IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 July 2024, accepted 3 August 2024, date of publication 6 August 2024, date of current version 19 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3439553

RESEARCH ARTICLE

Eye Tracking Study on Visual Search Performance of Automotive Human–Machine Interface for Elderly Users

SONGMAN LI AND SONG HAO
Department of Waikato, School of Mechanical Engineering, Shandong University, Jinan 250061, China
Corresponding author: Song Hao (haosong@sdu.edu.cn)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Academic Committee of the School of Mechanical Engineering, Shandong University.

ABSTRACT The advancement of Intelligent vehicle is leading to a growing prevalence and importance of automotive interactive interfaces, attracting considerable research focus. With the increasing trend of global aging, the number of elderly drivers is on the rise. Research on the search performance of the automobile interactive interface for elderly users has yet to be carried out. Based on this, this research adopts eye tracking technology to explore the effects of varying icon colors on the preferences and perceptions of elderly drivers within car-human interfaces, employing eye-tracking technology to gauge their impact. This work aims to improve driving experience by enhancing drivers' information processing capabilities and interaction comfort with in-car interfaces. In this study, we examined six distinct foreground colors against two background colors in icon designs, conducting eye-tracking experiments in standard indoor lighting. The analysis results show that elderly drivers have faster search speeds for orange icons and the slowest search speeds for yellow icons. Additionally, the variation in search times for different icon colors is more pronounced on a white background. These conclusions hold significant implications for future automotive interface designers. They can leverage these results to optimize the background and icon designs of interactive interfaces, thereby enhancing drivers' safety and driving experience, and contributing efforts to the transportation industry.

INDEX TERMS Human-computer interface, automobile, elderly users, visual search performance, eye tracking analysis.

I. INTRODUCTION

With the continuous emergence of advanced technologies, the field of automotive design is undergoing constant transformation, with increasing depth in the research of human-machine interaction (HMI) [1]. Due to the widespread adoption of autonomous driving technology and smart cockpits, large interactive screens will become standard components in future vehicles. Designers will focus on enhancing passenger comfort and in-vehicle experience [2], heralding a new era of driving experience. Simultaneously, the aging population is leading to an increasing number of elderly drivers [3], posing challenges for them in using smart screens. Consequently, user-centered automotive HMI design becomes crucial [4]. Elderly drivers often face numerous challenges while driving, such as slower reaction times and diminished vision [5], making them more prone to accidents than younger individuals [6], [7], [8]. With the advent of new vehicles, more mechanical control systems are being replaced by digital interfaces, leading to increased operational demands and cognitive load, thereby making driving more challenging for the elderly. The continuous increase in the number of elderly drivers highlights the urgency of research into age-friendly automotive interface design. Designers need to cater to the advancements in modern automotive technology while considering the special needs and limitations of the elderly, and

The associate editor coordinating the review of this manuscript and approving it for publication was Jingang Jiang.

explore suitable design principles to enhance their safety and comfort on the road [4].

Flöck et al. pointed out in 2008 that one of the future challenges is designing cars that meet the increasing specific needs of elderly drivers [9]. Despite elderly drivers not making up the majority of drivers, the rate of traffic accidents involving them remains disproportionately high [6], [10] This has been linked to the increased complexity of vehicle interface interactions, raising concerns over the elderly drivers' ability to adapt to the modern, increasingly complex digital interfaces of cars. In this context, enhancing the design of smart screens becomes particularly crucial, with a focus on improving accessibility and usability for elderly drivers. Significant progress has been made in research on vehicle interface design for elderly drivers, identifying key areas of difficulty and proposing various solutions. May et al. [11] focused on the elderly in their interface research, discovering significant differences between elderly and young drivers in providing landmarks in instructions provided by in-car navigation systems. They further discussed the impact of these results on the design of car interfaces for elderly drivers, laying the foundation for studying more complex HMI interfaces. Mourant et al. [12] and others found that elderly drivers need longer eye accommodation to retrieve information from vehicle displays and are less accurate, highlighting the importance of reducing the search load on the elderly through interface design. Jung et al. [5] emphasized the importance of designing systems that align with the sensory, cognitive, and physical abilities of elderly drivers, suggesting that car interface research for the elderly needs to consider their sensory capabilities more, exploring cognitive interfaces suitable for elderly users from the interface. Furthermore, some scholars have pointed out that automotive systems are rapidly transitioning from pure mechanization to intelligent assistants. However, the increasing complexity of these systems not only fails to provide support but also adds pressure and increases workload for drivers [13. Therefore, from the perspectives of usability and user experience, it is crucial to understand the impact of age-related workload differences. Young et al. [14] revealed that existing automotive HMI guidelines have not fully addressed the issues brought about by aging, emphasizing the importance of considering the limitations and capabilities of elderly drivers in the design process. On this basis, we reviewed a large number of relevant literatures on many factors of HMI usability, among which color is considered a key factor affecting information search and cognitive load [15], [16]. Appropriate use of color properties can not only alleviate visual fatigue but also enhance the visual appeal of the interface and the user's recognition speed [17]. Studies show that with aging, the human visual system's sensitivity to color decreases [18], reducing elderly people's ability to distinguish certain colors [19], thereby affecting the ability of elderly drivers to quickly and accurately process information displayed on vehicle-mounted screens [20]. Although the importance of HMI has been recognized, the ability for elderly users to interact quickly and accurately with automotive HMIs through color has not been fully explored. Therefore, considering the critical role of visual cues in the overall usability of digital interfaces inside vehicles, especially for users with declining vision and reaction capabilities, addressing this oversight becomes crucial.

In the visual search process of icons, color plays a crucial role, guiding users' visual behavior [21]. Studies on elderly individuals' color recognition have been extensive. Research indicates that the aging process typically leads to a decline in vision and color discrimination ability, making colors like blue and purple harder to perceive, while colors like red and yellow remain relatively stable [22], [23]. Some studies have also categorized elderly participants by gender, investigating differences in color discrimination, color naming, and color preference, ultimately uncovering gender-related disparities [24]. Subsequent research began to combine color with icons, advancing applications on smartphones [25]. However, these studies have not been applied to automotive interfaces. Indeed, research on the role of color in automotive HMI design has been gradually increasing in recent years. You et al. [26] tested three commonly used color and five contrasts on the vehicle interface, collecting driving behavior and user experience data to inform the color design of vehicle dashboards. Research has also explored how neural network systems for visual recognition and user experience analysis can investigate the color design of intelligent vehicle Human-Machine Interface (HMI). This research aims to meet users' physiological and psychological needs, enhancing their aesthetic experience of interface colors and providing conditions for designers and companies to improve HMI interfaces [27]. Xu et al. [28] found that high-contrast blue purple/cyan color combinations and icons with a high aspect ratio are more suitable for in-car human-machine interfaces. These findings contribute to enhancing drivers' information processing efficiency and comfort in various driving environments. However, while existing research on HMI design considers broad user demographics and specific color styles, it often overlooks the cognitive preferences and practical applicability for the elderly. Given the crucial role of customized, age-friendly interface design in ensuring the safety of elderly drivers, research in this area is essential.

Based on the above research status, this study focuses on enhancing the usability of automotive HMI designs for elderly drivers, specifically examining how icon colors [29], [30] and background colors affect identification speed and fatigue levels. By reducing cognitive strain and improving recognition efficiency, the aim is to make vehicle interactive interfaces more accessible to older adults. Utilizing eye-tracking technology [31], the study gathers data on the automotive HMI, with search time [32], [33], [34] and pupil diameter [35], [36] serving as quantitative metrics for assessing search performance. These metrics help evaluate the effectiveness of interactive interfaces. A mixed-method approach, combining quantitative evaluation and qualitative feedback, is employed to assess the performance and user experience of different color schemes on smart automotive screens.

II. METHODS AND MATERIALS

A. PARTICIPANTS

This study recruited a total of 46 participants, aged between 55 and 70 years, with a gender ratio of 31 males to 15 females (as males typically have a higher rate of traffic accidents [37]). All participants underwent vision screening to ensure their eyesight was normal or corrected to normal. Measurements were taken under the participants' ''habitual correction,'' meaning they wore the glasses or contact lenses they usually used in the designated viewing distance [38]. No subjects reported ongoing ophthalmic treatments or eye diseases (detected or declared) [39]. Participants were required to have no habits of smoking or excessive alcohol consumption and to have the ability to concentrate. Additionally, participants were required to be right-handed, have normal vision or corrected-to-normal vision, no significant astigmatism, and no cognitive impairments. In order to adhere to ethical standards, all participants filled out informed consent forms and all experiments were executed with their consent.

B. MEASUREMENT

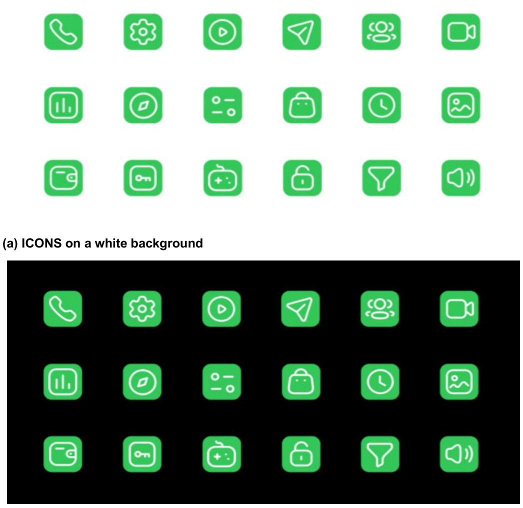
Participants sat at a distance of 24 inches from the monitor (1920 × 1200 pixels) at a distance of 45 centimeters (68.2 ° viewing angles). The Tobii Pro X3-120 eye tracker (sampling rate: 30 Hz) was mounted on the monitor.

C. STIMULI

The stimuli for the study consisted of 18 common operational icons found on car dashboard interfaces. All images were displayed on a screen with a resolution of 1920 × 1200 pixels. These icons were commonly encountered in daily life and possessed clear meanings during the experiment. To ensure the accuracy of the experiment, we processed these icons to ensure that each icon appeared on different colored backgrounds. Each icon had a size of 120 pixels × 120 pixels, with foreground colors of red, green, blue, yellow, orange, and purple, while the overall background of the images was set to black and white. The foreground colors of the 18 icons used in the experiment were paired with the overall background as shown in Table 1. These 18 icons were randomly arranged into three rows, with each row containing 6 icons. The distance between adjacent icons within each image was set to 45 millimeters, and the vertical center-to-center distance was 100 millimeters. The final layout of the complete icon interface is depicted in Figure 1, with all images stored in JPEG format as bitmaps.

TABLE 1. The foreground color of 18 ICONS is matched with the overall background.

Icon color rendering mode					
T1	T2	T3	T4	T5	T6
Linear red + white backgro und	Linear green + white backgro und	Linear blue + white backgro und	Linear yellow + white backgro und	Linear orange + white backgro und	Linear purple + white backgro und
T7	T8	T9	T10	T11	T12
Linear red + black backgro und	Linear green + black backgro und	Linear blue + black backgro und	Linear yellow + black backgro und	Linear orange + black backgro und	Linear purple + black backgro und



(a) ICONS on a white background

(b) ICONS on a black background

FIGURE 1. Icon layout and position display.

1) ICON STYLES

To investigate the impact of foreground/background color combinations of icons on visual search efficiency during driving tasks on real automotive manufacturer in-vehicle displays, we selected 18 common icons from the automotive manufacturer in-vehicle displays [40]. These icons were then converted into linear icons [41]. The final style of icons used in the experiment is shown in Figure 2. Additionally, to ensure the accuracy of the experiment, we set the target search icons to different icons. In the experiment, we provided users with the interface for searching targets as shown in Figure 3.

2) COLOR COMBINATION

To ensure the referential integrity of the experimental data, the colors for the 18 icons were all selected from the default palette provided by the CarPlay system. These colors were carefully chosen to maintain consistency and avoid any external influences that might arise from using non-standard colors. The icons in different colors were grouped

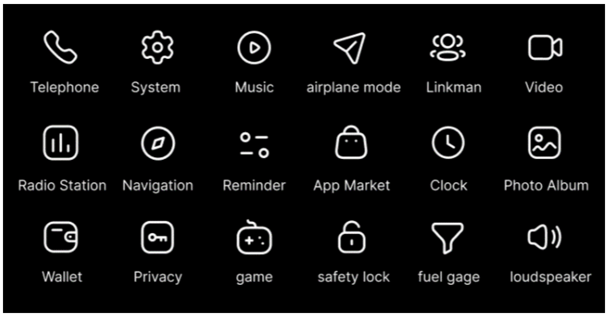


FIGURE 2. Display of icon graphic types (combined with the Tesla brand for effect processing).

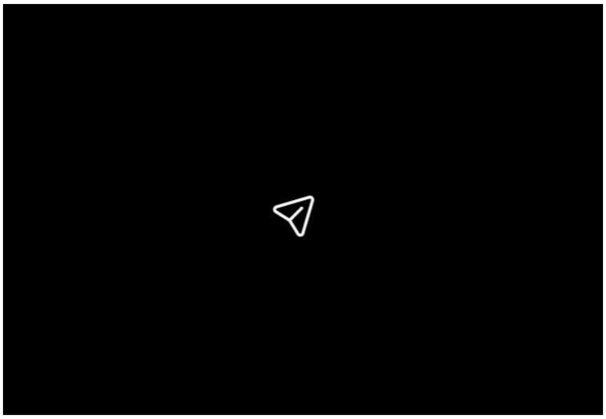


FIGURE 3. Interface for remembering target ICONS.

and contrasted against black and white backgrounds to examine their visibility and readability in varying conditions. The six colors chosen for the experiment are red, orange, yellow, green, blue, and purple. These colors represent a broad spectrum and are commonly used in various interface designs, providing a comprehensive basis for the study. The results of this research are detailed in Table 4, which includes the RGB values for each color. The RGB color model is an industry standard that creates various colors through variations and combinations of the red (R), green (G), and blue (B) color channels. This model encompasses almost all the colors perceivable by human vision and is one of the most widely used color systems.

D. EXPERIMENTAL DESIGN

The independent variables in this study include icons of different colors and background colors, both manipulated within subjects. Participants were tasked with performing search tasks targeting the search objectives on each image. Therefore, each experiment consisted of search tasks for a given image, with each participant completing a total of 72 experiments.

In Experiment 1, participants conducted 36 trials using icons of different colors against a white background, with each color having 6 sets (the order of experiments for the 6 colors was randomized). Following Experiment 1, participants took a 5-minute break.

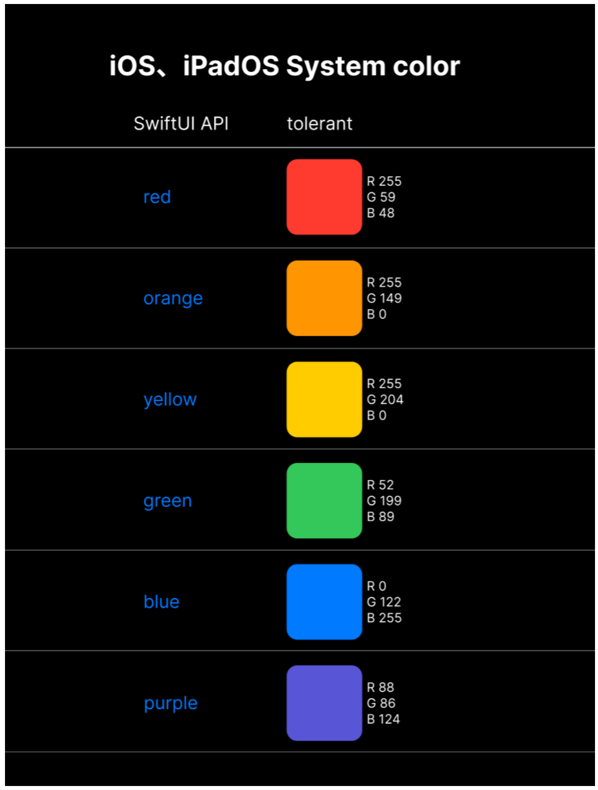


FIGURE 4. The six carplay system colors are the icon colors in the experiment.

In Experiment 2, participants conducted 36 trials using icons of different colors against a black background, with each color having 6 sets (the order of experiments for the 6 colors was randomized).

During the experiment, after each participant completed a trial, a gray blank screen appeared for 2 seconds to help the pupil diameter return to a relaxed state. Afterward, the participant proceeded to the next trial. In each trial, initially, an instruction page containing text appears for 5 seconds, informing participants of the target icon they need to search for in that trial. Then, an image with icons appears. At this point, participants begin searching for the target icon, with no time limit imposed on the search process, until they complete the search. The images in each set of trials are randomly ordered.

E. DEPENDENT VARIABLES

The dependent variables include search time and cognitive load. In this study, search time refers to the time taken from when the eyes move from the fixation area to completing the operation during the search process. It is generally believed that the size of the pupil diameter is related to the user's emotions and cognitive load. Typically, larger pupils indicate higher cognitive load [42], [43], [44], [45]. Any trials containing errors are not included in the calculation of search time, fixation duration, and fixation count. Two types of errors are considered: miss errors (selecting the wrong target) and abandon errors.

F. EXPERIMENTAL PROCEDURE

Before the experiment begins, participants are required to complete a personal information questionnaire (demographic information) and sign an informed consent form. They are then briefed on the relevant matters and procedures of the experiment. After completing 5 training tasks, the formal experiment begins.

First, participants were informed that they would be undergoing formal trials. All participants completed two control experiments: recognizing six colored icons against a black background and recognizing the same colored icons against a white background. This setup allowed for an in-group comparison of icon color data and enabled us to investigate the effect of background color on the experiment. There was no time limit for each trial, and a 5-minute break was provided between the two sessions. Subsequently, the eye tracker is calibrated using a nine-point grid. In our study, each target icon within the complete icon interface was designated as an Area of Interest (AOI). Participants gazed at the screen to locate the target icon, while the eye-tracker's camera continuously captured images of their eyes. By analyzing these sequentially captured images, the eye-tracker was able to pinpoint the exact locations of the gaze, assigning a timestamp to each captured fixation. Comparing the timestamps of consecutive fixation points allowed us to calculate the time spent by users moving between two points of interest. The captured images were then sent to a processing unit, where specific algorithms were employed to identify the pupil and corneal reflection points within the images. By comparing the size of the pupil against a known scale, the software was able to accurately calculate the pupil's diameter.

At the beginning of the experiment, participants must observe and memorize the target icons on the in-car display screen. When participants observe a complete icon interface on the in-car display screen (consisting of 18 icons arranged in a 3 rows $\times$ 6 columns format), they immediately begin the search. When participants find the target icon, they must click the mouse immediately to complete one task and wait for the next set of task targets to appear. This experimental process is repeated until the experiment is completed.

After completing the visual search tasks, participants are required to fill out a questionnaire to investigate their subjective feedback. They are asked to fill out a subjective feedback questionnaire based on their experimental experience. They need to rank their experimental experience on a nine-point Likert scale ranging from "describes very poorly (= 1)" to "describes very well (= 9)." This questionnaire includes basic information about the participants and their visual comfort. Throughout the entire experimental process, the average duration for each participant is approximately 20 minutes.

III. RESULTS

A. SEARCH TIME

In this study, we initially conducted a normality test on the reaction time data. Table 2 presents the results of the normality test, indicating that all groups had p-values greater than 0.05 according to the Jarque-Bera test. Consequently, the search time data for both the white and black background groups followed a normal distribution.

Subsequently, we performed a two-way analysis of variance (ANOVA) to examine the effects of icon color and background color on search time. Table 3 shows the ANOVA results, revealing that icon color significantly impacted search time ($F = 2.605$, $p = 0.034 < 0.05$). This suggests that different icon colors significantly affect the search performance of elderly participants on both light and dark backgrounds.

We also plotted the average search time for different icon colors on white and black backgrounds for elderly participants. Figure 5 illustrates the data for six colors on both backgrounds, with dashed lines representing the black background and solid lines representing the white background. The x-axis denotes the six icon colors, while the y-axis indicates the average search time for each color group on the white and black backgrounds.

From Figure 5, it can be observed that the range of fluctuations of the six colors is greater on the white background compared to the black background. The average search time for orange was the shortest on both backgrounds, whereas yellow had the longest average search time. This indicates that the interaction between background color and icon color has a differential impact on the visual search efficiency of elderly participants.

B. PUPIL DIAMETER

In this study, we initially conducted a normality test on the reaction time data obtained from the experiment. Table 4 presents the results of the normality test, indicating that all groups had p-values greater than 0.05 according to the Jarque-Bera test. This finding suggests that the data for average pupil diameter under both white and black background conditions followed a normal distribution.

Subsequently, a two-factor analysis of variance (ANOVA) was employed to investigate the impact of icon color and background color on average pupil diameter. Table 5 presents the variance analysis results, which revealed that color significantly affects mean pupil diameter ($F = 8.633$, $p = 0.000 < 0.05$). Additionally, background color also significantly impacts mean pupil diameter, underscoring the importance of our study on color.

The objective of this study was to assess the fatigue level of elderly individuals in recognizing different colors using pupil diameter as an indicator of cognitive load. Line graphs were constructed using the mean pupil diameter data obtained under both white and black background conditions. The dashed line represents the black background, while the solid line represents the white background. The horizontal axis represents the six icon colors, while the vertical axis depicts the average pupil diameter for each color group under both background conditions.

Figure 6 depicts the mean pupil diameter of elderly individuals for different colors under black and white

TABLE 2. Normal distribution test data under two background colors.

Table 2 (a) white background

Search Time											
Item	Sample Size	Mean	Standard Deviation (SD)	Skewness	kurtosis	Kolmogorov-Smirnov test		Shapiro-Wilk test		Jarque-Bera test	
						(D Value)	(p Value)	(W Value)	(p Value)	(χ2)	p Value
Undistinguished	36	3.755	1.235	0.867	1.532	0.097	0.538	0.956	0.163	6.179	0.046
Red	6	3.612	0.923	1.175	1.298	0.232	0.413	0.912	0.448	0.779	0.677
Yellow	6	4.995	1.681	0.674	-0.500	0.188	0.738	0.953	0.768	0.507	0.776
Orange	6	3.192	0.338	0.782	-1.734	0.324	0.049	0.823	0.094	0.853	0.653
Blue	6	3.935	1.443	-1.021	-0.618	0.284	0.151	0.859	0.186	0.842	0.656
Green	6	3.542	1.067	-1.126	-0.119	0.348	0.023	0.835	0.118	0.878	0.645
Purple	6	3.252	1.009	0.373	-1.652	0.220	0.490	0.893	0.332	0.581	0.748

Table 2 (b) black background

Search Time											
Item	Sample Size	Mean	Standard Deviation (SD)	Skewness	kurtosis	Kolmogorov-Smirnov test		Shapiro-Wilk test		Jarque-Bera test	
						(D Value)	(p Value)	(W Value)	(p Value)	(χ2)	p Value
Undistinguished	36	3.537	0.768	-0.269	0.923	0.095	0.558	0.972	0.483	1.008	0.604
Red	6	3.632	0.622	-0.035	-0.428	0.124	0.990	0.989	0.986	0.253	0.881
Yellow	6	3.899	0.973	0.547	1.810	0.242	0.356	0.949	0.733	0.174	0.917
Orange	6	3.041	0.817	-0.864	1.259	0.275	0.186	0.897	0.359	0.443	0.801
Blue	6	3.611	0.624	0.157	-1.584	0.201	0.638	0.920	0.505	0.503	0.778
Green	6	3.668	0.848	-1.703	3.808	0.359	0.017	0.817	0.083	1.597	0.450
Purple	6	3.372	0.701	-1.441	3.153	0.317	0.061	0.866	0.209	1.121	0.571

TABLE 3. The influence of color and background color on SearchTime.

Factorial ANOVA Results					
Source of Variation	Sum of Squares	df	Mean Square	F	p
Intercept	957.046	1	957.046	987.216	0.000**
Color	12.626	5	2.525	2.605	0.034*
BackgroundColor	0.852	1	0.852	0.879	0.352
Color*BackgroundColor	3.227	5	0.645	0.666	0.651
Residual	58.166	60	0.969		
R 2: 0.223					
* p<0.05 ** p<0.01					

backgrounds. The experimental data from both groups indicate that the pupil diameter for the six icon colors is generally higher under the black background compared to the white background.

C. USER QUESTIONNAIRE

1) COLOR PREFERENCE

The results of the study indicate that, with regard to background color, 61.7% of participants expressed a preference

TABLE 4. Normal distribution test data under two background colors.

(a) white background

Search Time											
Item	Sample Size	Mean	Standard Deviation (SD)	Skewness	kurtosis	Kolmogorov-Smirnov test		Shapiro-Wilk test		Jarque-Bera test	
						(D Value)	(p Value)	(W Value)	(p Value)	(χ2)	p Value
Undistinguished	36	2.498	0.028	0.875	-0.015	0.183	0.004	0.906	0.005	4.263	0.119
Red	6	2.489	0.035	1.454	2.786	0.313	0.067	0.853	0.167	1.130	0.568
Yellow	6	2.489	0.010	0.289	-1.191	0.171	0.850	0.953	0.764	0.445	0.801
Orange	6	2.477	0.016	1.139	1.441	0.205	0.608	0.921	0.511	0.725	0.696
Blue	6	2.510	0.022	-0.597	-1.945	0.217	0.512	0.843	0.139	0.771	0.680
Green	6	2.497	0.027	1.960	4.169	0.340	0.031	0.755	0.022	2.131	0.345
Purple	6	2.525	0.032	0.058	-3.056	0.280	0.168	0.797	0.055	0.909	0.635

(b) black background

Search Time											
Item	Sample Size	Mean	Standard Deviation (SD)	Skewness	kurtosis	Kolmogorov-Smirnov test		Shapiro-Wilk test		Jarque-Bera test	
						(D Value)	(p Value)	(W Value)	(p Value)	(χ2)	p Value
Undistinguished	36	2.843	0.053	0.249	-0.801	0.124	0.186	0.960	0.219	1.441	0.486
Red	6	2.829	0.028	-0.071	-2.101	0.223	0.475	0.928	0.566	0.625	0.732
Yellow	6	2.788	0.035	-0.101	-2.918	0.279	0.172	0.827	0.102	0.868	0.648
Orange	6	2.818	0.037	0.134	-1.714	0.179	0.797	0.917	0.487	0.531	0.767
Blue	6	2.874	0.053	-0.309	-1.526	0.207	0.591	0.908	0.426	0.528	0.768
Green	6	2.850	0.048	0.889	-1.036	0.314	0.065	0.866	0.211	0.789	0.674
Purple	6	2.902	0.036	-0.757	-1.173	0.225	0.458	0.903	0.395	0.702	0.704

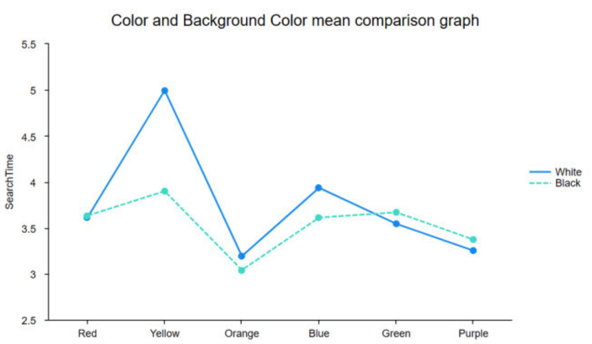


FIGURE 5. Average search time for different colors on black and white background.

for white, while 38.3% indicated a preference for black. These findings are illustrated in Figure 7. Figure 8 presents the preferences of all participants for six icon colors. On a white background, the preferences for red, yellow, blue, orange, green, and purple were 23.91%, 0%, 21.74%, 6.52%, 8.69%, and 2.17%, respectively. On a black background, the preferences for the same colors were 4.34%, 8.69%, 8.69%, 2.17%, 10.87%, and 2.17%, respectively.

In the context of a white background, red was identified as the most preferred color, while yellow was the least preferred. In contrast, on the black background, green was identified as the most preferred color, while orange was the least preferred. These subjective data are inconsistent with the previous physiological data, underscoring the necessity of this research.

2) SATISFACTION DATA

After the experiment, we conducted a subjective evaluation survey with the users, as shown in Table 6 (located in the Appendix A). All participants were asked to complete a questionnaire to indicate their experiences and feelings during the experiment.

The ratings were based on interest, perception, effort, pressure, and value. From the final proportions indicated by the affirmative values, it is evident that all five aspects received high affirmative ratings, suggesting a high level of user cognition across these dimensions.

TABLE 5. Analysis of variance results of color and background color for mean pupil diameter.

Factorial ANOVA Results					
Source of Variation	Sum of Squares	df	Mean Square	F	p
Intercept	513.537	1	513.537	453890.967	0.000**
Color	0.049	5	0.010	8.633	0.000**
BackgroundColor	2.149	1	2.149	1899.778	0.000**
Color*BackgroundColor	0.011	5	0.002	1.927	0.103
Residual	0.068	60	0.001		
R 2: 0.970					
* p<0.05 ** p<0.01					

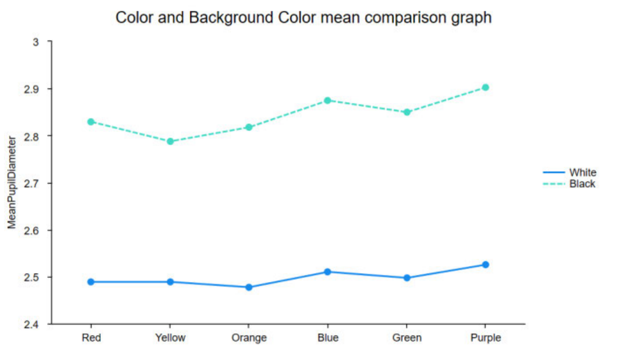


FIGURE 6. Average pupil diameter for different colors in the elderly group against a black and white background.

IV. DISCUSSION

A. SEARCH TIME

This study employed a rectangular layout based on 18 evenly distributed positions from existing automotive interface designs. Experiments were conducted using black and white background groups, with six different colors paired with either a white or black background for each experiment. Each group included six icons for each of the six colors, arranged in a cycle to prevent visual fatigue from affecting the results. The findings indicate that color recognition for users searching for targets is influenced by the background color, with significant impacts on search performance observed for all six color icon combinations under both background conditions [46].

From a fluctuation perspective, the variability in response times is more pronounced with a white background compared to a black background. Specifically, when users search on a white background, the color of the icons significantly influences their search times, leading to larger fluctuations in average search time. This phenomenon can be attributed to the white background enhancing color contrast, making colors more vivid and distinc.

Consequently, colors stand out more prominently against the background, increasing their ability to attract attentio—a phenomenon consistent with the Von Restorff effect [47].

Conversely, on a black background, the fluctuations in attention caused by icon colors are relatively smaller. This might be because black backgrounds absorb more light than they reflect, resulting in an overall lower brightness of the visual environment compared to brighter backgrounds [48]. Consequently, the distinction between colors is less pronounced on a black background than on a white one. Therefore, while color remains an important visual cue on a black background, its impact on user search efficiency is somewhat diminished.

In both black and white backgrounds, orange consistently has the shortest average search time, while yellow has the longest. A review of the literature on user interfaces for the elderly indicates that there is a decline in sensitivity to color discrimination with age, particularly for short-wavelength colors [19]. Consequently, warm tones such as orange are more easily recognized by older adults, which effectively reduces their search time. However, despite also being a warm tone, yellow has the slowest recognition speed. Studies have demonstrated that higher color contrast can enhance users' search performance [49]. This suggests that the slow recognition of yellow may be attributed to its high brightness value relative to white icons, which results in insufficient color contrast [50].

Both sets of experimental data underscore that orange minimizes search time effectively among elderly users, thereby enhancing search efficiency. In contrast, yellow consistently prolongs search times in this experimental context, significantly increasing both search duration and gaze duration for users. Therefore, yellow is not recommended for use in vehicle interface designs aiming to optimize user interaction.

B. PUPIL DIAMETER

This study investigated the impact of background color and icon color on pupil diameter as a measure of cognitive load in visual search tasks among elderly participants. The data indicate that the combination of background and icon colors significantly affects the ease with which users process information on the display interface [51]. Appropriate color combinations can reduce the cognitive effort required for searching and understanding information [52].

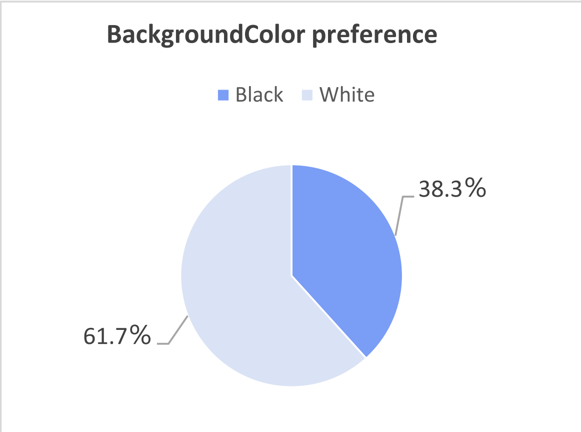


FIGURE 7. **Participants' preference for black and white background at the end of the experiment.**

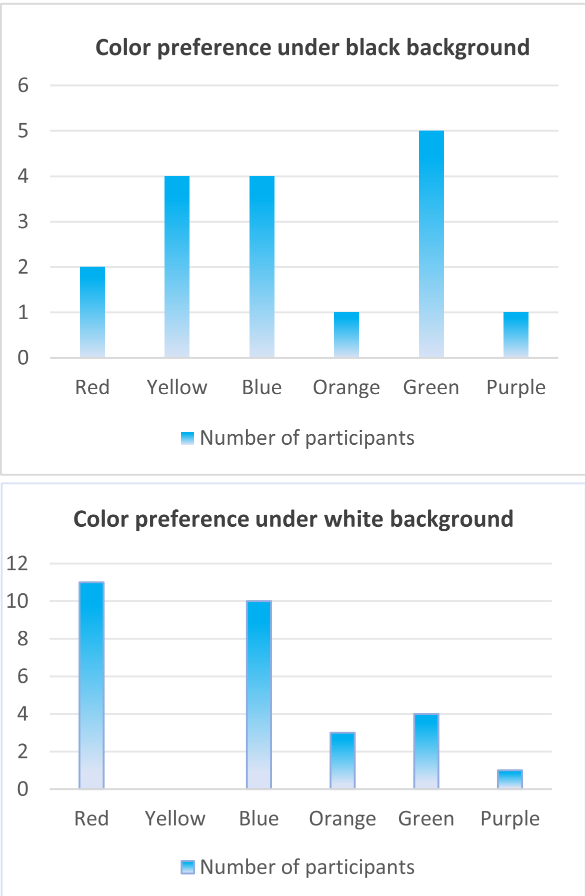


FIGURE 8. **Participants' preferences for six icon colors after the experiment.**

The data indicate that the average pupil diameter is larger when icon colors are displayed on a black background. This may be due to a higher level of arousal associated with darker colors [53]. The extant literature indicates that darkness can intensify individuals' startled responses, which may culminate in elevated cognitive load during recognition tasks on a black background due to augmented physiological strain [54].

In contrast, the average pupil diameter is smaller with a white background. Studies have shown that under artificial "daylight" white lighting, performance on short-term memory tasks improves [55]. This may explain the smaller pupil diameter observed in subjects when viewing a white background. The higher brightness of the white background likely improves visibility and enhances the overall harmony with the icon's brightness, making it more comfortable for subjects to view. This reduces the stress associated with recognizing colors and shapes.

Further research indicates that perceived cognitive load is influenced by readability, which in turn affects the target search process. Brighter displays cause pupil constriction, limiting the amount of light reaching the retina, whereas darker displays cause pupil dilation, allowing more light to reach the retina to view the information correctly. These findings align with our experimental results [50].

This insight can guide designers in creating more user-friendly interfaces for elderly users by considering color combinations that reduce visual stress. By doing so, they can enhance the safety and comfort of visual display environments.

C. USER QUESTIONNAIRE

1) COLOR PREFERENCE

In a survey assessing the impact of background color on search experience, the data clearly revealed user preferences. By surveying a substantial number of participants, the results indicated a pronounced preference for a white background, with the majority of participants favoring it. This preference underscores the widespread popularity of white backgrounds in providing a clear and bright visual experience. However, 36.95% of participants expressed a preference for a black background, indicating that while white backgrounds are generally preferred, black backgrounds should not be entirely disregarded by designers.

When examining preferences for six icon colors, blue was the most favored on a white background, and yellow the least favored. Conversely, on a black background, green was the most favored, and orange the least. However, objective eye-tracking experiments demonstrated that orange yielded the fastest search times on both black and white backgrounds. These findings suggest that subjective preferences do not align with experimental data, indicating that the differences in objective results are not influenced by participants' color preferences. Further analysis revealed that no participants preferred yellow icons on a white background, indicating that the combination of yellow and white is suboptimal for elderly users, both subjectively and objectively.

The data presented herein assist designers and developers in better understanding the preferences of older user groups and also emphasize the importance of considering user choices in the design process. In order to meet the individual needs of different user groups, the use of switchable color combination modes (e.g., day mode and night mode) is an effective strategy in interfaces that do not involve operational performance and cognitive load.

2) SATISFACTION DATA

To ensure that the difficulty of the experiment was appropriate and that it accurately reflected the normal psychological state of drivers, we included an evaluation scale during the study. The specific contents are detailed in Table 6 (located in the Appendix A). The initial question was employed to ascertain the degree of attention exhibited by the drivers throughout the experiment, evaluating the extent to which the experimental tasks aligned with their cognitive load and difficulty. The results indicated that 84.78% of participants reported a high level of concentration.

Secondly, questions 2, 3, and 4 were employed to ascertain whether the tasks induced undue pressure and whether the participants' experiences were consistent with our expectations. This allowed us to ascertain whether the subjects were able to complete the experimental tasks in a manner consistent with their psychological state while driving, and to identify any anomalous data that may have been caused by excessive stress.

Finally, we sought to ascertain the impact of our experiment on the target population through the fifth question. The final data indicated that 91.3% of users perceived a benefit from the intervention, thereby demonstrating that our study aligned with the needs of the elderly.

In conclusion, the aforementioned measures were implemented with the objective of validating the moderate difficulty of the experiment and ensuring its capacity to authentically represent the psychological state of drivers. This subjective assessment contributes to the exploration of the validity and reasonableness of the experimental results from a subjective perspective, thereby providing valuable insights for future experiments.

V. CONCLUSION AND FUTURE WORK

This study examines the effects of different background colors (white and black) and six color icons on user search efficiency. The findings demonstrate that, irrespective of background color, purple and orange icons demonstrably reduce search time, whereas yellow icons consistently result in the longest search times. Furthermore, the research shows that black and white background colors affect users' color recognition and visual search efficiency. The findings offer valuable recommendations for the design of automotive interfaces to enhance traffic safety.

Subjective surveys revealed a notable discrepancy between users' subjective color judgments and objective data. Accordingly, design processes must consider a variety of perspectives.

Additionally, this experiment has certain limitations. Currently, our study population may predominantly reflect users' experiences from a specific region in China. In the future, we plan to broaden both the quantity and scope of partic-

TABLE 6. A questionnaire used to survey participants' post-trial satisfaction.

Interest/Enjoyment			
topic		I was very focused when I was doing this activity.	
score	frequency	percentage	
1——2	0	0	strong disagreement
3——4	0	0	disagree
5——6	0	0	Neither agree nor disagree
7——8	7	15.22	agree
9——10	39	84.78	Firmly agree
Perceived Competence			
topic		I think I'm good at this activity.	
score	frequency	percentage	
1——2	0	0	strong disagreement
3——4	0	0	disagree
5——6	3	6.52	Neither agree nor disagree
7——8	12	26.09	agree
9——10	31	67.39	Firmly agree
Effort / Importance			
topic		I put a lot of effort into this activity.	
score	frequency	percentage	
1——2	0	0	strong disagreement
3——4	7	15.22	disagree
5——6	7	15.22	Neither agree nor disagree
7——8	13	28.26	agree
9——10	19	41.30	Firmly agree
Pressure / Tension			
topic		I feel pressured to do this activity.	
score	frequency	percentage	
1——2	19	41.30	strong disagreement
3——4	12	26.09	disagree
5——6	8	17.39	Neither agree nor disagree
7——8	4	8.70	agree

TABLE 6. ***(Continued.)*** **A questionnaire used to survey participants' post-trial satisfaction.**

9——10	3	6.52	Firmly agree
Value / Usefulness			
topic		I believe it is good for me to do this activity.	
score	frequency	percentage	
1——2	0	0.00	strong disagreement
3——4	0	0.00	disagree
5——6	4	8.70	Neither agree nor disagree
7——8	14	30.43	agree
9——10	28	60.87	Firmly agree

ipants to enhance the comprehensiveness of our research. Differences in diversity among the target groups, including age, cultural backgrounds, and cognitive abilities, can also impact the extrapolation of research results.

This study primarily focuses on the short-term graphical search speeds of the target group under different background colors, without evaluating the long-term effects of training and neural adaptation. This limitation restricts the applicability of our study results for practical applications. In addition to pure visual assessments, integrating cognitive evaluations is crucial. Future efforts could delve deeper into understanding the cognitive responses of elderly drivers interacting with a vehicle interfaces in real driving scenarios to derive more design principles aimed at enhancing safety and usability for older drivers.

In addition, environmental factors significantly influence color recognition during driving. Therefore, future experiments will encompass a broader spectrum of environments and incorporate a more diverse range of color variations. These studies will also explore non-uniformly distributed interactive interface layouts, moving beyond traditional car central control interfaces to investigate a wider array.

APPENDIX A
POST-EXPERIMENT QUESTIONNAIRE

To understand the participants' subjective evaluations after the experiment, we included the specific questionnaire items and their responses in Table 6.

REFERENCES

[1] Z. Tan, N. Dai, Y. Su, R. Zhang, Y. Li, D. Wu, and S. Li, "Human–machine interaction in intelligent and connected vehicles: A review of status quo, issues, and opportunities," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 13954–13975, Sep. 2022.

[2] D. Paddeu, G. Parkhurst, and I. Shergold, "Passenger comfort and trust on first-time use of a shared autonomous shuttle vehicle," *Transp. Res. Part C, Emerg. Technol.*, vol. 115, Jun. 2020, Art. no. 102604.

[3] F. Zeng, M. Wang, L. Li, and S. Cai, "Survey on the acceptance of smart cars amongst middle-aged and elderly in China," *Technol. Soc.*, vol. 73, May 2023, Art. no. 102234.

[4] M. Bradley, P. M. Langdon, and P. J. Clarkson, "An inclusive design perspective on automotive HMI trends," in *Universal Access in Human-Computer Interaction Users and Context Diversity*. Cham, Switzerland: Springer, 2016, pp. 548–555.

[5] S. Jung, H.-Y. Meng, and S.-F. Qin, "Exploring challenges in developing a smart and effective assistive system for improving the experience of the elderly drivers," *Chin. J. Mech. Eng.*, vol. 30, no. 5, pp. 1133–1149, Sep. 2017.

[6] D. S. Loughran, S. A. Seabury, and L. Zakaras. (2007). *What Risks Do Older Drivers Pose to Traffic Safety?*. Accessed: Jun. 23, 2024. [Online]. Available: https://policycommons.net/artifacts/978871/what-risks-do-older-drivers-pose-to-traffic-safety/1707999/

[7] R. Welsh, A. Morris, A. Hassan, and J. Charlton, "Crash characteristics and injury outcomes for older passenger car occupants," *Transp. Res. Part F, Traffic Psychol. Behaviour*, vol. 9, no. 5, pp. 322–334, Sep. 2006.

[8] B. Reimer, "Driver assistance systems and the transition to automated vehicles: A path to increase older adult safety and mobility?" *Public Policy Aging Rep.*, vol. 24, no. 1, pp. 27–31, Mar. 2014.

[9] C. Flock, J. Le Cardinal, and M. Mekhilef, "A systemic analysis of older drivers requirements to gather knowledge to be used by design engineers," in *Proc. 10th Int. Design Conf., Dubrovnik, Croatia*, vol. 2008, 2008, pp. 361–370.

[10] M. Wisch. (2017). *Road Traffic Crashes in Europe Involving Older Car Occupants, Older Pedestrians or Cyclists in Crashes With Passenger Cars-Results From Seniors*. Accessed: Jun. 23, 2024. [Online]. Available: https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/index/index/docId/1967

[11] A. May, T. Ross, and Z. Osman, "The design of next generation in-vehicle navigation systems for the older driver," *Interacting Comput.*, vol. 17, no. 6, pp. 643–659, Dec. 2005.

[12] R. R. Mourant, F.-J. Tsai, T. Al-Shihabi, and B. K. Jaeger, "Measuring divided-attention capability of young and older drivers," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1779, no. 1, pp. 40–45, Jan. 2001, doi: 10.3141/1779-06.

[13] J. Orlovska, F. Novakazi, C. Wickman, and R. Soderberg, "Mixed-method design for user behavior evaluation of automated driver assistance systems: An automotive industry case," in *Proc. Design Society, Int. Conf. Eng. Design*, 2019, pp. 1803–1812.

[14] K. L. Young, S. Koppel, and J. L. Charlton, "Toward best practice in human machine interface design for older drivers: A review of current design guidelines," *Accident Anal. Prevention*, vol. 106, pp. 460–467, Sep. 2017.

[15] J. Zhou, X. Miao, F. He, and Y. Miao, "Effects of font style and font color in news text on user cognitive load in intelligent user interfaces," *IEEE Access*, vol. 10, pp. 10719–10730, 2022.

[16] L. Deng, Z. Zhang, F. Zhou, and R. Liu, "Effects of app icon border form and interface background color saturation on user visual experience and search performance," *Adv. Multimedia*, vol. 2022, pp. 1–12, Jul. 2022, doi: 10.1155/2022/1166656.

[17] F. J. Varela, E. Thompson, and E. Rosch, *The Embodied Mind, Revised Edition: Cognitive Science and Human Experience*. Cambridge, MA, USA: MIT Press, 2017.

[18] M. Huang, Y. Xi, J. Pan, R. He, and X. Li, "Colorimetric observer categories for young and aged using paired-comparison experiments," *IEEE Access*, vol. 8, pp. 219473–219482, 2020.

[19] D. Hawthorn, "Designing effective interfaces for older users," Ph.D. dissertation, Univ. Waikato, Hamilton, New Zealand, 2006.

[20] S. Kim, A. K. Dey, J. Lee, and J. Forlizzi, "Usability of car dashboard displays for elder drivers," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, May 2011, pp. 493–502.

[21] M. Zhang, Y. Gong, R. Deng, and S. Zhang, "The effect of color coding and layout coding on users' visual search on mobile map navigation icons," *Frontiers Psychol.*, vol. 13, Dec. 2022, Art. no. 1040533.

[22] P. A. Cernin, B. K. Keller, and J. A. Stoner, "Color vision in Alzheimer's patients: Can we improve object recognition with color cues?" *Aging, Neuropsychology, Cognition*, vol. 10, no. 4, pp. 255–267, Dec. 2003.

[23] N. Nagarajan, L. Assi, V. Varadaraj, M. Motaghi, Y. Sun, E. Couser, J. R. Ehrlich, H. Whitson, and B. K. Swenor, "Vision impairment and cognitive decline among older adults: A systematic review," *BMJ Open*, vol. 12, no. 1, Jan. 2022, Art. no. e047929.

[24] H. Wijk, S. Berg, L. Sivik, and B. Steen, "Color discrimination, color naming and color preferences in 80-year olds," *Aging Clin. Experim. Res.*, vol. 11, no. 3, pp. 176–185, Jun. 1999, doi: 10.1007/bf03399660.

[25] C. Sha, R. Li, and K. Chang, "Color affects the usability of smart phone icon for the elderly," in *Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management, Health and Safety*. Cham, Switzerland: Springer, 2017, pp. 173–182.

[26] F. You, "Interface color design of intelligent vehicle central console," in *Proc. Advances in Usability, User Experience, Wearable and Assistive Technology*. Cham, Switzerland: Springer, 2021, pp. 784–792.

[27] D. Zhao, "Application of neural network based on visual recognition in color perception analysis of intelligent vehicle HMI interactive interface under user experience," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Oct. 2022, doi: 10.1155/2022/3929110.

[28] B. Xu, Z. Li, Z. Wu, and X. Ai, "A visual study of in-vehicle human-machine interaction interface," in *Proc. 3rd Int. Conf. Intell. Design (ICID)*, Oct. 2022, pp. 65–70. Accessed: Jun. 23, 2024.

[29] M. Nadin. (1988). *Interface Design: A Semiotic Paradigm*. Accessed: Jun. 23, 2024. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/semi.1988.69.3-4.269/html

[30] S.-M. Huang, K.-K. Shieh, and C.-F. Chi, "Factors affecting the design of computer icons," *Int. J. Ind. Ergonom.*, vol. 29, no. 4, pp. 211–218, Apr. 2002.

[31] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: Methods and constructs," *Int. J. Ind. Ergonom.*, vol. 24, no. 6, pp. 631–645, Oct. 1999.

[32] N. Moacdieh and N. Sarter, "Clutter in electronic medical records: Examining its performance and attentional costs using eye tracking," *Hum. Factors, J. Human Factors Ergonom. Soc.*, vol. 57, no. 4, pp. 591–606, Jun. 2015.

[33] N. M. Moacdieh and N. Sarter, "The effects of data density, display organization, and stress on search performance: An eye tracking study of clutter," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 6, pp. 886–895, Dec. 2017.

[34] N. M. Moacdieh and N. Sarter, "Using eye tracking to detect the effects of clutter on visual search in real time," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 6, pp. 896–902, Dec. 2017.

[35] S. Moresi, "Pupil dilation in response preparation," *Int. J. Psychophysiology*, vol. 67, no. 2, pp. 124–130, Feb. 2008.

[36] M. A. Recarte and L. M. Nunes, "Mental workload while driving: Effects on visual search, discrimination, and decision making," *J. Experim. Psychol., Appl.*, vol. 9, no. 2, pp. 119–137, 2003.

[37] C. Atombo and C. Wu, "Traffic safety climate factors in explaining driving behaviours and traffic crash involvement: Comparative study among male and female drivers," *J. Road Saf.*, vol. 33, no. 1, pp. 49–62, Feb. 2022.

[38] C. Owsley, G. McGwin, and K. Searcey, "A population-based examination of the visual and ophthalmological characteristics of licensed drivers aged 70 and older," *Journals Gerontology Ser. A, Biol. Sci. Med. Sci.*, vol. 68, no. 5, pp. 567–573, May 2013.

[39] W. Grundler and H. Strasburger, "Visual attention outperforms visual-perceptual parameters required by law as an indicator of on-road driving performance," *PLoS ONE*, vol. 15, no. 8, Aug. 2020, Art. no. e0236147.

[40] Y. Zuo, J. Qi, Z. Fan, Z. Wang, H. Xu, S. Wang, N. Zhang, and J. Hu, "The influence of target layout and target graphic type on searching performance based on eye-tracking technology," *Frontiers Psychol.*, vol. 14, Feb. 2023, Art. no. 1052488.

[41] T. Xi and X. Wu, "The influence of different style of icons on users' visual search in touch screen interface," in *Proc. AHFE Int. Conf. Ergonom*, 2017, pp. 222–232.

[42] E. H. Hess and J. M. Polt, "Pupil size in relation to mental activity during simple problem-solving," *Science*, vol. 143, no. 3611, pp. 1190–1192, Mar. 1964.

[43] D. Kahneman and J. Beatty, "Pupil diameter and load on memory," *Science*, vol. 154, no. 3756, pp. 1583–1585, Dec. 1966.

[44] J. Beatty, "Task-evoked pupillary responses, processing load, and the structure of processing resources," *Psychol. Bull.*, vol. 91, no. 2, pp. 276–292, 1982.

[45] P. van der Wel and H. van Steenbergen, "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychonomic Bull. Rev.*, vol. 25, no. 6, pp. 2005–2015, Dec. 2018.

[46] Z. Shen, L. Zhang, R. Li, J. Hou, C. Liu, and W. Hu, "The effects of color combinations, luminance contrast, and area ratio on icon visual search performance," *Displays*, vol. 67, Apr. 2021, Art. no. 101999.

[47] T. J. Bireta and C. M. Mazzei, "Does the isolation effect require attention?" *Memory Cognition*, vol. 44, no. 1, pp. 1–14, Jan. 2016.

[48] N. Ochiai and M. Sato, "Effects of surrounding brightness on visual search for safety colors," *Color Res. Appl.*, vol. 30, no. 6, pp. 400–409, Dec. 2005, doi: 10.1002/col.20152.

[49] K.-C. Huang, "Effects of computer icons and figure/background area ratios and color combinations on visual search performance on an LCD monitor," *Displays*, vol. 29, no. 3, pp. 237–242, Jul. 2008.

[50] D. Bhattacharyya, B. Chowdhury, T. Chatterjee, M. Pal, and D. Majumdar, "Selection of character/background colour combinations for onscreen searching tasks: An eye movement, subjective and performance approach," *Displays*, vol. 35, no. 3, pp. 101–109, Jul. 2014.

[51] Q. Wu, Z. Tan, and J. Liu, "Experimental study on the relationship between the harmony and cognitive load of business intelligence dashboard color combinations," *Color Res. Appl.*, vol. 47, no. 4, pp. 920–941, Aug. 2022.

[52] Y.-H. Ko, "The effects of luminance contrast, colour combinations, font, and search time on brand icon legibility," *Appl. Ergonom.*, vol. 65, pp. 33–40, Nov. 2017.

[53] M. L. Weijs, D. Jonauskaite, R. Reutimann, C. Mohr, and B. Lenggenhager, "Effects of environmental colours in virtual reality: Physiological arousal affected by lightness and hue," *Roy. Soc. Open Sci.*, vol. 10, no. 10, Oct. 2023, Art. no. 230432, doi: 10.1098/rsos.230432.

[54] C. Grillon, M. Pellowski, K. R. Merikangas, and M. Davis, "Darkness facilitates the acoustic startle reflex in humans," *Biol. Psychiatry*, vol. 42, no. 6, pp. 453–460, Sep. 1997.

[55] I. Knez, "Effects of colour of light on nonvisual psychological processes," *J. Environ. Psychol.*, vol. 21, no. 2, pp. 201–208, Jun. 2001.

SONGMAN LI is currently pursuing the bachelor's degree in product design with the School of Mechanical Engineering, Shandong University, Jinan, China.

Her research interests include inclusive design, particularly in catering to the needs of the elderly and other marginalized groups. She strives to provide more empathetic and human-centered design solutions.

SONG HAO received the Ph.D. degree in mechanical manufacture and automation from the School of Mechanical Engineering, Shandong University, Jinan, China, in 2012.

From 2010 to 2011, she conducted the Ph.D. research at the School of Materials Science and Engineering, The Pennsylvania State University, as a Visiting Scholar. She was a Visiting Scholar at the College of Design, North Carolina State University, from 2018 to 2019. She is currently a Senior Lecturer with the School of Mechanical Engineering, Shandong University. Her current research interests include usability research of rehabilitation therapy product, automobile HM systems, artificial intelligence for design, and Atlas of Chinese traditional cultural gene.