

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

Jingruo Chen*

Information Science
Cornell University
Ithaca, New York, USA
jc3564@cornell.edu

Kexin Nie*

The University of Sydney
Sydney, Australia
knie0519@uni.sydney.edu.au

Mingshan Zhang*

Department of Computer Science and
Technology
Tsinghua University
Beijing, China
zms21@tsinghua.org.cn

Chun Yu

Department of Computer Science and
Technology
Tsinghua University
Beijing, China
chunyu@tsinghua.edu.cn

Zhiqi Gao

School of Data Science
The Chinese University of Hong
Kong, Shenzhen
Shenzhen, China
zhiqigao@link.cuhk.edu.cn

Kun Yue

Tsinghua University
Beijing, Beijing, China
kunyue@infera.im

Yuanchun Shi

Tsinghua University
Beijing, China
Qinghai University
Qinghai, China
shiyc@tsinghua.edu.cn

Chen Liang[†]

Computational Media and Arts Thrust
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, Guangdong, China
liangchenc@163.com

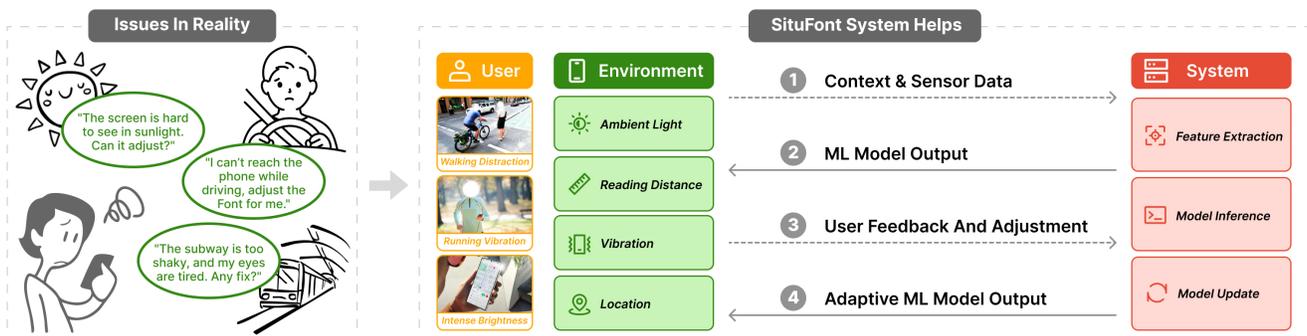


Figure 1: The SituFont system dynamically adapts font parameters based on user context and environmental conditions. The process involves four steps: (1) collecting context and sensor data, (2) generating ML model output, (3) incorporating user feedback for font parameter adjustments, and (4) producing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

Abstract

Situational visual impairments (SVIs) hinder mobile readability, causing discomfort and limiting information access. Building on prior work in adaptive typography and accessibility, this paper

presents SituFont, a context-aware and human-in-the-loop adaptive typography adjustment approach that enhances smartphone mobile 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). 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

*These authors contributed equally to this research.

[†] Corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License.
CHI '26, Barcelona, Spain

© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2278-3/26/04
<https://doi.org/10.1145/3772318.3791020>

SituFont’s effectiveness in improving smartphone mobile readability in terms of improved efficiency and reduced workload compared with a non-trivial manual adjustment baseline.

CCS Concepts

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

Keywords

Situational Visual Impairments, Just-in-time adaptive intervention, Human-in-the-loop, Font Parameters, Text Readability

ACM Reference Format:

Jingruo Chen, Kexin Nie, Mingshan Zhang, Chun Yu, Zhiqi Gao, Kun Yue, Yuanchun Shi, and Chen Liang. 2026. SituFont: A Just-in-Time Adaptive Intervention Interface for Enhancing Mobile Readability in Situational Visual Impairments. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 25 pages. <https://doi.org/10.1145/3772318.3791020>

1 Introduction

Mobile computing presents unique challenges due to its dynamic usage contexts [82]. Unlike desktop computing, mobile devices are used in varied environments, leading to Situationally Induced Impairments and Disabilities (SIIDs) [82, 92]. Among these, Situational Visual Impairments (SVIs) arise from factors like low lighting and user motion, significantly affecting text readability [56, 75, 76]. These challenges lead to visual fatigue, distraction, and difficulty processing information [90].

Current solutions for addressing SVIs include manual adjustments [29], such as increasing font size [31], using auditory substitutes [8, 28, 87], and automatic adjustments [29].

While prior work has explored static solutions such as UI brightness modeling [13], post-hoc coping strategies for low visibility [90], and the effects of font parameters on readability [7, 31], these approaches typically examine individual factors in isolation and rely on fixed or manually triggered adjustments [46]. However, real-world situational visual impairments (SVIs) often involve multiple dynamic contextual factors, such as motion, lighting variation, distraction, and fatigue, that interact and fluctuate over time. Under such conditions, static accessibility settings and simple rule-based interactions (e.g., manually increasing text size or pinch-and-zoom) place continuous attentional and interactional burden on users when visual conditions change rapidly or overlap [38]. Users need to repeatedly detect readability breakdowns and intervene manually, interrupting reading flow and increasing cognitive load.

This gap highlights the need for dynamic and personalized adaptation mechanisms that can respond to changing contexts and learn from user behavior over time. To address this need, we focus on smartphone reading in everyday dynamic environments, such as walking, commuting, or moving across diverse lighting and motion conditions [37, 38, 71, 114], a setting that is highly representative given the prevalence of mobile reading in daily life [41, 84, 85, 104]. Within this context, we focus on Chinese text as a deliberate design probe for examining mobile readability under situational visual impairments, given its dense character-based structure and widespread use on smartphones, while recognizing that typographic adaptation may differ across alphabetic or bidirectional scripts [17].

Building on prior research in adaptive typography [38] and situational visual impairments [46, 90], we propose SituFont, an adaptive typography approach that leverages context-aware sensing and human-in-the-loop personalization to adjust font parameters just in time. SituFont leverages smartphone sensing to infer situational context and supports incremental personalization through user-initiated adjustments, which are incorporated as feedback to refine adaptation over time.

To ground the design of SituFont, we conducted two user studies to investigate the behavior patterns and SVI factors in smartphone mobile reading. We first conducted a semi-structured interview (N=15) to identify users’ practices, challenges, coping mechanisms, and expectations in smartphone mobile reading. Results found three main categories (environmental, personal, and informational) of SVI factors and indicated typography adjustment as a general solution, but with personalized needs. We further conducted a controlled experiment (N=18) to examine how key environmental SVI factors influence the adjustment of key typography parameters. Results showed that different combinations of SVI factors led to desired preferences towards certain typography parameter trends.

Building upon these study results, we grounded the system design and implementation of SituFont. First, we employed a label tree structure as an efficient representation of context, using hierarchical contextual labels that can be customized by the user. Based on such label sets, a machine learning model is adopted to predict the optimal font parameters based on the current context representation. Moreover, a human-in-the-loop interface is designed to collect individual adaptation feedback that can be progressively used to personalize the prediction model. To evaluate SituFont’s effectiveness, we conducted a comparative study (N=12) simulating diverse SVI scenarios. Results show that SituFont improves reading efficiency, reduces perceived workload, and enhances users’ awareness of readability factors comparing with a non-trivial manual adjustment baseline where the user adjusts the parameters to a static optimal according to the scenario before the reading tasks.

Building on prior work in adaptive typography and situational visual impairments, this paper makes the following contributions:

- Empirical insights from qualitative and quantitative studies examining how environmental, personal, and informational factors jointly affect mobile reading under SVIs.
- The design and evaluation of SituFont, a population-informed, human-in-the-loop approach for just-in-time typographic adaptation that supports incremental personalization in dynamic contexts.
- Design considerations for SVI interventions that emphasize the timing, activation, and integration of adaptation into the reading flow.

2 Related Work

2.1 Situational Visual Impairments and Mitigation Strategies

Situation-induced disorders and disabilities (SIIDs), introduced by Sears et al. [82], describe contexts in which environmental, application-specific, or human factors impair user interaction. Sarsenbayeva et al. identified six key contributors to such barriers in

mobile use: ambient temperature [19, 75], ambient light [92], environmental noise [4], mobility status [16, 43, 55, 78], burden [58, 59], and pressure [43]. These factors disrupt interaction by degrading motor control and environmental awareness (e.g., movement, burden, temperature) or by hindering information acquisition (e.g., light, noise, attention). From a perceptual perspective, such situational factors reduce effective visual span and increase visual crowding, constraining reading rate and recognition under suboptimal viewing conditions [69, 73]. Among these, situational visual impairments (SVIs) are particularly impactful, directly reducing legibility and increasing visual strain on mobile displays [91, 92]. SVIs commonly arise from low lighting or motion, which impair reading comprehension and visual performance [51, 56, 95]. Ambient lighting [14, 90] and environmental noise [53, 77] further degrade text recognition and cognitive function.

To mitigate SVIs, prior work has explored compensatory strategies, particularly for “reading on the move,” where users must divide attention between navigation and reading. Auditory feedback has been proposed as a workaround [8, 36, 95, 112], but its linear structure and reliance on the auditory channel limit flexibility and usability [36]. Other systems emphasize safety via visual augmentation; for example, CrashAlert [28] uses depth cameras to detect obstacles. In more static contexts, design guidelines address SVIs under ambient light: Evans [13] proposed a brightness perception model for UI design, and Tigwell [90] identified common SVI scenarios and coping strategies through a model linking environment, device, and user interaction. Complementary work on alternative display modes (e.g., light and dark modes) examines how global UI color schemes are adopted to improve comfort and accessibility across luminance conditions [1]. Despite these contributions, most approaches rely on static settings, predefined rules, or post-hoc compensation, offering limited real-time adaptation to dynamic environments. They also tend to overlook individual differences and user preferences in text perception, constraining personalization and broader applicability.

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 [30, 31, 117]. Bold fonts enhance recognition but may cause discomfort if too thick [7]. Proper line and character spacing improve layout clarity and character differentiation, making text easier to read [49, 63, 103, 117]. These typographic effects can be interpreted through visual-span and crowding accounts of reading, which show that letter size and spacing jointly constrain maximum reading rate under limited visual resolution and increased clutter [11, 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 [106]. Similarly, text orientation and reading direction influence legibility,

but their effects are often less pronounced than expected [21, 109]. 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 [23, 54, 93]. 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, and they are often conducted in controlled laboratories rather than realistic SVI contexts. As a result, it remains unclear how established typographic guidelines should be adjusted when readers must cope with motion, glare, or divided attention on mobile devices.

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

Just-in-time adaptive interventions (JITAs) provide personalized support by dynamically adapting to a user’s internal state and external context [57], making them particularly relevant for mitigating SVIs where rapid environmental changes disrupt readability. While JITAs have been widely applied in health domains [18, 25, 50, 67, 72, 89], their use for addressing SVIs remains largely unexplored. JITAs can be categorized as rule-based or AI-based: rule-based systems like FOCUS [10] rely on predefined triggers and user input [22], while AI-driven JITAs personalize interventions using behavioral and contextual data [64, 74]. For example, Matthew et al. used a random forest model to predict nudge receptiveness from individualized context [74]. Recent systems such as Time2Stop [64] incorporate human-in-the-loop feedback, refining strategies over time while maintaining transparency through AI-generated explanations.

Outside JITAs, sensor-based adaptive mobile user interfaces have long explored how to personalize layout and interaction based on sensed context [32, 79], yet many of these systems have seen limited deployment, in part due to concerns around intrusiveness, transparency, and cost. Advances in situational awareness technologies further extend the potential of JITAs for SVIs. Mobile sensors can capture ambient light and motion data [7, 65, 118], while image recognition techniques extract semantic context from visual surroundings [15, 83, 88]. The integration of Large Language Models (LLMs), which excel at text understanding and generation [96], enables context-aware interventions such as automatic font adjustments and personalized reading recommendations [45, 62, 116]. At the same time, mobile sensing for adaptive interfaces raises well-documented privacy and acceptability concerns, prompting proposals for user-centered privacy controls and privacy-preserving processing of sensor data [5, 27]. Despite these advances, most JITA systems focus on behavioral outcomes and overlook perceptual and cognitive challenges like SVIs. Furthermore, current AI-driven solutions often lack effective mechanisms for balancing automation and user control, raising concerns about intervention timing, transparency, and personalization.

3 Formative study

To design a system that mitigates SVIs in mobile reading, we operationalize the “Understanding – Sensing – Modeling – Adapting”

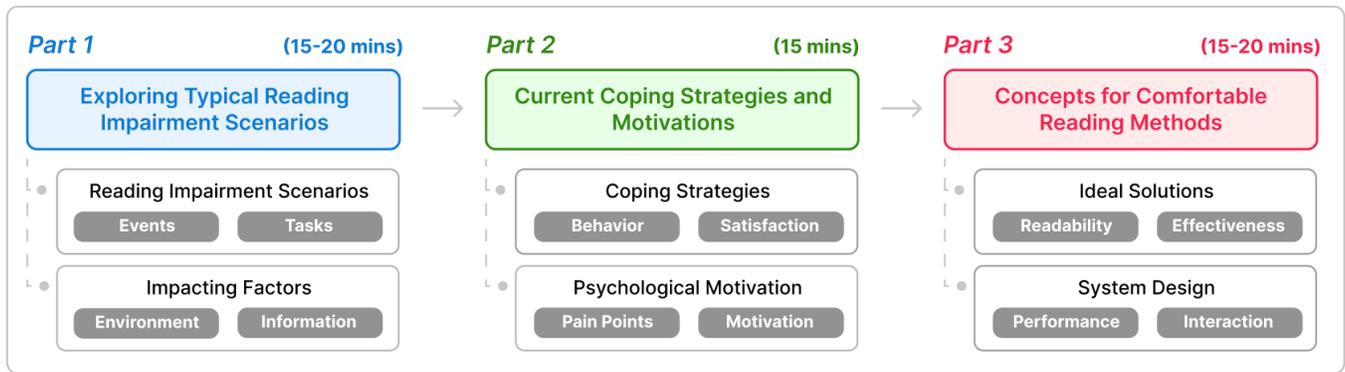


Figure 2: Workflow of the Semi-Structured Interviews.

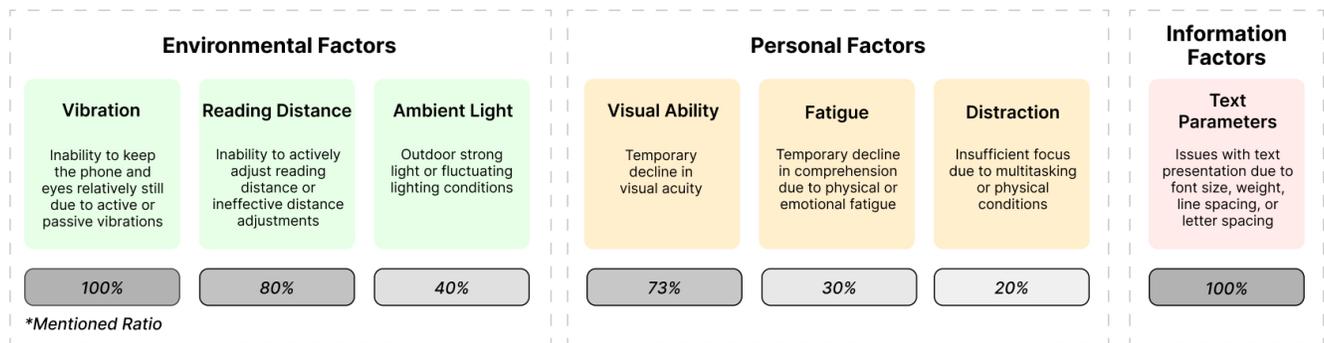


Figure 3: Interview Findings: Factors Affecting Situational Visual Impairments. The numbers inside the rectangles indicate the proportion of respondents who mentioned each factor.

cycle for situational visual impairments and related SIIDs in smartphone reading [92]. In this framework, the *Understanding* module identifies contextual factors that affect readability, which are then detected by the *Sensing* module. The *Modeling* module determines how these factors influence reading, enabling the *Adapting* module to dynamically adjust text presentation. In the following sections, we first contextualize key SVI-inducing factors and users' existing coping strategies for Chinese mobile reading through Study 1 (Section 3.1), then examine how to model these factors and the associated font adjustments for real-time adaptation in Study 2 (Section 3.2), and finally synthesize the findings into design implications that inform the sensing, modeling, and interaction mechanisms of SituFont.

3.1 Study 1: Contexts, Factors, and Coping Strategies

To identify environmental and personal factors that induce SVIs during mobile reading, as well as user strategies for mitigation, we conducted 15 semi-structured interviews to explore participants' challenges, coping mechanisms, and expectations for more supportive reading interfaces in everyday situations.

3.1.1 Research Methods. We interviewed 15 participants aged 19–52, including students and professionals with diverse visual conditions such as myopia, astigmatism, presbyopia, and normal vision (Appendix A, Table 9). Participants were recruited via university mailing lists and public social media posts. Interviews were conducted remotely using Tencent Meeting¹ and lasted 30–40 minutes. Each session began with an overview of SVIs, followed by discussions of participants' mobile reading experiences, coping strategies, and suggestions for improving readability (Figure 2; Appendix A). All interviews were audio-recorded, transcribed, and analyzed in MAXQDA [97] using inductive thematic analysis [6]. Two researchers independently conducted open coding, iteratively refined a shared codebook, and clustered conceptually related codes to generate final themes. Disagreements were resolved through discussion until consensus was reached.

3.1.2 Findings. Environmental Factors. Reading while in motion posed a major challenge, especially during walking (11/15) and running (9/15). Movement-induced vibration blurred text, impairing recognition (P3, P6, P8–9, P13, P15), and sometimes caused dizziness (P8, P10, P13) or slower reading speeds (P6, P12, P14–15). Some participants (P4) found it difficult to concentrate due to the need to

¹<https://meeting.tencent.com/>

monitor their surroundings, while others (P1, P5, P11) avoided reading while running altogether. In addition, all participants (15/15) reported difficulty reading in vehicles, especially during commutes (P6, P7, P10). Movement in private cars disrupted focus and sometimes led to motion sickness (P2, P7, P10, P12). Bright outdoor light was another common issue. Six participants (P1–2, P10–11, P13–14) noted that screen brightness adjustments were often insufficient under direct sunlight, necessitating increased font size and weight to maintain legibility. Reading at suboptimal viewing distances was also problematic, particularly when users could not freely adjust their phone position. This occurred in scenarios such as driving, where phones were mounted at a fixed distance (P5), or crowded environments like subways (P13), where users lacked space to reposition their device. These four primary SVI factors, which include intense brightness, high vibration, distraction, and fatigue, have also been widely reported in prior work on SVIs and mobile readability. Rather than proposing new SVI categories, our contribution lies in characterizing how these factors co-occur in everyday mobile reading, and in operationalizing their most realistic combinations into design-relevant scenarios for adaptive typography and JITAI-based interventions.

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 lengthy texts. Concentration difficulties were another barrier, as participants (P3, P10, P13) mentioned distractions and multitasking negatively affected their reading efficiency. These personal factors later informed the personalized branch of our context label hierarchy and the design of user-configurable preferences in SituFont.

Information Factors. Participants generally agree that text presentation played a crucial role 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 ($n = 2$). These observations motivated our focus on font parameters as the primary adaptation channel, while treating content difficulty as a separate factor that is more appropriately addressed through content selection or summarization rather than low-level typography alone.

Strategies and Challenges. Adjusting reading distance was the most commonly reported strategy, though it often caused physical discomfort. For instance, P6 described repeatedly moving their phones closer or farther, which led to neck strain during prolonged use. To reduce visual load, several participants (P8, P10, P13, P15) used auditory substitutes like voice assistants or text-to-speech tools, particularly when visual input was compromised. However, these were frequently impractical. P13 noted inefficiency while driving, and others (P9, P15) cited issues in noisy or private environments. All participants (15/15) increased font size as a workaround, but many, especially younger users (8/15), found that it disrupted formatting and required excessive scrolling. There was also widespread reluctance to manually adjust settings due to the time and effort involved; participants expressed a preference for systems

that could automatically adapt to changing conditions without user input. These findings highlight the demand for seamless, context-aware font adjustment that complements, rather than replaces, users' existing strategies, and they directly motivate our human-in-the-loop approach in later system design.

3.2 Study 2: Modeling Key Factors for Adaptation

Building on qualitative insights from Study 1, we identified three major categories of factors influencing SVIs: environmental conditions (e.g., motion, lighting), personal characteristics (e.g., visual acuity, comprehension, concentration), and information-related aspects (e.g., text size, spacing, content complexity, length) (Figure 3). These findings informed the design of Study 2, which aimed to quantify how environmental conditions affect preferred font parameters and to provide aggregate data that can be used as an initial model for JITAI-based systems before sufficient individual data become available. We conducted a controlled experiment examining how light intensity, motion states, and vibration levels impact adjustments to key font parameters, including font size, weight, line spacing, and character spacing.

3.2.1 Research Methods. Participants. We recruited 18 university students aged 18 to 25 with normal or corrected vision through public social media posts; none had participated in the prior interviews. Consistent with the scope of this work, participants reported daily experience reading Chinese text on smartphones. Each participant experienced six scenarios combining two light intensity levels (low, high) and three movement states (standing, walking, running) (Table 1). The indoor standing condition served as a low-SVI baseline, and the remaining five conditions combined increased motion and/or outdoor light to approximate common real-world SVI situations (e.g., commuting, walking outdoors).

Procedure. Participants used a custom-built mobile application preloaded with six isomorphic Chinese reading passages (550 characters each, high school level) and displayed in plain text. They read the passages in both indoor and outdoor environments while standing, walking, or running. Participants were instructed to stand still, walk at a natural pace, and run at a light jogging pace. Before the experiment, we instructed participants to prioritize legibility and comfort rather than speed, and to adjust font parameters until they felt the text was comfortably readable in each scenario. During reading, participants adjusted font parameters (font size, weight, line spacing, and character spacing) via a one-handed interface. Adjustments could be made freely at any time by double-tapping the screen. During the experiment, indoor lighting was strictly controlled to be below 200 lux (with an average of approximately 100 lux), while outdoor lighting was strictly maintained above 10,000 lux (with an average of approximately 30,000 lux).

Apparatus. To ensure consistency, we used two HUAWEI P40 smartphones (6.1-inch screen, 2340×1080 resolution) across all sessions. Built-in sensors recorded ambient light, acceleration (vibration), and reading distance. Display settings were standardized: black text on a white background, with brightness fixed at 100% for outdoor (corresponding to 531 nits) and 50% for indoor conditions. These settings were uniformly applied to both devices. Environmental data (light intensity, reading distance, vibration offset) were

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 1: Experimental Scenarios: Motion, Light, and Vibration Conditions

recorded alongside participants' font adjustments. Line and character spacing were logged in *em* units (e.g., $0.05em = 5\%$ of font size). Data cleaning involved removing incomplete or clearly erroneous entries, such as missing sensor data or implausible reading distances caused by blur or sensor errors. Thresholds for exclusion were empirically defined based on pilot testing. After cleaning, 497 valid datasets remained, averaging 28 per participant and 83 per scenario. Study 2 did not evaluate SituFont directly; instead, its role was to characterize how participants adjust font parameters under controlled manipulations of light and motion, and to provide design priors for later modeling and adaptation. Additional technical and procedural details are provided in Appendix B.

3.2.2 Findings. Significant differences were observed in light intensity and motion states across scenarios while reading distance remained stable (Table 2). 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.

Light Intensity: Low vs. High. Our analysis revealed that light intensity significantly influenced font weight. In low-light indoor conditions, participants maintained relatively consistent font weights ($M = 0.59$ indoor standing, $M = 0.79$ indoor walking). In contrast, in high-light conditions (outdoor), participants consistently increased font weight to enhance text contrast, with the strongest adjustments seen in outdoor running ($M = 1.30, SD = 0.87$). Font weight showed a strong positive correlation with light intensity ($r = 0.373, p < 0.01$). Font size showed relatively minor adjustments under varying light conditions ($r = 0.117, p < 0.01$), indicating that participants primarily relied on weight adjustments rather than size to maintain readability.

Motion: Low vs. High. Motion had a pronounced impact, particularly on font size. In low-motion scenarios, participants used smaller font sizes ($M = 20.60, SD = 3.99$ indoor). In high-motion scenarios, participants significantly increased font size ($M = 23.84, SD = 5.31$ indoor, $M = 23.64, SD = 4.92$ outdoor). Font size showed a strong positive correlation with motion ($r = 0.368, p < 0.01$), reflecting the need for larger text to maintain clarity during movement. Line spacing also exhibited moderate adjustments in high-motion scenarios, showing significant positive correlations with vibration

($r = 0.170, p < 0.01$). However, character spacing remained largely unaffected.

Vibration: Low vs. High. Vibration particularly impacts font size and font weight. In low-vibration scenarios (standing), participants used smaller font sizes ($M = 20.60, SD = 3.99$ indoor). However, in high-vibration scenarios (running), font size ($M = 23.84, SD = 5.31$ indoor, $M = 23.64, SD = 4.92$ outdoor) and font weight ($M = 1.04, SD = 0.95$ indoor, $M = 1.30, SD = 0.87$ outdoor) increased significantly. Vibration showed strong correlations with font size ($r = 0.360, p < 0.01$) and font weight ($r = 0.239, p < 0.01$).

High-Light + Motion. In scenarios combining high light and motion, participants made the most significant adjustments in both font size and weight. For instance, in outdoor running scenarios, font size reached $M = 23.64, SD = 4.92$, and font weight increased to $M = 1.30, SD = 0.87$ (Table 4). These patterns suggest that high luminance and motion jointly motivate participants to increase both the size and contrast of text to preserve legibility.

High-Light + Vibration. Under conditions combining high light and vibration, font weight was the most adjusted parameter ($r = 0.373, p < 0.01$), emphasizing contrast as the critical factor.

Individual Differences. Text adjustments varied significantly across participants, reflecting individual preferences (Table 5). Font size variability was highest in indoor running scenarios ($SD = 4.467$). Font weight variability peaked in outdoor walking ($SD = 0.827$) and outdoor running ($SD = 0.791$). Line spacing showed moderate variability, while character spacing exhibited minimal changes. These variations motivated us to treat the relationships between environmental conditions and font parameters derived from Study 2 as population-level priors that can be further personalized in later adaptive models, rather than as prescriptions for a single optimal configuration.

3.3 Design Implications

Findings from Study 1 (interviews) and Study 2 (controlled experiment) converge on several design implications for readability support under SVIs. These implications informed the sensing, modeling, and interaction mechanisms of SituFont.

Contextual Awareness and Automated Adaptation. Both studies highlight the importance of context-aware font adjustments grounded in sensed environmental conditions. In Study 1, participants expressed a preference for systems that adjust text presentation in response to changing conditions, reducing the need for repeated manual configuration. Study 2 quantified how motion and lighting are associated with systematic changes in preferred font size and weight, particularly in high-motion (e.g.,

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
<i>F</i> Value		261.146	121.825	258.395	125.669	1.496
<i>p</i> Value		0.000**	0.000**	0.000**	0.000**	0.189

* $p < 0.05$, ** $p < 0.01$.**Table 2: Variance Analysis Table of Sensor Data Features in 6 Experimental Scenarios**

	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**

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
<i>F</i> Value		15.108	7.275	4.666	1.301
<i>p</i> 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	Size (sp)	Weight (px)	Line (em)	Char (em)
1	Indoor Standing	2.70	0.65	0.14	0.09
2	Indoor Walking	2.94	0.66	0.16	0.08
3	Indoor Running	4.47	0.69	0.16	0.09
4	Outdoor Standing	2.08	0.61	0.13	0.06
5	Outdoor Walking	2.40	0.83	0.13	0.08
6	Outdoor Running	3.80	0.79	0.12	0.13

Table 5: Standard deviation of preferred typographic parameters across six experimental scenarios (N=18).

running) and high-light (e.g., outdoor) scenarios. Together, these findings suggest that real-time sensing via smartphone sensors (e.g.,

accelerometers, ambient light) can provide a basis for automated

adjustments that approximate users' preferred configurations under different SVIs.

Personalization and User Agency. Study 1 emphasized the need for customizable settings, with participants noting that "everyone's eyes are different." Study 2 confirmed this through large individual variations in font size and weight preferences, especially under demanding conditions. These observations indicate that population-level models derived from aggregate data are insufficient on their own. Systems should therefore offer multiple adaptation levels, such as adjustable sensitivity, optional prompts before changes, and the ability to override or refine suggestions, so that automated adaptations can be progressively personalized. A hybrid model that balances automation and user agency is likely to better accommodate diverse visual needs and preferences.

Addressing Limitations of Existing Coping Strategies. Study 1 participants reported using manual adjustments, larger fonts, and text-to-speech, but found them insufficient. Reading distance adjustments caused discomfort, large fonts disrupted layout, and auditory tools were impractical in noisy or private settings. Study 2 confirmed that no single adjustment resolved all challenges, highlighting the need for multifactorial adaptation strategies. These findings point toward combining adjustments to multiple font parameters (e.g., size, weight, spacing) in response to specific SVI patterns, rather than relying solely on single-parameter changes such as global zoom or font size increase.

Simplified Interaction and Reduced Cognitive Load. Participants in Study 1 found manual adjustments burdensome and often avoided them. To ease interaction, systems should minimize the need for deep navigation into settings menus and instead offer lightweight, context-aware interaction mechanisms. For example, gesture-based shortcuts (e.g., a long-press that triggers an adaptive suggestion) can provide quick access to adjustments, while adaptive interfaces can learn from users' confirmations and corrections over time, thereby reducing repeated manual configuration.

Targeted Adaptations for High-Variability Scenarios. Both studies point to the importance of tailoring interventions to high-variability scenarios, especially reading in motion, adapting to light changes, and accounting for user differences. Study 1 identified these as most disruptive; Study 2 confirmed that running and outdoor settings showed the greatest need for font adjustments. These patterns motivated us to treat such scenarios (e.g., outdoor running, outdoor walking) as distinct context nodes when structuring the label hierarchy and to prioritize them as primary targets for adaptive interventions in SituFont.

4 SituFont System Design and Implementation

As summarised in Figure 4, we used these formative insights to design SituFont, a JITAI-inspired system that enhances readability by dynamically adjusting font parameters based on real-time contextual changes and user feedback. 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.* SituFont initializes font parameter recommendations using an ML model trained on aggregated group-level adjustment data collected in Formative Study 2 (Section 3.2). This model captures common adaptation patterns observed across users as they adjusted font size, weight, letter spacing, and line spacing under varying environmental conditions such as reading distance, light intensity, and phone acceleration. The resulting predictions reflect population-level trends in how typography is adapted under SVIs, providing a reasonable cold-start behavior when users first interact with the system.

Developing this initial model serves two purposes. First, it enables the system to provide immediate, context-sensitive font recommendations without requiring prior personalization. Second, it establishes a foundation that can be incrementally refined with user-specific feedback over time, allowing the model to adapt to individual reading habits with minimal additional data. In this setup, environmental factors are used as input features, while adjusted typographic parameters serve as output variables for supervised learning.

(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 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 the 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 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. However, personalized needs require manual input

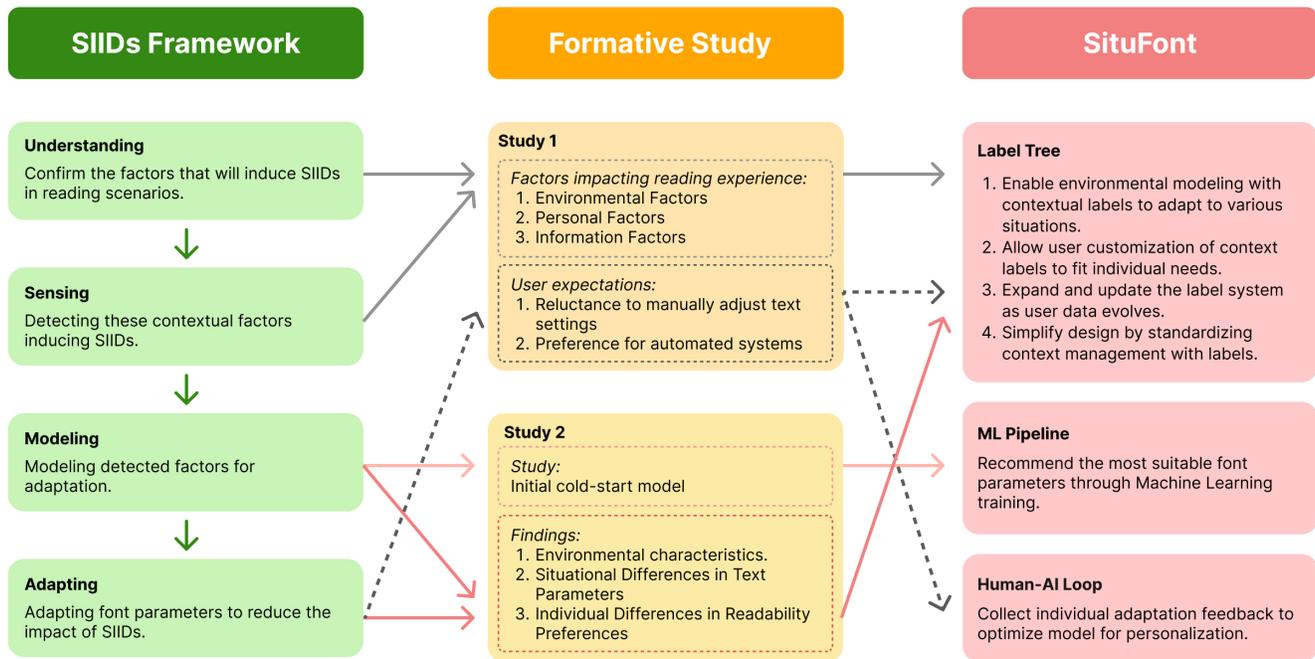


Figure 4: Formative study’s findings inspire the system design of SituFont, which mainly include the label tree, machine learning training, human-ai loop modules.

Data Name	Unit	Component and Method	Frequency
Ambient Light	<i>lux</i>	Light Sensor; Direct Call	10 times/sec
Reading Distance	<i>cm</i>	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	<i>sp</i>	Android function for size after each adjustment	Recorded on upload
Font Weight	<i>px</i>	Android function for weight after each adjustment	Recorded on upload
Line Spacing	<i>em</i>	Android function for spacing after each adjustment	Recorded on upload
Letter Spacing	<i>em</i>	Android function for spacing after each adjustment	Recorded on upload

Table 6: Environmental Sensor Data collected in SituFont

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

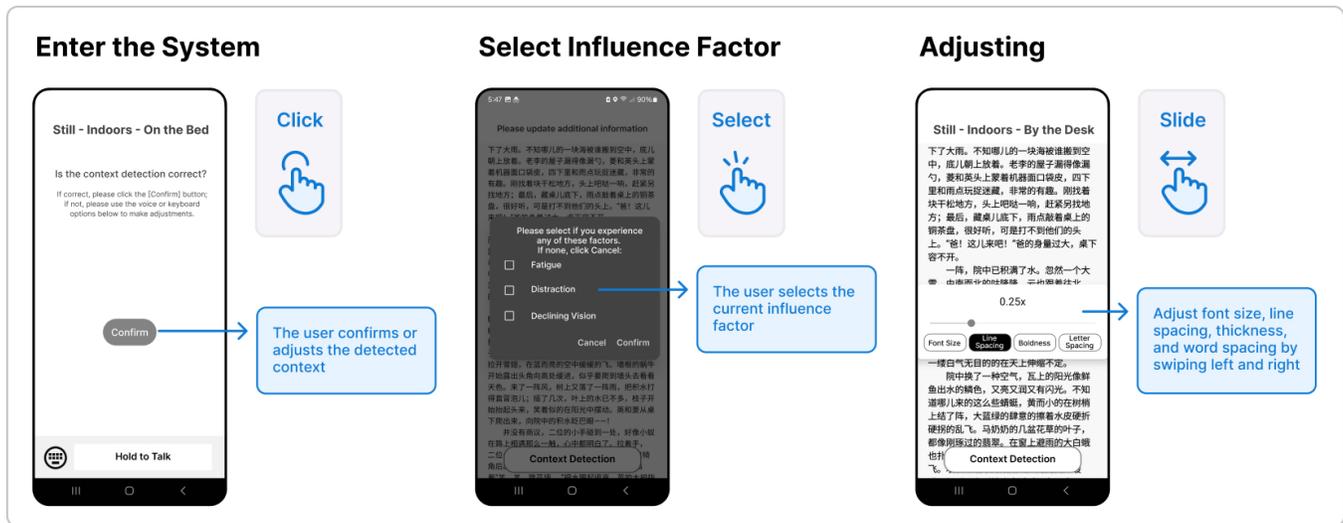


Figure 5: The user interface of SituFont involves three key interactions: (1) Entering the System – the user confirms or adjusts the detected reading context; (2) Selecting Influence Factors – the user specifies factors affecting readability, such as fatigue or distraction; and (3) Adjusting Text Parameters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

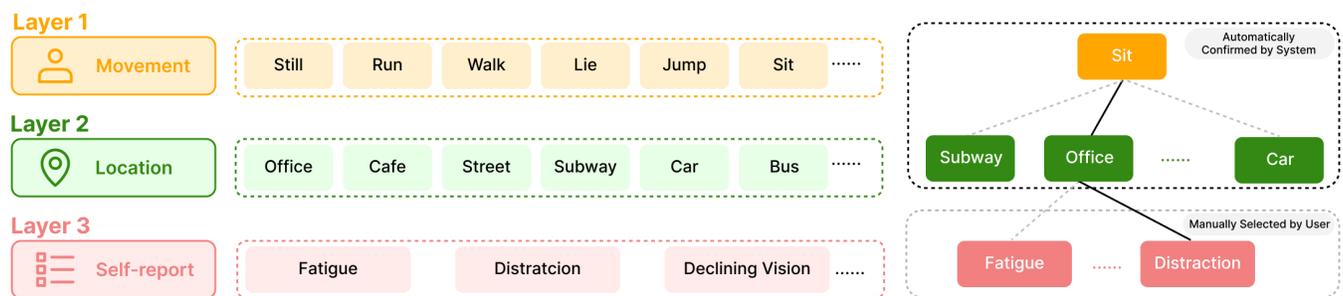


Figure 6: The left part of the figure describes a three-layer labeling system used to mark situations, where the priority decreases from top to bottom when constructing the label tree. The right part of the figure presents an example of a label tree, where the solid lines indicate the process of detecting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

(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-4o²). 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-3.5Turbo³),

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 [47], 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 [34].

²<https://openai.com/index/hello-gpt-4o/>

³<https://openai.com/index/gpt-3-5-turbo-fine-tuning-and-api-updates/>

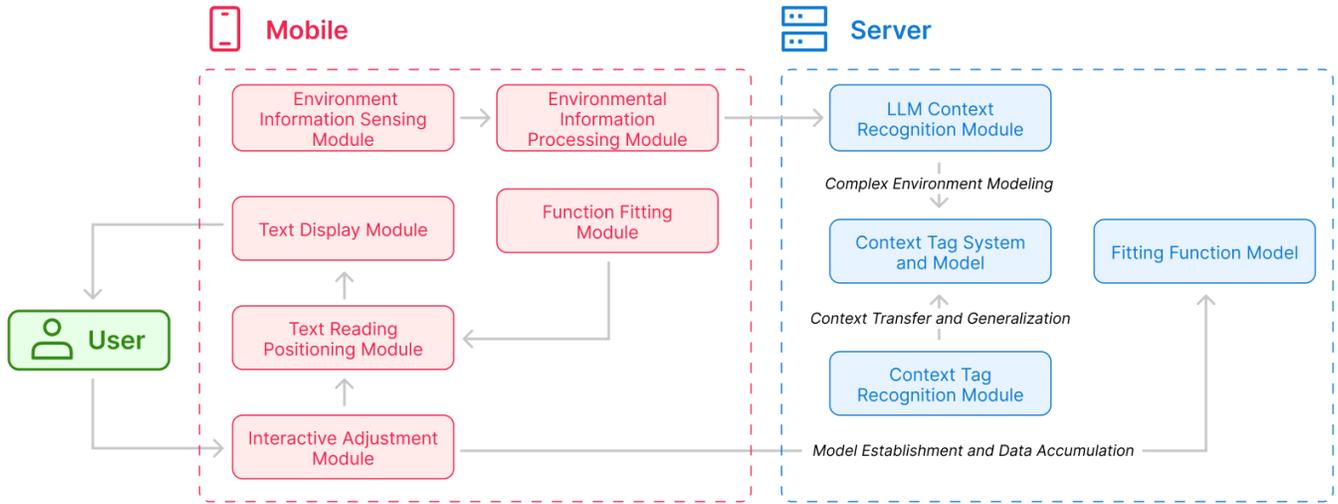


Figure 7: System architecture of SituFont, showing the flow of information from sensors to the adaptation module.

4.2.2 *Adaptive User Interface.* As described in Section 4.1.1 and illustrated in Figure 5, the controls to adjust the text parameters are developed in native Android, allowing dynamic adjustment of the font parameters based on environmental context.

SituFont’s adaptive user interface is designed to balance automation with user control. The system begins with a stable and familiar typographic presentation [105], which serves as an initial reference point rather than an assumed optimal configuration. By default, the interface displays black text on a white background, with a font size of **20sp**, a font weight of **400** (regular), a line spacing of **1.1 em**, and a letter spacing of **0.03 em**. These values were selected based on prior work on mobile legibility and established guidelines for Chinese text layout [30, 31, 103]. These parameters do not change automatically. Instead, typographic adaptation is initiated by the user through a long-press gesture, which triggers the Font Adjustment Model to generate context-sensitive recommendations. A brief vibration confirms the activation, after which the suggested font parameters are applied. This interaction design ensures that adaptation occurs at moments chosen by the user, reducing unintended disruptions while allowing the system to respond effectively to changing reading conditions.

During operation, environmental parameters such as ambient light, reading distance, and device acceleration are continuously sampled at a rate of 5Hz (every 200 milliseconds), but the font remains unchanged unless a long-press gesture is detected. Once triggered, the model infers a new set of font parameters, including size, weight, line spacing, and letter spacing, that are then applied to the TextView in real time using Android’s native UI methods. These changes are applied instantly, ensuring a smooth and responsive user experience.

This user-initiated, context-aware adjustment strategy offers a balance between automation and user control. It prevents intrusive or unnecessary UI changes while allowing the system to optimize readability. Such a design also enhances reproducibility: the adaptation logic is simple to implement with access to environmental

sensors and can be replicated using either rule-based logic or ML-based inference, as detailed in Sections 4.2.3 and 4.2.4.

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 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 C.

5 User study

To evaluate SituFont’s impact on reading performance and user experience compared to a traditional reading interface that requires 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 reading interface 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 and Apparatus

Participants were recruited via public social media posts. 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 Chinese speakers. Participants reported using a variety of devices for reading, including smartphones ($n = 12$), tablets ($n = 6$), and laptops ($n = 7$). Daily sustained reading time on smartphones, meaning the reading of longer and more cognitively demanding passages that require comprehension and focused engagement, ranged from 10 minutes to over 2 hours. Ten participants reported wearing corrective lenses. One 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.

All pre-test and experimental sessions used a HUAWEI P40 (6.1-inch screen, 2340×1080 resolution) standard Android phone to ensure consistency. During the adaptation period, participants used either their own Android devices or the provided Android phone.

5.2 Design

The overall study design is shown in Fig. 8 and consists of three phases: a pre-test, a four-day adaptation period, and a comparative experiment, allowing us to examine SituFont’s advantages over a traditional reading interface after substantial adaptation. The pre-test serves as a baseline for task completion under the traditional interface, enabling us to assess whether participants exhibit learning effects on the reading tasks themselves by comparing traditional-interface performance between the pre-test and post-test. The primary comparison of interest, however, is between SituFont and the traditional interface in the final phase. All controlled experiments, including the pre-test and both settings of the comparative experiment, were conducted under the eight SVI conditions. In each condition, participants completed two tasks: a read-aloud task and a reading comprehension task. Below, we describe the experimental conditions, corpus, tasks, and baseline configuration.

5.2.1 Conditions. Guided by prior literature on SVIs [56, 91, 92] and insights from our formative study, we selected eight experimental conditions (Table 7, Figure 9) to balance ecological validity with feasibility. We prioritized four primary SVI factors: intense brightness, high vibration, distraction, and fatigue, all frequently cited as impactful in mobile reading contexts. Intense brightness and high vibration were especially disruptive, often occurring outdoors (e.g., reading in sunlight or while moving), while distraction and fatigue were more relevant to indoor settings or prolonged use. Testing all 16 possible combinations was impractical due to participant burden, so we selected a representative subset that reflected the most common and realistic scenarios. For example, *Intense Brightness + High Vibration + Distraction* simulates running in crowded outdoor spaces, and *Intense Brightness + High Vibration + Fatigue* mirrors the

compounded strain of outdoor exercise. Scenarios lacking brightness or vibration were excluded, as the formative study indicated they were less frequent and less disruptive in real-world use.

5.2.2 Tasks and Corpus. All experimental sessions involved two tasks: a read-aloud task and a reading comprehension task. The read-aloud task required participants to read the entire passage aloud at a natural and comfortable pace. The reading comprehension task required participants to complete five multiple-choice questions designed to assess their understanding of the passage content. Participants completed these comprehension questions at a natural, self-paced rhythm without time constraints.

Each condition used two 50-character passages sourced from the HSK Level 5 examination. Although originally designed for advanced non-native learners, HSK Level 5 provides standardized Chinese passages of medium to high difficulty that remain appropriate for native speakers [70, 111]. The HSK corpus also includes predefined multiple-choice comprehension questions, ensuring consistent difficulty and minimizing memorization bias [42].

5.2.3 Baseline. To evaluate the effectiveness of SituFont, we selected a competitive and realistic baseline: a manually adjustable reading interface that matches SituFont in both interface elements and parameter degrees of freedom. This baseline operates as follows. Before each reading task, under a specific SVI condition, participants were allowed unlimited time to freely adjust all font parameters under the template corpus until they reached what they perceived as the optimal static configuration. **Adjustment time was not counted as part of the task, and no further adjustments were allowed once reading began.**

We argue that such a baseline setting is fair and non-trivial. First, using a fixed-layout baseline would be too weak, as it would substantially underestimate the upper bound of manual optimization and may introduce avoidable readability issues and unfairness for different SVI scenarios, thereby exaggerating the advantages of SituFont. Second, allowing participants to tune the interface before each task yields a near-optimal static layout for the given scenario. This enables two meaningful comparisons: whether SituFont’s predicted parameters approach this near-optimal configuration, and whether SituFont’s ability to adapt dynamically to changing conditions provides additional benefits beyond what static optimization can achieve. Third, this baseline reflects a fair and realistic usage pattern. In everyday reading, users typically adjust settings before reading rather than during the act of reading itself. By excluding adjustment time from task completion and keeping the reading process identical across conditions, we ensure that SituFont and the baseline differ only in their adaptation mechanisms rather than other operational interferences.

5.3 Procedures

5.3.1 Pre-test. Before the adaptation period, participants completed a baseline reading pre-test using a traditional mobile reading interface with manual font control, across eight SVI conditions (Table 7). Comprehension was assessed using five HSK-defined multiple-choice questions per passage, with accuracy calculated as the percentage of correct responses. Reading performance was measured via goodput (characters per minute, CPM), defined as

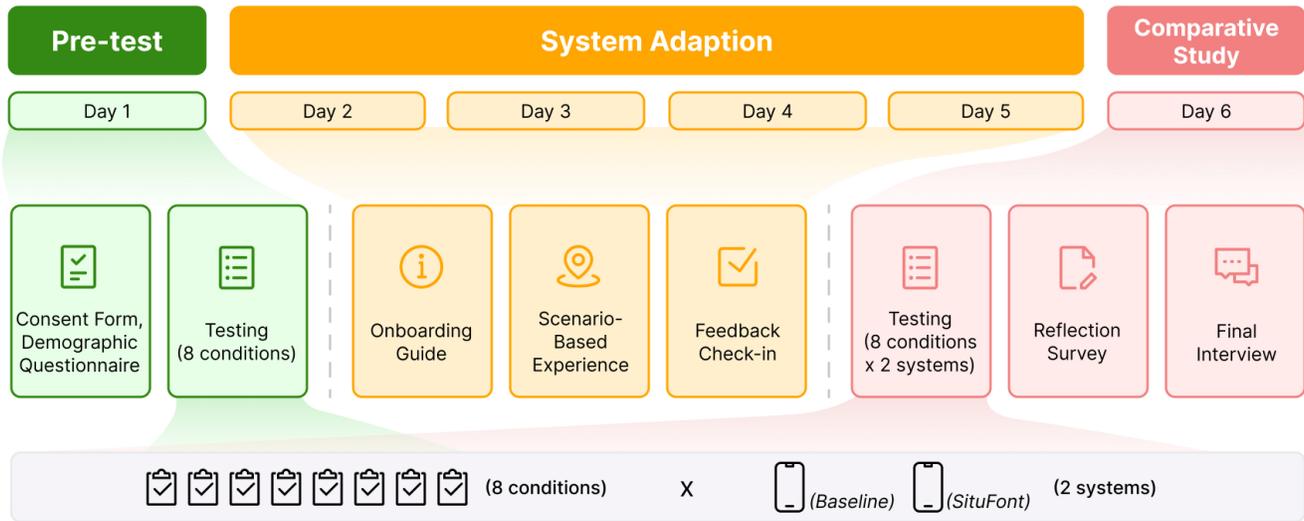


Figure 8: User study timeline.

the number of correctly read characters divided by total reading time [39]. Participants read aloud at their normal pace, with audio recorded for later analysis. A pre-test questionnaire also gathered demographic information, vision history, and reading habits (Appendix D.1).

5.3.2 Adaptation Period. Over four days, participants used SituFont in three to four different reading scenarios daily. A daily survey captured reading contexts, perceived fatigue, distraction levels, and usability issues (Appendix D.2). This phase allowed participants to label their personalized preferences, where the prediction model learns personalization adaptation strategies, and become familiar with the system before the main evaluation.

5.3.3 Post-Adaptation Reading Evaluation. Following the adaptation period, participants completed the same reading tasks under the eight SVI conditions, this time using both SituFont and a traditional display. Each condition included two new HSK Level 5 passages. Interface order was counterbalanced across participants and conditions. After each passage, participants answered five standardized comprehension questions and their reading goodput was calculated. They then completed the NASA-TLX questionnaire to assess perceived workload. After all conditions, participants completed UEQ-S and SUS questionnaires and participated in a semi-structured interview (Appendix D.3).

5.4 Measures

5.4.1 Reading Performance. Reading performance was assessed through goodput (characters per minute, CPM) and reading comprehension accuracy.

Goodput was defined as the number of characters correctly read aloud by the participant, divided by the total reading time, following and adapting the approach from Ku et al. [39]. While Ku et al. measured words correctly perceived in a rapid serial visual presentation setting using AR glasses, we extended this to naturalistic

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	✓	✓	✓	✓

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

read-aloud tasks on mobile devices under SVI conditions. Participants were instructed to read each passage aloud at a comfortable pace during each SVI condition. Reading time was measured from the moment the participant began reading until they completed the passage. A character was considered “correct” if it matched the corresponding character in the source text. This was done while ignoring case, punctuation, and minor mispronunciation. Goodput was calculated as:

$$\text{Goodput} = \frac{C}{T}$$

where C is the number of correctly read characters, and T is the total reading time in minutes. This metric intends to capture both speed and accuracy.

Reading comprehension was measured by the percentage of correct responses to five multiple-choice questions per passage. These questions were pre-defined based on the HSK Level 5 Mandarin proficiency examination [4, 115].

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 ^(d)	None
Intense Brightness + High Vibration + Distraction	50,000+ lux ^(a)	Running ^(d)	Navigation Task ^(c)
Intense Brightness + High Vibration + Fatigue	50,000+ lux ^(a)	Running ^(d)	Weighted Load ^(e)
Intense Brightness + High Vibration + Fatigue + Distraction	50,000+ lux ^(a)	Running ^(d)	Navigation Task ^(c) + Weighted Load ^(e)

^(a) Outdoor conditions used strong midday natural lighting (50,000+ lux, clear sky). ^(b) Walking occurred along a straight 50 m path (indoors: office hallway; outdoors: running track), no obstacles/pedestrians. ^(c) Navigation task: soccer training cones (30 cm tall) placed 5 m apart; pass yellow cones on the left and orange cones on the right while reading. ^(d) High-vibration condition: treadmill running at 6 km/h, no incline, no handrails. ^(e) Weighted load: 3 kg backpack (laptop + two books), straps adjusted snugly for even weight distribution.

Table 7: Experimental Conditions in User Study

5.4.2 Perceived Workload & User Experience. Perceived workload was evaluated using NASA-TLX [26], measuring mental demand, physical demand, temporal demand, performance, effort, and frustration. Overall usability was assessed through UEQ-S [80] and SUS [3], capturing users' subjective impressions of each interface.

5.4.3 Qualitative Feedback. Post-study semi-structured interviews were conducted to collect in-depth feedback on participants' experiences with both SituFont and the traditional display, including difficulties encountered, perceived benefits, and interface preferences. Interview transcripts were thematically analyzed following Braun and Clarke's guidelines [6]. Two researchers independently performed open coding on the first six transcripts using MAXQDA [97], after which a refined codebook was developed through discussion. This codebook was then applied to the remaining six transcripts. Final themes were generated by clustering related codes and refining them collaboratively.

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. Before conducting parametric tests (e.g., paired t-tests, ANOVA), we assessed data normality for each metric using the Shapiro-Wilk test. All variables used in parametric analyses (e.g., goodput, comprehension accuracy, UEQ-S, and SUS) met the normality assumption ($p > 0.05$). For metrics that violated normality (e.g., NASA-TLX workload subscales), we used the Wilcoxon signed-rank test as a non-parametric alternative.

6.1 Reading Goodput

Across conditions, SituFont consistently improved users' reading efficiency, yielding higher goodput scores than the baseline system. Paired sample t-tests 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 Baseline system in all other conditions (Figure 10). We interpret these two single-factor results as marginal trends that did not reach the conventional $p < .05$ threshold and are consistent with smaller or context-dependent effects when a single SVI factor is present in isolation. Across scenarios, effect sizes ranged from moderate in the two marginal single-factor cases (*Intense Brightness*: $d = 0.57$; *High Vibration*: $d = 0.79$) to large or very large in multi-factor scenarios (e.g., *Intense Brightness + High Vibration + Distraction*: $d = 1.63$; *Fatigue*: $d = 2.92$). The largest gains were observed in scenarios involving multiple simultaneous impairments, such as *Intense Brightness + High Vibration + Distraction* ($\Delta = 79.2$) and *Intense Brightness + High Vibration + Fatigue + Distraction* ($\Delta = 46.0$), suggesting that SituFont's dynamic adjustments are particularly beneficial when contextual factors interact in complex ways. In contrast, the improvements in the *Intense Brightness* and *High Vibration* single-factor scenarios did not reach statistical significance. This pattern suggests that when the impairment is dominated by a single, strong factor, participants can often approximate a reasonable static configuration during the baseline optimization phase, narrowing the performance gap between static and adaptive adjustments.

A comparison between pre-test and post-test performance under the traditional interface further shows that users exhibited some adaptation to the task itself. Goodput increased in several conditions after repeated exposure to similar reading tasks, indicating a post-learning effect. However, this improvement did not eliminate SituFont's advantage. Even after users had opportunities to learn optimal manual adjustments and become familiar with the reading tasks, SituFont continued to outperform the manually optimized baseline in the majority of SVI conditions. This finding reinforces that SituFont's benefit stems not only from parameter selection but also from its ability to respond dynamically to fluctuating SVI factors, which a static configuration cannot capture.

Scenarios ID [†]	1	2	3	4	5	6	7	8
Baseline Pre-test	254.17 (40.72)	252.59 (42.43)	208.15 (44.15)	206.84 (40.38)	272.69 (40.40)	289.52 (54.93)	277.61 (39.95)	226.04 (45.80)
SituFont	300.64 (56.80)	288.25 (40.72)	271.51 (50.05)	304.58 (39.36)	295.78 (48.91)	352.52 (51.06)	288.34 (31.97)	292.83 (47.65)
Baseline Post-test	285.32 (39.66)	241.95 (49.01)	256.10 (70.46)	249.14 (52.39)	273.26 (50.30)	273.33 (51.05)	256.77 (50.38)	246.83 (53.02)
Situ vs. Baseline (t)	-1.961	-7.741	-2.166	-5.634	-2.742	-10.122	-2.245	-3.256
Situ vs. Baseline (p)	0.076	0.000***	0.052	0.000***	0.019*	0.000***	0.046*	0.008**
Pre vs. Post (t)	-3.36	1.29	-3.09	-4.31	0.04	1.14	1.40	-2.73
Pre vs. Post (p)	0.006**	0.22	0.01*	0.001**	0.97	0.28	0.19	0.02*

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

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 8: Mean (standard deviation) of reading goodput (CPM) across different SVI scenarios. Paired t -tests show SituFont’s goodput improvement is significant in several scenarios.

Together, these results suggest that SituFont provides robust performance advantages across diverse and dynamically changing environments, particularly in scenarios where multiple SVI factors combine to create substantial visual and cognitive challenges.

6.2 Reading Comprehension

Overall, reading comprehension accuracy remained comparable between SituFont and the Baseline system across all task scenarios (Figure 10). Although SituFont showed slightly higher mean accuracy than the Baseline system in most conditions (except Scenario 7), none of these differences reached statistical significance. This stability is consistent with our task design, where participants completed comprehension questions in a natural manner, meaning that comprehension performance was driven primarily by text difficulty rather than interface differences. The overlapping error bars further indicate that both interfaces supported a similar level of understanding, suggesting that while SituFont improves reading efficiency, it does not alter users’ ability to comprehend the passage content.

6.3 Perceived Workload

Participants reported significantly lower mental and physical workload when using SituFont compared to the Baseline system (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 the *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 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.

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 Baseline system may cause greater frustration under distracting and multi-factor impairment conditions.

6.4 User Experience and Preferences

User experience evaluations revealed that SituFont was generally preferred over the Baseline system, 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).

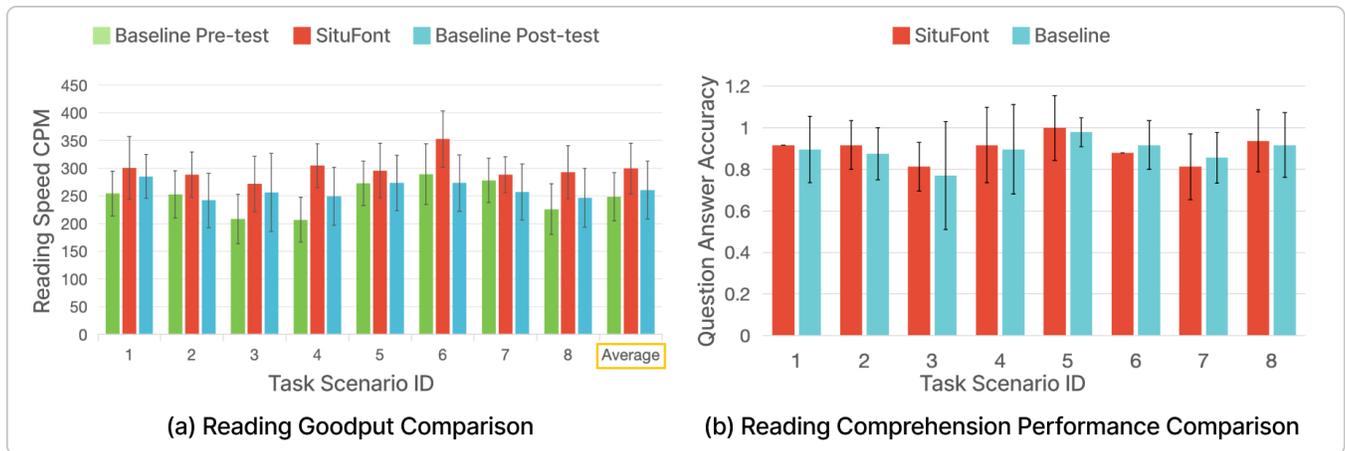


Figure 10: Comparison of reading goodput (a) and comprehension accuracy (b) across different SVI conditions. SituFont consistently demonstrated improved goodput, while comprehension accuracy remained stable across both interfaces.

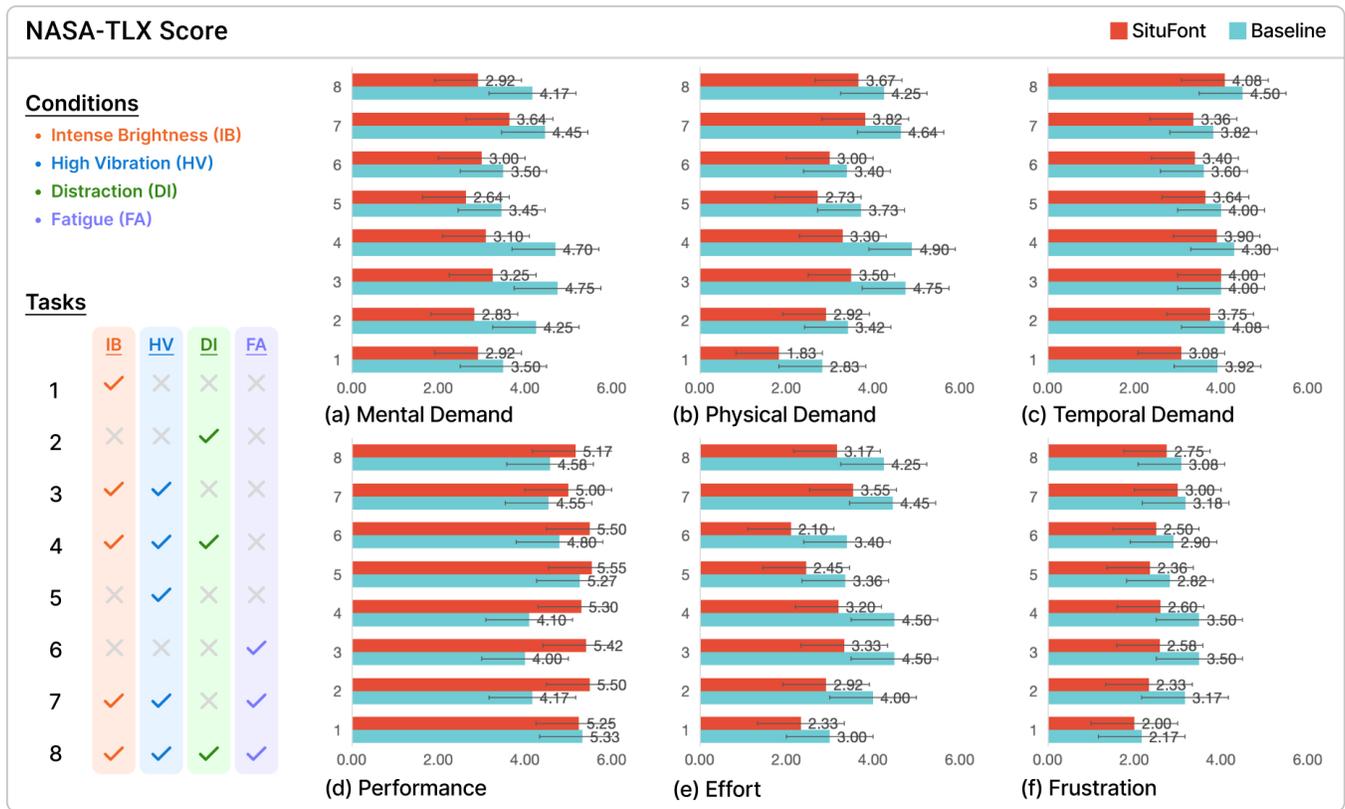


Figure 11: NASA-TLX scores across experimental conditions, illustrating that SituFont consistently reduced mental and physical workload compared to the Baseline system. No significant differences were observed in temporal demand, while frustration levels showed minor variations. For convenience of analysis,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

6.4.1 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 Baseline system ($M = 3.92, SD = 2.27$), $p = 0.0005$, suggesting a strong perception of

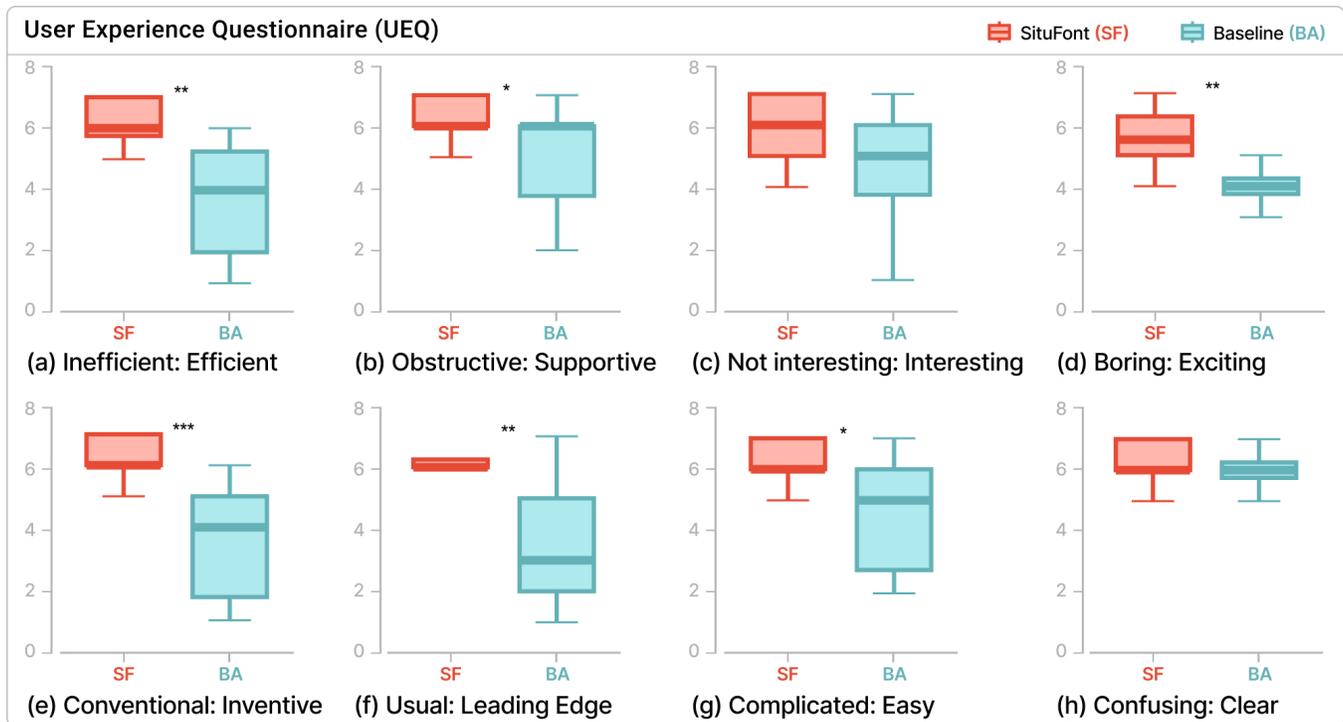


Figure 12: User Experience Questionnaire (UEQ) results, comparing SituFont and the Baseline system across multiple dimensions. SituFont scored significantly higher in efficiency, supportiveness, and novelty, with marginal differences in stimulation and perspicuity. *: $p < .001$, **: $p < .01$, *: $p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

efficiency. Similarly, on the *Obstructive: Supportive* scale, SituFont ($M = 6.17$, $SD = 0.94$) significantly outperformed the Baseline system ($M = 4.92$, $SD = 1.73$), $p = 0.039$, 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 Baseline system ($M = 4.50$, $SD = 1.83$) 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 Baseline system ($M = 4.08$, $SD = 2.23$), $p = 0.024$. Similarly, in the *Usual: Leading edge* dimension, SituFont ($M = 5.92$, $SD = 1.62$) scored significantly higher than the Baseline system ($M = 3.50$, $SD = 2.11$), $p = 0.002$, 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 Baseline system 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.

6.4.2 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 Baseline system ($M = 4.08$, $SD = 2.50$), $p = 0.045$, 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 Baseline system ($M = 4.67$, $SD = 2.43$), $p = 0.012$, 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 Baseline system ($M = 1.83$, $SD = 1.03$), $p = 0.014$, 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 Baseline system 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 Baseline system ($M = 2.08$, $SD = 0.67$), $p = 0.030$, indicating that SituFont required a slightly higher initial learning effort.

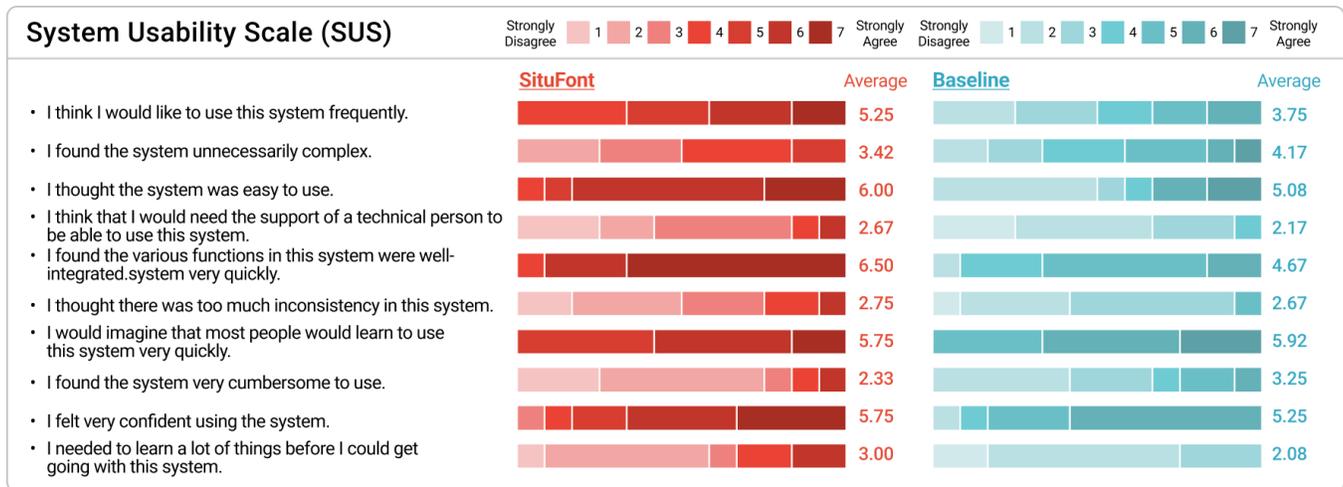


Figure 13: System Usability Scale (SUS) results. SituFont was rated significantly lower in complexity and higher in ease of use and consistency, though it required a slightly higher initial learning effort.

6.4.3 Qualitative Feedback. Participants valued the convenience of automatic font adjustment, particularly in dynamic contexts such as cycling or running. P2 noted that the system became especially useful once sufficient training data had accumulated, and P5 remarked, “When I’m running, long-pressing is much more convenient than manually selecting.” The hands-free long-press activation was widely appreciated: P3 described it as effortless, and P4 highlighted its single-handed operation and immersive reading experience. Several participants also reported increased awareness of reading environments; as P9 reflected, “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.”

Despite these benefits, participants identified several usability issues. Automatic scene detection occasionally misfired due to its reliance on “historically recorded descriptions” (P8), and long-press activation sometimes lacked precision, with P9 noting that it was “sometimes not sensitive enough” and other times triggered unintentionally. Manual adjustments via the horizontal slider were challenging in motion: P2 found it difficult to use “when the phone is very bumpy,” and P3 described the slider placement as awkward for one-handed use. Participants suggested improvements such as adding paragraph spacing (P5), reducing disruption from font changes that “interrupt the original reading rhythm” (P4), and making adaptation more proactive after calibration, particularly for users with larger font needs (P2). Finally, the post-adjustment vibration was perceived as intrusive; P5 commented that it “interferes with my reading experience.”

7 Discussion and Limitation

In this section, we discuss design implications for SVI interventions and JITAI systems with human-in-the-loop feedback, and reflect on ethical and practical considerations. We also discuss limitations and potential future research directions of SituFont.

7.1 Design Considerations for SVIs Intervention

Our user study compared SituFont with a realistic baseline where participants could freely tune all font parameters before each task until they reached a preferred static configuration, with adjustment time excluded from performance measures. Under these conditions, SituFont still led to higher goodput in most of the eight SVI scenarios, especially when intense brightness and high vibration co-occurred, while comprehension accuracy remained similar. These results suggest that adaptive typography adds value beyond one-off manual tuning by supporting perceptual and cognitive processes that are challenged under SVIs. Prior work on visual span [48], crowding [81, 108], and contrast sensitivity [52, 100] suggests that factors such as motion, glare, and vibration can reduce effective perceptual span and increase visual noise, making text harder to process efficiently. SituFont’s context-sensitive adjustments to font size, weight, and spacing help mitigate these perceptual constraints as situational conditions fluctuate, without requiring users to continuously anticipate or correct readability breakdowns. Although participants were able to learn effective manual adjustments over time, doing so required sustained attention and frequent context switching [44], particularly in multi-factor conditions such as walking under intense brightness or vibration. SituFont’s just-in-time adaptations reduced the need for repeated manual intervention, allowing participants to maintain reading flow under changing conditions.

The results also show that the success of SVI interventions depends not only on “what” parameters are chosen but also on “how” adaptation is triggered and applied. First, activation must be precise and predictable. From a cognitive perspective, mobile reading under SVIs often involves divided attention [2, 38, 44], as readers must allocate resources between text processing and environmental monitoring [102, 110]. Frequent interruptions or unintended activations can therefore impose additional attentional costs [9, 94]. Participants reported that the current long-press gesture often fired

unintentionally during scrolling or text selection, turning adaptation into an interruption. To support SVI scenarios, activation should be “disruption-safe”, input mechanisms that require deliberate intent and include basic safeguards (e.g., suppressing activation when finger velocity is high or when selection handles are active). A vertical edge slider can provide a more ergonomic, deliberate control for one-handed use [12, 35, 68], and allowing users to customize gestures [61] further aligns activation with individual reading habits. Voice commands can offer hands-free control in motion-heavy contexts [98, 99], but must be balanced against social and privacy concerns.

Second, adaptation must preserve layout stability. Sudden jumps in font size or spacing caused participants to lose their place, particularly under high motion. This observation aligns with perceptual accounts emphasizing the role of spatial continuity and stable reference frames in efficient reading [33]. Rather than updating typography immediately with every sensor change, adaptations should align with readers’ micro-rhythms, brief pauses in scrolling or task transitions, and be applied in modest, predictable steps. Simple buffering, such as aggregating small fluctuations into a single adjustment after a short pause, can maintain spatial continuity while still responding to SVIs. In practice, sensing should be treated as a signal that adaptation might be needed; whether and how to adapt should be coordinated with the ongoing reading flow.

7.2 Implications on SituFont towards the JITAI Paradigm

SituFont enables personalized font adaptation by incorporating user feedback, but several aspects require refinement to better capture individual needs. The current binary feedback mechanism (yes/no) is helpful for initial adaptation but lacks the granularity needed for truly personalized adjustments. A more detailed system, such as Likert-scale ratings for distraction, fatigue, and visual strain, or predefined reasons for discomfort, would yield richer data for more tailored interventions. Participants sometimes responded to a font change even though their discomfort began earlier or was caused by a different environmental factor, making it difficult for the system to interpret which condition the feedback referred to. This temporal mismatch suggests that JITAIs should explicitly capture “when” feedback refers to, such as by storing short interaction histories and allowing users to tag which moment or condition prompted their response.

Additionally, the cold-start problem limits early system performance due to insufficient user data during initial use, which some participants noted as a period when the system felt less helpful. This could be mitigated by pre-training on data from Study 2 (3.2) and fine-tuning with incoming user inputs. Few-shot learning techniques [20, 86] may further enhance rapid personalization with minimal data. Beyond model-level strategies, the cold-start period also requires interaction-level scaffolding. For example, early in the deployment, the system could: (a) present users with brief “initial preference snapshots” (e.g., selecting preferred base font size or sensitivity to brightness), (b) temporarily increase explanation frequency (“We enlarged spacing because ambient motion increased”),

or (c) provide optional confirmation prompts to reduce miscalibrated early moves. These measures help bootstrap a reliable personalization trajectory while avoiding user frustration during the model’s least accurate phase.

Long-term usage patterns also merit investigation. Although our study focused on short-term use, participants reported increased awareness of how font settings affect readability [66, 101]. This learning effect, along with improved reading goodput observed in the Baseline system after exposure to SituFont, suggests that adaptive typography may foster lasting improvements in digital reading strategies. This behavioral shift highlights an important distinction for JITAIs: interventions can serve both an immediate corrective function and a long-term educational function. Systems that recognize this dual role could incorporate mechanisms that gradually decrease intervention frequency as users develop stronger metacognitive control, or allow users to transition between “assistive,” “collaborative,” and “autonomous” modes depending on their evolving confidence and preferences. This pattern also resonates with perspectives on metacognitive awareness and self-regulation in reading, indicating that adaptive typography can function not only as an assistive tool for mitigating SVIs in the moment but also as a scaffold that helps readers learn to manage challenging visual contexts over time. Thus, longitudinal studies should examine whether user behavior continues to evolve and how personalization mechanisms should adapt over extended use.

7.3 Ethical Considerations

SituFont relies on continuous sensing of ambient light and motion and, in some configurations, on external APIs. Several participants questioned permissions and possible “always-on” sensing, echoing broader concerns about privacy and pervasive data collection. Ethical deployment, therefore, requires privacy by design, prioritizing on-device processing and inference, minimizing retention of raw sensor data, and clearly explaining what is collected and how it informs font adjustments. When cloud servers are used for real-time machine learning, techniques such as encryption, federated learning, and differential privacy should be considered to reduce the risk of data breaches [24, 60]. Designers must also address potential exclusion, bias, and discriminatory inferences that can arise from cameras, microphones, and large language models [40], and should recognize that opaque adaptation may undermine trust.

Meanwhile, we did not measure battery, latency, or monetary cost in our studies, but continuous sensing and model updates inevitably consume resources. These costs affect both user acceptance and who can benefit from such systems. Providing users with control over sensing frequency and adaptation intensity, as well as simple ways to pause or suspend adaptation, can help negotiate trade-offs between readability, privacy, and energy use. More transparent information about data practices and system behavior, as suggested in work on explainable and trustworthy interfaces [107, 113], may further foster informed adoption. Our findings show that human-in-the-loop adaptive typography can improve reading efficiency and reduce workload under SVIs; future deployments must ensure that these benefits are delivered in a way that remains transparent, controllable, and sustainable for everyday use.

7.4 Limitations and Future Work

Our studies involved young adult native Chinese speakers reading short HSK Level 5 passages on smartphones. As a result, we do not claim generalizability beyond this population and language setting. The HSK materials may also have been insufficiently challenging for some participants, which likely contributed to ceiling effects in comprehension. Future work should recruit older adults and readers with diagnosed visual impairments, and include more demanding and diverse reading tasks (e.g., academic or long-form content) to better probe comprehension and cognitive load under SVIs.

In addition, we evaluated eight ecologically common SVI conditions rather than all 16 possible combinations and did not include a non-SVI baseline. Scenarios such as *Distraction + Fatigue* without brightness or motion may still meaningfully affect reading. A more complete factorial design with a neutral baseline would help quantify interaction effects and express improvements relative to comfortable reading conditions. We reported only basic accuracy measures for context recognition and did not systematically compare the Label Tree hierarchy to simpler rule-based or flat-feature baselines. Nor did we evaluate feature importance or users' perceived appropriateness of model outputs. A more complete ML evaluation, including ablation studies and user-facing assessments of perceived fit, would strengthen the justification for our architectural choices and clarify when additional model complexity is warranted.

The current personalization mechanism also has limitations, such as binary comfort feedback being coarse, sometimes temporally ambiguous, and contributing to a cold-start period in the first days of use. Future versions should explore richer yet lightweight feedback structures (e.g., quick tags such as “too small” or “too dense”) and temporally anchored annotations that tie corrections to recent context windows. We also did not evaluate the optional voice-based context channel included in the design, leaving its usability and value an open question.

Finally, we did not directly measure privacy, energy, or monetary costs. SituFont relies on continuous sensing of ambient light and motion and, in some configurations, external APIs. These channels are important for adaptation but raise privacy concerns and consume battery and data. Future work should investigate privacy-preserving and resource-aware deployments, such as on-device processing, adjustable sensing frequency, user controls over adaptation intensity, and clear communication of data practices. Longer-term field studies will also be needed to examine how personalization strategies evolve over time, whether users continue to rely on adaptation once they internalize effective font settings, and how adaptive typography can be extended to other scripts and accessibility features such as contrast adjustments and screen reader support.

8 Conclusion

Rather than introducing a fundamentally new adaptive framework, this paper builds on existing accessibility and adaptive typography research to show how just-in-time, human-in-the-loop adaptation can support mobile reading under dynamic situational visual impairments through SituFont, and to articulate design considerations grounded in empirical evidence. 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. We believe that these insights can inform a broader class of adaptive systems that respond intelligently to environmental and user-centered variability.

Acknowledgments

This work is supported by the Natural Science Foundation of China under Grant No. 62502411. This work is also supported by the Natural Science Foundation of China under Grant No. 62132010, Key Research and Development Program of Ningbo City under Grant No. 2023Z062, Institute for Artificial Intelligence, Tsinghua University (THUAI).

References

- [1] Sarah Andrew, Chelsea Bishop, and Garreth W. Tigwell. 2024. Light and Dark Mode: A Comparison Between Android and iOS App UI Modes and Interviews with App Designers and Developers. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 32 (March 2024), 23 pages. doi:10.1145/3643539
- [2] Yunpeng Bai, Aleks Ikkala, Antti Oulasvirta, Shengdong Zhao, Lucia J Wang, Pengzhi Yang, and Peisen Xu. 2024. Heads-Up Multitasker: Simulating attention switching on optical head-mounted displays. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [3] 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. doi:10.1080/10447310802205776
- [4] 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. doi:10.1007/s00779-006-0063-x
- [5] Florian Bemmman. 2023. User-Centered Privacy to Improve User Quantification using Smartphone Sensing. In *Proceedings of the 25th International Conference on Mobile Human-Computer Interaction (Athens, Greece) (MobileHCI '23 Companion)*. Association for Computing Machinery, New York, NY, USA, Article 42, 4 pages. doi:10.1145/3565066.3609737
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa
- [7] 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. doi:10.1109/I2CT51068.2021.9418001
- [8] Scott Carter and Laurent Denoue. 2009. SeeReader: An (Almost) Eyes-Free Mobile Rich Document Viewer. *arXiv preprint arXiv:0909.2185* (2009).
- [9] Yue-Yuan Chen, Wei-Ning Fang, Hai-Feng Bao, and Bei-Yuan Guo. 2024. The effect of task interruption on working memory performance. *Human factors* 66, 4 (2024), 1132–1151.
- [10] 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).
- [11] Jonathan Dobres, Benjamin Wolfe, Nadine Chahine, and Bryan Reimer. 2018. The effects of visual crowding, text size, and positional uncertainty on text legibility at a glance. *Applied Ergonomics* 70 (2018), 240–246. doi:10.1016/j.apergo.2018.03.007
- [12] Augusto Esteves, Elizabeth Bouquet, Ken Pfeuffer, and Florian Alt. 2022. One-handed input for mobile devices via motion matching and orbits controls. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–24.
- [13] Jordan Evans. 2023. Measuring the effects of bright mobile device display illumination on achromatic colour perception. <https://hdl.handle.net/10214/27365>
- [14] Jordan Evans. 2023. Measuring the effects of bright mobile device display illumination on achromatic colour perception. <https://hdl.handle.net/10214/27365>

- [15] 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. doi:10.1109/iThings.2014.43
- [16] Mayank Goel, Leah Findlater, and Jacob Wobbrock. 2012. WalkType: using accelerometer data to accommodate 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. doi:10.1145/2207676.2208662
- [17] Yulia Goldenberg and Noam Tractinsky. 2021. Towards the right direction in bidirectional user interfaces. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–13.
- [18] 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.
- [19] 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. doi:10.1093/iwc/iww029
- [20] 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. doi:10.1145/3596256
- [21] 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. doi:10.1145/1240624.1240702
- [22] 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.
- [23] 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. doi:10.1007/978-3-031-05463-1_2
- [24] Andrew Hard, Kanishka Rao, Rajiv Mathews, Swaroop Ramaswamy, Françoise Beaufays, Sean Augenstein, Hubert Eichner, Chloé Kiddon, and Daniel Ramage. 2018. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604* (2018).
- [25] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. 2019. A systematic review of just-in-time adaptive interventions (JITAs) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 16 (2019), 1–21.
- [26] 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. doi:10.1016/S0166-4115(08)62386-9
- [27] Luis Hernández-Álvarez, José María de Fuentes, Lorena González-Manzano, and Luis Hernández Encinas. 2021. Privacy-Preserving Sensor-Based Continuous Authentication and User Profiling: A Review. *Sensors* 21, 1 (2021). doi:10.3390/s21010092
- [28] 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. doi:10.1145/2470654.2466463
- [29] 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.
- [30] 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. doi:10.1016/j.ergon.2008.09.004
- [31] 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. doi:10.1016/j.ergon.2018.10.002
- [32] Muhammad Waseem Iqbal, Nadeem Ahmad, Syed Khuram Shahzad, Irum Feroz, and Natash Ali Mian. 2018. Towards adaptive user interfaces for mobile-phone in smart world. *International journal of advanced computer science and applications* 9, 11 (2018).
- [33] Helena Eschricht Jensen, Chaudhary Muhammad Aqdu Ilyas, Ashkan Tashk, Bart Cooreman, Sofie Beier, and Per Bækgaard. 2025. Context Preservation Through Eye Tracking: Adaptive Reading Application Design for an Optimal Reading Experience. In *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications*. 1–8.
- [34] 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. doi:10.1145/3678525
- [35] Amy K Karlson, Benjamin B Bederson, and J Contreras-Vidal. 2006. Understanding single-handed mobile device interaction. *Handbook of research on user interface design and evaluation for mobile technology* 1 (2006), 86–101.
- [36] 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. doi:10.1145/3313831.3376641
- [37] Junhan Kong, Tianyuan Cai, and Zoya Bylinskii. 2023. Improving Mobile Reading Experiences While Walking Through Automatic Adaptations and Prompted Customization. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) (UIST '23 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 19, 3 pages. doi:10.1145/3586182.3616666
- [38] Junhan Kong, Jacob O. Wobbrock, Tianyuan Cai, and Zoya Bylinskii. 2025. Supporting Mobile Reading While Walking with Automatic and Customized Font Size Adaptations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 1007, 13 pages. doi:10.1145/3706598.3713367
- [39] 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. doi:10.1109/VR.2019.8798065
- [40] Charaka Vinayak Kumar, Ashok Urlana, Gopichand Kanumolu, Bala Mallikarjunarao Garlapati, and Pruthwik Mishra. 2025. No LLM is Free From Bias: A Comprehensive Study of Bias Evaluation in Large Language models. *arXiv preprint arXiv:2503.11985* (2025).
- [41] Anežka Kuzmíčová, Theresa Schilhab, and Michael Burke. 2020. m-Reading: Fiction reading from mobile phones. *Convergence* 26, 2 (2020), 333–349.
- [42] 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. doi:10.1145/3551708.3551772
- [43] 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. doi:10.1016/j.ijhcs.2007.04.001
- [44] Ming-I Brandon Lin and Shun-Wen Cheng. 2024. Different smartphone tasks and traffic complexity affect pedestrian awareness of co-existing road objects and cerebral oxygenation during shared space walking. *Transportation research part F: traffic psychology and behaviour* 103 (2024), 460–479.
- [45] Haotian Liu, Chunyuan Li, Yuheng Li, et al. 2023. Improved baselines with visual instruction tuning. doi:10.48550/arXiv.2310.03744 [A/OL].
- [46] Xingyu Bruce Liu, Jiahao Nick Li, David Kim, Xiang'Anthony' Chen, and Ruofei Du. 2024. Human i/o: Towards a unified approach to detecting situational impairments. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [47] 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> [A/OL].
- [48] Steven G Luke, Celeste Tolley, Adriana Gutierrez, Cole Smith, Toni Brown, Kate Woodruff, and Olivia Ford. 2024. The perceptual span in dyslexic reading and visual search. *Dyslexia* 30, 4 (2024), e1783.
- [49] Guojie Ma. 2017. Does interword spacing influence lexical processing in Chinese reading? *Visual Cognition* 25, 7-8 (2017), 815–824. doi:10.1080/13506285.2017.1338322
- [50] 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.
- [51] 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.
- [52] Marcello Maniglia, Steven M Thurman, Aaron R Seitz, and Pinakin G Davey. 2018. Effect of varying levels of glare on contrast sensitivity measurements of young healthy individuals under photopic and mesopic vision. *Frontiers in psychology* 9 (2018), 899.

- [53] 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. doi:10.1016/0749-596X(88)90063-0
- [54] 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. doi:10.1145/2393347.2396535
- [55] 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. doi:10.1145/1085777.1085798
- [56] 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. doi:10.1145/985921.986034
- [57] 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 (JITAI) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine* 52, 6 (May 2018), 446–462. doi:10.1007/s12160-016-9830-8
- [58] 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. doi:10.1007/978-3-642-40477-1_6
- [59] 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. doi:10.1145/2556288.2557312
- [60] Thông T Nguyễn, Xiaokui Xiao, Yin Yang, Siu Cheung Hui, Hyejin Shin, and Junbum Shin. 2016. Collecting and analyzing data from smart device users with local differential privacy. *arXiv preprint arXiv:1606.05053* (2016).
- [61] Uran Oh and Leah Findlater. 2013. The challenges and potential of end-user gesture customization. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1129–1138.
- [62] OpenAI. 2024. ChatGPT. <https://openai.com/blog/chatgpt> [EB/OL].
- [63] Gaisha Oralova and Victor Kupperman. 2021. Effects of spacing on sentence reading in Chinese. *Frontiers in Psychology* 12 (2021), 765335. doi:10.3389/fpsyg.2021.765335
- [64] 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*. doi:10.1145/3613904.3642747
- [65] 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. doi:10.1016/j.jnca.2016.03.013
- [66] Hilary Palmén, Michael Gilbert, and David Crossland. 2023. How bold can we be? The impact of adjusting font grade on readability in light and dark polarities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [67] 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.
- [68] Yong S Park and Sung H Han. 2010. One-handed thumb interaction of mobile devices from the input accuracy perspective. *International Journal of Industrial Ergonomics* 40, 6 (2010), 746–756.
- [69] Denis G Pelli, Katharine A Tillman, Jeremy Freeman, Michael Su, Tracey D Berger, and Najib J Majaj. 2007. Crowding and eccentricity determine reading rate. *Journal of vision* 7, 2 (2007), 20–20.
- [70] Yue Peng, Wei Yan, and Liying Cheng. 2021. Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test. *Language Testing* 38, 2 (2021), 326–337.
- [71] Stefan Poslad. 2011. *Ubiquitous computing: smart devices, environments and interactions*. John Wiley & Sons.
- [72] 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.
- [73] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372.
- [74] 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.
- [75] Zhanna Sarsenbayeva, Jorge Goncalves, Juan Garcia, 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. doi:10.1145/2971648.2971734
- [76] 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. doi:10.1145/3152771.3156161
- [77] 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. doi:10.1145/3214285
- [78] 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. doi:10.1145/1851600.1851619
- [79] Albrecht Schmidt, Michael Beigl, and Hans-Werner Gellersen. 1999. Sensor-based adaptive mobile user interfaces. In *HCI (2)*. 251–255.
- [80] 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. doi:10.9781/ijimai.2017.09.001
- [81] Lisa Schwetlick and Michael H Herzog. 2025. Visual crowding. *Annual Review of Vision Science* 11 (2025).
- [82] 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.
- [83] 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. doi:10.1109/PARC49193.2020.236619
- [84] Somipam R Shimray, Chennupati Keerti, Chennupati K Ramaiah, et al. 2015. An overview of mobile reading habits. *DESIDOC Journal of Library & Information Technology* 35, 5 (2015), 343–354.
- [85] Anna Shutaleva, Ekaterina Kuzminykh, and Anastasia Novgorodtseva. 2023. Youth Practices of Reading as a Form of Life and the Digital World. *Societies* 13, 7 (2023). doi:10.3390/soc13070165
- [86] Yisheng Song, Ting Wang, Puyu Cai, Subrota K Mondal, and Jyoti Prakash Sahoo. 2023. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *Comput. Surveys* 55, 13s (2023), 1–40.
- [87] 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. doi:10.1007/978-3-030-29513-4_12
- [88] 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. doi:10.1007/s11042-017-5236-2
- [89] Nada 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. doi:10.1145/3132635.3132648
- [90] Garrett 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) (NordCHI '18)*. Association for Computing Machinery, New York, NY, USA, 338–351. doi:10.1145/3240167.3240207
- [91] Garrett 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. doi:10.1145/3196709.3196760
- [92] Garrett 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. doi:10.1145/3290607.3290929
- [93] 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. doi:10.3389/fpsyg.2017.01390

- [94] Soner Ülkü, Stephan Getzmann, Edmund Wascher, and Daniel Schneider. 2025. EEG Correlates of Cognitive Dynamics in Task Resumption After Interruptions: The Impact of Available Time and Flexibility. *European Journal of Neuroscience* 61, 4 (2025), e70027.
- [95] 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. doi:10.1145/1152215.1152262
- [96] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).
- [97] VERBI Software. 2024. *MAXQDA 2024: Software for qualitative and mixed methods data analysis*. VERBI Software, Berlin, Germany. Available from: <https://www.maxqda.com>.
- [98] Minh Duc Vu, Han Wang, Jieshan Chen, Zhuang Li, Shengdong Zhao, Zhenchang Xing, and Chunyang Chen. 2024. Gptvoicetasker: Advancing multi-step mobile task efficiency through dynamic interface exploration and learning. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–17.
- [99] Minh Duc Vu, Han Wang, Zhuang Li, Gholamreza Haffari, Zhenchang Xing, and Chunyang Chen. 2023. Voicify your ui: Towards android app control with voice commands. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–22.
- [100] John Vukich, Satish Modi, Bret L Fisher, Karl Stonecipher, Ling Lin, and Magda Michna. 2024. Clinical comparison of a small-aperture intraocular lens versus a monofocal control. *Journal of Refractive Surgery* 40, 11 (2024), e824–e835.
- [101] Shaun Wallace, Zoya Bylinskii, Jonathan Dobres, Bernard Kerr, Sam Berlow, Rick Treitman, Nirmal Kumawat, Kathleen Arpin, Dave B Miller, Jeff Huang, et al. 2022. Towards individuated reading experiences: Different fonts increase reading speed for different individuals. *ACM Transactions on Computer-Human Interaction (TOCHI)* 29, 4 (2022), 1–56.
- [102] Chenzhu Wang, Mingyu Hou, Said M Easa, and Jianchuan Cheng. 2024. Safety analysis of pedestrians distracted by mobile phones at street crossings: Field study in Nanjing. *Accident Analysis & Prevention* 200 (2024), 107563.
- [103] 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. doi:10.1080/03601270802491122
- [104] Peng Wang, Dickson KW Chiu, Kevin KW Ho, and Patrick Lo. 2016. Why read it on your mobile device? Change in reading habit of electronic magazines for university students. *The Journal of Academic Librarianship* 42, 6 (2016), 664–669.
- [105] Wei Wang, Hourieh Khalajzadeh, John Grundy, Anuradha Madugalla, and Humphrey O Obie. 2024. Adaptive user interfaces for software supporting chronic disease. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Society*. 118–129.
- [106] 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. doi:10.1145/2541016.2541087
- [107] Jonas Wanner, Lukas-Valentin Herm, Kai Heinrich, and Christian Janiesch. 2022. The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. *Electronic Markets* 32, 4 (2022), 2079–2102.
- [108] David Whitney and Dennis M Levi. 2011. Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in cognitive sciences* 15, 4 (2011), 160–168.
- [109] 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.
- [110] Ankit Kumar Yadav, Nishant Mukund Pawar, and Nagendra R Velaga. 2024. Modeling the influence of smartphone distraction and pedestrian characteristics on pedestrian road crossing behavior. *Transportation research record* 2678, 5 (2024), 346–361.
- [111] Yang Yang and Yanchinda Jirawit. 2025. Navigating HSK level 5 challenges: dual perspectives from Thai undergraduates and Chinese lecturers. *Language Testing in Asia* 15, 1 (2025), 24.
- [112] 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. doi:10.1145/1979742.1979845
- [113] John Zerilli, Umang Bhatt, and Adrian Weller. 2022. How transparency modulates trust in artificial intelligence. *Patterns* 3, 4 (2022).
- [114] Xiaochen Zhang, Xiaoyu Huang, Yiran Ding, Liumei Long, Wujing Li, and Xing Xu. 2024. Advancements in Smart Wearable Mobility Aids for Visual Impairments: A Bibliometric Narrative Review. *Sensors* 24, 24 (2024), 7986.
- [115] 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. doi:10.1145/3544548.3581430

- [116] Deyao Zhu, Jun Chen, Xiaoqian Shen, et al. 2023. Minigt-4: Enhancing vision-language understanding with advanced large language models. doi:10.48550/arXiv.2304.10592 [A/OL].
- [117] Shangshang Zhu, Xinyu Su, and Yanan 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. doi:10.1093/iwc/iwab020
- [118] 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. doi:10.1504/ijnsnet.2019.100218

A Study 1 Interview Participant Demographics and Interview Protocol

This appendix provides the participants' demographic and semi-structured interview guide used in Formative Study 1.

Part 1: Exploring Typical Reading Impairment Scenarios (15–20 minutes)

1. Can you describe a situation where you found it difficult to read on your phone?
2. What were you doing at that time? (e.g., walking, riding a bus, lying down)
3. What type of content were you trying to read? (e.g., articles, messages, navigation)
4. What environmental factors made reading difficult? (e.g., lighting, movement, distractions)
5. Did the content's presentation (e.g., layout, font, content density) contribute to the difficulty?

Part 2: Current Coping Strategies and Motivations (15 minutes)

6. What do you find most frustrating about reading under difficult conditions?
7. How do you typically deal with reading difficulties on your phone?
8. Have you tried changing settings like font size or brightness? Why or why not?
9. How effective do you find these coping strategies?
10. What keeps you motivated to continue reading even when it is challenging?

Part 3: Concepts for Comfortable Reading Methods (15–20 minutes)

11. In an ideal situation, how could your phone help you read more comfortably?
12. What features or changes would make reading easier for you?
13. Would you prefer the system to adjust settings automatically, or would you like manual control?
14. How would you like the system to interact with you? (e.g., quietly in the background, notify you, ask for input)
15. What would make you feel comfortable using such a system?

B Study 2 Experimental System and Equipment

We developed an application with six preloaded 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

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; uses 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	Graduate Student	Myopia 500; wears corrective glasses
P9	Male	23	Graduate Student	Normal vision
P10	Female	22	Graduate Student	Myopia 450; wears corrective glasses
P11	Female	21	Undergraduate	Myopia 300; wears corrective glasses
P12	Female	21	Graduate 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 9: Demographic Overview of Interview Participants

preloaded, 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.

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.

Data Collection Smartphone sensors recorded environmental data throughout the experiment. The reading distance was measured using the front camera, the intensity of the ambient light was captured by the light sensor, and the vibration displacement of the three axes was tracked by the accelerometer. The adjustments of the text parameters were recorded using the built-in functions of Android, 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.

C Label Tree Prompt

C.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)

C.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.)

D User Study Questionnaires and Interview Protocol

This appendix provides the full set of questionnaires and the interview protocol used in the user study. Standardized scales (NASA-TLX, UEQ-S, and SUS) were used without modifications and are not included here. Participants completed these surveys to provide demographic information, context data during the adaptation period, and qualitative feedback in a post-study interview.

D.1 Pre-test Demographic and Vision Questionnaire

1. Age: a. 18–24 b. 25–34 c. 35–44 d. 45–54 e. 55–64 f. 65+
2. Gender: a. Male b. Female c. Non-Binary d. Prefer not to say
3. Occupation: a. Student b. Academic/Researcher c. Professional d. Retired e. Unemployed f. Other: (please specify)
4. Is Chinese your native language? a. Yes b. No
5. Education level: a. High School b. Undergraduate c. Graduate d. Doctorate e. Other: (please specify)
6. Vision Condition (select all): a. Normal Vision b. Myopia c. Hyperopia d. Astigmatism e. Presbyopia f. Other: (please specify)
7. Do you wear corrective lenses? a. Yes b. No
8. Have you been diagnosed with any other eye conditions? a. Yes (please specify:) b. No
9. How often do you read on a mobile device? a. Daily b. Several

times a week c. Weekly d. Rarely e. Never

10. Daily reading time on mobile devices: a. <10 min b. 10–30 min c. 30–60 min d. 1–2 hrs e. >2 hrs

11. Primary devices for reading (select all): a. Smartphone b. Tablet c. Laptop d. E-Reader e. Desktop f. Printed Books g. Other: please specify)

12. Do you often adjust text settings? a. Yes b. No

13. If yes, which settings do you adjust most? (Select all): a. Font Size b. Font Style c. Line Spacing d. Brightness e. Color Contrast f. Other: (please specify)

D.2 Daily Survey during Adaptation Period

1. How many different reading scenarios did you use SituFont in today? a. 3 b. 4 c. 5 d. >5
2. Where did you use SituFont today? (Select all): a. Indoors b. Outdoors c. Public Transport d. Walking e. Stationary
3. Lighting conditions? (Select all): a. Bright b. Normal Indoor c. Low Light d. Other: (please specify)
4. Did you experience distractions? a. Yes b. No
5. If yes, what are the main sources of distraction? (Select all): a. Noise b. Physical Movement c. Notifications d. Visual Distractions e. Other: (please specify)
6. Did you experience eye strain or discomfort? a. None b. Slight c. Moderate d. Severe
7. Did you manually adjust text settings in SituFont? a. Yes b. No
8. If yes, which settings did you adjust? (select all): a. Font Size b. Font Weight c. Line Spacing d. Brightness e. Other: (please specify)
9. Did you encounter any problems with SituFont today? a. Yes (please describe:) b. No

D.3 Post-Test Semi-Structured Interview Protocol

1. Could you tell me about your overall experience using SituFont over the past week?
2. Did you encounter any difficulties adjusting the font settings in SituFont? If so, please describe.
3. What did you find most helpful about using SituFont?
4. Were there any situations where SituFont did not help or made reading harder?
5. How did SituFont compare to the traditional display you used?
6. Did you feel you had enough control over SituFont's text adjustments?
7. Would you prefer more options to manually customize text settings?
8. Did SituFont ever make adjustments that you disagreed with? If so, which ones?
9. Would you continue using SituFont on your device? Why or why not?
10. What improvements would you suggest for SituFont?