately following the action. Designing data collection in this way guarantees precise and real-time ground truth. This accurate UX data could be used to teach the vehicle the driver’s preferred driving pattern and will increase driver satisfaction. In the first scenario participants preferred to overtake the slow car rather than slowing behind it. The reasons they gave for this selection included saving time, preference for driving fast, or they wanted the vehicle to take quicker actions. For the second scenario, most participants preferred to change the route rather than change lanes to avoid an accident. They said changing the route is safer and helps to avoid an expected traffic jam. Involving drivers in the loop and asking their opinions about vehicle action will increase feelings of control, which decreases by default when driving an AV [17,18]. Responding to a 5-point Likert scale question, participants rated the importance of using IESM to teach the vehicle their driving preferences at ($M = 4.18$, $SD = 1.16$), which is relatively high.

6. Limitations

In the presented work, G-HI and touch input were technically implemented, which can lead to moderate time delays and errors. Conversely, voice input was implemented with WoZ by the experimenter, which could guarantee an error rate close to zero and influence positively reaction time and scores for subjective preferences and lower workload.

This study did not discriminate between participants’ ages nor degree of attention participants paid to the cellphone. Our results also showed no significant differences in reaction time, SUS and perceived workload for each input interaction between female and male participants. Testing G-HI, however, may show varying performance in different age groups, i.e., it may perform better with younger participants than older participants. Also testing based on attention level may indicate a higher preference for using G-HI as it is intuitive, fast and requires low physical demand compared to voice and touch inputs.

As our study was implemented using a driving simulator, testing the proposed input interactions in a real AV is necessary to confirm our results. Testing in an on-road AV will provide greater insight into how input interaction preferences may be affected by different driving environments, possible risk factors, and a driver's perception of assessing AV driving actions on a real road. We tested our input interactions via IESM by rating participant's UX toward AV driving actions. However, we did not include other in-vehicle interactions systems, such as the infotainment systems (e.g., music player, navigation, vehicle setting), which influence where gaze-head may perform better. These interactions are also important to test with IESM.

7. Conclusions and Future Work

This study proposed and tested gaze-head input (G-HI) in an autonomous vehicle (AV) compared to the most common input interactions (touch and voice inputs). Reaction time, user's perceived workload, system usability, and subjective preferences were the evaluation criteria. G-HI showed faster reaction time and lower physical load index and appeared as a promising candidate when a driver is engaged in an NDRT to assist interaction. G-HI created higher workload in general and scored lower in system usability, however as drivers become more accustomed to it, we expect G-HI workload to decrease and usability ratings to improve. To assess the performance of the three input interactions, the immediate experience sampling method (IESM) was proposed where drivers could accurately report their satisfaction and expectations about vehicle decision making in real-time. IESM was found to be acceptable by participants, as they believed the system could enhance vehicle performance over time to match their preferences and increase feelings of control over AV actions. Generally, we believe G-HI can be a promising candidate to be used in AVs, and IESM is essential to enhance drivers' experience for the same vehicle category.

We examined drivers' UX on AV actions through two simulated scenarios. Real-world driving scenarios are more varied and more complicated, however, and many are unknown. Therefore, future studies should include more scenarios, focusing on various NDRTs as well as automated driving scenarios. Similarly, we could seek to understand the most appropriate input interactions according NDRTs. Finally, conducting multiple tests with the same participants across various NDRTs would increase their experience and familiarity with G-HI and thus affect positively system usability results. These approaches would help establish human-centered services in future AVs.

Author Contributions: Conceptualization, A.A.; writing—original draft preparation A.A., W.K., A.E.; writing—review and editing, A.A., W.K., A.E., S.K.; Supervision, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry and Energy (MOTIE) (No. 20204010600340) in 2020, and in part by the GIST Research Institute (GRI) granted funded by the GIST in 2020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. SAE International. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Technical Report No. J3016. 2018. Available online: https://www.sae.org/standards/content/j3016_201806 (accessed on 15 June 2020).
2. Pfleging, B.; Rang, M.; Broy, N. Investigating user needs for non-driving-related activities during automated driving. In Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Rovaniemi, Finland, 13–15 December 2016; pp. 91–99. [CrossRef]
3. Reimer, B.; Pettinato, A.; Fridman, L.; Lee, J.; Mehler, B.; Seppelt, B.; Park, J.; Iagnemma, K. Behavioral impact of drivers' roles in automated driving. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16), Ann Arbor, MI, USA, 24–26 October 2016; pp. 217–224. [CrossRef]

4. Hecht, T.; Feldhütter, A.; Draeger, K.; Bengler, K. What do you do? An analysis of non-driving related activities during a 60 minutes conditionally automated highway drive. In Proceedings of the International Conference on Human Interaction and Emerging Technologies, Nice, France, 22–24 August 2019; Springer: Cham, Switzerland, 2019; pp. 28–34.
5. Stevens, G.; Bossauer, P.; Vonholdt, S.; Pakusch, C. Using Time and Space Efficiently in Driverless Cars: Findings of a Co-Design Study. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–14.
6. Kim, H.S.; Yoon, S.H.; Kim, M.J.; Ji, Y.G. Deriving future user experiences in autonomous vehicle. In Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Nottingham, UK, 1–3 September 2015; pp. 112–117. [\[CrossRef\]](#)
7. Yeo, D.; Lee, J.; Kim, W.; Kim, M.; Cho, K.; Ataya, A.; Kim, S. A Hand-Over Notification System of Vehicle Driving Control according to Driver's Condition based on Cognitive Load in Autonomous Driving Situation. In Proceedings of the HCI Korea, Seogwipo-si, Korea, 13–16 February 2019; pp. 495–500.
8. Kim, W.; Kim, S. A New Approach to Studying Sleep in Autonomous Vehicles: Simulating the Waking Situation. In Proceedings of the Adjunct Publication of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19), New Orleans, LA, USA, 20–23 October 2019; pp. 10–12. [\[CrossRef\]](#)
9. Large, D.; Burnett, G.; Morris, A.; Muthumani, A.; Matthias, R. Design Implications of Drivers' Engagement with Secondary Activities During Highly-Automated Driving—A Longitudinal Simulator Study. In Proceedings of the Road Safety and Simulation International Conference (RSS), The Hague, The Netherlands, 17–19 October 2017; pp. 1–10.
10. Detjen, H.; Geisler, S.; Schneegass, S. Maneuver-based Driving for Intervention in Autonomous Cars. In *CHI'19 Work-Shop on "Looking into the Future: Weaving the Threads of Vehicle Automation"*; ACM: New York, NY, USA, 2019.
11. Tscharn, R.; Latoschik, M.E.; Löffler, D.; Hurtienne, J. "Stop over there": Natural gesture and speech interaction for non-critical spontaneous intervention in autonomous driving. In Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, UK, 13–17 November 2017; pp. 91–100.
12. Detjen, H.; Faltaous, S.; Geisler, S.; Schneegass, S. User-Defined Voice and Mid-Air Gesture Commands for Maneuver-based Interventions in Automated Vehicles. In Proceedings of the Mensch und Computer (MuC'19), Hamburg, Germany, 8–11 September 2019; pp. 341–348.
13. Walch, M.; Mühl, K.; Baumann, M.; Weber, M. Click or Hold: Usability Evaluation of Maneuver Approval Techniques in Highly Automated Driving. In Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–6. [\[CrossRef\]](#)
14. Walch, M.; Sieber, T.; Hock, P.; Baumann, M.; Weber, M. Towards cooperative driving: Involving the driver in an autonomous vehicle's decision making. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26 October 2016; pp. 261–268.
15. Kim, M.; Seong, E.; Jwa, Y.; Lee, J.; Kim, S. A Cascaded Multimodal Natural User Interface to Reduce Driver Distraction. *IEEE Access* **2020**, *8*, 112969–112984. [\[CrossRef\]](#)
16. Lee, S.C.; Nadri, C.; Sanghavi, H.; Jeon, M. Exploring User Needs and Design Requirements in Fully Automated Vehicles. In Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–9. [\[CrossRef\]](#)
17. Frison, A.K.; Wintersberger, P.; Riener, A.; Schartmüller, C. Driving Hotzenplotz: A Hybrid Interface for Vehicle Control Aiming to Maximize Pleasure in Highway Driving. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17), Oldenburg, Germany, 24–27 September 2017; pp. 236–244. [\[CrossRef\]](#)
18. Rödel, C.; Stadler, S.; Meschtscherjakov, A.; Tscheligi, M. Towards autonomous cars: The effect of autonomy levels on acceptance and user experience. In Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14), Seattle, WA, USA, 17–19 September 2014; pp. 1–8. [\[CrossRef\]](#)
19. Meschtscherjakov, A.; Trösterer, S.; Döttlinger, C.; Wilfinger, D.; Tscheligi, M. Computerized experience sampling in the car: Issues and challenges. In Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '13), Eindhoven, The Netherlands, 27–30 October 2013; pp. 220–223.

20. Meschtscherjakov, A.; Wilfinger, D.; Osswald, S.; Perterer, N.; Tscheligi, M. Trip experience sampling: Assessing driver experience in the field. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '12), Portsmouth, NH, USA, 17–19 October 2012; pp. 225–232. [\[CrossRef\]](#)
21. Karapanos, E. Technology-assisted reconstruction: A new alternative to the experience sampling method. *Behav. Inf. Technol.* **2020**, *39*, 722–740. [\[CrossRef\]](#)
22. Walch, M.; Jaksche, L.; Hock, P.; Baumann, M.; Weber, M. Touch screen maneuver approval mechanisms for highly automated vehicles: A first evaluation. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct (Automotive'UI 17), Oldenburg, Germany, 24–27 September 2017; pp. 206–211. [\[CrossRef\]](#)
23. Kauer, M.; Schreiber, M.; Bruder, R. How to conduct a car? A design example for maneuver based driver-vehicle interaction. In Proceedings of the 2010 IEEE Intelligent Vehicles Symposium, San Diego, CA, USA, 21–24 June 2010; IEEE: Piscataway Township, NJ, USA, 2010; pp. 1214–1221.
24. Manawadu, U.E.; Kamezaki, M.; Ishikawa, M.; Kawano, T.; Sugano, S. A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; IEEE: Piscataway Township, NJ, USA, 2016; pp. 1785–1790.
25. Sidenmark, L.; Gellersen, H. Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, New Orleans, LA, USA, 20–23 October 2019; pp. 1161–1174.
26. Elepfandt, M.; Grund, M. Move it there, or not? The design of voice commands for gaze with speech. In Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction, Santa Monica, CA, USA, 26 October 2012; pp. 1–3.
27. Chatterjee, I.; Xiao, R.; Harrison, C. Gaze + Gesture: Expressive, Precise and Targeted Free-Space Interactions. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15), Seattle, WA, USA, 9–13 November 2015; pp. 131–138. [\[CrossRef\]](#)
28. Špakov, O.; Majaranta, P. Enhanced gaze interaction using simple head gestures. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; pp. 705–710.
29. Creed, C.; Frutos-Pascual, M.; Williams, I. Multimodal Gaze Interaction for Creative Design. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13. [\[CrossRef\]](#)
30. Hyrskykari, A.; Istance, H.; Vickers, S. Gaze gestures or dwell-based interaction? In Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, CA, USA, 28–30 March 2012; pp. 229–232. [\[CrossRef\]](#)
31. Roider, F.; Gross, T. I See Your Point: Integrating Gaze to Enhance Pointing Gesture Accuracy While Driving. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; pp. 351–358.
32. Roider, F.; Reisig, L.; Gross, T. Just Look: The Benefits of Gaze-Activated Voice Input in the Car. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; pp. 210–214.
33. Špakov, O.; Isokoski, P.; Majaranta, P. Look and lean: Accurate head-assisted eye pointing. In Proceedings of the Symposium on Eye Tracking Research and Applications, Safety Harbor, CA, USA, 26–28 March 2014; pp. 35–42.
34. Qian, Y.Y.; Teather, R.J. The eyes don't have it: An empirical comparison of head-based and eye-based selection in virtual reality. In Proceedings of the 5th Symposium on Spatial User Interaction, Brighton, UK, 16–17 October 2017; pp. 91–98.
35. Mardanbegi, D.; Hansen, D.W.; Pederson, T. Eye-based head gestures. In Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, CA, USA, 28–30 March 2012; pp. 139–146. [\[CrossRef\]](#)
36. Yan, Y.; Yu, C.; Yi, X.; Shi, Y. Headgesture: Hands-free input approach leveraging head movements for hmd devices. In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, New York, NY, USA, 27 December 2018; pp. 1–23.
37. Kytö, M.; Ens, B.; Piumsomboon, T.; Lee, G.A.; Billinghurst, M. Pinpointing: Precise head-and eye-based target selection for augmented reality. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–14. [\[CrossRef\]](#)

38. Hansen, J.P.; Trudslev, A.K.; Harild, S.A.; Alapetite, A.; Minakata, K. Providing access to VR through a wheelchair. In Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–8. [\[CrossRef\]](#)
39. Špakov, O.; Istance, H.; Räihä, K.J.; Viitanen, T.; Siirtola, H. Eye gaze and head gaze in collaborative games. In Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (ETRA '19), Denver, CO, USA, 25–28 June 2019; pp. 1–9. [\[CrossRef\]](#)
40. Sidenmark, L.; Clarke, C.; Zhang, X.; Phu, J.; Gellersen, H. Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13. [\[CrossRef\]](#)
41. Kurauchi, A.; Feng, W.; Morimoto, C.; Betke, M. HMAGIC: Head movement and gaze input cascaded pointing. In Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '15), Corfu, Greece, 1–3 July 2015; pp. 1–4. [\[CrossRef\]](#)
42. Niforatos, E.; Karapanos, E.; Langheinrich, M.; Wurhofer, D.; Krischkowsky, A.; Obrist, M.; Tscheligi, M. eMotion: Retrospective in-car user experience evaluation. In Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 15), Nottingham, UK, 1–3 September 2015; pp. 118–123.
43. Park, S.Y.; Moore, D.J.; Sirkin, D. What a Driver Wants: User Preferences in Semi-Autonomous Vehicle Decision-Making. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13.
44. Van Berkel, N.; Ferreira, D.; Kostakos, V. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* **2018**, *50*, 1–40. [\[CrossRef\]](#)
45. Braunagel, C.; Geisler, D.; Stolzmann, W.; Rosenstiel, W.; Kasneci, E. On the necessity of adaptive eye movement classification in conditionally automated driving scenarios. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications (ETRA'16), Charleston, SC, USA, 14–17 March 2016; pp. 19–26.
46. Gerber, M.A.; Schroeter, R.; Xiaomeng, L.; Elhenawy, M. Self-Interruptions of Non-Driving Related Tasks in Automated Vehicles: Mobile vs Head-Up Display. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20), Honolulu, HI, USA, 25–30 April 2020; pp. 1–9. [\[CrossRef\]](#)
47. Angelini, L.; Baumgartner, J.; Carrino, F.; Carrino, S.; Caon, M.; Khaled, O.; Sauer, J.; Lalanne, D.; Mugellini, E.; Sonderegger, A. Comparing gesture, speech and touch interaction modalities for in-vehicle infotainment systems. In Proceedings of the IHM'16, Fribourg, Switzerland, 25–28 October 2016; pp. 188–196.
48. Large, D.R.; Burnett, G.; Anyasodo, B.; Skrypchuk, L. Assessing Cognitive Demand during Natural Language Interactions with a Digital Driving Assistant. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16), Ann Arbor, MI, USA, 24–26 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 67–74. [\[CrossRef\]](#)
49. Bangor, A.; Kortum, P.; Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.* **2009**, *4*, 114–123.
50. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Adv. Psychol.* **1988**, *52*, 139–183. [\[CrossRef\]](#)

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).