International Symposium on World Ecological Design
F. Ying et al. (Eds.)

doi:10.3233/FAIA240005

Research on the Usability of Automotive Human-Computer Interaction Interface System for Elderly Drivers

Yao XUE[a,1], Yueting JIANG[a], Jing CHAI[b], Guangdong LI[a], and Chao ZHANG[a]
[a] *Faculty of Art and Design, Anhui Institute of Information Technology, China*
[b] *Faculty of Art and Design, Hubei University of Arts and Science, China*

Abstract. In order to investigate the system availability of the human-computer interaction interface for elderly drivers during driving, this study employs questionnaires and in-depth interviews to identify common interactive tasks performed by elderly drivers while driving. Experimental models were selected, and elderly drivers were enlisted to partake in real vehicle tests. Data on task completion times during human-computer interactions were collected, and a multiple linear regression model was constructed by integrating the SUS scale. This model was used to analyze the impact of each independent variable's task completion time on the availability of the dependent variable system. Based on this foundation, additional screen samples were expanded for verification. Control performance across four typical vehicle screens was horizontally compared, and causal analysis was conducted using eye movement data. The findings reveal that the completion time of each human-computer interaction task for elderly drivers during driving significantly affects system availability, with the exception of air conditioning settings. Notably, the completion time of the navigation setting task exhibits the greatest impact on system availability, followed by broadcast radio. Observational speed and call-related tasks have a slightly smaller impact. The human-computer interaction system in vehicles equipped with dual-screens demonstrates the highest availability, suggesting its suitability for elderly drivers.

Keywords. Elderly drivers, System availability, Human-computer interaction interface, Car screen

1. Introduction

According to the latest data released by the Traffic Management Bureau of the Ministry of Public Security of China, the number of motor vehicle drivers in China had surged to an unprecedented 502 million by the end of 2022, thereby securing its position as the world's largest. Among these figures, the age distribution data of Chinese drivers in 2022 reveals a notable presence of over 75 million drivers aged 51 and above, constituting approximately 17%. Notably, the percentage of elderly drivers is experiencing an escalating year-on-year growth trend.In recent years, research pertaining to the aging effects on automobile-human-computer interaction has primarily concentrated within the realm of automobile information interface interaction

[1] Corresponding author. Yao XUE, Faculty of Art and Design, Anhui Institute of Information Technology, Wuhu 241000, Anhui Province, China; E-mail: xueyao1990@126.com

design. Predominantly, these studies have sought to harness the characteristics exhibited by the elderly demographic through static analyses of human-computer interactions, with the intent of refining the driving experience for this group. A paradigmatic instance is the work of Wang Xiguang and Fan Ying et al. [1]. which delved into the genuine challenges arising from aging in China. They probed the physiological and psychological attributes along with the travel requisites of elderly drivers, ultimately positing tailored navigation design strategies. In an innovative approach, Lv Shuai et al [2]. employed an evaluation model hinging on entropy weight methodology and grey correlation analysis to quantify elderly users' preferences and aptly assess interface layouts of information terminals. Their findings bear strategic implications for optimizing interface arrangements. Li Yongfeng et al. [3], adopting a research framework grounded in failure mode and effect analysis as well as fault tree analysis, have effectively evaluated and synthesized the operational dynamics of the car interface for elderly drivers. This effort has led to heightened contentment among elderly drivers employing the car interface. Meanwhile, Li Junhua et al [4]. dissected the intricacies of vehicle interface aging design by meticulously investigating both the deterioration of physical functionality and social cognitive faculties among the elderly. Through such an approach, they successfully elevated the safety and experiential facets of vehicle interfaces employed by elderly drivers during their journeys.

Nonetheless, the process of vehicle-human-computer interaction is characterized by its real-time, dynamic, and systematic nature, closely intertwined with the execution of driving tasks. The aforementioned research has predominantly centered on the static analysis of human-computer interaction indices within vehicular contexts. Investigations concerning the level of human-computer interaction in vehicles, along with its determining factors, have predominantly revolved around descriptive analyses of theoretical models, with quantitative and empirical data analyses and verifications receiving comparatively less attention. Against the backdrop of Internet technology's advancement propelling the evolution of vehicle connectivity systems, screen interactions and system visual designs undergo continual updates and iterations.

Consequently, this study employs a combination of questionnaires and in-depth interviews to glean insights into the common interactive task requirements of elderly drivers during car journeys. Experimental models have been judiciously selected, and elderly drivers were engaged in real-world vehicle tests. Temporal data corresponding to each interaction task were meticulously collected. The subsequent step involved constructing a multiple linear regression model that integrates the System Usability Scale (SUS) in order to scrutinize the relationship between completion times of independent variables' tasks and the resultant system's availability, the dependent variable. Consequently, this study employs a combination of questionnaires and in-depth interviews to glean insights into the common interactive task requirements of elderly drivers during car journeys. Experimental models have been judiciously selected, and elderly drivers were engaged in real-world vehicle tests. Temporal data corresponding to each interaction task were meticulously collected. The subsequent step involved constructing a multiple linear regression model that integrates the System Usability Scale (SUS) in order to scrutinize the relationship between completion times of independent variables' tasks and the resultant system's availability, the dependent variable.

2. Research methods

2.1. Interactive behavior analysis

In pursuit of understanding the prevalent interactive behaviors of elderly drivers engaged in vehicle-mounted human-computer interaction, this study undertook a comprehensive approach. A combination of questionnaire surveys and in-depth interviews was administered among a cohort of 100 elderly drivers aged over 60, all possessing more than two years of driving experience within the designated region. Statistical analysis of the collected data elucidated that the most frequently employed human-computer interaction tasks by elderly drivers during vehicle operation include speed observation, navigation configuration, air conditioning settings, phone calls, and radio playback.

Subsequent interviews yielded insights into the challenges frequently encountered by elderly drivers during these interaction episodes. Notably, these difficulties often manifest as operational errors, protracted task durations, and convoluted sequences of actions. These complexities potentially exert a bearing on the overall system availability of vehicle-mounted human-computer interaction. Motivated by these findings, the research deems it imperative to conduct real-world vehicle experiments aimed at delving into the intricate interplay between the aforementioned five interactive tasks and the system's overall availability.

The structural organization of papers must reflect a discernible, content-driven framework. Each paper must include a title along with a series of section headings. Additionally, references must be included at the document's conclusion. Figures and tables should be seamlessly integrated into the text. Please adhere to the guideline's prescription of employing a single space after sentence-ending periods. Conformity to the specified style guidelines within this compendium will expedite the subsequent adaptation of your work into InDesign templates, thus facilitating the publication process.

2.2. System availability experiment

The focus of this experiment centers on the secondary interactive tasks concurrently undertaken by elderly drivers during the primary driving task. The examination encompasses two primary metrics: the completion time data of each task (independent variable) and the availability score data of the System Usability Scale (SUS) system (dependent variable).

In the process of selecting an experimental model, the 2021 Volkswagen Maiteng was chosen as the test vehicle for both quantitative and qualitative evaluation. This choice adheres to several guiding principles: firstly, the vehicle's screen interface represents the prevailing mainstream standard, catering to a broad spectrum of drivers; secondly, the system interface embodies comprehensive functionality, along with representative architecture, interaction, and interface features. Lastly, the vehicle's aesthetics, driving experience, and overall comfort align closely with the preferences of the middle-aged and elderly driving demographic. As for participant selection, the study enlisted 20 elderly drivers aged 60 to 70 years, possessing good health and more than two years of accident-free driving experience. These participants were smartphone users accustomed to interacting with the in-car touchscreen.

For data collection, the experiment predominantly relied on the gopro panoramic camera, a device employed to capture both driving interaction times and images of interactive behaviors within the vehicle.

To ensure driver safety, the experiment was conducted on a wide, closed section of road during daylight hours, aiming to closely simulate real-world driving conditions and acquire accurate data. The comprehensive driving route spans approximately 3 kilometers, encompassing straight stretches, turns, U-turns, and stops at traffic lights (Table 1). The configuration of secondary tasks involved the selection of five aforementioned interactive tasks as independent variables, forming the basis for the experiment. To ensure seamless execution, participants were provided with a preliminary period to acquaint themselves with both the primary driving task and the vehicle's secondary tasks.

During the driving process, all five tasks were triggered by voice prompts at predetermined waypoints along the route. Under the guidance of trained staff, participants operated the test vehicle to successfully complete the designated tasks. Due to the extensive participant count and the relatively prolonged duration needed to gather data from the interactive tasks, the data collection phase spanned a total of 2 months. Throughout the testing period, variables such as driver height, seating posture, driving style, and weather conditions might minimally impact in-vehicle interaction behavior. However, these factors, on the whole, exerted negligible influence on the experiment, warranting their exclusion from the scope of this study.

In the course of investigating product user experiences, the System Usability Scale (SUS) has emerged as a pivotal instrument for assessing the efficacy of interactive product systems. Notably, within the domain of automotive human factors engineering, the SUS scale has demonstrated its pertinence in quantifying the efficacy of human-computer interaction interfaces in vehicles. A comprehensive study involving a large sample substantiates the SUS scale's robustness, revealing an impressive reliability coefficient of 0.91 alongside excellent internal consistency [5].

Consequently, this study capitalizes on the SUS scale to ascertain system availability. Post completion of a series of designated tasks, participants are prompted to complete the SUS scale, thereby providing a metric for their subjective evaluation. Additionally, the experiment meticulously records the time elapsed for participants to accomplish each specific task.

Table 1. Route and task arrangement

Route type	Task Description
Straight Driving	Task 1 : Observe the dashboard and say the current speed Task 2 : Set up navigation. From the current navigation to the specified destination Task 3 : Turn on the air conditioner and adjust the warm air to 27 °. Task 4 : Set the radio FM to FM90.8
Parking Waiting	Task 5 : Enter the specified phone number and make a successful call

3. Experimental Results and Analysis

3.1. Interactive behavior analysis

Given the diverse nature of tasks within this study, multiple independent variables collectively influence the system availability of the interactive interface, consequently impacting the interaction dynamics experienced by elderly drivers. To comprehensively address this complexity, the study employs a multiple linear regression model for data analysis. Before engaging in regression analysis, a preliminary assessment of the correlation between independent variables and the dependent variable is conducted.

As delineated in Table 2, the outcomes reveal that factors including observation speed, navigation setting, phone calls, and radio playback manifest significantly negative correlations at the 0.01 significance level. Notably, the correlation coefficients for navigation setting, phone calls, and radio playback range from 0.3 to 0.8, indicating a substantial degree of correlation. Conversely, air conditioning setting demonstrates significance at the 0.1 level with a coefficient of 0.2, suggesting a comparatively weaker correlation. Consequently, the imperative arises to further develop and validate a multiple linear regression equation to unveil deeper insights into the relationships.

Table 2. Analysis of Relationship

		Observation speed	Navigation settings	Air conditioning settings	Call the phone	Play radio station
SUS	pearson correlation	-0.519**	-0.733**	-0.246	-0.714**	-0.731**
	Sig.	0.000	0.000	0.085	0.000	0.000
	Number of cases	50	50	50	50	50
Note :**.At the 0.01 level (two-tailed), the correlation is significant.						

3.2. Model construction and verification

The multiple linear regression analysis was conducted using SPSS 26.0 software, with the resultant regression outcomes, employing the SUS score as the dependent variable, presented as follows:

Table 3. Correlation Coefficient of Regression Model

Model	Non-standardized coefficient		Standardized coefficient	T	Significance	Colinearity statistics	
	B	Standard error	Beta			Allowance	VIF
(Constant)	120.039	5.136		23.373	0.000		
Observation speed	-6.420	1.996	-0.248	-3.217	0.002	0.791	1.265
Navigation settings	-1.048	0.253	-0.387	-4.150	0.000	0.539	1.856
Air conditioning settings	0.961	0.499	0.147	1.926	0.061	0.806	1.240
Call the phone	-1.201	0.529	-0.228	-2.270	0.028	0.463	2.161
Play radio station	-0.840	0.228	-0.348	-3.687	0.001	0.526	1.901

Table 4. Model Fit

R	R^2	Adjusted R^2	Errors in standard estimates
0.891a	0.794	0.770	4.81063

Table 5: Analysis of Variance

Model	Quadratic Sum	Degree of Freedom	Mean Square	F	Significance
Regression	2487.738	5	497.548	15.834	0.000
Residual error	754.128	24	31.422		
Grand total	3241.867	29			

Initial observations from Table 3 reveal that the P-values corresponding to observation speed, navigation setting, phone calls, and radio playback all fall below the 0.05 threshold, indicating statistical significance at the 0.05 level. These findings further underscore a negative correlation: shorter task completion times for the aforementioned four activities correspond to higher SUS scores, reflective of an enhanced user experience pertaining to system availability. The culmination of these observations is encapsulated within the multiple linear regression equation, depicted as follows:

$$\hat{Y} = 120.039 - 6.42X_1 - 1.048X_2 + 0.961X_3 - 1.201X_4 - 0.84X_5$$

Taking the task completion time, X1, for observation speed as an illustrative example, it becomes evident that in the context of unchanging external factors, a 1% increase in task completion time precipitates a 6.42% decrease in the SUS value.

"Subsequently, the standardized coefficients for these four variables are as follows: observation speed (-0.248), navigation setting (-0.387), phone calls (-0.228), and radio playback (-0.348). These coefficients underscore the degree of influence on the SUS score. Notably, the completion time of the navigation setting task exerts the most pronounced impact on the SUS value, trailed by radio playback. Meanwhile, observation speed and phone calls demonstrate relatively milder effects. This pattern is attributable to the substantial demand among elderly drivers for navigation and radio station selection, tasks characterized by intricate operations and a pronounced influence on user experience. The observation speed task, conversely, involves straightforward information transmission, leading to simplified operation and shorter completion times, ultimately engendering a comparatively lower impact. For elderly drivers, the preference for completing phone call tasks through mobile devices leads to diminished screen-human interaction, modestly affecting system availability without reaching prominence.

Conversely, the commonplace function of making and receiving calls indirectly underscores the need for enhancing the current central control screen's phone function utilization. Intriguingly, the air conditioning setting's operation time does not significantly affect the SUS value. This phenomenon might be attributed to the test drive's air conditioning setting entailing a repeated interplay between screen interactions and the physical knob, thereby mitigating the observed impact on the SUS value.

Subsequently, the model undergoes a comprehensive diagnostic examination. As delineated in Table 4, the R^2 value of the adjusted model stands at 77%, signaling a high level of goodness of fit. This outcome reinforces the model's efficacy in elucidating variations in SUS values based on the completion times of the five tasks. Referring to Table 5, the model is endowed with a highly significant fit, as evidenced by an F-statistic of 15.834 with a p-value of less than 0.001.

Visual inspection, as illustrated in Fig. 1, reveals that the residuals for the majority of independent variables adhere to a normal distribution, with scatter points predominantly dispersed around the diagonal within the first quadrant. This distribution pattern substantiates the conformity of residuals to the normal distribution.

In conclusion, the empirical results garnered from the diagnostic evaluations substantiate their validity and robustness.

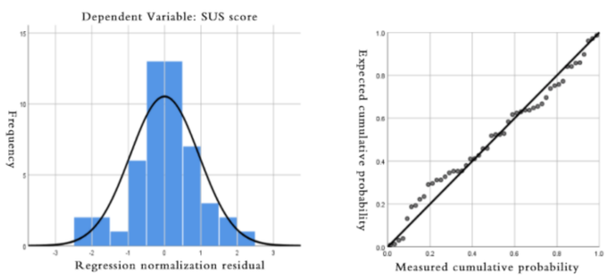

Figure 1. Histogram and normal probability plot of standardized residuals (P-P plot).

4. Multi-sample Test

Through the above research, it can be known that except that the completion time of the air conditioning task and the SUS value are not significant, the other independent variables show a significant negative effect, and the insignificance of the air conditioning task may be caused by the operation error, which does not mean that there is no effect at all. Therefore, all five tasks need to be included in the subsequent design considerations. Due to the limited representativeness of the screen style, interaction mode and interface design of the experimental vehicle. Therefore, based on the model conclusions, to explore the impact of changes in the screen interaction interface of different models on the task completion time of elderly drivers, it is necessary to expand the vehicle sample for further testing and horizontally compare the effectiveness of the interaction interface with typical screen features in the market.

4.1. Experimental content and data acquisition

In the expansion of the sample, a questionnaire was set up to count the purchase rate and satisfaction of the elderly driving population, and 10 models with the highest satisfaction were selected. Four models with typical screen interaction characteristics were selected as test driving. They are suspended horizontal screen(a), embedded horizontal screen(b), dashboard-central control double screen(c) and vertical screen in turn(d). The experiment used a real car but only did a static experience this time.20 elderly drivers with rich driving experience were selected to conduct a screen interaction experiment in the car. Existing studies have shown that the process of human-computer interaction is also accompanied by search performance and cognitive load. The more the number of fixations, the more inefficient the search, the greater the cognitive load, and the larger the pupil diameter [6]. Therefore, in addition to collecting the completion time of each task and SUS score data, this experiment also carries an eye tracker to collect the average pupil diameter and observation times of elderly drivers to complete each task, so as to facilitate multi-aspect evaluation, see Tables 6,7,8.

Table 6. Four types of screen SUS scores and average time units for various tasks (seconds)

Screen type	Observation speed	Navigation settings	Air conditioning settings	Call the phone	Play radio station	SUS score
a	1.53	20.19	7.23	5.91	17.16	85.208
b	2.55	18.15	6.91	5.61	15.42	86.3479
c	1.48	22.63	6.34	6.14	10.53	87.23956
d	2.23	26.48	6.32	7.23	18.08	73.06945

Table 7. Mean pupil diameter at completion of four screens (mm)

Screen Type	Observation speed	Navigation settings	Air conditioning settings	Call the phone	Play radio station
a	2.95	3.32	3.51	3.08	3.22
b	3.13	2.67	2.55	3.31	3.32
c	2.61	2.93	3.18	3.18	3.21
d	3.46	3.45	3.03	3.45	3.78

Table 8. Average number of observations on four screens (times)

Vehicle Model	Observation Speed	Navigation Settings	Air Conditioning Settings	Call the Phone	Play Radio Station
a	1.78	3.52	4.23	7.31	4.08
b	1.97	3.02	3.12	6.24	4.51
c	1.83	3.13	3.09	6.22	3.58
d	2.45	3.61	3.16	7.23	4.28

4.2. Results and Discussion

First of all, among the four typical screens, according to the SUS score, the order is suspended horizontal screen, embedded horizontal screen, double screen and vertical screen. It can be known that for the elderly driver, the dual screen system has the highest availability and the best experience.

Secondly, the analysis of each sub-task is as follows: In the navigation setting task, the elderly driver completes the navigation task on the embedded horizontal screen for 18.15 s, which takes far less time than other models, and has a small pupil diameter and the least number of observations. It can be known that the task load is low and the completion efficiency is high. The main reason is that the navigation function in the vehicle interface is streamlined and the priority is the highest. It occupies a large click area in the screen and is close to the driver 's line of sight to facilitate the driver to set the navigation task. The navigation functions of other models are mostly arranged on the second or third level, and the number of operation steps is more, which prolongs the search time to a certain extent.

In the broadcast radio task, the elderly driver spends the least time on the dual screen, which is 10.53 s, and the pupil diameter is small and the number of observations is the least. It can be known that the task load is low and the operation efficiency is high. The reason may be that the touch control method is simple and easy to use, and the frequency modulation switching is carried out by button clicking, which is in line with the habits of elderly drivers, and the layout of the interface is prominent, the primary and secondary are clear, and the search efficiency is also improved. In the speed observation task, the time spent by the elderly driver on the dual screen was 1.48 s, which was significantly faster than that of other remaining models, and the pupil diameter was small and the number of observations was small. The reason may be that the dashboards carried in this model are all analog digital displays and are in front of the driver, and the remaining models are displayed using traditional pointer scales. For older drivers, pure digital displays have a higher degree of visualization, and digital displays are superior to pointer scale displays in recognition efficiency and cognitive load. Although the suspension screen displays the speed per hour, it may affect the observation performance because it is not in front of the driver 's line of sight ; in the call task, the time spent on the embedded screen is 5.61 s, which is significantly smaller than that of other models and the number of observations is small. It can be concluded that this task has the highest search efficiency.

By comparing the screen interaction and interface design of several vehicles, the advantages of the interaction layer and the visual layer are also obvious. The specific performance is that the digital dial interaction button has a large size, high color contrast, and obvious feedback during the interaction process. Compared with the above tasks, the air conditioning setting task shows little difference in several models. From the eye movement trajectory and the collected images, it can also be known that the driver observes the screen less, which reflects the driver 's operating habits to a certain extent, and the elderly driver 's operating experience plays a leading role in the human-computer interaction interface. Finally, the tasks and their influencing factors are summarized with the corresponding modules, see Table 9, and provide the basis for the design strategy.

Table 9: The influencing factors of various tasks and their corresponding module relationships

All the Tasks	Influencing Factor	Corresponding Module
	Functional division	Functions
Navigation Settings	Level setting	Functions
	Text, icon, button recognition	Interface
	Interactive Form	Interaction
Play Radio Station	Touch Mode	Interaction
	Icon Characteristics, Image Degree	Interface
	Text, Icons, Button Size	Interface
Call the Phone	Color, Contrast	Interface
	Interactive Feedback	Interaction
Observation Speed	Information Visualization	Interface
Air Conditioning Settings	Experience, Habits	Interaction, Interface

5. Results, Discussion and Future Considerations

1. Apart from air conditioning settings, the completion time of various human-computer interaction tasks by elderly drivers during driving significantly affects system availability. Among them, the completion time of navigation setting tasks has the greatest impact on system availability, followed by the influence of radio playback, with slightly less impact observed in observing speed and making phone calls.

2. The vehicle model equipped with a dual-screen human-computer interaction system achieves the highest level of availability. Furthermore, a detailed summary of the influencing factors and corresponding module relationships for each task has been provided.

Acknowledgments.

We thank the Humanities and Social Science Fund of Ministry of Education (22YJC760110) Research on Age-Friendly Design of Automobile Human-Machine Interaction Interfaces.

References

[1] Fan Ying, Wang Xiguang, Design of human-machine interface for vehicle navigation devices suitable for middle-aged and elderly people, Heilongjiang Science and Technology Information 16 (2009),17-20.

[2] Lv Shuai, Li Yongfeng, Research on the layout evaluation of elderly information terminal interfaces based on entropy weight and grey correlation , Packaging Engineering44(2023),128-136.

[3] Li Yongfeng, Chen Zeyan, Research on the Interaction design of the elderly car human-computer interface based on FMEA and FTA , Packaging Engineering42(2021), 98-105.

[4] Li Junhua, Shen Keyu, Xu Yuan, Zhou Yingbin, Research on aging friendly design of vehicle human-computer interaction interface, Packaging Engineering44(2023),145-151.

[5] Li R.,Chen Y. V.,Sha C.F.,Effects of Interface Layout on the Usability of In-Vehicle Information Systems and Driving Safety, Displays,49(2017),124-132.

[6] Wang Chao, Yu Chaobo,Comprehensive assessment of Cognitive load of air traffic controllers based on multiple physiological parameters,Science, Technology and Engineering 14(2014), 295-300.