SituFont: A Just-in-Time Adaptive Intervention System for Enhancing Mobile Readability in Situational Visual Impairments

KUN YUE*, Tsinghua University, China

JINGRUO CHEN*, Cornell University, USA

MINGSHAN ZHANG*, Tsinghua University, China

CHUN YU, Tsinghua University, China

KEXIN NIE, The University of Sydney, Australia

ZHIQI GAO, Nankai University, China

JINGHAN YANG, Tsinghua University, China

CHEN LIANG, The Hong Kong University of Science and Technology (Guangzhou), China

YUANCHUN SHI, Tsinghua University, China

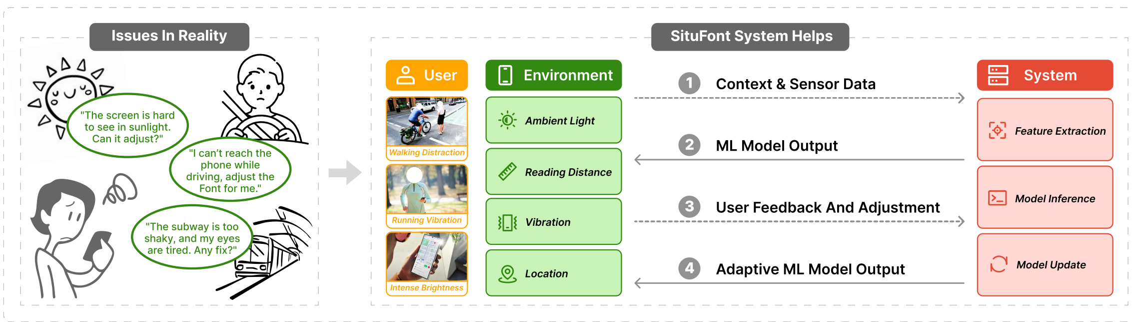

Fig. 1. The SituFont system dynamically adapts font parameters based on user context and environmental conditions. The process involves four steps: (1) collecting context and sensor data, (2) generating ML model output, (3) incorporating user feedback for font parameters adjustments, and (4) producing adaptive model output for continuous improvement.

Situational visual impairments (SVIs) hinder mobile readability, causing discomfort and limiting information access. This paper presents SituFont, a just-in-time adaptive intervention (JITAI) system that enhances readability by dynamically adjusting font parameters based on real-time contextual changes. Using smartphone sensors and a human-in-the-loop approach, SituFont personalizes text presentation to accommodate personal factors (e.g., fatigue, distraction) and environmental conditions (e.g., lighting, motion, location).

*These authors contributed equally to this research.

Authors' addresses: Kun Yue, yuek20@mails.tsinghua.edu.cn, Tsinghua University, Beijing, China; Jingruo Chen, jc3564@cornell.edu, Cornell University, Ithaca, New York, USA; Mingshan Zhang, Tsinghua University, Beijing, China; Chun Yu, Tsinghua University, Beijing, China; Kexin Nie, The University of Sydney, Sydney, Australia; Zhiqi Gao, Nankai University, Tianjin, China; Jinghan Yang, Tsinghua University, Beijing, China; Chen Liang, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, Guangdong, China; Yuanchun Shi, Tsinghua University, Beijing, China.

Manuscript submitted to ACM

To inform its design, we conducted formative interviews (N=15) to identify key SVI factors and controlled experiments (N=18) to quantify their impact on optimal text parameters. A comparative user study (N=12) across eight simulated SVI scenarios demonstrated SituFont's effectiveness in improving readability. Our findings demonstrate that SituFont successfully mitigates SVIs, highlighting the potential of JITAI systems to enhance mobile accessibility.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Situational Visual Impairments, Just-in-time adaptive intervention, Human-in-the-loop, Font Parameters, Text Readability

ACM Reference Format:
Kun Yue, Jingruo Chen, Mingshan Zhang, Chun Yu, Kexin Nie, Zhiqi Gao, Jinghan Yang, Chen Liang, and Yuanchun Shi. 2025. SituFont: A Just-in-Time Adaptive Intervention System for Enhancing Mobile Readability in Situational Visual Impairments. 1, 1 (February 2025), 29 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Mobile computing presents unique challenges due to its dynamic usage contexts [52]. Unlike desktop computing, mobile devices are used in varied environments, leading to Situationally Induced Impairments and Disabilities (SIIDs) [52, 59]. Among them, Situational Visual Impairments (SVIs) arise from factors like low lighting and user motion, significantly affecting text readability [36, 47, 48]. These challenges lead to visual fatigue, distraction, and difficulty processing information [57].

Current solutions for addressing SVIs include manual adjustments [20], such as increasing font size [22], using auditory substitutes [5, 19, 54], and automatic adjustments [20]. However, these methods often lack the contextual adaptability required to respond to the needs of diverse users in real-time. Traditional approaches in HCI research focus on optimizing user interfaces for specific impairments, but often do not account for the dynamic and multifaceted nature of real-world contexts [20]. As a result, there is a need for more comprehensive solutions that can automatically adapt text presentation based on changing environmental conditions and user preferences.

Just-in-Time Adaptive Interventions (JITAIs) offer a promising alternative by dynamically adapting to users' changing environments and needs [17, 31, 37, 45, 46]. A JITAI system for SVIs could leverage real-time data from mobile sensors to detect changes in environmental factors such as ambient lighting or user motion and automatically adjust text parameters, such as font size or line space. Furthermore, incorporating a "human-in-the-loop" approach can enhance the personalization and adaptability of JITAIs. By integrating user preferences and feedback, JITAIs can dynamically adjust to individual visual needs in real-time.

This study, guided by the principles of JITAIs and human-in-the-loop systems, proposes a novel approach to dynamically enhancing text readability on mobile devices under varying conditions. We introduce SituFont, a system that dynamically adjusts text parameters, including font size, weight, line spacing, and letter spacing, based on real-time contextual factors such as movement, lighting, and individual user preferences to enhance readability across different environments.

The research focuses on three key areas: (1) identifying the SVI factors affecting text readability through qualitative interviews with 15 participants, (2) quantifying the relationship between environmental factors and readability through controlled experiments with 18 participants, and (3) designing and evaluating the effectiveness of SituFont through comparative studies in eight SVI scenarios with 12 participants.

In summary, we contribute:

- Insights from qualitative and quantitative studies identifying key SVI factors affecting text readability.
- The design and implementation of SituFont, a human-in-the-loop JITAI system that personalizes text presentation by automatically adapting to the user's environment.
- A comparative study of SituFont's effectiveness and user experience with traditional display, demonstrating improvements in text readability across simulated SVI scenarios.

2 RELATED WORK

2.1 Situational Visual Impairments and Overcome

Situation-induced disorders and disabilities (SIIDs), introduced by Sears et al. [52], refer to scenarios where environmental, application-specific, or human factors impair the user's interaction with a system. Sarsenbayeva et al. identified six key factors that contribute to these barriers in the use of mobile devices: ambient temperature [12, 47], ambient light [59], environmental noise [2], user mobility status [10, 27, 35, 50], burden [38, 39], and pressure [27]. These factors disrupt normal user-system interaction by affecting fine motor control and environmental awareness (e.g. movement, burden, temperature) or impairing information gathering (e.g., light, noise, attention). Among these, situational visual impairments (SVIs) are particularly significant, as they directly impact a user's ability to read and process information on mobile screens. SVIs arise when environmental conditions, such as low lighting or motion, reduce text legibility and strain visual perception [58, 59]. Research shows that movement negatively affects visual performance, affecting reading comprehension [32, 36, 61, 61]. Similarly, ambient lighting [7, 57] and environmental noise [33, 49] can hinder text recognition and cognitive functions.

To mitigate SVIs, researchers have explored various approaches. One common challenge is "Reading on the Move," where users divide attention between navigation and screen content. Auditory feedback has been proposed as a compensatory method [5, 24, 61, 66], though its effectiveness is limited when the auditory channel is occupied, and its linear nature reduces flexibility[24]. Other solutions enhance safety by addressing visual concerns, such as CrashAlert[19], which uses depth cameras to detect obstacles and prevent collisions. SVIs also arise in static environments, particularly under ambient lighting. Research in this area has primarily focused on design guidelines. Evans [8] developed a brightness perception model for UI designers. Tigwell [57] examined SVI contexts, causes, and coping strategies, proposing a model linking environment, devices, and user interaction to inform design improvements. While these studies provide valuable insights, existing solutions often focus on static adjustments, predefined rules, or post-hoc compensatory measures rather than real-time, adaptive interventions. Most approaches lack contextual adaptability and fail to account for the dynamic nature of SVIs, where environmental factors continuously change. Furthermore, user preferences and individual differences in text perception are often overlooked.

2.2 Font Characteristics and Mobile Phone Text Legibility

Optimizing text for SVIs requires understanding how font characteristics influence legibility on smartphones. Research highlights font size, weight, line spacing, and character spacing as key factors. Larger fonts improve readability, particularly for users with impaired vision, but excessively large sizes reduce efficiency by increasing scrolling demands [21, 22, 69]. Bold fonts enhance recognition but may cause discomfort if too thick [4]. Proper line and character spacing improve layout clarity and character differentiation, making text easier to read [30, 41, 63, 69].

Beyond these font characteristics, additional factors such as screen size, text orientation, and reading direction also affect legibility. While larger screens generally facilitate easier reading, research suggests they do not significantly impact reading efficiency [64]. Similarly, text orientation and reading direction influence legibility, but their effects are often less pronounced than expected [14, 65]. Dynamic text presentation methods, such as Rapid Serial Visual Presentation (RSVP) and peripheral vision considerations, have also been explored to enhance reading speed and efficiency [16, 34, 60]. While prior research has explored how font characteristics affect readability, most studies focus on general legibility principles rather than their impact under situational constraints such as motion or lighting. Furthermore, existing findings are often derived from controlled environments that fail to reflect real-world SVI conditions, limiting their applicability to mobile use. The interaction between multiple font attributes across varying situational factors remains underexplored.

2.3 Just-in-time Adaptive Interventions with Human in the Loop

Just-in-time adaptive interventions (JITAIs) provide personalized support by dynamically adapting to an individual's internal state and external context [37]. This approach is particularly relevant for addressing SVIs, where external contextual changes can disrupt readability. While JITAIs have been successfully applied in various health domains [11, 17, 31, 44, 45, 56], their use in mitigating SVIs remains largely unexplored.

JITAIs can be categorized as rule-based or AI-based. Rule-based systems, such as the FOCUS intervention for schizophrenia management [6], rely on predefined rules and user input to trigger responses [6, 15]. In contrast, AI-based JITAIs leverage machine learning to personalize interventions based on user behavior and contextual data [42, 46]. For example, Matthew et al. developed a random forest model to predict nudge receptiveness by analyzing individualized and contextualized data [46]. Recent systems have also integrated human-in-the-loop approaches, such as Time2Stop [42], which continuously refines intervention strategies using user feedback while providing AI-driven, transparent explanations.

Advancements in situational awareness technologies further enhance the potential of JITAIs for SVIs. Mobile sensors can track environmental data, such as ambient light and motion [4, 43, 70], while image recognition can extract semantic information about the surroundings [9, 53, 55]. The integration of Large Language Models (LLMs), which excel in text understanding and generation [62], expands possibilities for context-aware systems. These technologies could enable interventions such as automatic font adjustments or real-time reading recommendations tailored to the user's immediate environment [28, 40, 68]. While JITAIs have shown promise in various domains, most existing JITAI applications focus on behavior-based interventions rather than perceptual and cognitive challenges like SVIs. Additionally, current AI-driven solutions often lack mechanisms for balancing automation with user control, raising concerns about intervention timing and user preferences.

3 FORMATIVE STUDY

To design a system that mitigates the impact of SVIs on mobile reading, we adopt the **Understanding-Sensing-Modeling-Adapting** framework, which serves as an instance of JITAIs in the domain of SIIDs [59]. Through the collaborative interaction of these four modules, the challenges posed by SIIDs can be more comprehensively addressed, leading to the design of more adaptive and inclusive mobile interaction systems. According to this framework, we need to identify the contextual impacting factors for the *Understanding* module, which will be detected by the *Sensing* module. Additionally, we need to figure out the specific methods for *Modeling* module to enable the *Adapting* module to reduce SVIs effects on mobile reading. In the following sections, we first identify the key factors inducing SVIs in various mobile reading scenarios through **Study 1**, and then in **Study 2**, we focus on how to model these factors for adaptation. Finally, we combine the main findings from theses 2 studies to depict the **Design Implications**.

ID	Gender	Age	Occupation	Vision Status
P1	Male	52	Public Servant	Myopia 300, Presbyopia 50, wears corrective glasses
P2	Female	50	University Teacher	Presbyopia 100, does not wear glasses daily, wears reading glasses for screens
P3	Female	46	Bank Employee	Normal vision
P4	Male	39	Engineer	Myopia 300, wears corrective glasses
P5	Male	30	Programmer	Myopia 600, wears corrective glasses
P6	Female	28	University Teacher	Myopia 200, wears corrective glasses
P7	Female	24	Bank Employee	Myopia 200, Astigmatism 150, wears corrective glasses
P8	Male	24	Grad Student	Myopia 500, wears corrective glasses
P9	Male	23	Grad Student	Normal vision
P10	Female	22	Grad Student	Myopia 450, wears corrective glasses
P11	Female	21	Undergraduate	Myopia 300, wears corrective glasses
P12	Female	21	Grad Student	Myopia 375, Astigmatism 200, wears corrective glasses
P13	Male	20	Undergraduate	Normal vision
P14	Male	19	Undergraduate	Myopia 600, wears corrective glasses
P15	Female	19	Undergraduate	Normal vision

Table 1. Demographic Overview of Interview Participants

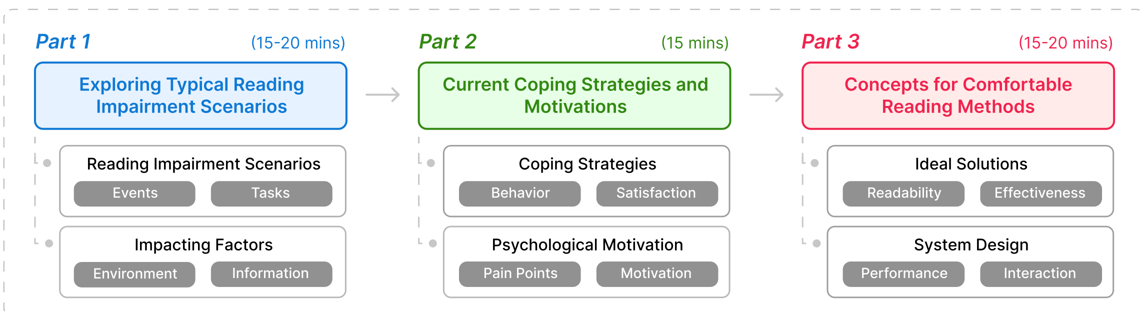

Fig. 2. Workflow of the Semi-Structured Interviews

3.1 Study 1: Potential Impacting & Adaptive Factors

To identify the potential impacting factors inducing SVIs and adaptive factors for mitigation during mobile reading, we conducted semi-structured interviews to explore the different scenarios, strategies, and challenges that participants use to cope with SVIs.

3.1.1 Research Methods. We conducted semi-structured interviews with 15 participants, aged 19 to 52, including students and professionals with varying visual conditions such as myopia, astigmatism, presbyopia, and normal vision (Table 1). The interview sessions, lasting 40 to 60 minutes (Figure 2), began with an introduction to SVIs, followed by discussions about participants' experiences, challenges, and coping strategies. Participants were also asked to suggest features and improvements that could enhance their reading comfort under such conditions. After thematic analysis [3], we identified several potential impacting factors that cause SVIs in mobile reading scenarios.

3.1.2 Environmental Factors. We found that movement and lighting conditions are most frequently mentioned by participants. ***Reading while in motion*** was a significant challenge, particularly while walking (11/15) or running (9/15). Movement-induced vibrations blurred text, making recognition difficult (P3, P6, P8-9, P13, P15) and often causing dizziness (P8, P10, P13) or slower reading speeds (P6, P12, P14-15). Some participants (P4) found it hard to concentrate due to the need to monitor their surroundings, while others (P1, P5, P11) avoided reading while running due to the difficulty of interaction. ***Reading in a moving vehicle*** was another common issue, with all participants (N=15) citing reading difficulties, especially during commutes (P6, P7, P10). Vehicle movement, particularly in private cars, impaired focus and sometimes caused motion sickness (P2, P7, P10, P12). Eight participants noted that reading was most uncomfortable in cars, followed by buses, with the subway being the least disruptive (P15).

In addition, ***Reading in strong light*** was frequently problematic, particularly in outdoor environments. Six participants (P1-2, P10-11, P13-14) struggled with excessive sunlight, which made screen brightness adjustments insufficient. *"Sometimes the sunlight is so strong that even when I max out the brightness, I can barely see anything on the screen"* (P12). Finally, ***Reading at an inappropriate viewing distance*** was challenging when users couldn't adjust their screen position. This was common in scenarios such as driving, where the phone was mounted at a fixed distance (P5), or in crowded spaces like subways, where participants (P13) lacked the flexibility to hold their phones at an optimal distance.

3.1.3 Personal Factors. ***Visual acuity*** impacted readability, especially for participants who wore glasses or contact lenses (9/11). Challenges were exacerbated when they were not wearing corrective lenses, particularly before bed (P4, P6, P7, P10, P11, P12). ***Comprehension ability*** influenced ease of reading, with some participants (P8, P14) reporting increased visual fatigue when reading complex or lengthy texts. ***Concentration difficulties*** were another barrier, as participants (P3, P10, P13) mentioned distractions and multitasking negatively affected their reading efficiency.

3.1.4 Information Factors. Participants generally agree that text presentation and complexity played crucial roles in readability. ***Text parameters*** such as excessively small font sizes ($n = 13$), thin fonts ($n = 6$), overly large line spacing ($n = 4$), and small character spacing ($n = 2$) were frequently cited as readability obstacles. ***Information difficulty*** was another key factor as highly specialized or information-dense content was difficult to comprehend ($n = 2$). ***Information length*** was an issue for overly long text passages, which some participants found challenging to read without interruption.

3.1.5 Strategies & Challenges. We found several strategies that participants used to cope with SVIs. The most common strategy was to ***adjust the reading distance***, though it often caused discomfort. *"I find myself constantly moving my phone closer or farther away, it's tiring, and after a while, my neck starts hurting"* (P6). While effective to some extent, this approach often led to physical strain, especially during prolonged reading. Another strategy was the ***use of auditory substitutes***, such as voice assistants or text-to-speech functions (P8, P10, P13, P15). However, this method was not always practical. *"When I'm driving, I'll use my phone's voice assistant to read texts aloud. But it's not always efficient, especially if the text is long or complicated."* (P13). While this method allowed participants to continue consuming information without relying on visual input, some participants (P9, P15) found voice assistants inconvenient in noisy settings or when privacy was a concern.

Using larger fonts was a universal strategy ($n = 15$), but not all participants favored it. *"I sometimes make the text bigger, but then I have to scroll a lot more, and it messes up the page layout. It's not ideal, but it's better than not being able to read at all"* (P10). Younger participants (8/15) disliked how larger fonts disrupted formatting and required more scrolling, while the oldest participant (P1) noted that even with the largest font settings, reading remained difficult due to age-related vision decline.

Participants also expressed ***reluctance to manually adjust text settings*** due to the inconvenience and time required. *"Honestly, I don't want to fiddle with settings every time I'm in a new situation. It's just too much hassle. I wish my phone*

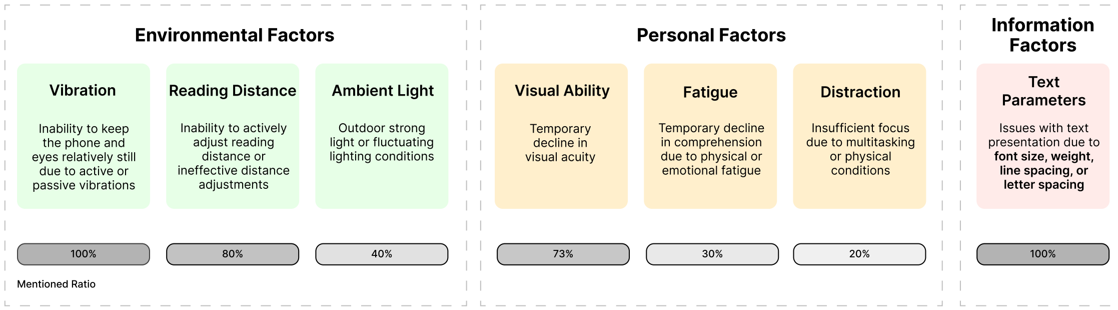

Fig. 3. Interview Findings: Factors Affecting Situational Visual Impairments. The numbers inside the rectangles indicate the proportion of respondents who mentioned each factor.

ID	Scenario	Motion State	Light Level	Vibration Level
1	Indoor Corridor	Standing	Low	Low
2	Indoor Corridor	Walking	Low	Medium
3	Indoor Corridor	Running	Low	High
4	Outdoor Playground	Standing	High	Low
5	Outdoor Playground	Walking	High	Medium
6	Outdoor Playground	Running	High	High

Table 2. Experimental Scenarios: Motion, Light, and Vibration Conditions

could just figure it out for me" (P8). This suggests a preference for automated systems that can adapt text parameters dynamically based on the user's current context.

3.1.6 Summary of Factors. Our findings highlight three major categories of factors influencing SVIs: environmental conditions (motion, lighting, viewing distance), personal characteristics (visual acuity, comprehension, concentration), and information-related aspects (text size, spacing, content complexity, and length) (Figure 3). These findings informed the design of Study 2, where we quantitatively examined how environmental conditions impact text readability adjustments.

3.2 Study 2: Modeling Key Factors for Adaptation

Building on Study 1's qualitative insights, Study 2 aims to **quantify the relationships between environmental conditions and text legibility.** By identifying measurable correlations, this study addresses the **cold-start challenge**, allowing JITAI systems to adapt effectively to diverse real-world scenarios without extensive initial user data.

To achieve this, we conducted a controlled experiment to analyze how light levels, motion states, and vibration conditions influence text parameter adjustments (e.g., font size, weight, line spacing, character spacing). The results provide empirical validation of Study 1's findings and establish guidelines for adaptive readability solutions.

3.2.1 Research Methods. We recruited 18 university students aged 18 to 25 with normal or corrected vision. Each participant experienced six scenarios combining two light intensity levels (low, high) and three movement states (standing, walking, running) (Table 2). The latter five scenarios were designed to replicate real-world conditions where

	Reading Distance	Light Intensity	Vibration Offset		
			X Axis	*Y* Axis	*Z* Axis
Font Size	0.404**	0.117**	0.368**	0.333**	0.381**
Font Weight	0.228**	0.373**	0.230**	0.263**	0.225**
Line Spacing	0.132**	-0.038	0.170**	0.168**	0.173**
Character Spacing	0.057	0.042	-0.036	-0.066	-0.037

* $p < 0.05$, ** $p < 0.01$

Table 3. Spearman Correlation Coefficients of Environmental Factors and Readable Text Parameters

SVIs commonly occur. Detailed information on the experimental design and data collection process is provided in Appendix A.

Participants used a mobile application to adjust text parameters (font size, weight, line spacing, and character spacing) while reading materials of equal length and difficulty under six environmental conditions. Mobile phone sensors recorded environmental data(light intensity, reading distance, vibration offset) alongside participants' text adjustments. After data cleaning, 497 valid data sets were collected, averaging 28 sets per participant and 83 per scenario. The data were analyzed to identify correlations between environmental factors and participants' readability adjustments.

3.2.2 Findings. Environmental characteristics. **Significant differences were observed in light intensity and motion states across scenarios while reading distance remained stable (Appendix A Table 9).** Light intensity was substantially higher in outdoor ($M = 41998.37, SD = 26421.78$) compared to indoor environments ($M = 93.90, SD = 79.66$). Motion states also exhibited significant differences, confirmed by three-axis acceleration data. Standing showed the lowest values ($X : M = 0.17, SD = 0.10; Y : M = 0.10, SD = 0.07; Z : M = 0.25, SD = 0.18$), followed by walking ($X : M = 0.53, SD = 0.14; Y : M = 0.67, SD = 0.14; Z : M = 0.79, SD = 0.24$) and running ($X : M = 1.48, SD = 0.57; Y : M = 1.48, SD = 0.87; Z : M = 1.65, SD = 0.55$). Reading distance remained consistent across scenarios.

Reading distance, light intensity, and vibration levels influenced text readability adjustments, particularly font size and font weight (Table 3). Using the Spearman correlation coefficient, we found significant positive correlations between reading distance and font size ($r = 0.404$), font weight ($r = 0.228$), and line spacing ($r = 0.132$). Light intensity was most strongly correlated with font weight ($r = 0.373$) and moderately with font size ($r = 0.117$). Vibration levels correlated positively with font size ($\bar{r} = 0.360$), font weight ($\bar{r} = 0.239$), and line spacing ($\bar{r} = 0.170$). Character spacing showed weak correlations with all environmental factors.

Situational Differences in Text Parameters **Text parameters varied significantly across scenarios, with font size and weight increasing in high-motion and high-light conditions (Table 4).** Running scenarios required the largest font size and boldest weight, followed by walking and standing, indicating that greater motion intensity necessitates increased text size for readability. Outdoor conditions required greater font weight than indoor scenarios, likely due to the need for increased contrast in high-light environments. Line spacing also exhibited significant differences across scenarios, though no clear pattern emerged, suggesting that individual preference rather than environmental conditions may more influence its adjustments. Character spacing remained largely unaffected, further reinforcing its weak correlation with external factors.

Individual Differences in Readability Preferences **Font size and font weight showed the most variability across participants, particularly in dynamic and high-light scenarios (Table 5).** Font size variability was highest in

ID	Scenario	Font Size (*sp*)	Font Weight (*px*)	Line Spacing (*em*)	Character Spacing (*em*)
1	Indoor Standing	20.60±3.99	0.59±0.65	0.25±0.14	0.10±0.10
2	Indoor Walking	21.29±3.98	0.79±0.76	0.27±0.17	0.08±0.09
3	Indoor Running	23.84±5.31	0.93±0.80	0.29±0.16	0.10±0.09
4	Outdoor Standing	19.28±3.63	0.93±0.79	0.25±0.16	0.09±0.08
5	Outdoor Walking	21.08±3.26	1.04±0.95	0.22±0.13	0.09±0.09
6	Outdoor Running	23.64±4.92	1.30±0.87	0.31±0.14	0.12±0.14
F Value		15.108	7.275	4.666	1.301
p Value		0.000**	0.000**	0.000**	0.262

* $p < 0.05$, ** $p < 0.01$

Table 4. ANOVA Table of Readable Text Parameters in 6 Experimental Scenarios

ID	Scenario	Font Size (*sp*)	Font Weight (*px*)	Line Spacing (*em*)	Character Spacing (*em*)
1	Indoor Standing	2.698	0.654	0.138	0.089
2	Indoor Walking	2.940	0.663	0.156	0.077
3	Indoor Running	4.467	0.690	0.157	0.088
4	Outdoor Standing	2.075	0.613	0.134	0.059
5	Outdoor Walking	2.398	0.827	0.127	0.076
6	Outdoor Running	3.796	0.791	0.123	0.131

Table 5. Standard Deviation Table of Readable Parameters for 18 Participants in 6 Experimental Scenarios

indoor running scenarios ($SD = 4.467$), suggesting that individuals adjust text size differently under rapid movement. Font weight variability peaked in outdoor walking ($SD = 0.827$) and outdoor running ($SD = 0.791$), highlighting individual differences in text boldness preferences under varying light conditions. Line spacing showed less variability across participants, while character spacing exhibited significant differences only in outdoor running scenarios.

3.3 Design Implications

Our findings from Study 1 (interviews) and Study 2 (experiments) converge on key design implications that inform the development of readability solutions for SVIs.

Contextual Awareness and Automated Adaptation Both Study 1 and Study 2 emphasize the need for context-aware font adjustments. In ***Study 1***, participants expressed a strong preference for systems that automatically adapt text based on their reading conditions, avoiding the burden of manual modifications. ***Study 2*** experimentally validated this need, demonstrating that motion and lighting variability significantly impact readability. Specifically, font size and weight increased in high-motion scenarios (e.g., running) and strong light conditions (e.g., outdoor environments). These findings highlight the necessity of real-time environmental sensing using smartphone sensors (e.g., accelerometers, ambient light sensors) to enable contextual adaptations.

Personalization and User Agency ***Study 1*** highlighted the importance of customization, with participants noting that "*everyone's eyes are different*" and that adaptive systems should allow users to define and store their font preferences. ***Study 2*** reinforced this, showing significant individual variability in font size and weight preferences, particularly under high-motion and high-light conditions. Given the high standard deviations observed in user preferences, systems should provide multiple levels of adaptation intensity, enabling users to fine-tune automation sensitivity or receive prompts before changes occur. A hybrid approach that balances automation with user control can improve both adaptability and usability.

Addressing Limitations of Existing Coping Strategies ***Study 1*** participants reported relying on manual adjustments, larger fonts, and text-to-speech functionalities, but found these strategies inadequate. Adjusting reading distance led to discomfort, increasing font size disrupted page layouts, and auditory substitutes were impractical in noisy or private settings. ***Study 2*** quantitatively confirmed these limitations by showing that no single text adjustment optimally addresses all readability challenges. These findings reinforce the need for multifactorial adaptation strategies.

Simplified Interaction and Reduced Cognitive Load ***Study 1*** participants found manual font adjustments cumbersome, often avoiding them due to the effort required to navigate settings. To streamline interactions, a context-aware shortcut system could allow users to quickly toggle settings based on detected conditions. Additionally, an adaptive user interface that learns and refines user preferences over time would reduce the need for repeated manual interventions.

Targeted Adaptations for High-Variability Scenarios Both studies highlight that adaptive interventions should prioritize scenarios where readability challenges are highest: reading while in motion, adjusting to changing lighting conditions, and accommodating individual readability preferences. ***Study 1*** participants described these as the most disruptive factors, and ***Study 2*** experimentally confirmed that running and outdoor environments exhibit the most significant variations in required text adjustments. These findings emphasize the need for targeted interventions tailored to these conditions.

By integrating insights from Study 1 (interviews) and Study 2 (experiments), these design implications guide the development of our system, introduced in the next section.

4 SITUFONT SYSTEM DESIGN & IMPLEMENTATION

As Figure 4 shows, inspired by the formative study in Section 3, we proposed SituFont, a just-in-time adaptive intervention (JITAI) system that enhances readability by dynamically adjusting font parameters based on real-time contextual changes. In the sections below, we outline the design (Section 4.1) and implementation (Section 4.2) of the SituFont system.

4.1 SituFont System Design Overview

SituFont's core components include the ML Training Pipeline, the Label Tree of Reading Scenario, and the Human-AI loop Workflow. Each module plays a key role in the system's font adaptation and effectiveness.

4.1.1 Machine Learning for Font Parameters Recommendation. Constructing an ML-driven JITAI system for recommending suitable font parameters involves two steps:

(1) Initial Model from Group Data Developing the initial model from the group data has two key objectives. First, it gives the system an initial adjustment capability when the user first interacts with it. Second, it enables the model to be fine-tuned with minimal user data over time, allowing the system to adapt to the user's reading habits. To establish this cold-start initial model, supervised learning is applied using group data collected from Formative Study 2 (Section 3.2). Formative Study 2 gathered data on adjusting text parameters—such as font size, weight, letter spacing, and line spacing—to suit varying environmental factors like reading distance, light intensity, and phone acceleration. In this setup, the text parameters are used as output variables, while the environmental factors serve as input features.

(2) Collecting Data from Users' Daily Usage Data is collected through the interface shown in the system (figure 5). When users double-tap the screen, a control panel for adjusting text parameters appears near the tapped location to

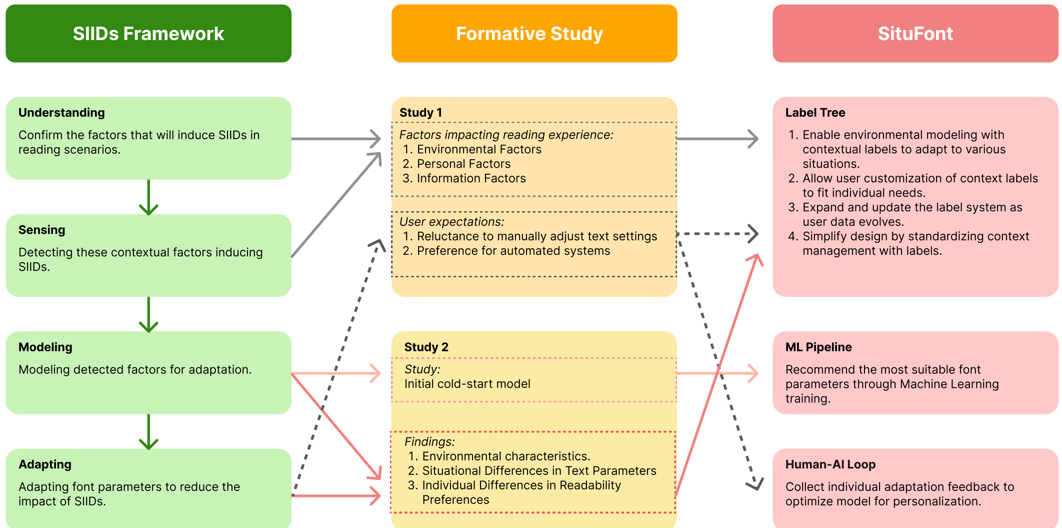

Fig. 4. Formative study 's findings inspire the system design of SituFont, which mainly include the label tree, machine learning training, human-ai loop modules.

Data Name	Unit	Component and Method	Frequency
Ambient Light	*lux*	Light Sensor; Direct Call	10 times/sec
Reading Distance	*cm*	Front Camera; Calculated from pupil distance	10 times/sec
Vibration Offset	m/s^2	Accelerometer; sensorEvent.values[0/1/2] for $x/y/z$ axis	10 times/sec
Font Size	*sp*	Android function for size after each adjustment	Recorded on upload
Font Weight	*px*	Android function for weight after each adjustment	Recorded on upload
Line Spacing	*em*	Android function for spacing after each adjustment	Recorded on upload
Letter Spacing	*em*	Android function for spacing after each adjustment	Recorded on upload

Table 6. Environmental Sensor Data collected in SituFont

make it easier to adjust settings. While users modify parameters, the system automatically detects reading distance, light intensity, and phone acceleration (Table 6). Users' current cognitive factors are collected by prompting them to choose whether any factors related to fatigue, distraction, and temporary decrease in visual ability exist. Once users click on a blank area of the screen, the data is sent to the backend to be stored in the corresponding contextual dataset, which is used to train and update the model for future recommendations.

4.1.2 Label Tree of Reading Scenario. **Label Tree Structure** Based on findings from the formative study, we designed a hierarchical data structure for reading context labels (figure 6), categorized as "[Movement/Posture] - [Environmental Scene] - [Personalized Needs]." The system automatically determines the first two layers of context labels by sensing the user's current movement state, environment, and location. However, personalized factors (such as visual ability, fatigue, and attention state) that represent the user's individual conditions cannot be directly detected by the system. For example, "Has the user's vision changed?", "Is the user feeling fatigued?", and "Is the user focused?" are factors that

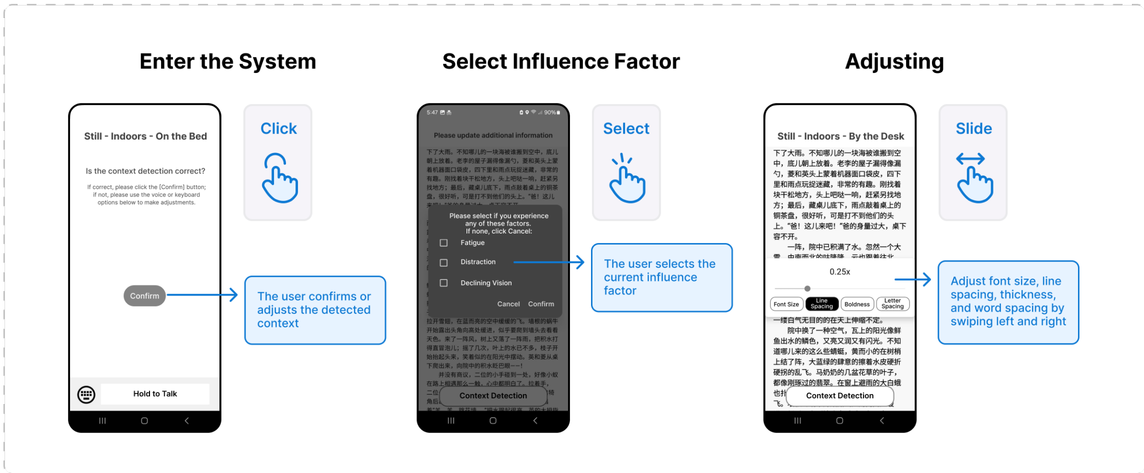

Fig. 5. The user interface of SituFont involves three key interactions: (1) Entering the System – the user confirms or adjusts the detected reading context; (2) Selecting Influence Factors – the user specifies factors affecting readability, such as fatigue or distraction; and (3) Adjusting Text Parameters – the user refines font size, line spacing, thickness, and word spacing through swipe gestures for a personalized reading experience.

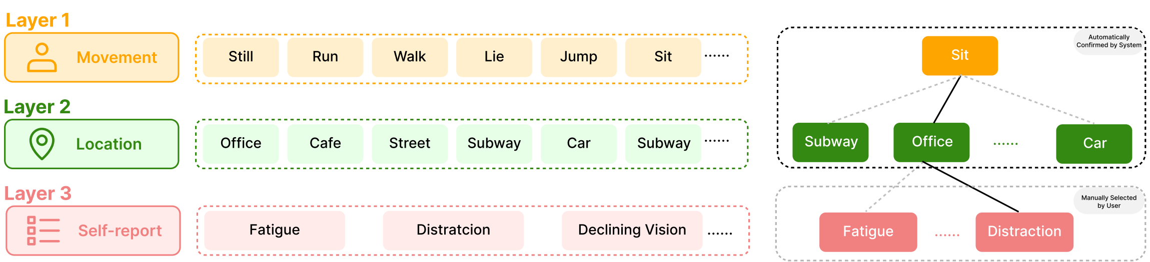

Fig. 6. The left part of the figure describes a three-layer labeling system used to mark situations, where the priority decreases from top to bottom when constructing the label tree. The right part of the figure presents an example of a label tree, where the solid lines indicate the process of detecting situational label combinations.

are difficult for the system to assess automatically. Therefore, the third layer of labels requires the user to manually indicate whether their vision, comprehension, or attention is currently affected.

The hierarchy follows the order of [Movement/Posture] - [Environmental Scene] - [Personalized Needs] because the first two categories can be automatically detected by the system. Among these, Movement/Posture is considered more influential than Environmental Scene on reading behavior, so it is placed at the top of the hierarchy. Personalized Needs, however, require manual input from the user and should be avoided as much as possible, which is why it is placed as the final layer in the label tree.

Label Tree Functions The Label Tree is designed to maximize the utility of small datasets by structuring them hierarchically based on contextual factors such as lighting conditions, user states, and task demands. This organization ensures that each dataset retains its relevance within its specific environment, allowing fine-tuned models to be applied effectively without requiring extensive data collection. Moreover, the Label Tree facilitates context transfer and generalization, enabling the system to identify the most relevant existing dataset when encountering a new but similar context. This reduces redundancy and enhances adaptability, allowing small datasets to be leveraged efficiently across multiple scenarios. By systematically preserving contextual distinctions and enabling knowledge transfer, the Label Tree significantly improves the effectiveness of small datasets, making them more impactful while minimizing the need for extensive user input or additional data collection.

4.1.3 Human-AI Loop in SituFont. When using SituFont, users can perceive the system's current font adjustments. If the font parameters are not optimal, they can manually adjust the font to better suit the environment. These new adjustments, along with the environmental data, align more closely with the user's personalized reading needs. The Human-AI Loop accelerates data accumulation and updates the model based on user feedback, enhancing the system's adaptability.

4.2 SituFont System Implementation Overview

Based on the system design in Section 4.1, we then introduce the implementation details of SituFont. We instantiated SituFont on Android OS (end-user side) and a server (cloud side), as shown in Figure 7. The SituFont system includes the context sensing module (Section 4.2.1), font adaptation user interface (Section 4.2.2), the ML pipeline (Section 4.2.3), and label tree module (Section 4.2.4).

4.2.1 Context Sensing. The contextual detection system has two key components: context recognition for confirming labels and input for the font adjustment model.

For context recognition and label confirmation, the system utilizes several data sources. It determines location based on GPS POI data and visual input from the rear camera using Vision-Language Models (GPT-4o[1]). Additionally, it assesses the user's movement state by analyzing 3-axis vibration data through a pre-trained machine-learning model designed to recognize specific movements. Furthermore, by leveraging Large Language Models (GPT-3.5Turbo[2]), users can actively describe or modify the current recognized context by typing or using voice input, allowing for more accurate or personalized adjustments to the detected environment.

The font adjustment model relies on various sensor inputs to optimize the reading experience. It monitors ambient light intensity using the mobile phone's light sensor and takes into account 3-axis vibration data collected from the mobile phone's accelerometer sensor. The user's reading distance refers to the distance between their eyes and the screen. To calculate this, MediaPipe's face recognition functionality is utilized [29], specifically leveraging the face landmarks to determine the proportion of the eyes in the image, which is then converted into the actual reading distance by factoring in the user's real interpupillary distance (IPD) before detection. A similar method is used in AngleSizer to detect the distance between two hands [23].

4.2.2 Adaptive User Interface. As described in Section 4.1.1 and illustrated in Figure 5, the controls for adjusting text parameters are developed using native Android. During the data collection process, text adjustments are made by dynamically controlling various properties of the TextView in real-time. When the user long-presses the screen for 1 second, the phone briefly vibrates, indicating that the Font Adjustment Model has optimized the font settings based on the current environmental parameters.

4.2.3 ML Pipeline. Since SituFont utilizes a simple Regression Machine Learning model with a small amount of fit data, the delay for both training and inference is negligible. The data for training and inference is collected directly on the

[1]https://openai.com/index/hello-gpt-4o/
[2]https://openai.com/index/gpt-3-5-turbo-fine-tuning-and-api-updates/

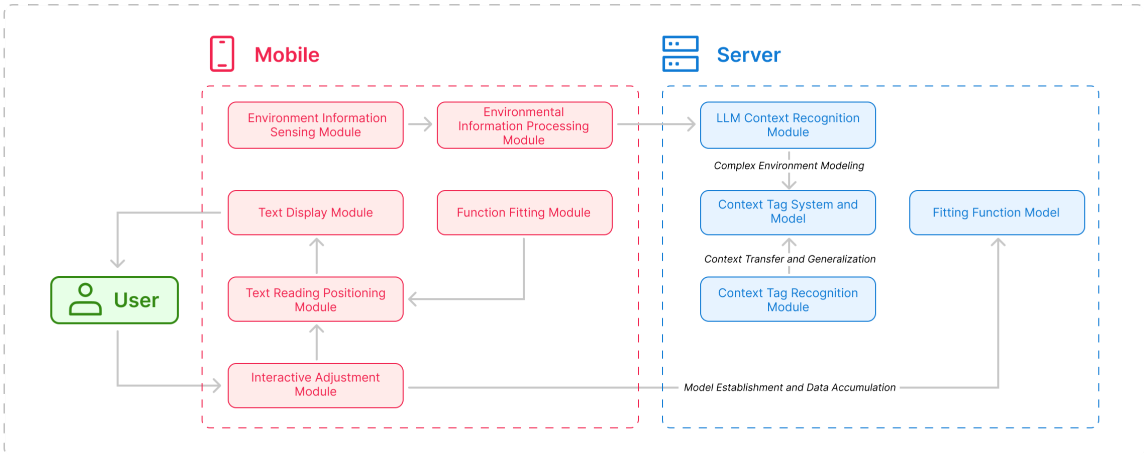

Fig. 7. System architecture of SituFont, showing the flow of information from sensors to the adaptation module

user's Android mobile phone. Once transmitted to the cloud backend, the model is typically updated or the inference results are returned to the front end within 2 seconds, allowing for prompt font parameter adjustments.

4.2.4 Label Tree Implementation. The implementation of the Label Tree consists of two main components: Label Generation and Label Selection. After gathering the context information outlined in Section 4.2.1, the system first uses an LLM to select the most appropriate label from the existing labels stored in the cloud database. If no suitable label is found, or the selected label does not match the user's current context, the system generates a new label based on the Label Tree Structure described in Section 4.1.2 and the newly provided context information. Once the user confirms the newly generated label, it is stored in the user's cloud database for future use. The prompts used in Label Generation and Label Selection are listed in Appendix B.

5 USER STUDY

To evaluate SituFont's impact on reading performance, comprehension, and user experience compared to a traditional display that manual font adjustments, we conducted a within-subject study with 12 participants under eight simulated SVI scenarios. The study consisted of three phases: (1) a pre-test establishing baseline reading performance under eight SVI conditions, (2) a four-day adaptation period using SituFont, and (3) experimental reading tasks comparing SituFont and a traditional display under identical conditions (Figure 8).

The user study aimed to address two research questions:

RQ1: Does SituFont improve reading performance compared to traditional displays under varying SVI conditions?

RQ2: How do users perceive SituFont's workload and overall experience compared to traditional displays?

5.1 Participants

Participants were recruited via online questionnaires. 12 participants took part in the study (5 male, 6 female, 1 non-binary, $M = 22.3$, $SD = 4.1$, age range = 18 to 34). All participants were native Mandarin speakers. Participants reported using a variety of devices for reading, including smartphones ($n = 12$), tablets ($n = 6$), and laptops ($n = 7$). Daily reading time on smartphones ranged from 10 minutes to over 2 hours. Ten participants reported wearing corrective lenses. One

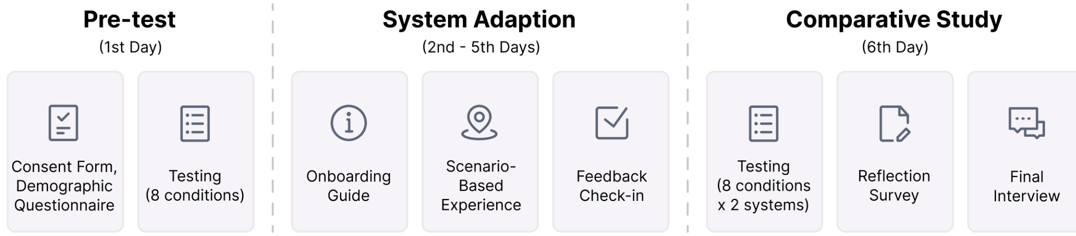

Fig. 8. User study timeline

participant reported difficulty reading standard-sized text on mobile screens due to eye strain, while another reported general eye fatigue. All participants provided written informed consent.

5.2 Study Procedures

5.2.1 Pre-test. Before the adaptation period, participants completed a baseline reading pre-test mirroring the main experimental sessions (Table 7) under eight SVI conditions, using a traditional mobile display with manual font control. Each condition included two 50-character reading passages sourced from the HSK Level 5 examination, a standardized Mandarin proficiency test for non-native speakers. The HSK test system includes predefined comprehension questions, ensuring standardized difficulty while minimizing memorization biases for native speakers [26].

Reading comprehension was assessed using five multiple-choice questions per passage, pre-defined by the HSK test system, with comprehension accuracy calculated as the percentage of correctly answered questions. Reading performance was also measured based on goodput (characters per minute, CPM), calculated as the total number of correctly read characters divided by total reading time [25]. Participants read passages aloud at their normal pace while maintaining accuracy, with audio recordings collected for later analysis. Additionally, a pre-test questionnaire gathered demographic data, reading habits, and vision details.

5.2.2 Adaptation Period. Participants used SituFont for four days, engaging with the system in at least three to four different scenarios daily. A daily survey recorded reading contexts, perceived fatigue, distraction levels, and usability issues based on their experiences in different SVI conditions. The adaptation period ensured participants became familiar with the system and personalized their settings before the experimental phase.

5.2.3 Post-Adaptation Reading Evaluation. Following the adaptation period, participants completed reading tasks under the same eight SVI conditions as in the pre-test. Each condition involved reading two new HSK Level 5 passages, using both SituFont and a traditional display. The order of interface use was counterbalanced across participants and conditions to control for order effects. After each reading passage, participants completed a comprehension test with five standardized multiple-choice questions. Reading performance was also evaluated based on reading goodput, measured in CPM. After each condition, participants completed a NASA-TLX questionnaire to assess the perceived workload. Upon completing all reading tasks, participants filled out a final survey assessing overall user experience (UEQ-S, SUS) and participated in semi-structured interviews to provide qualitative feedback.

Condition	Location	Intense Brightness	High Vibration	Distraction	Fatigue
Task 1	Outdoor	✓	✗	✗	✗
Task 2	Indoor	✗	✗	✓	✗
Task 3	Outdoor	✓	✓	✗	✗
Task 4	Outdoor	✓	✓	✓	✗
Task 5	Indoor	✗	✓	✗	✗
Task 6	Indoor	✗	✗	✗	✓
Task 7	Outdoor	✓	✓	✗	✓
Task 8	Outdoor	✓	✓	✓	✓

Fig. 9. An example order of the eight experiment conditions for comparative study

Condition	Lighting	Motion	Task Load
Intense Brightness	50,000+ lux[(a)]	Static	None
Distraction	Normal (Indoor)	Walking [(b)]	Navigation Task[(c)]
High Vibration	Normal (Indoor)	Running[(d)]	None
Fatigue	Normal (Indoor)	Walking[(b)]	Weighted Load[(e)]
Intense Brightness + High Vibration	50,000+ lux[(a)]	Running[(b)]	None
Intense Brightness + High Vibration + Distraction	50,000+ lux[(a)]	Running[(b)]	Navigation Task[(c)]
Intense Brightness + High Vibration + Fatigue	50,000+ lux[(a)]	Running[(b)]	Weighted Load[(e)]
Intense Brightness + High Vibration + Fatigue + Distraction	50,000+ lux[(a)]	Running[(b)]	Navigation Task[(c)] + Weighted Load[(e)]

Table 7. Experimental Conditions in User Study

[(a)] Outdoor conditions occur under strong midday natural lighting (50,000+ lux, clear sky).
[(b)] Motion occurs along a straight 50m path, either indoors (office hallway, no obstacles) or outdoors (running track, no pedestrians).
[(c)] Participants navigated through soccer training cones (30 cm tall) placed 5 meters apart with no external distractions, passing yellow cones on the left and orange cones on the right while reading.
[(d)] High-vibration treadmill condition involves running at 6 km/h with no incline or handrails.
[(e)] Weighted load consists of a 3 kg backpack (containing a laptop and two books), worn with straps adjusted snugly for even weight distribution.

5.3 Conditions

Guided by the literature on factors contributing to SVIs [36, 58, 59] and insights from our formative study, The eight SVI conditions (see Table 7 and Figure 9) were designed to simulate various real-world reading scenarios, manipulating lighting, movement, and cognitive load.

5.4 Apparatus

All pre-test and experimental sessions used a standardized Android phone to ensure consistency. During the adaptation period, participants used either their own Android devices or the provided Android phone. Read-aloud sessions were audio-recorded for later analysis. Passage order, conditions, and app usage were counterbalanced using a Latin square design to minimize order effects.

5.5 Measures

5.5.1 Reading Performance. Reading performance was assessed through goodput (characters per minute, CPM) and reading comprehension accuracy. Goodput was calculated as the total number of correctly read characters divided by reading time, following Ku et al. [25]. Reading comprehension was measured by the percentage of correctly answered questions, using five multiple-choice questions per passage, pre-defined by the HSK Level 5 examination [2, 67].

5.5.2 Perceived Workload & User Experience. Perceived workload was evaluated using NASA-TLX [18], measuring mental demand, physical demand, temporal demand, performance, effort, and frustration. Overall usability was assessed through UEQ-S [51] and SUS [1], capturing users' subjective impressions of each interface.

5.5.3 Qualitative feedback. Post-study semi-structured interviews were conducted to gather in-depth qualitative feedback. These interviews explored participants' overall experiences with both SituFont and the traditional display, including encountered difficulties, perceived benefits, and comparative preferences. Thematic analysis of the interview transcripts was conducted following the guidelines outlined by Braun and Clarke [3].

6 RESULT

Over the four-day adaptation period, participants generated 490 valid data entries, with an average of 7 recorded reading scenarios per participant. The following sections evaluate SituFont's effectiveness in improving reading performance, reducing workload, and enhancing user experience across the eight SVI conditions.

6.1 Reading Performance

6.1.1 Reading Goodput. **SituFont led to a significant increase in reading goodput compared to traditional manual font adjustments.** A paired sample t-test confirmed this improvement, showing that, except for the Intense Brightness Scenario ($t = -1.961, p = 0.076$) and High Vibration Scenario ($t = -2.166, p = 0.052$), SituFont significantly outperformed the traditional app in all other conditions (Table 8). To further contextualize this improvement, an ANOVA analysis incorporating the pre-experiment baseline was conducted (Figure 10). The observed trend, ***SituFont (comparative study)*** > ***Traditional app (comparative study)*** > ***Traditional app (pre-test)***, suggests that SituFont not only enhances reading goodput but that the adaptation period itself may contribute to improved reading efficiency. This effect could be attributed to increased user awareness of optimal font parameters for different SVI conditions.

6.1.2 Reading Comprehension. **Overall, reading comprehension accuracy remained comparable between SituFont and the traditional app across all task scenarios (Figure 10).** While minor variations were observed, such as SituFont slightly outperforming the traditional app in the Intense Brightness + High Vibration + Distraction Scenario and High Vibration Scenario, and the traditional app performing marginally better in the Intense Brightness + High Vibration Scenario and Fatigue Scenario, these differences were not statistically significant. The overlapping error bars suggest that both applications provided a consistent comprehension experience.

Scenarios ID [1]	1	2	3	4	5	6	7	8
SituFont	300.64	288.25	271.51	304.58	295.78	352.52	288.34	292.83
	(56.80)	(40.72)	(50.05)	(39.36)	(48.91)	(51.06)	(31.97)	(47.65)
Traditional App	285.32	241.95	256.10	249.14	273.26	273.33	256.77	246.83
	(39.66)	(49.01)	(70.46)	(52.39)	(50.30)	(51.05)	(50.38)	(53.02)
T Value	-1.961	-7.741	-2.166	-5.634	-2.742	-10.122	-2.245	-3.256
p Value	0.076	0.000**	0.052	0.000**	0.019*	0.000**	0.046*	0.008*

* $p < 0.05$, ** $p < 0.01$

Table 8. Mean (standard deviation) of reading goodput (CPM) across different SVIs scenarios. The results of paired t-tests demonstrated that SituFont's observed goodput improvement was statistically significant

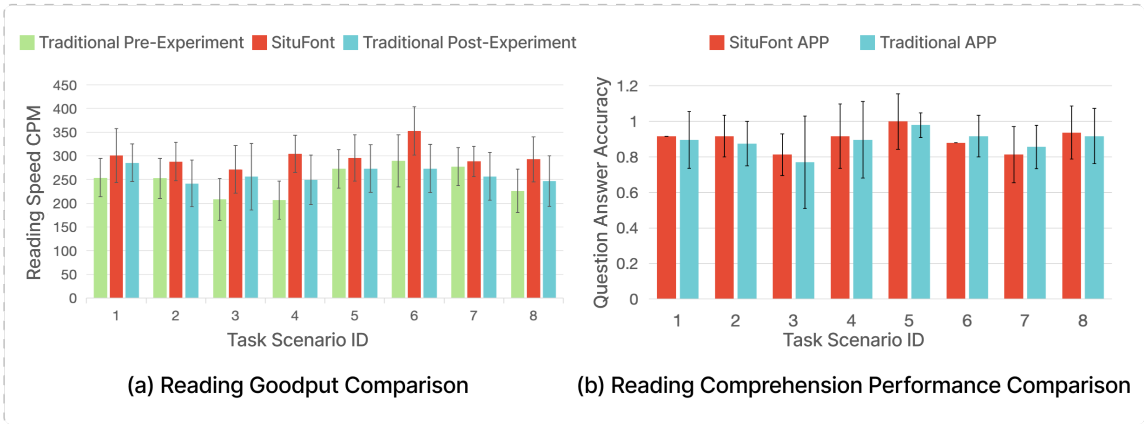

Fig. 10. Comparison of reading goodput (a) and comprehension accuracy (b) across different SVI conditions. SituFont consistently demonstrated improved goodput, while comprehension accuracy remained stable across both interfaces.[1]

6.2 Perceived Workload

Participants reported significantly lower mental and physical workload when using SituFont compared to the traditional app (Figure 11). Since the Shapiro-Wilk test indicated that workload measures did not follow a normal distribution ($p < 0.05$ for all), Wilcoxon signed-rank tests were used for pairwise comparisons between the two app conditions within each experimental scenario.

SituFont significantly reduced both mental demand and physical demand in multiple conditions. The Wilcoxon signed-rank test showed that SituFont resulted in significantly lower **mental demand** under Intense Brightness condition ($W = 2.0$, $p = 0.068$), Distraction condition ($W = 0.0$, $p = 0.010$), and Fatigue condition($W = 0.0$, $p = 0.011$). **Physical demand** was significantly lower in Intense Brightness condition ($W = 9.0$, $p = 0.055$), Intense Brightness + High Vibration condition ($W = 8.5$, $p = 0.050$), and High Vibration condition ($W = 0.0$, $p = 0.016$). These results

[1]For the convenience of analysis, scenarios are labeled as: 1 - Intense Brightness + Normal, 2 - Distraction + Normal, 3 - Intense Brightness + High Vib + Normal, 4 - Intense Brightness + High Vib + Distraction, 5 - High Vib + Normal, 6 - Fatigue + Normal, 7 - Intense Brightness + High Vib + Fatigue + Normal, 8 - Intense Brightness + High Vib + Distraction + Fatigue.

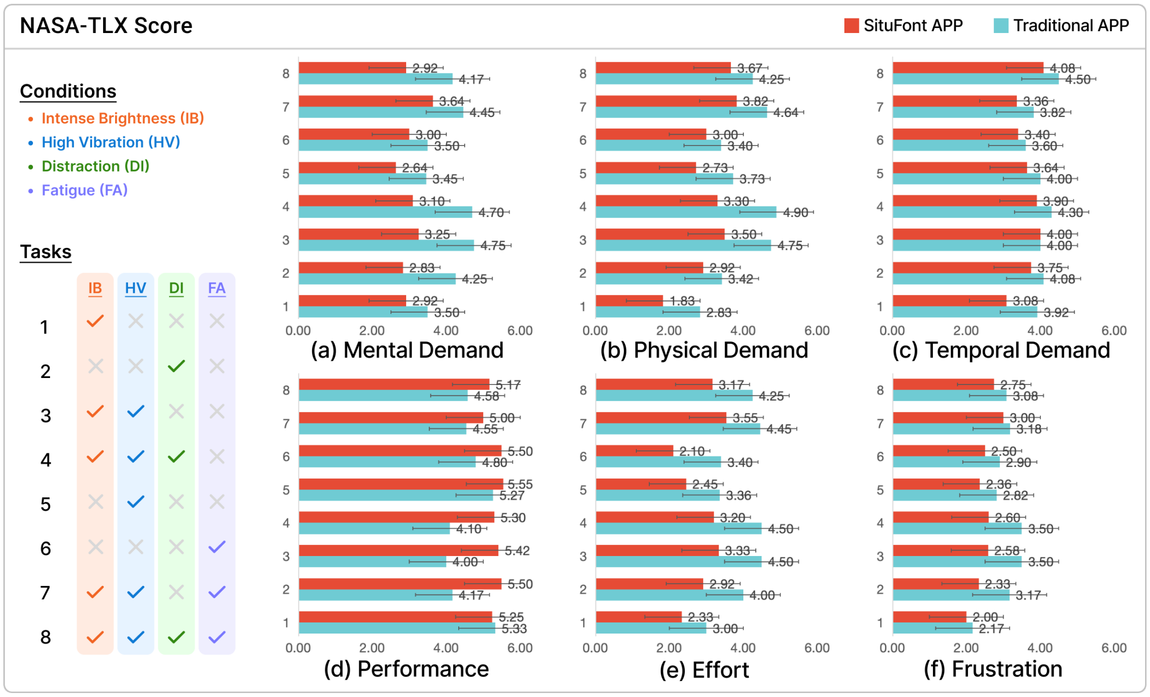

Fig. 11. NASA-TLX scores across experimental conditions, illustrating that SituFont consistently reduced mental and physical workload compared to the traditional app. No significant differences were observed in temporal demand, while frustration levels showed minor variations. For convenience of analysis, we labeled different experimental conditions with numbers, with each number representing specific conditions as shown on the left.

suggest that SituFont effectively reduces cognitive and physical workload under bright lighting, high vibration, and fatigue-related impairments.

Differences in **temporal demand** were less pronounced. A statistically significant reduction in temporal demand was found only in the High Vibration condition ($W = 9.0$, $p = 0.380$), indicating that SituFont marginally improved time efficiency under vibration stress.

Users perceived higher **performance** with SituFont under Distraction condition ($W = 0.0$, $p = 0.004$), Intense Brightness + High Vibration condition ($W = 0.0$, $p = 0.004$), and Intense Brightness + High Vibration + Distraction condition ($W = 1.5$, $p = 0.058$). Additionally, **effort** was significantly lower in Distraction condition ($W = 5.5$, $p = 0.041$), Intense Brightness + High Vibration + Distraction condition ($W = 0.0$, $p = 0.026$), and Fatigue condition ($W = 0.0$, $p = 0.011$). These findings indicate that SituFont enhances reading performance under high-intensity brightness and vibration while reducing effort in distracting and fatigue-inducing environments.

Frustration SituFont significantly reduced **frustration levels** in Distraction condition ($W = 5.5$, $p = 0.075$) and Intense Brightness + High Vibration + Distraction + Fatigue condition ($W = 6.0$, $p = 0.084$). This suggests that the traditional app may cause greater frustration under distracting and multi-factor impairment conditions.

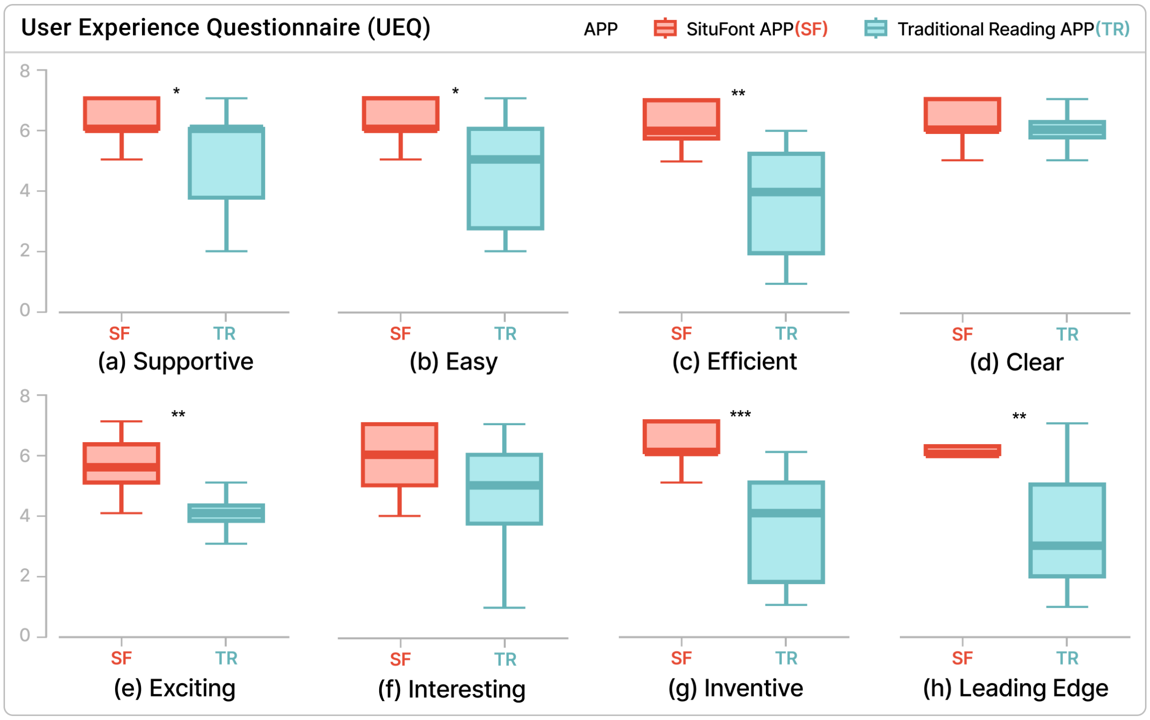

Fig. 12. User Experience Questionnaire (UEQ) results, comparing SituFont and the Traditional App across multiple dimensions. SituFont scored significantly higher in efficiency, supportiveness, and novelty, with marginal differences in stimulation and perspicuity. ***: p<.001, **: p<.01, *: p<.05

6.3 User Experience and Preferences

User experience evaluations revealed that SituFont was generally preferred over the traditional app, with higher ratings in efficiency, ease of understanding, and novelty. While some aspects of stimulation showed marginal differences, participants perceived SituFont as significantly more supportive, efficient, and easy to use (Figure 12).

6.3.1 User Experience Questionnaire (UEQ). **Attractiveness and Efficiency were key aspects where SituFont demonstrated a clear advantage.** In the *Inefficient: Efficient* scale, SituFont scored significantly higher (M = 6.08, SD = 1.00) compared to the Traditional App (M = 3.92, SD = 2.27), p = 0.0005, suggesting a strong perception of efficiency. Similarly, on the *Obstructive: Supportive* scale, SituFont (M = 6.17, SD = 0.94) significantly outperformed the Traditional App (M = 4.92, SD = 1.73), p = 0.039, indicating that users found SituFont more supportive in facilitating their reading experience.

The **Stimulation** scale measured engagement and interest. While SituFont scored higher in *Not interesting: Interesting* (M = 5.75, SD = 1.36) and *Boring: Exciting* (M = 5.50, SD = 1.38) compared to the Traditional App (M = 4.50, SD = 1.83) and (M = 4.58, SD = 1.88), respectively, the differences did not reach statistical significance (p = 0.063 and p = 0.160).

Novelty was another category where SituFont was rated higher, indicating a perception of innovation. On the *Conventional: Inventive* scale, SituFont (M = 5.58, SD = 1.56) significantly outperformed the Traditional App (M =

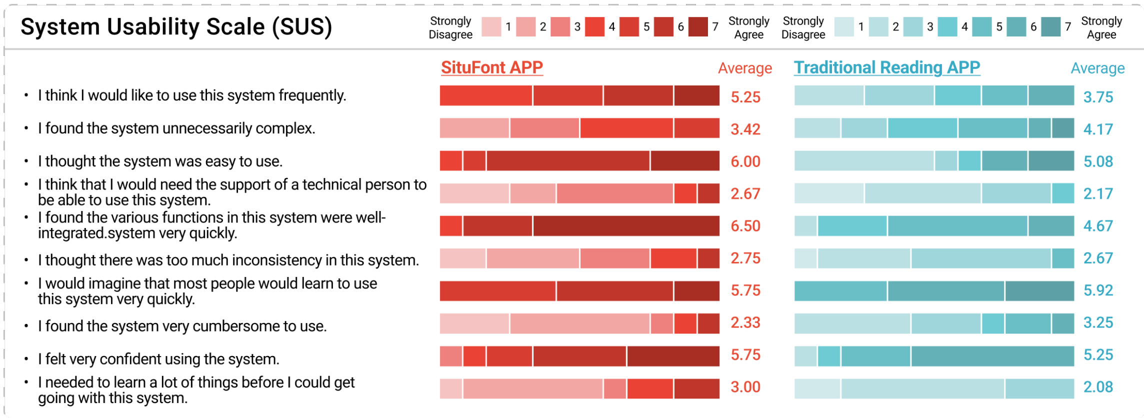

Fig. 13. System Usability Scale (SUS) results. SituFont was rated significantly lower in complexity and higher in ease of use and consistency, though it required a slightly higher initial learning effort.

4.08, SD = 2.23), p = 0.024. Similarly, in the *Leading edge* dimension, SituFont (M = 5.92, SD = 1.62) scored significantly higher than the Traditional App (M = 3.50, SD = 2.11), p = 0.002, reinforcing the perception that SituFont introduces an innovative and modern approach.

Perspicuity, which measures ease of understanding, also showed notable differences. SituFont scored a mean of (M = 6.25, SD = 0.75) for *Confusing: Clear* and (M = 6.08, SD = 0.90) for *Complicated: Easy*, while the Traditional App scored (M = 6.00, SD = 0.74) and (M = 4.50, SD = 1.98), respectively. Although the difference in *Confusing: Clear* was not significant (p = 0.421), the difference in *Complicated: Easy* was statistically significant (p = 0.019), suggesting that SituFont was perceived as easier to use.

6.3.2 System Usability Scale (SUS). System usability evaluations indicated that **SituFont was generally perceived as more user-friendly and consistent, with some minor learning challenges reported (Figure 13)**.

For the statement *"I found the system unnecessarily complex,"* SituFont (M = 2.58, SD = 1.62) was rated significantly lower than the Traditional App (M = 4.08, SD = 2.50), p = 0.045, indicating that users perceived SituFont as less complex.

In terms of ease of use, SituFont scored significantly higher on the statement *"I thought the system was easy to use"* (M = 6.42, SD = 0.90) compared to the Traditional App (M = 4.67, SD = 2.43), p = 0.012, reinforcing that users found SituFont easier to interact with.

For system consistency, SituFont received a mean score of (M = 1.08, SD = 0.29) for *"I thought there was too much inconsistency in this system"*, significantly lower than the Traditional App (M = 1.83, SD = 1.03), p = 0.014, suggesting that users found SituFont to be more consistent in its functionality.

Confidence in using the system was also measured, with SituFont scoring (M = 5.75, SD = 1.29) and the Traditional App scoring (M = 5.25, SD = 1.22), but this difference was not statistically significant (p = 0.402).

Lastly, for *"I needed to learn a lot of things before I could use this system,"* SituFont scored higher (M = 3.00, SD = 1.65) than the Traditional App (M = 2.08, SD = 0.67), p = 0.030, indicating that SituFont required a slightly higher initial learning effort.

6.3.3 Qualitative Feedback. Participants highlighted the **convenience of automatic font adjustment**, particularly in dynamic contexts like cycling and running. P2 emphasized it was especially useful after sufficient training data. P5 noted, "*When I'm running, long-pressing is much more convenient than manually selecting.*" The hands-free activation via long-press was well-received with P3 describing it as effortless, and P4 appreciating the single-handed operation and quick immersion in reading. Several users also noted that SituFont made them more aware of their reading environment and its impact, as P9 mentioned, "*It helped me consciously change bad reading scenarios; without this experiment, I wouldn't have realized the impact of reading fonts on the reading experience.*"

However, some usability challenges emerged. P8 remarked that automatic scene detection was sometimes inaccurate due to *"reliance on historically recorded descriptions."* The long-press activation, though generally useful, occasionally lacked sensitivity, as P9 explained, *"Sometimes it's not sensitive enough, sometimes I don't want to... but it's triggered."* Additionally, manual adjustments using the horizontal slider were reported as difficult during movement, as P2 commented: "*when the phone is very bumpy*". P3 found the slider placement awkward for one-handed use, saying, *"It's difficult to reach the adjustment cursor."* Further improvements are needed, including adding paragraph spacing adjustment as noted by P5, and addressing the disruption to reading flow caused by font changes as P4 stated, *"(It) interrupts the original reading rhythm after adjusting the font."*

SituFont's adaptability also requires further refinement. P2 highlighted the need for larger font sizes due to impaired vision, suggesting a post-calibration proactive adjustment inquiry. *"For example, after a long press to trigger manual adjustment, it could actively ask if I need to increase the font size based on my vision situation."* The post-adjustment vibration was also considered disruptive. P5 mentioned, *"The phone vibrates after each successful adjustment, that interferes with my reading experience."*

7 DISCUSSION

We present SituFont, a system for personalized font adjustments based on real-time user context. While our findings demonstrate the potential of SituFont to improve reading efficiency, its impact on reading comprehension was less pronounced. This section discusses key design considerations for both SVIs and JITAI with a human-in-the-loop, as well as ethical implications and future research directions.

7.1 Design Considerations for SVIs Intervention

SituFont was designed to enhance reading efficiency by automating font adjustments based on contextual factors. The significant improvement in reading goodput across most experimental conditions strongly supports this goal, demonstrating that automation reduces the cognitive and physical effort associated with manual font adjustments. Participants' qualitative feedback confirmed that the ability to automate font changes contributed to a more seamless reading experience.

SituFont's current activation method, which relies on **a long-press gesture, posed usability challenges** as participants frequently activated the font adjustment unintentionally while scrolling or highlighting text, leading to frustration and disrupting their reading experience. This issue necessitates exploring alternative activation methods that offer more precise and user-friendly control.

Integrating voice control would allow users to activate and adjust font parameters hands-free. Users could simply utter a command like "Increase font size" or "Adjust for low light" to trigger the desired changes. This approach would be particularly beneficial in situations where hands-free operation is preferred, such as while commuting or exercising.

Providing users with the ability to **customize the activation gesture** could further enhance usability. Allowing users to choose from a range of gestures, such as double-tap, swipe, or a custom combination of touches, would cater to individual preferences and minimize accidental triggers based on their typical reading habits.

Implementing a vertical slider accessible from the edge of the screen, would potentially provide a more ergonomic and deliberate activation method, especially for one-handed phone use. The slider's position and size could be customized to minimize accidental touches while maintaining accessibility.

7.2 Design Considerations for JITAI with Human-in-the-loop

SituFont's JITAI component enables personalized font adaptation by incorporating user feedback into the adjustment process. However, several design aspects require further refinement.

The **current binary feedback mechanism** (yes/no) for user feedback on personal factors limits the system's understanding of their specific needs. While useful for initial adaptation, it lacks the nuance required for truly personalized adjustments. Implementing a more detailed feedback system, such as Likert-scale ratings for distraction, fatigue, and visual strain, or an option to select predefined reasons for discomfort, would provide richer data and enable a more tailored font adjustment experience.

The **cold-start problem** remains a challenge in personalized adaptive systems like SituFont. During the first two days of the adaptation period, the system's performance during this period is suboptimal due to the lack of personalized information. One potential solution is to pre-train a model using data from the formal study 2 (3.2) and fine-tuning it with incoming user data. This approach leverages existing knowledge to bootstrap the personalization process. Further exploration of few-shot learning techniques[13], which aim to train effective models with limited data, could enhance the system's ability to adapt rapidly to new users.

Long-Term usage patterns also warrant further investigation. Although the study focused on short-term adaptation, participants reported increased awareness of how font settings influenced their readability experience. This learning effect suggests that, over time, users may develop strategies for optimizing font settings even outside the SituFont system. The improvement in reading goodput within the traditional app between the pre-experiment and the comparative study suggests that exposure to font adjustments may contribute to better reading efficiency over time. Longitudinal studies could explore whether SituFont induces lasting behavioral changes and whether adaptation strategies should evolve as users gain more experience with the system.

7.3 Ethical Considerations

SituFont's reliance on sensor data raised **privacy concerns** among participants, particularly regarding the system's permission requirements. Maintaining ethical standards and protecting user privacy is critical when designing and deploying systems involving active cameras and microphones like ours. Techniques such as anonymization, encryption, on-device federated learning, or differential privacy, should be explored to prevent breaches when using cloud servers for real-time machine learning inferences.

Furthermore, the potential for exclusion, bias, and discriminatory inferences arising from the use of cameras, microphones, and LLMs must be carefully addressed. Lastly, increasing **transparency about data collection practices and the rationale behind system actions**, such as explaining how data is collected and used to inform font adjustments can potentially foster trust and encourage user adoption.

8 LIMITATIONS AND FUTURE WORK

While this study demonstrates SituFont's potential, several limitations must be acknowledged, guiding future research. First, the participant pool was limited to young adults, restricting generalizability. Future studies should include older adults and individuals with diagnosed visual impairments to assess SituFont's effectiveness across diverse populations.

Second, SituFont's reliance on sensor data raises privacy concerns, particularly regarding the collection of contextual information such as lighting and motion. While essential for adaptation, these data sources may raise security and ethical challenges. Future work should explore privacy-preserving techniques to ensure personalization while minimizing data collection. Additionally, increasing transparency in data processing and giving users greater control over their information could foster trust and improve adoption.

Third, while SituFont significantly improved reading efficiency, its effect on comprehension remains unclear. The HSK reading tasks used in this study may have been too simple for native speakers, limiting the ability to detect meaningful differences. Future studies should include more complex reading materials, such as technical, academic, and long-form texts, to assess SituFont's impact under varied cognitive loads.

Fourth, the adaptive font adjustments may disrupt page layouts and increase scrolling frequency, potentially affecting reading flow. While optimization of typography enhances readability, abrupt changes in text presentation may introduce usability issues. Future iterations should explore layout-aware adaptation strategies that preserve text stability while accommodating font adjustments.

Finally, this study was conducted in Chinese, offering insights mainly on logographic script readability. While Chinese characters generally maintain uniform character block alignment, alphabetic languages involve variable word lengths and text reflow. Future research could extend these findings to bidirectional scripts, inflected languages, and dynamic text rendering, addressing cross-linguistic readability challenges to enhance multilingual usability.

In addition to these limitations, future work should also examine longitudinal effects, as users' exposure to adaptive typography may lead to improved self-adjustment strategies over time. Studying extended user interactions with JITAI systems could provide insights into whether personalization strategies evolve with prolonged use. Refining personalization mechanisms and adapting Situfont for diverse linguistic and accessibility needs (e.g. screen readers, contrast adjustments) will be crucial for its continued development as an intelligent reading interface.

9 CONCLUSION

This paper introduced SituFont, a just-in-time adaptive intervention system designed to enhance mobile readability under SVIs. Findings from our formative (N=15) and exploratory (N=18) studies identified key SVI factors affecting readability, informing the integration of contextual cues such as motion, lighting, and user preferences into SituFont's design. Using smartphone sensors and a human-in-the-loop approach, the system dynamically personalizes text presentation to adapt to changing reading conditions. A comparative evaluation (N=12) demonstrated SituFont's effectiveness in improving readability and reducing task workload across simulated SVI scenarios. The system significantly enhanced reading efficiency, lowered mental and physical effort, and increased user awareness of font readability factors, as highlighted in qualitative feedback. Our findings suggest that real-time, personalized font adjustments hold promise for improving the reading experience in dynamic environments. We hope SituFont inspires future research on intelligent mobile interfaces and broader accessibility solutions that enhance usability across diverse contexts.

A STUDY 2 EXPERIMENTAL SYSTEM AND EQUIPMENT

We developed an application with six pre-loaded reading materials in Chinese, ensuring isomorphic language, structure, and difficulty (high school level). Each 550-character passage was presented in plain text, with no additional visual elements. Passages were pre-loaded, each appearing once during the experiment.

The application allowed real-time text adjustments via a one-handed interface: double-tapping opened a menu for font size, weight, line spacing, and letter spacing. Smartphone sensors recorded light intensity, three-axis acceleration, and reading distance. After participants completed their adjustments, they could upload the text parameter data and environmental data by tapping any other area on the screen.

To avoid the impact of screen size and resolution differences, we used two HUAWEI P40 smartphones as the experimental devices, each with a 6.1-inch screen and a resolution of 2340×1080 pixels.

A.1 Experimental Procedure

Before the experiment, the lead researcher assigned participant IDs, explained procedures and demonstrated a trial. Participants read passages in indoor and outdoor settings while standing, walking, and running, with 3-minute reading durations per passage. During reading, participants encountered SVIs caused by environmental factors such as strong light, vibration, or reading distance, and they were instructed to adjust text parameters to mitigate these issues. Devices were held one-handed, with adjustments permitted anytime.

A.2 Data Collection

Smartphone sensors recorded environmental data throughout the experiment. Reading distance was measured using the front camera, ambient light intensity was captured by the light sensor, and three-axis vibration displacement was tracked via the accelerometer. Text parameter adjustments were logged using Android's built-in functions, with line spacing and letter spacing expressed in *em* units. For example, a line spacing of $1em$ equaled the text height, while $0.05em$ represented 5% of the font size.

A.3 Detailed Result

ID	Scenario	X-Axis (m/s^2)	Y-Axis (m/s^2)	Z-Axis (m/s^2)	Light (lux)	Reading Distance (cm)
1	Indoor Standing	0.18 ± 0.13	0.11 ± 0.08	0.26 ± 0.22	111.69 ± 89.14	33.00 ± 11.46
2	Indoor Walking	0.53 ± 0.15	0.65 ± 0.15	0.80 ± 0.28	86.08 ± 81.86	31.84 ± 9.69
3	Indoor Running	1.41 ± 0.59	1.42 ± 0.85	1.58 ± 0.55	82.42 ± 62.78	30.99 ± 8.73
4	Outdoor Standing	0.16 ± 0.08	0.10 ± 0.05	0.24 ± 0.13	51208.67 ± 27990.79	29.96 ± 10.98
5	Outdoor Walking	0.52 ± 0.14	0.68 ± 0.14	0.78 ± 0.19	38263.15 ± 23249.25	30.84 ± 9.03
6	Outdoor Running	1.54 ± 0.55	1.52 ± 0.90	1.70 ± 0.55	36482.24 ± 25716.72	29.33 ± 8.57
F Value		261.146	121.825	258.395	125.669	1.496
p Value		0.000**	0.000**	0.000**	0.000**	0.189

* $p < 0.05$, ** $p < 0.01$

Table 9. Variance Analysis Table of Sensor Data Features in 6 Experimental Scenarios

B LABEL TREE PROMPT

B.1 Context Label Generation & Selection

system_role:

(You need to determine my current status based on the environmental information and photos I provide, including movement, environment, and personalized description.

- *Movement* refers to my current physical activity,

- *Environment* refers to the setting I am currently in, and
- *Personalized Description* refers to any specific details related to my status or environment.

In order to obtain labels that describe my status, you should follow the steps below based on the environmental information and two photos I provide:)

step1:
(*Step 1: Based on the location information, movement status, and photos I provide, choose my 'movement status' from the following three options: 1. Still 2. Walking 3. Running*)

step2:
(*Step 2: Combine the provided location information and images to determine my current 'environment name.'*)

step3:
(*Step 3: Based on the judgments from the first two steps and the photos, select the most appropriate status label from the provided label options:*)

example:
(*For example: If my location and photos show that I am in an office, and my movement status is still, the status label would be "Still - Office";*)

attention:
(*Note: Ensure the result format is 'Movement-Environment-Personalized Description' or 'Movement-Environment.'*
!!! Please strictly return results in the following format:
Movement-Environment-Personalized Description or **Movement-Environment**)

B.2 Prompt for Editing Label

user_description_prompt:
(*This is the current label used to describe my Movement-Environment-Personalized Description:*)

user_step1:
(*I first want to modify some parts of the movement, environment, or personalized description.*
For example, if the user says, "I am in an office," the user wants to change the environment.
If the user says, "I am running," then the user wants to modify the movement.
If the user says, "I am wearing a hat," the personalized description should add "wearing a hat."
If the user says, "I am not wearing glasses," the personalized description should emphasize "no glasses.")

user_step2:
(*Based on the following requirements, update the current Movement-Environment-Personalized Description to reflect the new status.*)

format_attention:
(*!!! Remember: only output the modified label, nothing else.*)

select_label_role:
(*You need to select the closest matching label from an existing label list based on the environmental information, photos, and current status label I provide.*)

REFERENCES

[1] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *Int. J. Hum.-Comput. Interact.* 24, 6 (2008), 574–594. https://doi.org/10.1080/10447310802205776

[2] Leon Barnard, Ji Soo Yi, Julie A Jacko, and Andrew Sears. 2007. Capturing the effects of context on human performance in mobile computing systems. *Personal and Ubiquitous Computing* 11 (2007), 81–96. https://doi.org/10.1007/s00779-006-0063-x

[3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[4] Rachel Calvin and Shravya Suresh. 2021. Image captioning using convolutional neural networks and recurrent neural network. In *2021 6th International Conference for Convergence in Technology (I2CT)*. IEEE, 1–4. https://doi.org/10.1109/I2CT51068.2021.9418001

[5] Scott Carter and Laurent Denoue. 2009. SeeReader: An (Almost) Eyes-Free Mobile Rich Document Viewer. *arXiv preprint arXiv:0909.2185* (2009).

[6] Anthony DeFulio, Hayley D Brown, Rosemarie M Davidson, Sean D Regnier, Navdeep Kang, and Melissa Ehart. 2014. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophrenia Bulletin* (2014).

[7] Jordan Evans. 2023. Measuring the effects of bright mobile device display illumination on achromatic colour perception. https://hdl.handle.net/10214/27365

[8] Jordan Evans. 2023. Measuring the effects of bright mobile device display illumination on achromatic colour perception. https://hdl.handle.net/10214/27365

[9] Stefan Forsstrom and Victor Kardeby. 2014. Estimating contextual situations using indicators from smartphone sensor values. In *2014 IEEE International Conference on Internet of Things (iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom)*. IEEE, 243–250. https://doi.org/10.1109/iThings.2014.43

[10] Mayank Goel, Leah Findlater, and Jacob Wobbrock. 2012. WalkType: using accelerometer data to accomodate situational impairments in mobile touch screen text entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2687–2696. https://doi.org/10.1145/2207676.2208662

[11] Stephanie P Goldstein. 2018. *Comparing effectiveness and user behaviors of two versions of a just-in-time adaptive weight loss smartphone app.* Doctoral dissertation. Drexel University.

[12] Jorge Goncalves, Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Simo Hosio, Sirkka Risanen, Hannu Rintamäki, and Vassilis Kostakos. 2017. Tapping task performance on smartphones in cold temperature. *Interacting with Computers* 29, 3 (2017), 355–367. https://doi.org/10.1093/iwc/iww029

[13] Taesik Gong, Yewon Kim, Adiba Orzikulova, Yunxin Liu, Sung Ju Hwang, Jinwoo Shin, and Sung-Ju Lee. 2023. DAPPER: Label-Free Performance Estimation after Personalization for Heterogeneous Mobile Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 55 (June 2023), 27 pages. https://doi.org/10.1145/3596256

[14] Tovi Grossman, Daniel Wigdor, and Ravin Balakrishnan. 2007. Exploring and reducing the effects of orientation on text readability in volumetric displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 483–492. https://doi.org/10.1145/1240624.1240702

[15] David H. Gustafson, Fiona M. McTavish, Ming-Yuan Chih, Amy K. Atwood, Roberta A. Johnson, Michael G. Boyle, Michael S. Levy, Hilary Driscoll, Steven M. Chisholm, Lisa Dillenburg, Andrew Isham, and Dhavan Shah. 2014. A smartphone application to support recovery from alcoholism: A randomized clinical trial. *JAMA Psychiatry* 71, 5 (2014), 566–572.

[16] Chihiro Hantani, Airi Tsuji, and Kaori Fujinami. 2022. A study on projection-based public displays that attract people with peripheral vision. In *International Conference on Human-Computer Interaction*. Springer, 13–27. https://doi.org/10.1007/978-3-031-05463-1_2

[17] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. 2019. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 16 (2019), 1–21.

[18] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology: Volume 52*. Elsevier, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[19] Juan David Hincapié-Ramos and Pourang Irani. 2013. CrashAlert: enhancing peripheral alertness for eyes-busy mobile interaction while walking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3385–3388. https://doi.org/10.1145/2470654.2466463

[20] K S Hong, J Park, K H Kim, et al. 2020. Implementation of automatic adjustment of font size system on smartphone. In *Advances in Computer Science and Ubiquitous Computing*. Springer Singapore, 159–164.

[21] Ding-Long Huang, Pei-Luen Patrick Rau, and Ying Liu. 2009. Effects of font size, display resolution and task type on reading chinese fonts from mobile devices. *International Journal of Industrial Ergonomics* 39, 1 (2009), 81–89. https://doi.org/10.1016/j.ergon.2008.09.004

[22] Shih-Miao Huang. 2019. Effects of font size and font style of traditional chinese characters on readability on smartphones. *International Journal of Industrial Ergonomics* 69 (2019), 66–72. https://doi.org/10.1016/j.ergon.2018.10.002

[23] Xiaoqing Jing, Chun Yu, Kun Yue, Liangyou Lu, Nan Gao, Weinan Shi, Mingshan Zhang, Ruolin Wang, and Yuanchun Shi. 2024. AngleSizer: Enhancing Spatial Scale Perception for the Visually Impaired with an Interactive Smartphone Assistant. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 108 (Sept. 2024), 31 pages. https://doi.org/10.1145/3678525

[24] Taslim Arefin Khan, Dongwook Yoon, and Joanna McGrenere. 2020. Designing an Eyes-Reduced Document Skimming App for Situational Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376641

[25] Pin-Sung Ku, Yu-Chih Lin, Yi-Hao Peng, and Mike. Chen. 2019. PeriText: Utilizing Peripheral Vision for Reading Text on Augmented Reality Smart Glasses. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 630–635. https://doi.org/10.1109/VR.2019.8798065

[26] Yongquan Li. 2022. An Empirical Study on Wrongly Written or Mispronounced Characters in Hanyu Shuiping Kaoshi Compositions from the Perspective of Readability Research. In *Proceedings of the 6th International Conference on Education and Multimedia Technology* (Guangzhou, China)

(ICEMT '22). Association for Computing Machinery, New York, NY, USA, 252–257. https://doi.org/10.1145/3551708.3551772

[27] Min Lin, Rich Goldman, Kathleen J. Price, Andrew Sears, and Julie Jacko. 2007. How do people tap when walking? An empirical investigation of nomadic data entry. *Int. J. Hum.-Comput. Stud.* 65, 9 (Sept. 2007), 759–769. https://doi.org/10.1016/j.ijhcs.2007.04.001

[28] Haotian Liu, Chunyuan Li, Yuheng Li, et al. 2023. Improved baselines with visual instruction tuning. https://doi.org/10.48550/arXiv.2310.03744 [A/OL].

[29] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-Teh Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. 2019. MediaPipe: A Framework for Building Perception Pipelines. arXiv:1906.08172 [cs.DC] https://arxiv.org/abs/1906.08172

[30] Guojie Ma. 2017. Does interword spacing influence lexical processing in Chinese reading? *Visual Cognition* 25, 7-8 (2017), 815–824. https://doi.org/10.1080/13506285.2017.1338322

[31] Jacqueline Louise Mair, Lawrence D Hayes, Amy K Campbell, Duncan S Buchan, Chris Easton, and Nicholas Sculthorpe. 2022. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. *JMIR formative research* 6, 4 (2022), e34662.

[32] Khalid Majrashi. 2022. Performance of mobile users with text-only and text-and-icon menus in seated and walking situations. *Behaviour & Information Technology* 41, 1 (2022), 32–50.

[33] Randi C Martin, Michael S Wogalter, and Janice G Forlano. 1988. Reading comprehension in the presence of unattended speech and music. *Journal of memory and language* 27, 4 (1988), 382–398. https://doi.org/10.1016/0749-596X(88)90063-0

[34] Kazushi Maruya, Miki Uetsuki, Hideyuki Ando, et al. 2012. "Yu bi yomu": Interactive reading of dynamic text. In *Proceedings of the 20th ACM International Conference on Multimedia*. 1499–1500. https://doi.org/10.1145/2393347.2396535

[35] Sachi Mizobuchi, Mark Chignell, and David Newton. 2005. Mobile text entry: relationship between walking speed and text input task difficulty. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services* (Salzburg, Austria) *(MobileHCI '05)*. Association for Computing Machinery, New York, NY, USA, 122–128. https://doi.org/10.1145/1085777.1085798

[36] Terhi Mustonen, Maria Olkkonen, and Jukka Hakkinen. 2004. Examining mobile phone text legibility while walking. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1243–1246. https://doi.org/10.1145/985921.986034

[37] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2018. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine* 52, 6 (May 2018), 446–462. https://doi.org/10.1007/s12160-016-9830-8

[38] Alexander Ng, Stephen A Brewster, and John Williamson. 2013. The impact of encumbrance on mobile interactions. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part III*. Springer, 92–109. https://doi.org/10.1007/978-3-642-40477-1_6

[39] Alexander Ng, Stephen A. Brewster, and John H. Williamson. 2014. Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1981–1990. https://doi.org/10.1145/2556288.2557312

[40] OpenAI. 2024. ChatGPT. https://openai.com/blog/chatgpt [EB/OL].

[41] Gaisha Oralova and Victor Kuperman. 2021. Effects of spacing on sentence reading in Chinese. *Frontiers in Psychology* 12 (2021), 765335. https://doi.org/10.3389/fpsyg.2021.765335

[42] Adiba Orzikulova, Han Xiao, Zhipeng Li andYukang Yan, Yuntao Wang, Yuanchun Shi, Marzyeh Ghassemi, Sung-Ju Lee, Anind K Dey, and Xuhai Xu. 2024. Time2Stop: Adaptive and Explainable Human-AI Loop for Smartphone Overuse Intervention. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613904.3642747

[43] Abayomi Moradeyo Otebolaku and Maria Teresa Andrade. 2016. User context recognition using smartphone sensors and classification models. *Journal of Network and Computer Applications* 66 (2016), 33–51. https://doi.org/10.1016/j.jnca.2016.03.013

[44] Junghwan Park, Meelim Kim, Mohamed El Mistiri, Rachael Kha, Sarasij Banerjee, Lisa Gotzian, Guillaume Chevance, Daniel E Rivera, Predrag Klasnja, and Eric Hekler. 2023. Advancing understanding of just-in-Time States for supporting physical activity (Project JustWalk JITAI): protocol for a System ID Study of just-in-Time adaptive interventions. *JMIR Research Protocols* 12, 1 (2023), e52161.

[45] I Wayan Pulantara. 2017. *Model, architecture and application of cross-platform just-in-time adaptive intervention (JITAI): An implementation in behavioral sleep intervention.* Doctoral dissertation. University of Pittsburgh.

[46] Matthew Saponaro, Ajith Vemuri, Greg Dominick, and Keith Decker. 2021. Contextualization and individualization for just-in-time adaptive interventions to reduce sedentary behavior. In *Proceedings of the conference on health, inference, and learning*. 246–256.

[47] Zhanna Sarsenbayeva, Jorge Goncalves, Juan García, Simon Klakegg, Sirkka Rissanen, Hannu Rintamäki, Jari Hannu, and Vassilis Kostakos. 2016. Situational impairments to mobile interaction in cold environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 85–96. https://doi.org/10.1145/2971648.2971734

[48] Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Vassilis Kostakos, and Jorge Goncalves. 2017. Challenges of situational impairments during interaction with mobile devices. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (Brisbane, Queensland, Australia) *(OzCHI '17)*. Association for Computing Machinery, New York, NY, USA, 477–481. https://doi.org/10.1145/3152771.3156161

[49] Zhanna Sarsenbayeva, Niels van Berkel, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. 2018. Effect of Distinct Ambient Noise Types on Mobile Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 82 (July 2018), 23 pages. https://doi.org/10.1145/3214285

[50] Bastian Schildbach and Enrico Rukzio. 2010. Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (Lisbon, Portugal) *(MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/1851600.1851619

[51] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4 (01 2017), 103. https://doi.org/10.9781/ijimai.2017.09.001

[52] A Sears, M Lin, J Jacko, et al. 2003. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *HCI international: volume 2*. 1298–1302.

[53] Himanshu Sharma, Manmohan Agrahari, Sujeet Kumar Singh, et al. 2020. Image captioning: a comprehensive survey. In *2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)*. IEEE, 325–328. https://doi.org/10.1109/PARC49193.2020.236619

[54] Hélène Soubaras. 2020. Voice recognition based system to adapt automatically the readability parameters of a user interface. In *Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys) Volume 2*. Springer, 166–178. https://doi.org/10.1007/978-3-030-29513-4_12

[55] Hai-Tao Zheng T, Zhe Wang, Ningning Ma, et al. 2018. Weakly-supervised image captioning based on rich contextual information. *Multimedia Tools and Applications* 77 (2018), 18583–18599. https://doi.org/10.1007/s11042-017-5236-2

[56] Nađa Terzimehić, Christina Schneegass, and Heinrich Hußmann. 2017. Exploring Challenges in Automated Just-In-Time Adaptive Food Choice Interventions. In *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care (MMHealth '17)*. Association for Computing Machinery, 81–84. https://doi.org/10.1145/3132635.3132648

[57] Garreth W. Tigwell, David R. Flatla, and Rachel Menzies. 2018. It's not just the light: understanding the factors causing situational visual impairments during mobile interaction. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction* (Oslo, Norway) *(NordiCHI '18)*. Association for Computing Machinery, New York, NY, USA, 338–351. https://doi.org/10.1145/3240167.3240207

[58] Garreth W. Tigwell, Rachel Menzies, and David R. Flatla. 2018. Designing for Situational Visual Impairments: Supporting Early-Career Designers of Mobile Content. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 387–399. https://doi.org/10.1145/3196709.3196760

[59] Garreth W. Tigwell, Zhanna Sarsenbayeva, Benjamin M. Gorman, David R. Flatla, Jorge Goncalves, Yeliz Yesilada, and Jacob O. Wobbrock. 2019. Addressing the Challenges of Situationally-Induced Impairments and Disabilities in Mobile Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3290607.3299029

[60] Miki Uetsuki, Junji Watanabe, Hideyuki Ando, et al. 2017. Reading traits for dynamically presented texts: comparison of the optimum reading rates of dynamic text presentation and the reading rates of static text presentation. *Frontiers in Psychology* 8 (2017), 269747. https://doi.org/10.3389/fpsyg.2017.01390

[61] Kristin Vadas, Nirmal Patel, Kent Lyons, Thad Starner, and Julie Jacko. 2006. Reading on-the-go: a comparison of audio and hand-held displays. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services* (Helsinki, Finland) *(MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 219–226. https://doi.org/10.1145/1152215.1152262

[62] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[63] Lin Wang, Hitomi Sato, Pei-Luen Patrick Rau, et al. 2008. Chinese text spacing on mobile phones for senior citizens. *Educational Gerontology* 35, 1 (2008), 77–90. https://doi.org/10.1080/03601270802491122

[64] You Wang, Zhihao Zhao, Danni Wang, et al. 2013. How screen size influences chinese readability. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. 277–280. https://doi.org/10.1145/2541016.2541087

[65] Daniel Wigdor and Ravin Balakrishnan. 2005. Empirical investigation into the effect of orientation on text readability in tabletop displays. In *ECSCW 2005: Proceedings of the Ninth European Conference on Computer-Supported Cooperative Work, 18–22 September 2005, Paris, France*. Springer, 205–224.

[66] Chen-Hsiang Yu and Robert C. Miller. 2011. Enhancing mobile browsing and reading. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 1783–1788. https://doi.org/10.1145/1979742.1979845

[67] Chen Zhou, Katherine Fennedy, Felicia Fang-Yi Tan, Shengdong Zhao, and Yurui Shao. 2023. Not All Spacings are Created Equal: The Effect of Text Spacings in On-the-go Reading Using Optical See-Through Head-Mounted Displays. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 720, 19 pages. https://doi.org/10.1145/3544548.3581430

[68] Deyao Zhu, Jun Chen, Xiaoqian Shen, et al. 2023. Minigpt-4: Enhancing vision-language understanding with advanced large language models. https://doi.org/10.48550/arXiv.2304.10592 [A/OL].

[69] Shangshang Zhu, Xinyu Su, and Yenan Dong. 2021. Effects of the font size and line spacing of simplified chinese characters on smartphone readability. *Interacting with Computers* 33, 2 (2021), 177–187. https://doi.org/10.1093/iwc/iwab020

[70] Wei Zhuang, Yi Chen, Jian Su, Baowei Wang, and Chunming Gao. 2019. Design of human activity recognition algorithms based on a single wearable IMU sensor. *Int. J. Sen. Netw.* 30, 3 (Jan. 2019), 193–206. https://doi.org/10.1504/ijsnet.2019.100218