scientific reports

Check for updates

OPEN

# Assessing driver distraction from in-vehicle information system: an on-road study exploring the effects of input modalities and secondary task types

Qi Zhong✉, Jinyi Zhi✉, Yongsheng Xu, Pengfei Gao & Shu Feng

**In-vehicle information system (IVIS) use is prevalent among young adults. However, their interaction with IVIS needs to be better understood. Therefore, an on-road study aims to explore the effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior. A 2 × 4 within-subject design was undertaken. The independent variables are input modalities (auditory-speech and visual-manual) and secondary task types (calls, music, navigation, and radio). The dependent variables include secondary task performance (task completion time, number of errors, and SUS), driving performance (average speed, number of lane departure warnings, and NASA-TLX), and visual glance behavior (average glance duration, number of glances, total glance duration, and number of glances over 1.6 s). The statistical analysis result showed that the main effect of input modalities is significant, with more distraction during visual-manual than auditory-speech. The main impact of secondary task types was also substantial across most metrics, aside from average speed and average glance duration. Navigation and music were the most distracting, followed by calls, and radio came in last. The distracting effect of input modalities is relatively stable and generally not moderated by the secondary task types, except radio tasks. The findings practically benefit the driver-friendly human–machine interface design, preventing IVIS-related distraction.**

The in-vehicle information system (IVIS) is an essential human–machine interface (HMI) providing drivers with both driving-related information (e.g., navigation or fuel level) and non-driving-related information (e.g., music or calls)[1]. IVIS is commonly recognized for its extensive entertainment features[2]. For the past few years, IVIS's sales have witnessed exponential growth. According to Statista[3], the global shipment of units will surpass 200 million by 2022. Compared to traditional in-vehicle devices, the advances in intelligent networking technology have begun to erode the boundary between mobile phones and IVIS[4]. Given the explicit prohibition surrounding mobile phone use while driving in most countries or regions[5] and the growing incorporation of mobile phone technology in IVIS[6], the issue of IVIS-related distraction has drawn widespread concerns worldwide. Therefore, it is essential to understand IVIS's challenges for road safety and public health[7–9].

## IVIS use while driving among young drivers

Driver distraction is increasingly recognized as a significant source of motor vehicle injuries and fatalities[10]. IVIS use while driving is dangerous and even illegal when it endangers safe driving, as it diverts attention away from the driving primary task and towards the IVIS secondary tasks[11]. It can involve visual, auditory, manual, cognitive, and temporal demands, often requiring a combination of all simultaneously[12]. There is extensive evidence from simulators[13–15], on-road studies[16–18], naturalistic[19–21], roadside observations[22–24], office accident reports[25–27], and self-report interviews or questionnaires[28–30] jointly demonstrating the negative consequences of IVIS-related distraction on the roadway. For example, Peng et al.[13] identified drivers exhibiting considerable visual attention on text-related IVIS tasks within a driving simulator test. Zhong et al.[16] conducted a field driving

Department of Industrial Design, School of Design, Southwest Jiaotong University, Chengdu 611756, China. ✉email: 18238836586@163.com; zhijinyi@swjtu.edu.cn

study insisting that IVIS use behind the wheel worsened driving performance and caused a potential accident. Dingus et al.[19], using naturalistic driving data, reported that interactions with IVIS increase the crash odds by 4.6 times in Americans, which was higher than fatigued driving (odds ratio = 3.4) and mobile phone use while driving (odds ratio = 3.6). A systematic review of observational studies on secondary task engagement while driving[22] documented that IVIS is the second most common in-car distraction after mobile phones. Beanland et al.[25] demonstrated the feasibility of using in-depth crash data to investigate driver inattention in casualty crashes. Furthermore, Oviedo-Trespalacios et al.[28] found that drivers can be aware of the distraction risk of IVIS, especially the more visually demanding glance behavior through interviews and questionnaires. Unfortunately, there is no clear legislation regarding IVIS-related distractions compared to mobile phones[31–33].

A recent meta-analysis conducted by Ziakopoulos et al.[34] revealed that young drivers exhibit the highest usage of IVIS and are more prone to distraction-related injuries. Similarly, Romer et al.[35] emphasized the importance of focusing on young drivers, considering their underdeveloped ability to allocate attention and limited driving experience, particularly regarding technology-based distractions. Lansdown[36] found that individuals aged 18–29 years were more prone to experiencing IVIS-related distractions compared to other age groups. Young drivers may face heightened driver errors due to their use of IVIS, such as running a yellow light and failing to detect a pedestrian[37]. When considering the combined influences of limited experience, the inclination to use their IVIS, and the subsequent rise in driver error, it becomes evident that determining the effects of IVIS use on young drivers' distracted behavior is of utmost importance.

## The distracting effects of input modalities and secondary task types

According to multi-resource theory[38], the primary driving task and secondary IVIS task can compete in each resource pool (e.g., visual, auditory, manual, and cognitive). Thus, the distraction effect caused by various input modalities may be different. Besides, given that the difficulty, steps, and duration can vary, the level of distraction may also be affected by secondary task types[39,40]. Several research studies have concentrated on the distracting effects of input modalities and secondary task types, some of which are mentioned in the current work.

For instance, Maciej and Vollrath[41], using a simple driving simulation, investigated the distracting effects of input modalities (auditory-speech and visual-manual) and secondary task types (music, calls, address, and area-of-interest navigation). As a result, auditory-speech improved driving performance, off-road glance behavior, and subjective experience, except for area-of-interest navigation. Garay-Vega et al.[42] examine the input modalities within in-vehicle music systems (auditory-speech and visual-manual) on driver distraction through a virtual world. Consequently, auditory-speech reduced the total glance duration, prolonged glances, number of glances, and perceived workload compared to the visual-manual. Zhong et al.[43] launched a medium-fidelity simulating experiment to directly compare the effect of four input modalities (QWERTY, hand-writing, shape-writing, and auditory-speech) within in-vehicle navigation systems on young drivers' distracted behavior. In general, auditory-speech significantly outperformed the other three visual-manual input modalities. Ma et al.[44] implemented a high-fidelity driving simulator to evaluate the effects of input modalities (steering wheel buttons, knobs and buttons, touchscreen, and auditory-speech) and secondary task classes (basic, medium, and advanced) on distracted driving behavior. Conclusively, steering wheel buttons and auditory-speech were better matched for basic tasks. Auditory-speech was appropriate for medium tasks. Both visual-manual and auditory-speech were proper for advanced tasks. Zhang et al.[45] explore the effects of input modalities (touchscreen, auditory-speech, and gesture) and secondary task types (navigation, music, and browsing) on IVIS performance, driving performance, and visual behavior. The simulator study reported that touchscreen led to the worst influence, gesture followed, and auditory-speech came in last. The interaction effects between input modalities and secondary task types are generally insignificant.

While increasing research on the distracting effects of input modalities and secondary task types, most are limited to simulated driving environments. Although it can provide safe and controllable conditions, and the relative validity for in-vehicle HMI visual distraction testing is gradually confirmed[46,47], it is still necessary to understand the distracting effect in production vehicles and actual driving conditions, which further increases external validity. Early, Chiang et al. compare driver use of the auditory-speech and visual-manual interface for navigation entry tasks based on over-the-road evaluations[48]. Similarly, Mehler and colleagues assess the on-road demand of the two types of interfaces in two production vehicles, which documented that auditory-speech interfaces can reduce visual demand but do not eliminate it[49,50]. Recently, the on-road study of Strayer et al. also provides empirical evidence that the distracting effects of IVIS systematically varied as a function of the input modalities, secondary task types, and vehicle models[17,18]. Another recent work by their research team also points out that age has a noteworthy moderating effect on IVIS-related distraction[51]. However, to our knowledge, few on-road studies reported the effects of input modalities and secondary task types on young drivers' distracted behavior.

## The current study

In conclusion, the literature on the distracting effects of input modalities and secondary task types is mainly limited to simulator research. With the maturity of intelligent network technology, field driving experiments with production vehicles have become possible. Besides, young people are the heaviest IVIS users and account for a significant share of distraction-related accidents, but how young drivers interact with IVIS under actual driving conditions needs to be better examined. Therefore, an on-road study was undertaken on an instrumented vehicle to explore the effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior. Theoretically, it supplements the current literature on IVIS-based distractions among young drivers. Practically, it is helpful to develop a driver-friendly IVIS. Specifically, three hypotheses are proposed as follows.

**Hypotheses 1:** The main effects of input modalities on young drivers' secondary task performance, driving performance, and visual glance behavior are significant.

**Hypotheses 2:** The main effects of secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior are significant.

**Hypotheses 3:** The interaction effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior are significant.

# Method

## Design

The 2 × 4 within-subject repeated measures experiment was designed, with input modalities and secondary task types as the independent variable. The interaction modalities cover auditory-speech and visual-manual. The secondary task types include calls, music, radio, and navigation. The dependent variable, distracted driving behavior, is represented by three dimensions (secondary task performance, driving performance, and visual glance behavior), the detailed information of which can be seen in Table 1. Specifically, secondary task performance includes task completion time, number of errors, and system usability scale (SUS). Driving performance includes average speed, number of lane departure warnings, and NASA-Task Load Index (NASA-TLX). Visual glance behavior includes average glance duration, number of glances, total glance duration, and number of glances over 1.6 s.

## Participant

The convenience sampling method was employed to recruit 28 participants (14 males) who resided in Chengdu, China, aged 18–25 years (M = 21.66; SD = 1.88), and had a full driver's license and proof of vehicle insurance. Over 50% (N = 17) of the participants were enlisted via the classroom or professional lectures from the School of Design, Southwest Jiaotong University. All the remaining participants (N = 13) were enlisted via a snowball sampling approach among friends and classmates. Individuals who did not have regular or corrected normal visual acuity were excluded. Each participant must also complete a 15 min online safety driving training and pass the accompanying certification examination following the policy of the Ethics Committee of the School of Design, Southwest Jiaotong University (Approval Number: 2022120904). All participants had signed written informed consent forms before participating.

As shown in Table 2, most participants had a college education or higher on average. They reported driving an average of 3.6 h per week (M = 3.60; SD = 1.43). Of the participants, 67.9% (N = 19) drove automatic vehicles, while 32.1% (N = 9) reported driving manual vehicles. 50% of the participants (N = 14) mainly drove on urban roads. The top four functions commonly utilized on their IVIS, in descending order, are as follows: calls (N = 26), music (N = 25), radio (N = 23), and navigation (N = 21).

| Dimension | Metrics | Definitions | Measurements | Correlation |
|---|---|---|---|---|
| Secondary task performance | Task completion time (s) | Task completion time refers to the interval from the issue of the pre-recorded voice command to the participant successfully completing the corresponding IVIS task[1,14,15,43,52–56] | Post-video analysis | − |
| | Number of errors (n) | The number of errors refers to the counts that deviated from the correct operation path during the task completion time[1,14,15,43,52–56] | Post-video analysis | − |
| | SUS (points) | The SUS is a subjective assessment tool of IVIS usability (0 ~ 100 points) on ten statement items[14,52] | SUS | + |
| Driving performance | Average speed (km/h) | The average speed is the mean driving speed during dual-task and baseline conditions[1,14,15,43,52,54,56] | Post-video analysis | + |
| | Number of lane departure warnings (n) | The number of lane departure warnings refers to the alarm counts automatically issued by the lane departure warning system of the instrumented vehicle during dual-task and baseline conditions[18,54] | Post-video analysis | − |
| | NASA-TLX (points) | The NASA-TLX is a subjective assessment tool deriving an overall workload score (0 ~ 100 points) that needs a weighted average of the ratings on six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration[14–16,52–56] | NASA-TLX | − |
| Visual glance behavior | Total glance duration (s) | The total glance duration refers to the total amount of glance time entering the area of interest (IVIS screen in this study) during dual-task conditions[1,14–16,43,52–56] | Tobii Pro Lab | + |
| | Number of glances (n) | The number of glances refers to the number of glances entering the AOI during dual-task conditions[1,14–16,43,52–55] | Tobii Pro Lab | + |
| | Average glance duration (s) | The average glance duration refers to the mean time for each glance entering the AOI during dual-task conditions[1,14–16,43,52–54,56] | Tobii Pro Lab | + |
| | Number of glances over 1.6 s (n) | The number of glances over 1.6 s refers to the mean glance duration exceeding 1.6 s during dual-task conditions, increasing near-crash and crash risks by at least two times than baseline driving[1,15,43,52,54] | Tobii Pro Lab | + |

**Table 1.** Definition, source, measurement, and correlation of dependent variables. " − " denotes the negative correlation. " + " denotes that the positive correlation.

| Demographics | Classification | N | % |
|---|---|---|---|
| Gender | Male | 14 | 50.0 |
| | Female | 14 | 50.0 |
| Education level | High school or below | 4 | 14.3 |
| | Undergraduate or junior college | 18 | 64.3 |
| | Postgraduate or above | 6 | 21.4 |
| Transmission type | Automatic | 19 | 67.9 |
| | Manual | 9 | 32.1 |
| Driving location | Largely urban area | 14 | 50.0 |
| | Both urban and rural areas | 11 | 39.3 |
| | Largely rural area | 3 | 10.7 |
| Top four IVIS functions | Calls | 26 | 92.9 |
| | Music | 25 | 89.3 |
| | Radio | 23 | 82.1 |
| | Navigation | 21 | 75.0 |

**Table 2.** Descriptive statistics of demographic characteristics (N = 28).

## Location

According to the driver behavioral adaptation theory[57], drivers generally do not interact with IVIS at particular road alignments (e.g., intersections, turns, and congested roads) to ensure safe driving benefits. Besides, a complex road environment will cause internal interference to the experimental result. Therefore, a two-way, four-lane urban road test route (see Fig. 1) located in Pidu District, Chengdu, China, was selected to ensure the experiment's external and internal validity as much as possible. The test route consists of two straight sections, a right angle turn, 14 intersections, and 12 traffic lights, starting at 999 Xian Road and ending at 555 Campus Road. The lane is 3.75 m in width and 4.5 km long, isolating the central green belt, the low-demand traffic flow, and no horizontal curves and vertical gradients. Roadside infrastructure includes traffic signs, buildings, streetlights, and tree-lined landscapes. All the work was carried out on weekdays from 8:30 ~ 10:30 and 15:00 ~ 17:00 to avoid the rush commuting hours.

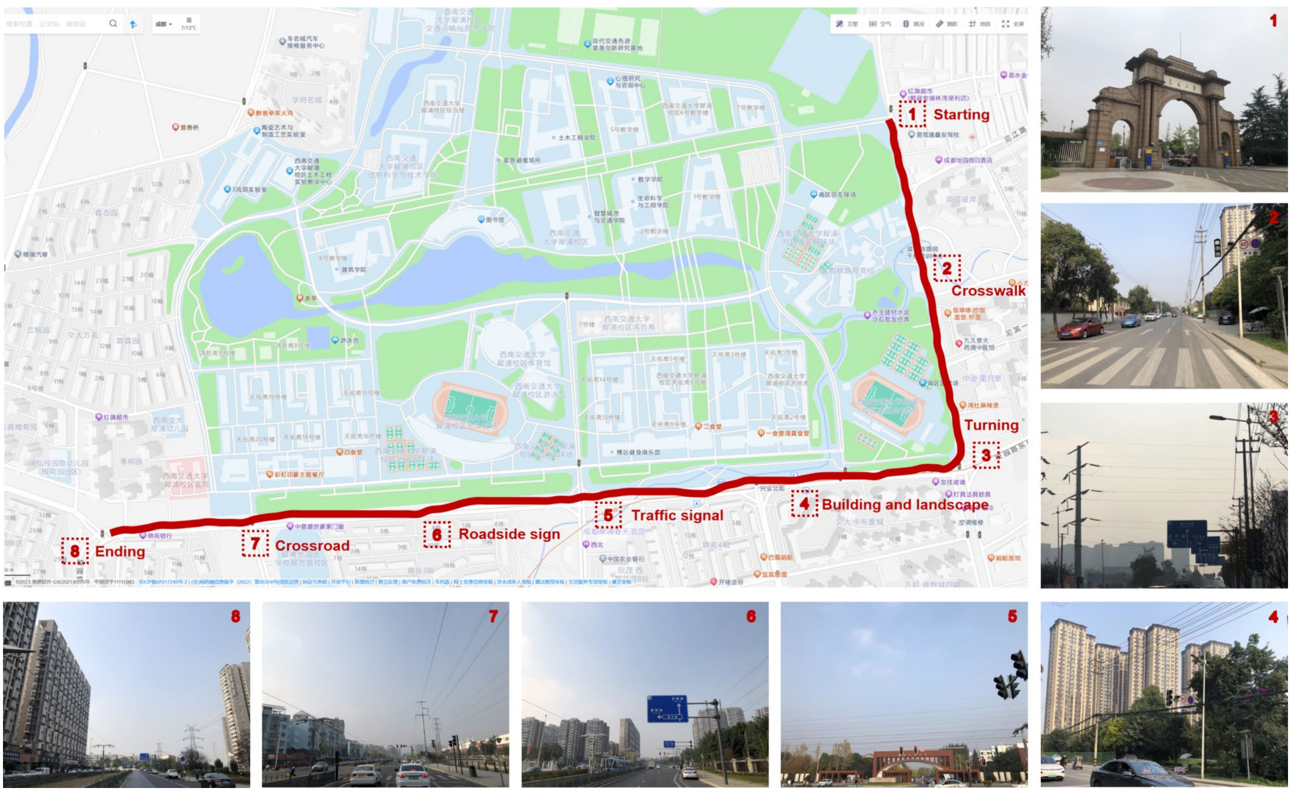



**Fig. 1.** The Amap application, version 13.0, created the map of the experimental route (URL:https://map.amap.com).

## Apparatus

The field driving experiments occurred in a left-hand Nissan SUV model—Trail 2019, representing modern intelligent connected vehicles to the market. The instrument vehicle has a 9.7-in touchscreen-based IVIS with a resolution of 2048 × 1536, equipped with various applications (e.g., calls, music, radio, and navigation) for subsequent secondary task comparisons. In addition, the lane departure warning system based on 77 GHz millimeter wave radar is sensitive to record the number of lane departure warnings. A USB-based GoPro Hero 12 Black motion camera connected to a mobile phone was mounted to the side of the IVIS screen by a bracket, providing a 90° view to capture the participants' interaction performance. The Tobii Glasses 3 is the latest wearable eye tracking device, using binocular stereo dark pupil tracking technology, sampling rate of 50 Hz, accuracy of 0.6°, very suitable for driving behavior research[14–18]. When the integrity and quality of automatic mapping are not reliable in the eye-tracking analysis, the frame-by-frame manual check and mapping are used to correct it. In necessity, we can also create the 'Snapshots' and mark the 'Events' types in the Tobii Pro Lab software to help process the eye-tracking data. A USB-based 70mai dash cam M500 was installed in the top right of the front windshield to record the driving speed, traffic video, and trajectory with a resolution of 2592 × 1944 and a wide angle of 170°. The snapshot is shown in Fig. 2, including the front side view Fig. 2a and trunk view Fig. 2b of the instrumented vehicle.

## Task

The dual-task paradigm was adopted; that is, participants need to perform additional IVIS tasks while driving. The primary driving task is lane-keeping with the upper-speed limit of 60 km/h for the selected naturalistic driving route[1,14–16,43]. No lower speed limit is set in this study so that the participants can self-control driving speed to perform IVIS tasks, just as they do in daily driving. The participants must pay attention to the road ahead, drive carefully, and complete all IVIS tasks to ensure safety. As a reference, three 30 s-baseline tasks (no distraction) were randomly recorded during the trial to allow a direct comparison of baseline and dual-task driving performance. According to the results in Table 2, four everyday secondary tasks of calls, music, radio, and navigation were selected, whose user interface and step-by-step procedures are provided in Fig. 3 and Table 3, respectively. In other words, each participant was required to participate in four IVIS activities through two input modalities.

Specifically, the call task required participants to open the call app, dial a predefined contact, and close the app 3 s later (no speaking). The music task requires participants to open a music app, play a predefined song, and close the app 3 s later. The radio task requires participants to open the radio app, find a predefined FM channel, and close the app 3 s later. The navigation task requires participants to open a navigation app, enter a predefined destination, and close the app 3 s later. The baseline task and eight IVIS tasks were performed alternately in a single drive, and each task was repeated three times. All phone contacts, music songs, radio FM, and navigation destinations are manipulated into the same for each participant.

## Procedure

The experiment lasts about 30 min and mainly includes the five stages: preparation, debugging, practice, formal, and questionnaire. The whole experiment took place from October 15 to 30, 2023.

(1) Preparation stage. Upon arrival, the assistant informs the participant about the purpose, procedure, precautions, benefits, risks, and anonymity. Then, the basic demographics were collected and re-confirmed according to the recruitment criteria, including gender, age, driving habits, and experience. Finally, each participant signed the written informed consent form.


(a) Front side view of the vehicle

(b) Trunk view of the vehicle

**Fig. 2.** The snapshot of the experimental apparatus.


**Fig. 3.** The user interface of the calls (top-left), music (top-right), radio (bottom-left), and navigation (bottom-right).

(2) Debugging stage. Participants performed pre-driving preparation and underwent a series of free tasks to familiarize themselves with the IVIS. Then, the Tobii glasses were calibrated to pre-check and ensure eye-tracking quality. In addition, the debugging of other instruments (e.g., cameras, dash cams, and recordings) also needs to be completed.

(3) Practice stage. Participants were provided a free driving scene to familiarize themselves with the daily vehicle maneuvers (e.g., start and stop, accelerating, steering, and braking) and to practice the IVIS task under dynamic driving conditions. The IVIS tasks in the practice stage are similar but different from those in the formal stage. After proficiency, proceed directly to the next stage.

(4) Formal stage. Participants performed different IVIS tasks according to the pre-recorded voice commands played by the experimenter in the passenger seat. Meanwhile, the assistant in the back seat should record and check all experimental data to prevent omissions. The experimenter and assistant accompanied all participants throughout the trial. However, their exchanges were limited to initial instructions and urgent safe questions.

(5) Questionnaire stage. At the end of the trial, all participants were asked to pull over safely. Then, the experimenter asked them to complete the online SUS and NASA-TLX according to their real feelings. After confirmation, each participant was reimbursed a movie ticket for their contributions.

## Results

A total of 4 out of the 28 participants were excluded due to less than 80% of gaze samples. Consequently, 24 participants $\times$ 9 tasks (baseline and 8 IVIS tasks) $\times$ 3 replicates formed the analytical sample. Data was imported into SPSS software, version 29.0, for further processing and analysis. In order to reduce the measurement error, the data from repeated tasks were superimposed and averaged[14–16]. Statistical assumptions, including outlier, normality, and heteroscedasticity, were evaluated to determine the suitability for the planned analyses. The number of errors and glances over 1.6 s were approximately subject to the Gaussian distribution by Arcsine Square Root conversion. Greenhouse–Geisser correction is used to correct the result of spherical violation. A series of $2 \times 4$ repeated measurement analyses of variance (R-ANOVA) and paired sample *t*-tests were performed to analyze the distracting effects of input modalities and secondary task types. Adding a baseline task for data analysis is necessary for driving performance. If the main effect was significant, the Student Newman Keuls (SNK) method was used for post hoc comparison. The chosen level of significance for all analyses was set at 0.05.

## Secondary task performance

Figure 4 shows the mean and standard deviation of task completion time, number of errors, and SUS. First, as shown in Fig. 4a, the two-factor R-ANOVA found that input modalities had a significant effect on task completion time [$F(1,23) = 52.75$, $\eta^2 = 0.692$, $p < 0.001$]. Secondary task types also had a substantial impact on task completion

| Task | Step | Visual-manual input | Auditory-speech input | IVIS out |
|---|---|---|---|---|
| Calls | Turn on the phone app | Tap the phone app button | Speak 'Turn on the phone app' | Instruction of device is opening app |
| | Find contact 'xxx' | Swipe to find the contact 'xxx' | Speak 'xxx' | Visual feedback + device displayed |
| | Start a call | Click the start button | Speak 'Please start a call' | Instruction of device is starting the call |
| | Turn off the phone app 3 s later | Click the end button | Speak 'Turn off the phone app' | Instruction of device has closed the app |
| | Back to the home page | Click the back button | Speak 'Please return to the home page' | Instruction of device is displaying the home page |
| Music | Turn on the music app | Tap the music app button | Speak 'Turn on the music app' | Instruction of device is opening app |
| | Search song 'xxx' | Enter the 'xxx' in the search box | Speak 'xxx' | Visual feedback + device displayed |
| | Start playing | Click the start button | Speak 'Please start playing' | Instruction of device is starting the playing |
| | Turn off the music app 3 s later | Click the end button | Speak 'Turn off the music app' | Instruction of device has closed the app |
| | Back to the home page | Click the back button | Speak 'Please return to the home page' | Instruction of device is displaying the home page |
| Radio | (1) Turn on the radio app | Tap the radio app button | Speak 'Turn on the radio app' | Instruction of device is opening app |
| | (2) Find FM 'xxx' MHZ | Swipe to FM 'xxx' MHZ | Speak FM 'xxx' MHZ | Visual feedback + device displayed |
| | Start playing | Click the start button | Speak 'Please start playing' | Instruction of device is starting the playing |
| | Turn off the radio app 3 s later | Click the end button | Speak 'Turn off the radio app' | Instruction of device has closed the app |
| | Back to the home page | Click the back button | Speak 'Please return to the home page' | Instruction of device is displaying the home page |
| Navigation | Turn on the navigation app | Tap the navigation app button | Speak 'Turn on the navigation app' | Instruction of device is opening app |
| | Import destination 'xxx' | Enter the 'xxx' in the search box | Speak 'xxx' | Visual feedback + device displayed |
| | Start navigating | Click the start button | Speak 'Please start navigation' | Instruction of device is starting the navigation |
| | Turn off the navigation app 3 s later | Click the end button | Speak 'Turn off the navigation app' | Instruction of device has closed the app |
| | Back to the home page | Click the back button | Speak 'Please return to the home page' | Instruction of device is displaying the home page |

**Table 3.** The step-by-step procedures for the IVIS secondary task. The smartphone was connected with the IVIS via Bluetooth in advance.

time [$F(3,69) = 40.38$, $\eta^2 = 0.422$, $p < 0.01$]. The interaction effect between input modalities and secondary task types was insignificant [$F(3,69) = 15.33$, $\eta^2 = 0.117$, $p > 0.05$]. The SNK test suggested that the task completion time of the visual-manual (M = 40.95; SD = 7.5) was significantly longer than auditory-speech (M = 20.05; SD = 4.80). Navigation (M = 37.12; SD = 8.01) and music (M = 36.15; SD = 4.25) together took the most time, followed by calls (M = 30.65; SD = 6.44), and radio (M = 18.12; SD = 5.94) came in last.

Second, as shown in Fig. 4b, the two-factor R-ANOVA manifested that the input modalities had a significant effect on the number of errors[$F(1.12,25.76) = 33.12$, $\eta^2 = 0.352$, $p < 0.01$]. The secondary task types also had a substantial impact on number of errors [$F(3.58,75.27) = 35.50$, $\eta^2 = 0.383$, $p < 0.01$]. The interaction effect between the input modalities and secondary task types was marginally significant [$F(4.09,115.44) = 20.66$, $\eta^2 = 0.149$, $p = 0.048$]. The SNK test showed that errors in the visual-manual condition (M = 4.05; SD = 1.23) were significantly higher than in auditory-speech (M = 3.15; SD = 1.04). However, the opposite is true for radio. Music (M = 4.58; SD = 0.44) and navigation (M = 4.54; SD = 1.76) had the highest errors, followed by calls (M = 3.04; SD = 0.73), and radio (M = 2.27; SD = 1.20) came in last.

Third, as shown in Fig. 4c, the *t*-test showed that interaction modalities had a significant effect on perceived usability [$t(23) = 49.75$, $\eta^2 = 0.605$, $p < 0.001$], and the SUS score of visual-manual (M = 58.41; SD = 5.48) was significantly lower than auditory-speech (M = 76.07; SD = 6.90).

## Driving performance

Figure 5 shows the mean and standard deviation of average speed, number of lane departure warnings, and NASA-TLX. First, as shown in Fig. 5a, the two-factor R-ANOVA displayed that input modalities had a significant effect on average speed [$F(1,23) = 50.12$, $\eta^2 = 0.650$, $p < 0.001$]. Secondary task types had no significant impact on average speed [$F(3,69) = 13.85$, $\eta^2 = 0.096$, $p > 0.05$]. There was no significant interaction effect between

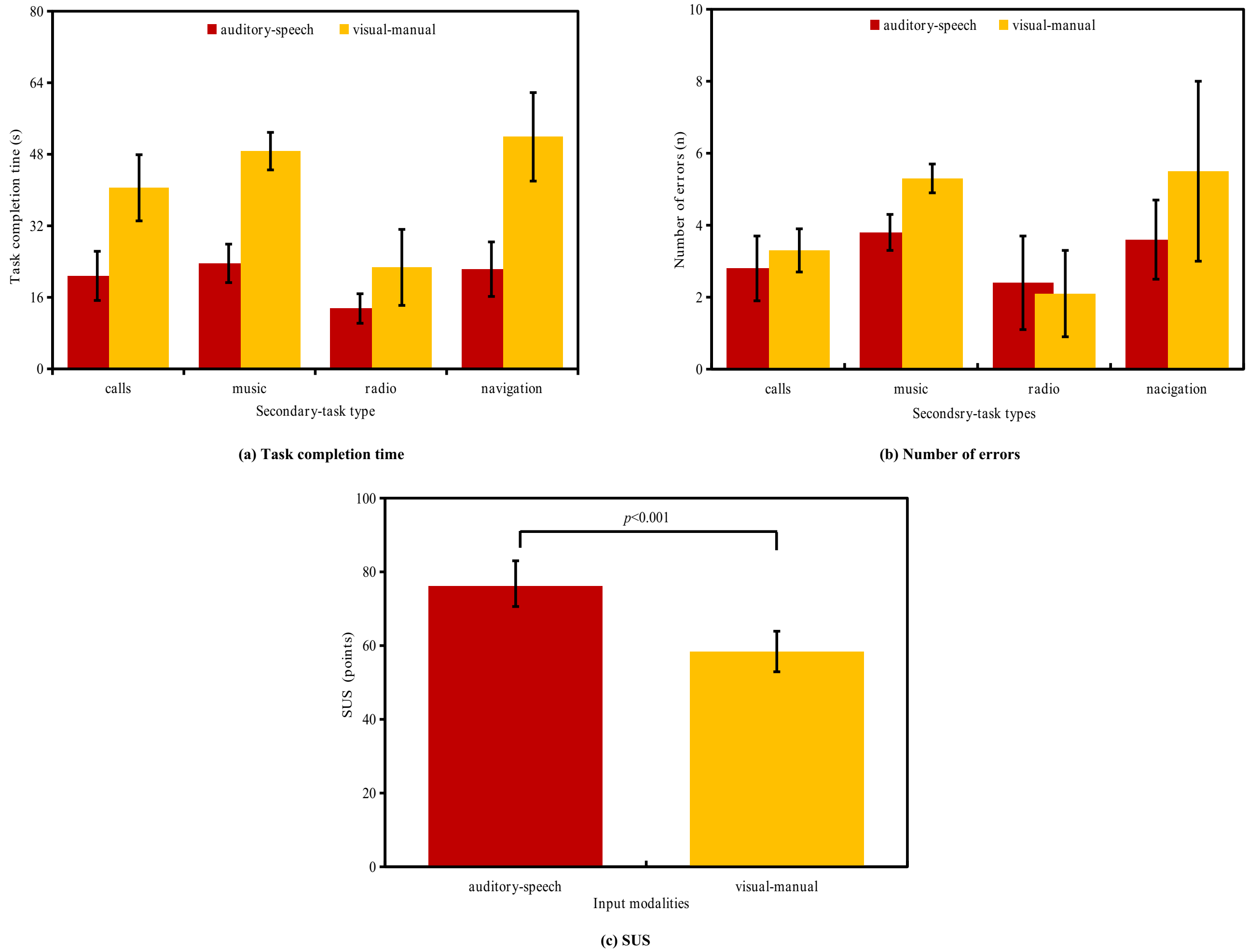



**Fig. 4.** The result of secondary task performance.

input modalities and secondary task types [$F(3,69) = 12.74$, $\eta^2 = 0.091$, $p > 0.05$]. The SNK test proved that the average speed of visual-manual (M = 37.55; SD = 3.50) was significantly slower than auditory-speech (M = 43.37; SD = 3.98), and the average speed of auditory-speech was markedly slower than baseline conditions (M = 51.64; SD = 5.49).

Second, as shown in Fig. 5b, the two-factor R-ANOVA showed that input modalities had an insignificant effect on number of lane departure warnings [$F(1,23) = 10.75$, $\eta^2 = 0.083$, $p > 0.05$]. Secondary task types had a substantial impact on number of lane departure warnings [$F(3,69) = 32.50$, $\eta^2 = 0.348$, $p < 0.01$]. There was an insignificant interaction effect between input modalities and secondary task types [$F(3,69) = 15.08$, $\eta^2 = 0.110$, $p > 0.05$]. The SNK test revealed that lane departure warnings under the vision-manual (M = 1.54; SD = 0.48) and auditory-speech (M = 1.42; SD = 0.51) conditions were significantly higher than the baseline condition (M = 0.81; SD = 0.39). Navigation (M = 1.72; SD = 0.47) and music (M = 1.63; SD = 0.45) cause the most lane departure warnings, followed by radio (M = 1.27; SD = 0.50) and calls (M = 1.24; SD = 0.57).

Third, the weighted total and six sub-scales scores of NASA-TLX under the baseline, auditory-speech, and visual-manual conditions are shown in Fig. 5c and d. The $t$-test exhibited that the score of visual-manual (M = 6.67; SD = 0.44) was significantly higher than auditory-speech (M = 5.62; SD = 0.41) [$t(23) = 38.52$, $\eta^2 = 0.405$, $p < 0.01$], and the score of visual-manual was significantly higher than the baseline condition (M = 4.29; SD = 0.50) [$t(23) = 50.18$, $\eta^2 = 0.610$, $p < 0.001$]. The score of auditory-speech was also significantly higher than the baseline condition [$t(23) = 40.77$, $\eta^2 = 0.425$, $p < 0.01$].

## Visual glance behavior

Figure 6 shows the mean and standard deviation of average glance duration, number of glances, total glance duration, and number of glances over 1.6 s. First, as shown in Fig. 6a, the two-factor R-ANOVA reported that input modalities had a significant effect on average glance duration [$F(1,23) = 22.76$, $\eta^2 = 0.204$, $p < 0.05$]. Secondary task types had no significant impact on average glance duration [$F(3,69) = 16.33$, $\eta^2 = 0.122$, $p > 0.05$]. There was no significant interaction effect between input modalities and secondary task types [$F(3,69) = 12.81$, $\eta^2 = 0.093$, $p > 0.05$]. The SNK test argued that the average glance duration under the visual-manual condition (M = 1.50; SD = 0.41) was significantly longer than the auditory-speech condition (M = 1.31; SD = 0.29).

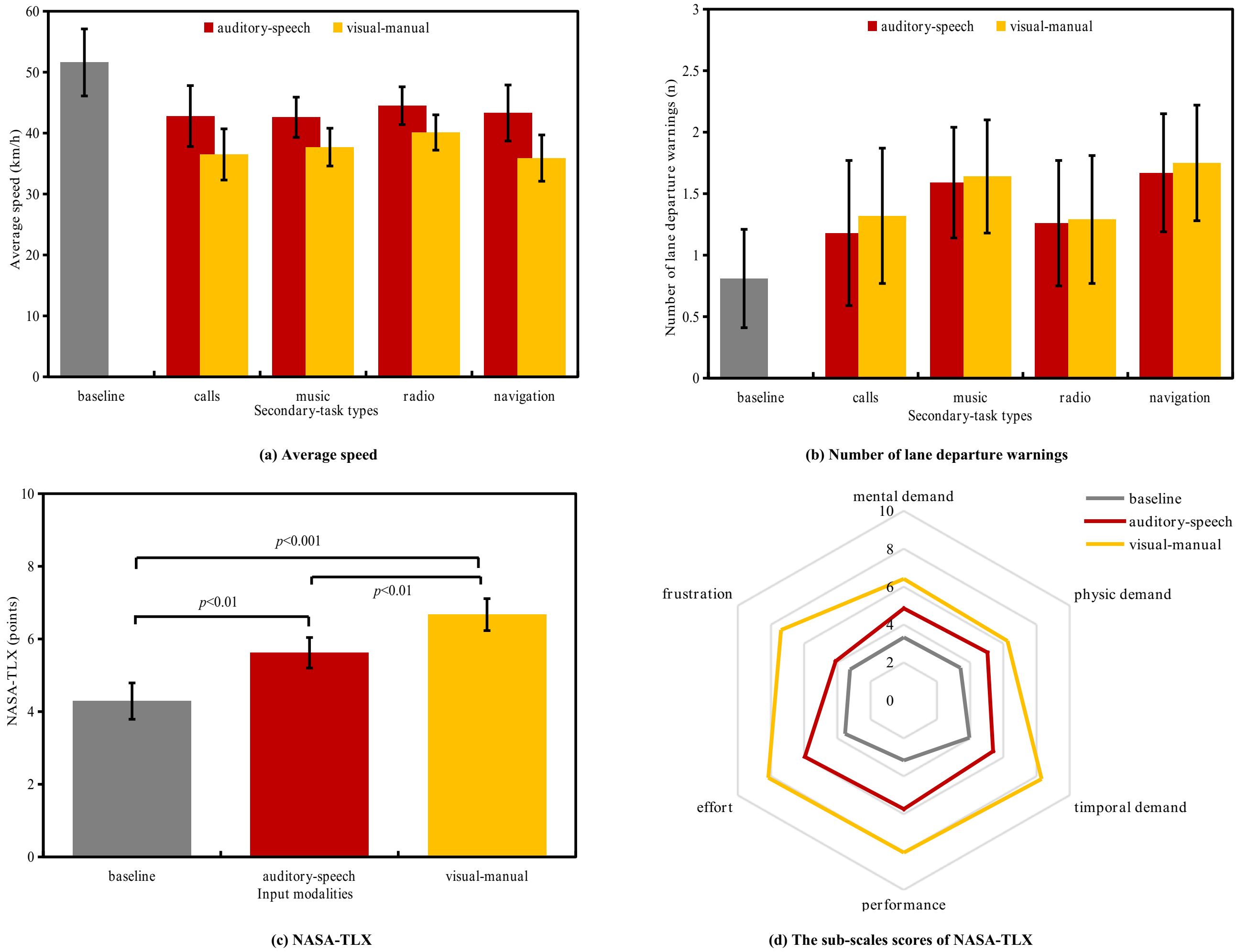
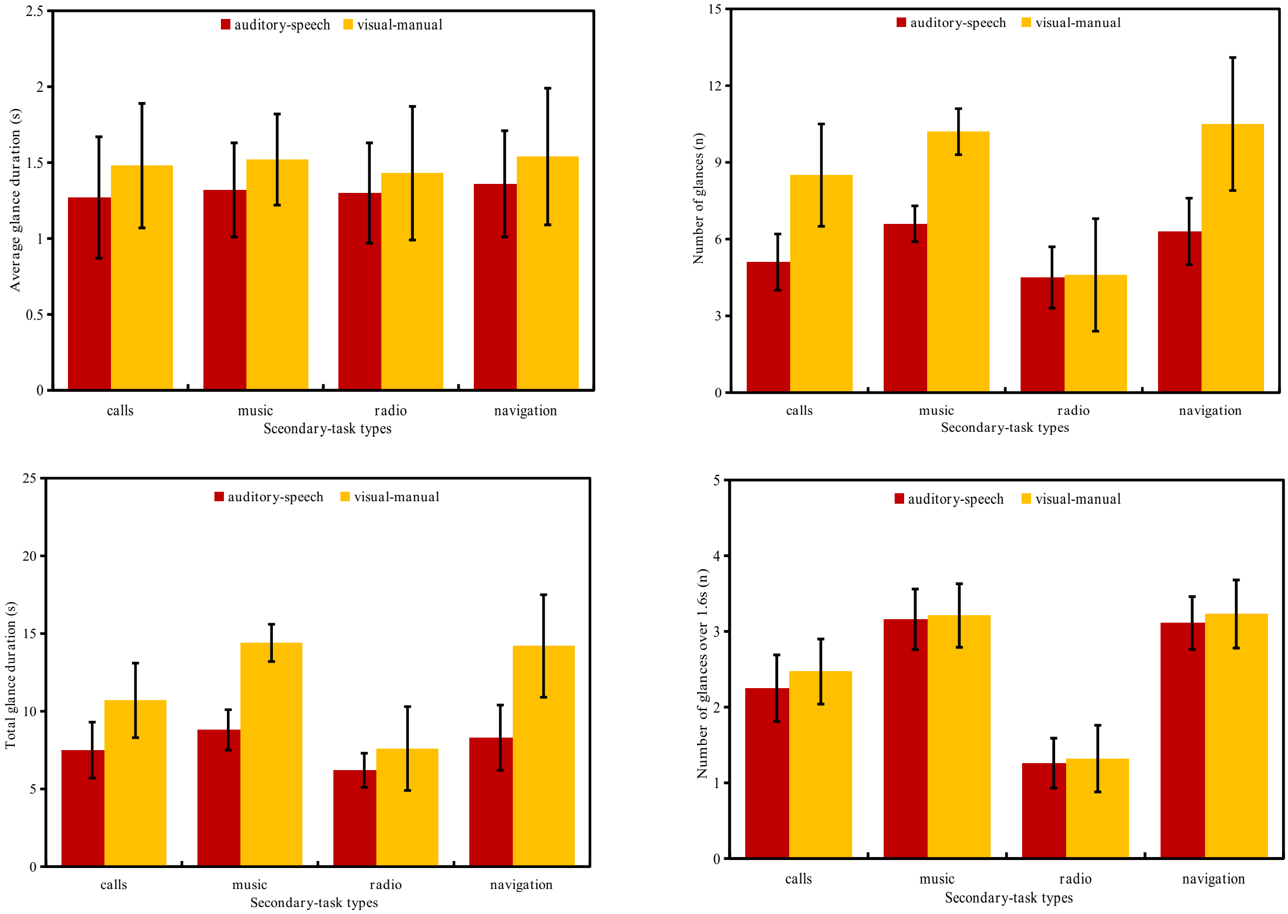



**Fig. 5.** The result of driving performance.

Second, as shown in Fig. 6b, the two-factor R-ANOVA showed that input modalities had a substantial effect on the number of glances [$F$(1.15,27.35) = 48.07, $\eta^2$ = 0.590, $p < 0.001$]. Secondary task types also had a significant impact on the number of glances [$F$(3.62,78.62) = 46.99, $\eta^2$ = 0.572, $p < 0.001$]. Besides, the interaction effects on the number of glances are significant [$F$(4.22,118.47) = 24.48, $\eta^2$ = 0.233, $p < 0.05$]. The SNK test found that the number of glances under the visual-manual condition (M = 8.45; SD = 1.90) was significantly more frequent than the auditory-speech condition (M = 5.62; SD = 1.07), except for radio. Navigation (M = 8.41; SD = 1.94) and music (M = 8.39; SD = 0.78) cause the most frequent glances, followed by calls (M = 6.78; SD = 1.55), and radio (M = 4.56; SD = 1.71) came in last.

Third, as shown in Fig. 6c, the two-factor R-ANOVA messaged that input modalities had a significant effect on total glance duration [$F$(1,23) = 46.84, $\eta^2$ = 0.376, $p < 0.001$]. Secondary task types also had a substantial impact on total glance duration [$F$(3,69) = 38.91, $\eta^2$ = 0.390, $p < 0.01$]. There was an insignificant interaction effect between input modalities and secondary task types [$F$(3,69) = 17.48, $\eta^2$ = 0.129, $p > 0.05$]. The SNK test informed that the total glance duration of the visual-manual (M = 11.72; SD = 2.38) was significantly greater than auditory-speech (M = 7.70; SD = 1.64). Music (M = 11.69; SD = 1.27) and navigation (M = 11.22; SD = 2.66) cause the most prolonged total glance duration, followed by calls (M = 9.12; SD = 2.07), and radio (M = 6.88; SD = 1.90) came in last.

Fourth, as shown in Fig. 6d, the two-factor R-ANOVA showed that the main effect of input modalities on the number of glances over 1.6 s was insignificant [$F$(1.22,28.06) = 9.82, $\eta^2$ = 0.075, $p > 0.05$]. The secondary task types have a significant impact on the number of glances over 1.6 s [$F$(3.95,81.64) = 30.64, $\eta^2$ = 0.322, $p < 0.01$]. Input modalities and secondary task types had an insignificant interaction effect [$F$(4.66,128.91) = 15.02, $\eta^2$ = 0.108, $p > 0.05$]. The SNK test messaged that the number of glances over 1.6 s occurred the most in music (M = 3.19; SD = 0.42) and navigation (M = 3.17; SD = 0.40), followed by calls (M = 2.36; SD = 0.44), and radio (M = 1.29; SD = 0.38) came in last.




(c) Total glance duration

(d) Number of glances over 1.6s

**Fig. 6.** The result of visual glance behavior.

## Discussion

The on-road driving experiment aims to explore the effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior. The main impact of input modalities was significant across most metrics, aside from the number of lane departure warnings and glances over 1.6 s. Compared with baseline conditions, both vision-manual and auditory-voice conditions caused different distraction levels, but auditory-speech was smaller than visual-manual. The main effect of secondary task types was also significant across most metrics, aside from average speed and average glance duration. Navigation and music were most distracted, followed by calls, and radio came in last. The distracting effect of input modalities is relatively robust and is generally not moderated by the secondary task types, except radio tasks.

**Hypotheses 1** The main effect of input modalities

First, the results prove that the visual-manual interface in this car will cause a decline in IVIS performance. Specifically, compared with auditory-speech interaction, when young drivers use vision-manual interaction to perform IVIS tasks, the task completion time significantly increases by about 104.2%, the number of errors significantly increases by about 28.6%, and the perceived usability decreases by 23.3%. This result is similar to those obtained in the previous literature under simulated driving conditions[41,42]. It is inferred that auditory-speech interaction is more valuable than visual-manual interaction for IVIS in simulated and actual driving conditions. In addition, more specifically, compared to the simulator experiment[41–45], this study unexpectedly found a growing difference between auditory-speech and visual-manual interaction. In other words, the on-road study further exacerbates the significant differences between the two input modalities. On the one hand, this is mainly due to the current advances in natural language processing technology[58]. The external technical factors (delay time and recognition accuracy rate) and internal design factors (e.g., menu depth and style strategy of machine dialogue) of auditory-speech interaction in modern vehicles are much improved compared with the original fixed voice command processing[58,59]. On the other hand, as various smartphone functions are integrated into IVIS, the interface design of smartphone applications is also followed. However, it may not meet driver's needs in safety–critical situations[28].

Then, the results confirmed that auditory-speech and visual-manual interfaces severely negatively impacted driving performance, but the visual-manual interaction was relatively worse. Specifically, compared with baseline conditions, when drivers performed IVIS tasks in auditory-speech and visual-manual input modalities, the average speed decreased by about 16.1% and 27.2%, respectively, and the subjective driving workload increased by 25.5% and 59.1%, respectively. In dual-task conditions, regardless of the input modalities, the young driver slows down to compensate for the reduced attention resources devoted to IVIS so that lane departure does not occur. This behavioral adaptation has been reported in the literature as a common operational-level strategy for coping with distraction[57]. However, even if such safety behavior adaptation exists, it is still detrimental to the driving experience characterized by a higher workload. This finding is similar to a recent on-road study by Strayer and their colleagues[17,18,51]. In other words, self-regulating behaviors adopted by young drivers did not fully offset the adverse impact of IVIS-related distractions. Besides, the current in-car auditory-speech and visual-manual technology has yet to improve and still falls short of baseline levels. According to Wickens' multi-resource theory[38], when interacting with the IVIS, the driver's eyes are away from the road ahead, compromising access to road information necessary for safe driving. At the same time, the driver has to take one hand off the wheel to operate the IVIS, which may also contribute to poor driving performance. Although the motor distraction is alleviated to a certain extent by not taking hands off the steering wheel during auditory-speech interaction, it still requires the driver to perform multiple visual feedback and confirmation, occupying specific visual and cognitive resources. The results of the on-road study are basically consistent with those of the existing simulator[43–45], which indirectly verifies the relative validity of the simulated driving method[46,47].

Finally, the experimental results further demonstrate the adverse effects of auditory-speech and visual-manual interactions on visual distraction. The distracting effect of visual-manual is the most prominent, inducing longer average glance duration, more frequent glances, and more total glance duration. Specifically, when the young driver performs the IVIS task in the vision-manual interaction, the average glance duration increases by about 13.7%, the number of glances increases by about 50.2%, and the total glance duration increases by about 52.3%, compared to the auditory-speech interaction. However, the number of glances over 1.6 s remained the same, which is inconsistent with the recent results of Zhong et al.[43] and Zhang et al.[45]. The possible reason is that they adopted simulated driving, while this study is field driving. Once drivers realize the lack of severe safety consequences, to ensure safe driving, they will not only induce driving behavior adaptation but also may adopt adaptive visual glance behavior, such as reducing the number of long glances[12,13]. The previous literature associated with IVIS-related distractions reported that the number of long glances of 1.6 s or 2.0 s is closely related to the near crash or crash risk[1,15,43,52,54].

**Hypotheses 2** The main effect of secondary task types

First, regarding IVIS performance, secondary task types significantly affect the task completion time and number of errors. Specifically, navigation and music had the most errors and completion times, followed by calls, and radio came in last. This result is partially similar to the study of Ma et al.[9]. The difference is that the music-playing task takes less time than the navigation setting. The possible reason is that their music playback is in push-button-based CD and MP3 formats. However, in this study, the touchscreen-based IVIS requires additional visual feedback and confirmation[16].

Then, regarding driving performance, the main effect of secondary task types on average speed is not significant, but the main effect on the number of lane departure warnings is significant. In the dual-task driving process, regardless of the secondary task, the driver will adopt the behavioral adaptation strategy of slowing down, and the average speed is very similar. This finding complements the lack of reports on longitudinal behavioral adaptation of vehicles by young drivers in similar literature[57]. Unfortunately, lateral driving performance varies significantly between the four secondary task types, even with longitudinal driving behavior adaptations. Specifically, the number of lane departure warnings for music and navigation was similar and significantly more extensive than that of calls and radio, and there was no statistical difference between the latter two tasks. This finding is roughly similar to the results of the on-road study by Zhong et al.[16]. Put differently, although the driver has taken self-regulation measures to reduce the speed, it still can not compensate for the increased lateral position deviation of the vehicle in some tasks, which may be related to young drivers' less driving experience and insufficient defensive driving awareness.

Finally, regarding visual glance features, the main effect of secondary task types on number of glances, total glance duration, and number of glances over 1.6 s is significant. The distribution patterns of these three visual glance measures were similar among the four tasks; music and navigation occupied most visual demand, followed by calls, and radio came in last. According to the task procedure analysis (see Table 3), music and navigation tasks involve texting and reading secondary tasks, during which the driver's eye duration significantly increased[12,13]. In contrast, calls and radio mainly involve the basic actions of touch button clicks and scroll-bar swipes. Therefore, music and navigation programs take up relatively large visual resources. In addition, the main effect of secondary task types on average glance duration was insignificant, which may be a valuable finding because the average glance duration is sensitive to vehicle crash risk[1,15,43,52,54]. When encountering complex secondary tasks (e.g., music and navigation in this study), the driver will interrupt in time and return his eyes to the road ahead to ensure safe driving[7,16]. In summary, drivers will likely adopt adaptive visual glances behavior that increases the number of glances and total glance duration to obtain enough information to complete corresponding IVIS tasks.

**Hypotheses 3** The interaction effect between input modalities and secondary task types

Input modalities and secondary task types have significant interaction effects on the number of errors and glances but not on other indicators. As for the calls, music, and navigation tasks, the number of errors and glances of the visual-manual interface was more significant than the auditory-speech interface. However, regarding radio, the number of errors and glances are similar, which is consistent with the results of the simulator experiment by Zhang et al.[45]. In general, the influence of input modalities on distracted driving behavior is relatively stable and only in a few indicators (e.g., task error times and scanning times), whose distracting effect is moderated by some tasks. The possible reason is that the radio task is relatively easy to operate using the visual-manual, and the auditory-speech still requires multiple visual confirmations. Compared with a simple button click or sliding, its advantages cannot be reflected[44], further illustrating the importance of integrating multimodal interaction technologies in modern vehicles. For example, relatively complex actions (e.g., text reading and typing) can use auditory-speech interaction, while simple tap and swipe actions can use visual-manual interaction. However, it should be noted that these interaction effects should be interpreted with caution since some previous works of literature[7,34–37,57] have shown that driver characteristics (e.g., age, sex, and personality) and road environment factors (e.g., road width, curvature, and gradient) are likely to affect the test results collectively.

### Practical implications

Practically, this study can offer valuable insights for interaction designers in developing driver-friendly IVIS to prevent distraction-related road injuries. First, it is necessary to acknowledge that visual-manual is still the dominant input mode for IVIS today. The possible reason is that the environment rarely affects visual-manual performance, while auditory-speech recognition is sensitive to ambient noise[58]. Another possibility is that drivers must learn and memorize standard voice commands in the first use of auditory-voice interaction, which can lead to a higher initial cognitive workload[59]. In contrast, visual-manual interaction is more intuitive and easy to learn. Therefore, the distraction effects induced by IVIS can be mitigated in the following ways.

(a) Strengthening the design components (e.g., menu depth, delay times, and recognition accuracy) of in-vehicle voice interaction, as it still needs to catch up to baseline conditions[58,59].
(b) Optimizing the visual HMI design of IVIS. IVIS should consider driver safety as the primary criterion, unlike the smartphone application[52–56].
(c) In contrast to head-down displays (HDD), augmented reality-based head-up displays (AR-HUD) presenting information on the windshield provide a less visually demanding display technology[60].
(d) Developing other input modalities (e.g., mid-air gesture and eye movement interaction) or multimodal interactions may provide new solutions for driver distraction[45,61].
(e) Locking the IVIS function while the vehicle is in motion may be a simple and practical approach[62]. However, it is necessary to be aware that many drivers will pick up their smartphones if the IVIS option is unavailable.

### Strength, limitations, and further study

This study represents a limited number of on-road studies focusing on the effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior. Meanwhile, the study emphasizes the significant practical relevance of focusing on young drivers, who are the most frequent users of IVIS services and suffer from a heightened risk of distraction-related crashes.

The study had certain limitations and claims for further research. Firstly, our research instructed participants to perform the IVIS tasks in a counterbalanced order across participants, which is still essentially a controlled experiment. However, in real-world settings, drivers can interact with IVIS if, when, and where they choose. Future research should conduct a naturalistic driving study to examine this matter and explicitly explore potential self-regulation behavior[19–21]. Second, it would be interesting to determine the applicability of these findings to other cars, age cohorts, and regions[34–37] or whether the outcomes are specific to the young drivers aged 18–25 years driving a Nissan SUV model—Trail 2019 in Chengdu, China. Future research needs to increase the participant pools to observe the influence of demographic characteristics and whether the findings prove this generalizes to all vehicles. Thirdly, for driving safety purposes, it is necessary to acknowledge that the driving route setting in this on-road study is relatively simple, and the experimental result would vary as a function of road environments[7,57,63]. Future research is also necessary to uncover the distraction mechanisms of IVIS use while driving in challenging traffic conditions, such as tunnels, bends, and night-time conditions.

## Conclusion

The current study employs on-road research to investigate the effects of input modalities and secondary task types on young drivers' secondary task performance, driving performance, and visual glance behavior, specifically those aged 18–25. The main effect of input modalities is significant. Compared with the baseline condition, both vision-manual and auditory-voice conditions caused different distraction levels, but auditory-speech was smaller than visual-manual. In addition to average speed and average glance duration, secondary task types also have significant main effects. Among them, navigation and music were most distracted, followed by calls, and radio came in last. The distracting effect of input modalities is relatively stable and is generally not moderated by the secondary task types, except radio tasks. These findings help understand how IVIS use affects young drivers' behavior, contributing to developing driver-friendly IVIS to prevent distraction-related road injuries and fatalities.

## Data availability

All data generated or analyzed during this study are included in this published article.

Received: 30 December 2023; Accepted: 26 August 2024
Published online: 31 August 2024

## References


1. Zhong, Q., Zhi, J. & Guo, G. Dynamic is optimal: Effect of three alternative auto-complete on the usability of in-vehicle dialing displays and driver distraction. *Traffic Inj. Prev.* **23**, 51–56 (2022).
2. Kim, J., Kim, S. & Nam, C. User resistance to acceptance of in-vehicle infotainment (IVI) systems. *Telecomm. Policy.* **40**, 919–930 (2016).
3. Statista. Shipments of the in-vehicle IVISs worldwide from 2015 to 2022 (in million units). https://www.statista.com/statistics/784966/in-car-infotainment-systems-shipments-worldwide (2023).
4. Kim, G. Y., Kim, S. R., Kim, M. J., Shim, J. M. & Ji, Y. G. Effects of animated screen transition in in-vehicle infotainment systems: Perceived duration, delay time, and satisfaction. *Int. J. Hum. Comput. Interact.* **39**, 203–216 (2023).
5. Lipovac, K., Deric, M., Tesic, M., Andric, Z. & Maric, B. Mobile phone use while driving-literary review. *Transp. Res. Part F Traffic Psychol. Behav.* **47**, 132–142 (2017).
6. Wang, L. & Ju, D. Y. Concurrent use of an in-vehicle navigation system and a smartphone navigation application. *Soc. Behav. Pers.* **43**, 1629–1640 (2015).
7. Oviedo-Trespalacios, O. Getting away with texting: Behavioural adaptation of drivers engaging in visual-manual tasks while driving. *Transp. Res. Part A Policy Pract.* **116**, 112–121 (2018).
8. Simmons, S. M., Caird, J. K. & Steel, P. A meta-analysis of in-vehicle and nomadic voice-recognition system interaction and driving performance. *Accid. Anal. Prev.* **106**, 31–43 (2017).
9. Ma, Y. *et al.* Support vector machines for the identification of real-time driving distraction using in-vehicle information systems. *J. Transp. Saf. Secur.* **14**, 232–255 (2022).
10. Kohl, J., Gross, A., Henning, M. & Baumgarten, T. Driver glance behavior towards displayed images on in-vehicle information systems under real driving conditions. *Transp. Res. Part F Traffic Psychol. Behav.* **70**, 163–174 (2020).
11. Ebel, P., Lingenfelder, C. & Vogelsang, A. On the forces of driver distraction: Explainable predictions for the visual demand of in-vehicle touchscreen interactions. *Accid. Anal. Prev.* **183**, 106956 (2023).
12. Peng, Y. & Boyle, L. N. Driver's adaptive glance behavior to in-vehicle information systems. *Accid. Anal. Prev.* **85**, 93–101 (2015).
13. Peng, Y., Boyle, L. N. & Lee, J. D. Reading, typing, and driving: How interactions with in-vehicle systems degrade driving performance. *Transp. Res. Part F Traffic Psychol. Behav.* **27**, 182–191 (2014).
14. Zhong, Q., Zhi, J. & Guo, G. Effect of the complexity of in-vehicle information interface on visual search and driving behavior. *J. Saf. Environ.* **22**, 2003–2010 (2022).
15. Zhong, Q., Guo, G. & Zhi, J. Chinese handwriting while driving: Effects of handwritten box size on in-vehicle information systems usability and driver distraction.Traffic Inj. *Prev.* **24**, 26–31 (2023).
16. Zhong, Q., Zhi, J. & Guo, G. Influence of in-vehicle information system interaction modes on driving behavior. *J. Saf. Environ.* **22**, 1406–1411 (2022).
17. Strayer, D. L. *et al.* Assessing the visual and cognitive demands of in-vehicle information systems. *Cogn. Res. Princ. Implic.* **4**, 1–22 (2019).
18. Strayer, D. L. *et al.* Visual and cognitive demands of CarPlay, Android Auto, and five native infotainment systems. *Hum. Factors* **61**, 1371–1386 (2019).
19. Dingus, T. A. *et al.* Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proc. Natl. Acad. Sci. USA.* **113**, 2636–2641 (2016).
20. Singh, H. & Kathuria, A. Analyzing driver behavior under naturalistic driving conditions: A review. *Accid. Anal. Prev.* **150**, 105908 (2021).
21. Wang, X., Xu, R., Zhang, S., Zhuang, Y. & Wang, Y. Driver distraction detection based on vehicle dynamics using naturalistic driving data. *Transp. Res. Part C Emerg. Technol.* **136**, 103561 (2022).
22. Huemer, A. K., Schumacher, M., Mennecke, M. & Vollrath, M. Systematic review of observational studies on secondary task engagement while driving. *Accid. Anal. Prev.* **119**, 225–236 (2018).
23. Prat, F., Planes, M., Gras, M. E. & Sullman, M. J. M. An observational study of driving distractions on urban roads in Spain. *Accid. Anal. Prev.* **74**, 8–16 (2014).
24. Kidd, D. G. & Chaudhary, N. K. Changes in the sources of distracted driving among Northern Virginia drivers in 2014 and 2018: A comparison of results from two roadside observation surveys. *J. Safety Res.* **68**, 131–138 (2019).
25. Beanland, V., Fitzharris, M., Young, K. L. & Lenné, M. G. Driver inattention and driver distraction in serious casualty crashes: Data from the Australian national crash in-depth study. *Accid. Anal. Prev.* **54**, 99–107 (2013).
26. Talbot, R., Fagerlind, H. & Morris, A. Exploring inattention and distraction in the safety net accident causation database. *Accid. Anal. Prev.* **60**, 445–455 (2013).
27. Wundersitz, L. Driver distraction and inattention in fatal and injury crashes: Findings from in-depth road crash data. *Traffic Inj. Prev.* **20**, 696–701 (2019).
28. Oviedo-Trespalacios, O., Nandavar, S. & Haworth, N. L. How do perceptions of risk and other psychological factors influence the use of in-vehicle information systems (IVIS)?. *Transp. Res. Part F Traffic Psychol. Behav.* **67**, 113–122 (2019).
29. Yao, X. *et al.* Analysis of psychological influences on navigation use while driving based on extended theory of planned behavior. *Transp. Res. Rec.* **2673**, 480–490 (2019).
30. Chen, H. W. & Donmez, B. What drives technology-based distractions? A structural equation model on social-psychological factors of technology-based driver distraction engagement. *Accid. Anal. Prev.* **91**, 166–174 (2016).
31. Parnell, K. J., Stanton, N. A. & Plant, K. L. Exploring the mechanisms of distraction from in-vehicle technology: The development of the PARRC model. *Saf. Sci.* **87**, 25–37 (2016).
32. Parnell, K. J., Stanton, N. A. & Plant, K. L. What's the law got to do with it? Legislation regarding in-vehicle technology use and its impact on driver distraction. *Accid. Anal. Prev.* **100**, 1–14 (2017).
33. Parnell, K. J., Stanton, N. A. & Plant, K. L. What technologies do people engage with while driving and why?. *Accid. Anal. Prev.* **111**, 222–237 (2018).
34. Ziakopoulos, A., Theofilatos, A., Papadimitriou, E. & Yannis, G. A meta-analysis of the impacts of operating in-vehicle information systems on road safety. *IATSS Res.* **43**, 185–194 (2019).
35. Romer, D., Lee, Y. C., McDonald, C. C. & Winston, F. K. Adolescence, attention allocation, and driving safety. *J. Adolesc. Health* **54**, S6–S15 (2014).
36. Lansdown, T. C. Individual differences and propensity to engage with in-vehicle distractions - A self-report survey. *Transp. Res. Part F Traffic Psychol. Behav.* **15**, 1–8 (2012).
37. Klauer, S. G. *et al.* Distracted driving and risk of road crashes among novice and experienced drivers. *N. Engl. J. Med.* **370**, 54–59 (2014).
38. Wickens, C. D. Multiple resources and mental workload. *Hum. Factors* **50**, 449–455 (2008).
39. Bamney, A., Pantangi, S. S., Jashami, H. & Savolainen, P. How do the type and duration of distraction affect speed selection and crash risk? An evaluation using naturalistic driving data. *Accid. Anal. Prev.* **178**, 106854 (2022).

40. Jin, L., Xian, H., Niu, Q. & Bie, J. Research on safety evaluation model for in-vehicle secondary task driving. *Accid. Anal. Prev.* **81**, 243–250 (2015).
41. Maciej, J. & Vollrath, M. Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accid. Anal. Prev.* **41**, 924–930 (2009).
42. Garay-Vega, L. *et al.* Evaluation of different speech and touch interfaces to in-vehicle music retrieval systems. *Accid. Anal. Prev.* **42**, 913–920 (2010).
43. Zhong, Q., Guo, G. & Zhi, J. Address inputting while driving: A comparison of four alternative text input methods on in-vehicle navigation displays usability and driver distraction. *Traffic Inj Prev.* **23**, 163–168 (2022).
44. Ma, J., Li, J. & Gong, Z. Evaluation of driver distraction from in-vehicle information systems: A simulator study of interaction modes and secondary tasks classes on eight production cars. *Int. J. Ind. Ergon.* **92**, 103380 (2022).
45. Zhang, T. *et al.* Input modality matters: A comparison of touch, speech, and gesture based in-vehicle interaction. *Appl. Ergon.* **108**, 103958 (2023).
46. Wang, Y. *et al.* The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics* **53**, 404–420 (2010).
47. Large, D. R., Pampel, S. M., Merriman, S. E. & Burnett, G. A validation study of a fixed-based, medium fidelity driving simulator for human-machine interfaces visual distraction testing. *IET Intell. Transp. Syst.* **17**, 1104–1117 (2023).
48. Chiang, D. P., Brooks, A. M. & Weir, D. H. Comparison of visual-manual and voice interaction with contemporary navigation system HMIs. *SAE Trans.* **114**, 436–443 (2005).
49. Mehler, B. *et al.* Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics* **59**(3), 344–367 (2016).
50. Reimer, B. *et al.* Patterns in transitions of visual attention during baseline driving and during interaction with visual-manual and voice-based interfaces. *Ergonomics* **64**(11), 1429–1451 (2021).
51. Cooper, J. M. *et al.* Age-related differences in the cognitive, visual, and temporal demands of in-vehicle information systems. *Front. Psychol.* **11**, 1154 (2020).
52. Harvey, C., Stanton, N. A., Pickering, C. A., McDonald, M. & Zheng, P. A usability evaluation toolkit for in-vehicle information systems (IVISs). *Appl. Ergon.* **42**, 563–574 (2011).
53. Kim, H., Kwon, S., Heo, J., Lee, H. & Chung, M. K. The effect of touch-key size on the usability of in-vehicle information systems and driving safety during simulated driving. *Appl. Ergon.* **45**, 379–388 (2014).
54. Kujala, T. Browsing the information highway while driving: Three in-vehicle touch screen scrolling methods and driver distraction. *Pers. Ubiquit. Comput.* **17**, 815–823 (2013).
55. Jung, S. *et al.* Effect of touch button interface on in-vehicle information systems usability. *Int. J. Hum. Comput. Interact.* **37**, 1404–1422 (2021).
56. Mitsopoulos-Rubens, E., Trotter, M. J. & Lenné, M. G. Effects on driving performance of interacting with an in-vehicle music player: A comparison of three interface layout concepts for information presentation. *Appl. Ergon.* **42**, 583–591 (2011).
57. Oviedo-Trespalacios, O., Haque, M. M., King, M. & Washington, S. Self-regulation of driving speed among distracted drivers: An application of driver behavioral adaptation theory. *Traffic Inj. Prev.* **18**, 599–605 (2017).
58. Miller, E. E., Boyle, L. N., Jenness, J. W. & Lee, J. D. Voice control tasks on cognitive workload and driving performance: Implications of modality, difficulty, and duration. *Transp. Res. Rec.* **2672**, 84–93 (2018).
59. Biondi, F. N., Getty, D., Cooper, J. M. & Strayer, D. L. Examining the effect of infotainment auditory-vocal systems' design components on workload and usability. *Transp. Res. Part F Traffic Psychol. Behav.* **62**, 520–528 (2019).
60. Kim, H. & Gabbard, J. L. Assessing distraction potential of augmented reality head-up displays for vehicle drivers. *Hum Factors* **64**, 852–865 (2022).
61. Graichen, L., Graichen, M. & Krems, J. F. Effects of gesture-based interaction on driving behavior: A driving simulator study using the projection-based vehicle-in-the-loop. *Hum Factors* **64**, 324–342 (2022).
62. Jung, T., Kass, C., Zapf, D. & Hecht, H. Effectiveness and user acceptance of infotainment-lockouts: A driving simulator study. *Transp. Res. Part F Traffic Psychol. Behav.* **60**, 643–656 (2019).
63. Onate-Vega, D., Oviedo-Trespalacios, O. & King, M. J. How drivers adapt their behaviour to changes in task complexity: The role of secondary task demands and road environment factors. *Transp. Res. Part F Traffic Psychol. Behav.* **71**, 145–156 (2020).

## Acknowledgements
The authors would like to thank all the participants in this field driving experiment.

## Author contributions
Conceptualization, Q. Z.; Data curation, Q. Z.; Funding acquisition, J. Z.; Investigation, Y. X., P. G., and S. F.; Methodology, Q. Z.; Resources, J. Z., Y. X., P. G., and S. F.; Supervision, J. Z.; Visualization, Y. X., P. G., and S. F.; Writing – original draft, Q. Z.; Writing – review & editing, Q. Z., and J. Z.

## Funding
This work was supported by the National Natural Science Foundation of China [Grant number 52175253], the National Key Research and Development Program of China [Grant number 2022YFB4301202-20], the Project of Sichuan Natural Science Foundation (Youth Science Foundation) [Grant number 22NSFSC0865], the Interdisciplinary Research Project of Southwest Jiaotong University [Grant number 2682023ZTPY042], and the New Interdisciplinary Cultivation Program of Southwest Jiaotong University [Grant number YG2022006].

## Competing interests
The authors declare no competing interests.

## Ethical approval
This study was conducted following the Declaration of Helsinki and approved by the Ethics Committee in the School of Design, Southwest Jiaotong University, China [Approval Number: 2022120904], where both authors were formerly affiliated.

## Informed consent
Written informed consent was obtained from all subjects for participation and publication of identifying information/images in an online open-access publication.

## Additional information
**Correspondence** and requests for materials should be addressed to Q.Z. or J.Z.