GlassMessaging: Towards Ubiquitous Messaging Using OHMDs

NUWAN JANAKA, National University of Singapore, Singapore
JIE GAO[*], Singapore University of Technology and Design, Singapore
LIN ZHU[*], Tsinghua University, China
SHENGDONG ZHAO, **LAN LYU**, and **PEISEN XU**, National University of Singapore, Singapore
MAXIMILIAN NABOKOW[*], University of North Carolina at Chapel Hill, United States
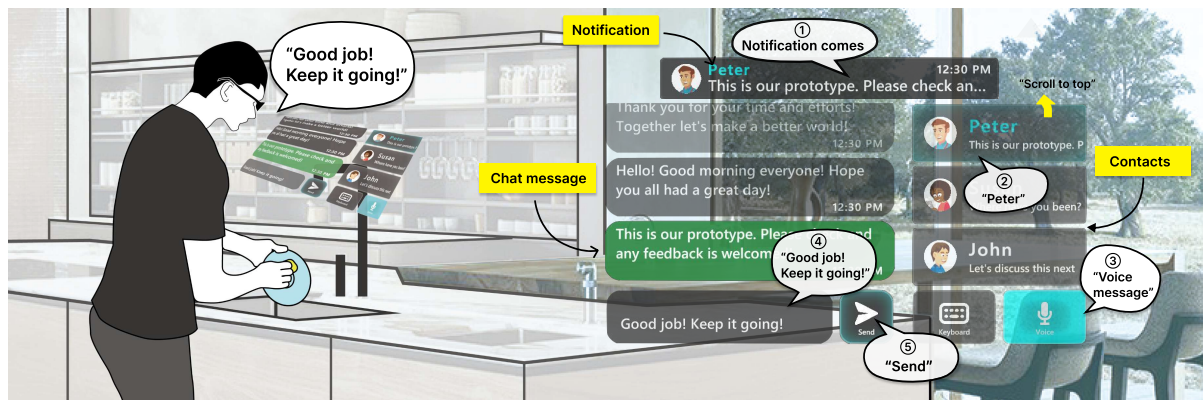**SILANG WANG** and **YANCH ONG**, National University of Singapore, Singapore

Fig. 1. Interactions in *GlassMessaging*. The user undertakes a sequence of steps: **1) receives a message notification** from Peter; **2) utters the voice command 'PETER'**, which prompts the contact list to automatically scroll to and open the chat interface of "Peter"; **3) invokes the voice command 'VOICE MESSAGE'**, thereby activating voice dictation (or alternatively, opening the keyboard by pressing the button in mid-air); **4) starts dictating** the reply message; **5) issues the voice command 'SEND'**, resulting in the message being sent.

Communicating with others while engaging in simple daily activities is both common and natural for people. However, due to the hands- and eyes-busy nature of existing digital messaging applications, it is challenging to message someone while performing simple daily activities. We present *GlassMessaging*, a messaging application on Optical See-Through Head-Mounted Displays (OHMDs), to support messaging with voice and manual inputs in hands- and eyes-busy scenarios.

---

[*]This work was done while the author was interning with NUS-HCI Lab at the National University of Singapore

---

Authors' addresses: Nuwan Janaka, nuwanj@comp.nus.edu.sg, National University of Singapore, NUS-HCI Lab, School of Computing, Singapore; Jie Gao, gaojie056@gmail.com, Singapore University of Technology and Design, Information Systems Technology and Design, Singapore; Lin Zhu, zhu-l20@mails.tsinghua.edu.cn, Tsinghua University, Information Systems Technology and Design, Beijing, China; Shengdong Zhao, zhaosd@comp.nus.edu.sg; Lan Lyu, lanlyu@andrew.cmu.edu; Peisen Xu, peisen.xu@cnrsatcreate.sg, National University of Singapore, NUS-HCI Lab, School of Computing, Singapore; Maximilian Nabokow, mnabokow@cs.unc.edu, University of North Carolina at Chapel Hill, Computer Science, North Carolina, United States; Silang Wang, wangsilangCHANGE@gmail.com; Yanch Ong, ongyc1999@gmail.com, National University of Singapore, NUS-HCI Lab, School of Computing, Singapore.

GlassMessaging is iteratively developed through a formative study identifying current messaging behaviors and challenges in common *multitasking with messaging* scenarios. We then evaluated this application against the mobile phone platform on varying texting complexities in eating and walking scenarios. Our results showed that, compared to phone-based messaging, *GlassMessaging* increased messaging opportunities during multitasking due to its hands-free, wearable nature, and multimodal input capabilities. The affordance of *GlassMessaging* also allows users easier access to voice input than the phone, which thus reduces the response time by 33.1% and increases the texting speed by 40.3%, with a cost in texting accuracy of 2.5%, particularly when the texting complexity increases. Lastly, we discuss trade-offs and insights to lay a foundation for future OHMD-based messaging applications.

CCS Concepts: • **Human-centered computing** → **Mixed / augmented reality**; **Empirical studies in ubiquitous and mobile computing**; *User studies*; **Interaction techniques**.

Additional Key Words and Phrases: texting, messaging, heads-up computing, smart glasses, OST-HMD

1 INTRODUCTION

The pervasive use of mobile devices has substantially augmented our communication abilities, facilitating instant connectivity irrespective of time and location. As a result, messaging applications have grown immensely popular, with WhatsApp, Telegram, Snapchat, and Messenger among the top 10 most downloaded mobile applications [20].

However, seamlessly integrating such applications into daily routines remains a challenge. Engaging with mobile messaging during daily activities such as preparing food, walking, or jogging can be difficult. The inherently handheld design of current messaging applications often necessitates considerable visual and manual engagement, thereby interfering with other tasks that demand similar attention. Despite these challenges, people persist in messaging while participating in various activities, as underscored by a study showing that 13% of text messages are sent while on the move [6]. This indicates an unwillingness among users to relinquish the benefits of mobile messaging even when multitasking.

Despite the introduction of alternative input methods, such as dictation, to mobile devices, these haven't entirely remedied the issue. The need to physically hold the device and concerns about convenience, privacy, and social perception have led most users to favor touchscreen typing over voice input [17, 26].

How can we improve mobile messaging for efficient communication during daily multitasking? Optical See-Through Head-Mounted Displays (OST-HMDs, OHMDs) or AR smart glasses [42] enable heads-up interactions with digital information while preserving awareness of the surrounding environment [52, 59, 84]. Moreover, since there is a lack of comfortable touch typing mechanisms [1], OHMD is more natural to afford hands-free voice input, which reduces the potential conflicts with daily activities. Therefore, examining the use of OHMDs for *multitasking with messaging* and their advantages over conventional phones is a promising avenue for research.

In this paper, we first conducted a survey and an observational study to gain a deeper understanding of the current needs, practices, challenges, and limitations users face with mobile messaging while multitasking. Based on these findings, we iteratively designed a messaging application for OHMDs, named *GlassMessaging*, and compared it with the Telegram[1] application on mobile phones in a controlled study during daily multitasking situations. Our results suggest that, despite the current technological constraints of the OHMD platform, *GlassMessaging* improved access to voice input and facilitated more seamless interactions compared to phones, reducing response

---

[1] https://telegram.org/

time by 33.1% and boosting texting speed by 40.3%. This points to the considerable potential of OHMDs as a valuable supplement for mobile phone-based messaging during multitasking.

Nonetheless, several limitations need to be addressed before the full potential of this platform can be realized. For instance, the use of *GlassMessaging* led to a 2.5% decrease in texting accuracy, particularly when the complexity of the text increased. Additionally, current OHMDs have certain inherent disadvantages (e.g., underdeveloped hardware capabilities, unfamiliar usage, lack of interactions [42, 50, 73]) compared to the extensively tested and well-established mobile phones presently available on the market.

The contribution of this work is threefold:

- Enhanced understanding of mobile multitasking messaging behavior through survey and observation study.
- An iteratively designed OHMD messaging application, *GlassMessaging*, with specific design features and guidelines that can inspire future OHMD-based messaging interfaces.
- An empirical evaluation comparing *GlassMessaging* to mobile phone-based messaging that reveals their trade-offs and demonstrates that *GlassMessaging* — despite the current technological limitations — has great potential to better accommodate people's daily communication needs in ubiquitous multitasking scenarios.

2 RELATED WORK

2.1 Multitasking with Messaging

People commonly talk to each other and sustain social interactions while performing daily tasks face-to-face. This social interaction is also mirrored when using messaging applications such as WhatsApp, Telegram, Slack, WeChat, etc., facilitating togetherness, intimacy, and support among users [13, 31]. For example, research has shown that using mobile instant messaging can improve student engagement and interactions during group discussions [72]. Nevertheless, the requirement for hands and eyes in the use of messaging apps can limit this type of social behavior, especially during multitasking activities.

Despite these challenges, people often engage with these applications while undertaking other activities such as walking, eating, or commuting. For example, a study of 60,000 text messages in the US found that 13% of the messages were sent while on the move [6]. Simultaneous conversations with multiple contacts and consumption of different media types are also common during messaging [6]. Furthermore, the expectation of a prompt response exerts social pressure on users, leading to frequent and unavoidable multitasking [4, 58, 60].

Previous research has investigated messaging/texting behaviors in various contexts, but limited research exists on understanding user requirements and improving behaviors during multitasking [6, 15, 31, 32, 58]. Hence, this study examines the messaging behaviors and requirements in everyday multitasking scenarios.

2.2 OHMD and Multitasking

The mobile phone, with numerous messaging applications available such as WhatsApp, Telegram, and WeChat, is a prevalent and convenient platform for multitasking [5]. However, it may be unsuitable for multitasking in scenarios requiring attention to surroundings or ongoing activities. For instance, texting while walking can lead to distractions and increased falling risk [36]. In driving scenarios, it can impair drivers' attention and increase accident risk [10, 61, 69, 71].

Conversely, Optical Head-Mounted Displays (OHMDs) can potentially enable better multitasking due to their hands-free nature and enhanced situational awareness [9, 11, 23, 33, 42, 43, 50–52, 59]. For example, using OHMDs for messaging can mitigate the distracting cognitive demands during driving compared to mobile phones [38, 39, 68].

However, current messaging apps for OHMDs, like Vuzix Blade's WeChat[2], are adaptations from mobile interfaces and lack design specificity for OHMD usage. These apps often feature a compact layout and an opaque interface, making them difficult for OHMD usage. Google Glass XE (2013-2017) [29, 30, 77] was designed for OHMDs but lacked formal evaluation in the literature and did not provide substantial contextual information for messages (see details at **sec 6.3**). This study explores the unexplored use case of OHMD messaging to support multitasking in daily scenarios.

2.3 Hands-free Text Entry on OHMD

The hands-free nature of OHMDs provides opportunities to support multitasking through hands-free texting inputs, such as voice input, head movement, and gaze input [50]. Among these, voice input is among the most promising strategy, proving especially useful when the availability of hands, eyes, keyboard, or screen is limited, or when natural language interaction is preferred [16]. It causes less visual distraction and offers a higher input speed than other methods, such as typing [49, 66]. Voice input can be a voice recording (recording voice messages) or voice dictation (transferring voice messages to text in real-time) [64]. Studies suggest that voice dictation improves the response rate and task completion efficiency without affecting multitasking levels compared to recording [64].

While head, tongue, and gaze inputs are potential methods for multitasking, ergonomic fatigue makes head movements unsuitable as a primary text input source [46, 50]. Current gaze input implementations on OHMDs are error-prone and require excessive calibration, making them impractical for daily *multitasking with messaging* scenarios [8, 23]. Thus, among hands-free input methods, voice input is the most reliable for enhancing multitasking support, notwithstanding potential drawbacks such as misrecognition and accidental triggering [48, 81].

3 STUDY 1: FORMATIVE STUDY ON UNDERSTANDING USERS' MESSAGING BEHAVIORS

We conducted a formative study to comprehend current messaging behaviors and related pain points during multitasking, guided by the following research questions.

- *RQ1: In what daily scenarios is messaging commonly performed?*
- *RQ2: What are the common features, existing practices, and difficulties of messaging in these scenarios?*

To answer our research questions, we conducted a survey to identify common messaging scenarios and then used an observational study to identify messaging behaviors.

3.1 Survey

We designed a questionnaire to assess messaging behaviors during daily activities (ADLs [22, 25]) and identify common scenarios. The questionnaire, consisting of 18 Likert-scale questions on messaging frequency during ADLs and an open-ended question on other multitasking scenarios, was posted on a university forum. We received 43 responses (17M, 26F; mean age = 26.1 years), with an average completion time of around 5 minutes. We excluded two incomplete responses.

3.1.1 Finding: Most Common Messaging Scenarios. After analyzing the survey, we narrowed down the most frequent scenarios to the top five (see Figure 2). However, we excluded the "online meeting" scenario since users tend to multitask on their laptops during the meeting, and our focus was on supporting mobility scenarios, specifically messaging on phones. We combined "walking" with the "transportation" scenario and disregarded

---

[2]https://apps.vuzix.com/app/wechat

"toileting" due to privacy concerns, despite its frequency. Hence, we selected two messaging-intensive multitasking scenarios, **Eating Alone** and **On Transportation**, for further observation.
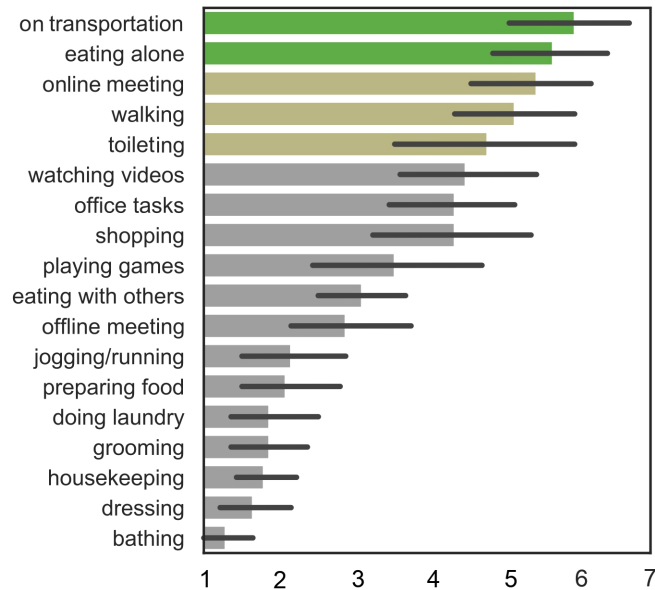

Fig. 2. Users' self-reported messaging frequencies during various daily activities. The data indicate that the two most common scenarios for daily messaging in conjunction with multitasking are **On Transportation** and **Eating Alone**. The Y-axis represents various activities, while the X-axis signifies a 7-point Likert score, with 1 representing "Do not message at all" and 7 being "Do most messaging". The error bar represents the 95% confidence interval.

3.2 Observational Study

3.2.1 Participants. After identifying the two most common messaging scenarios, we conducted a series of interviews and observations. We invited ten (6M, 4F; mean age = 25.3 years) university students (P1-P10) who frequently use messaging apps for social and learning purposes. All had over nine years of experience with WhatsApp, WeChat, or Telegram for mobile and desktop messaging. Each session lasted 30-60 minutes, with participants receiving a reward equivalent to USD 7.

3.2.2 Process. We observed users' messaging behaviors and attention division between screens and surroundings while they ate alone or commuted by bus, to understand their usage of familiar applications. We encouraged verbalization of actions and self-reporting. We also conducted semi-structured interviews for further insight into their messaging habits and real-life usage. These interviews were audio-recorded, transcribed, and qualitatively analyzed.

3.2.3 Data Analysis. One co-author analyzed interview transcripts and observation notes using thematic analysis described in Braun and Clarke's methodology [7]. After familiarizing with the data and generating initial codes, the coauthor grouped the codes into common themes based on the content. Next, with the help of an additional co-author, they discussed, interpreted, and resolved discrepancies or conflicts during the grouping process. Finally, they reviewed the transcripts and audio recordings to extract specific quotes relevant to each theme (See Appendix A-Figure 11 for the themes and codes).

3.2.4 *Findings.* We identified users' messaging and multitasking behaviors, messaging practices and features, and difficulties faced when messaging and multitasking.

Common Messaging and Multitasking Behaviors. People's messaging and multitasking behaviors differ in the two identified scenarios, which represent general hands-busy and mobility scenarios in which people cannot exclusively concentrate on messaging as their main focus lies on the primary tasks. To limit our scope in the transportation scenario, we only observed participants walking, standing, and sitting on the bus. Firstly, when standing on public transport, users typically lean on handrails for support while using their phones: 1) with both hands for viewing and composing messages, or 2) with one hand for viewing and the other alternating between holding the handrail and messaging. Once feeling safe or stable, they prefer using both hands for quicker typing. Similarly, while seated, users typically use both hands for messaging, resting one or both elbows on their waists for comfort (P1-P10).

Secondly, when people use messaging apps while eating alone, they often eat with one hand and use the other for the phone. They use both hands for efficient messaging when necessary and accept slower, less accurate one-handed use for non-urgent messages. "When using one hand [to type], I feel the screen is a bit wide, so I have to lean my phone to my side a bit and touch the keys on another side of the keyboard (P9)".

In both scenarios, people like to complement their phone usage with a smartwatch, if available. They check the notification on the smartwatch, determine the message's urgency, and pull out their phone if they choose to reply.

We observed that for formal messages, users typically use two hands to compose well-structured sentences and avoid errors, whereas for casual messages, they prefer quick, shorter replies split into multiple messages.

Common Features and Practices for Messaging. By identifying the most important and frequently used features of mobile messaging applications, we could prioritize them in future designs, ensuring that users can easily access all needed functions and not change their fundamental messaging style. The messaging process generally involves three key interfaces: notification, chat screen (one-to-one or group), and chat selection. Users primarily engage with the **notification interface**. Here, users determine whether to open the corresponding chat by viewing the sender's name and avatar in the pop-up. In **one-to-one chat screens**, users typically send texts and use stickers or emojis to express feelings. In **group chat screens**, users recognize members by names instead of avatars, given the frequent avatar changes. For **chat selection**, users generally scroll through chats or use the search bar for less familiar contacts.

We noted that **viewing and replying** to messages were the most common activities. Users, when focused on their primary tasks, tended to avoid high-attention-demanding activities like initiating new conversations. The process of viewing and replying, as described by participants, is detailed in Appendix A-Figure 10.

Common Difficulties Faced when Messaging and Multitasking. Apart from their behaviors and common practices, people face several difficulties when messaging and multitasking:

- **Hands-busy: Primary tasks may slow down or cease due to occupied hands.** Because people need to stop eating to compose messages with two hands, they may slow down or pay less attention to the food, thus reducing their enjoyment of the meal. In addition, some might need to use both hands while handling utensils (P3, P6); thus, they find it hard to free their hands for messaging. Similarly, messaging on a crowded bus is challenging, as one or both hands are needed for stability (P1-P10), restricting messaging. This temporarily discourages bus passengers from messaging (P9).
- **Eyes-busy: Switching limited attention between device and environment.** When occupied by their primary task, participants had little capacity to think thoroughly about their messaging. For example, while waiting for a bus, frequent attention shifts between the screen and the arriving bus were necessary, causing some participants to fear missing the bus due to excessive messages (P2, P3, P9). Similarly, messaging while

walking on a busy street often requires individuals to slow down or halt in a safe place, diverting their attention from their surroundings to their phones.

- **Other difficulties.** Participants also reflected on other issues: prolonged screen time in moving vehicles induced dizziness (P1, P3, P9). Participants subtly tilted their phones to protect their privacy when viewing personal messages or photos (P3, P4). During meals, screen glare and reflection, especially when phones were on a table, compounded visual difficulties. This effect was amplified for those with poor eyesight (P1).

4 STUDY 2: ITERATIVE DEVELOPMENT OF A MESSAGING APPLICATION FOR OHMDS

Study 1 identified limitations in messaging capability while multitasking on existing mobile platforms. In response, we iteratively developed a messaging application specifically for OHMDs, demonstrating promise for multitasking across various contexts (sec 2.2). This section focuses on evaluating the feasibility of **one-to-one text messaging on OHMDs during multitasking** (excluding multimedia content or group conversations), as it is the most common format (sec 3.2.4, [74]).

4.1 Iteration 1

4.1.1 Exploration of Existing Applications. Since we did not find any existing *messaging* applications designed explicitly for OHMDs, we utilized two popular mobile *messaging* apps (Telegram and WhatsApp) as proxies to inform the design of *messaging* apps on OHMDs. We installed Telegram on Vuzix Blade[3] and Epson BT-300[4], and mirrored Telegram and WhatsApp on Microsoft HoloLens 2 (HL2)[5] (using Mirage[6]) and Nreal Light[7] (using USB-C display connection[8]). Figure 3 shows the interfaces.

We asked four volunteers (2M, 2F; mean age = 22.9) with more than five years of mobile messaging experience to compose, view, and reply to messages using each OHMD while sitting, standing, and walking. They used the think-aloud protocol to identify usability issues and shared their messaging experiences on each device. Before each device testing, the participants were trained on the supported interactions, such as the 2D touchpad on the right temples for Vuzix Blade, the hand-held 2D touchpad for BT-300, the 3D hand controller for Nreal Light, and the 3D mid-air gestures for HL2.

4.1.2 Insights for Iteration. Overall, participants found **the UI of the messaging apps to be generally intuitive but not well-tailored to OHMDs**, resulting in some usability issues. The non-transparent and cluttered UI layout with long chat history blocked eyesight, the color scheme was unsuitable for near-eye displays (e.g., the light mode was too bright, the dark mode was too transparent), and the font and buttons were too small. Interactions were not intuitive (e.g., 3D hand controller) and caused fatigue (e.g., mid-air gestures with large hand movements), and text entry with virtual keyboards was challenging.

4.2 Iteration 2

The focus of Iteration 2 was on adapting the UI and interactions of existing mobile messaging apps for OHMDs. To tackle the difficulties associated with virtual keyboards and support hands-busy situations, we added **voice dictation** for text entry and **voice commands** for hands-free UI navigation. We also incorporated **ring mouse interaction** for quicker and more accurate scrolling and selection [65, 67], while retaining **mid-air gestures**

---

[3]https://www.vuzix.com/products/vuzix-blade-smart-glasses
[4]https://tech.moverio.epson.com/en/bt-300/
[5]https://www.microsoft.com/en-us/hololens/hardware
[6]https://www.mirage-app.com/
[7]https://www.nreal.ai/light/
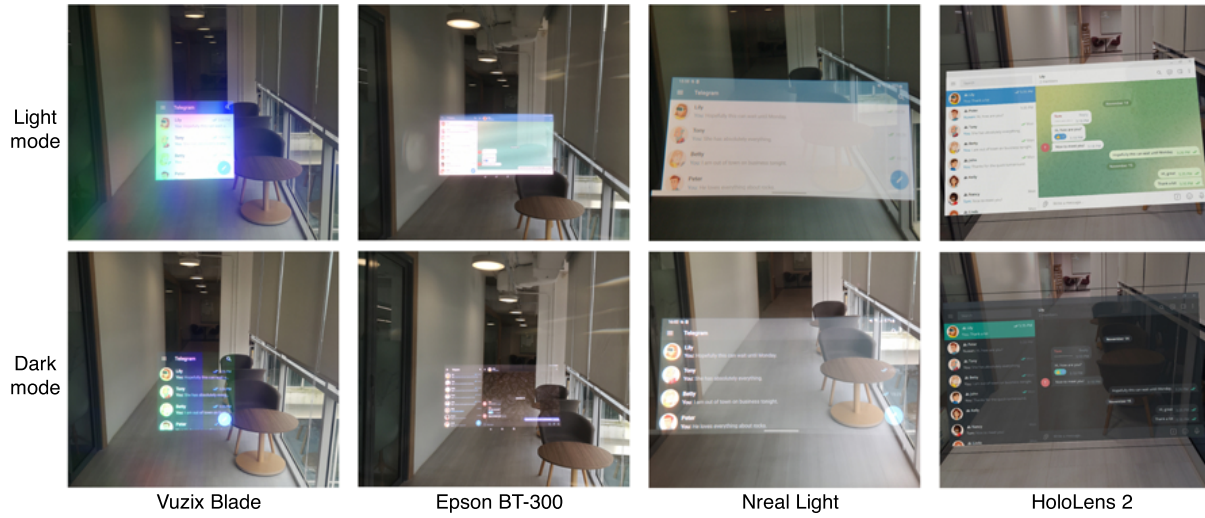[8]https://www.nreal.ai/compatibility-list/

Fig. 3. Interfaces of the Telegram messaging app on four OHMD devices (Vuzix Blaze, Epson BT-300, Nreal Light, HoloLens 2) used in **Iteration 1**, each presented in both Light and Dark Display Modes. Note: For optimal visibility and detail, please view this content on a device with a color display.

for their "intuitive" touch-like content manipulation paradigm. These updates aimed to support non-invasive multitasking scenarios on OHMDs.

4.2.1 Apparatus. We selected Microsoft HoloLens 2 (HL2), an OHMD with hand-tracking, voice commands, and world-scale positioning (2k resolution, 52° diagonal FoV), to develop *GlassMessaging*, our messaging app designed for OHMDs. A wireless ring mouse (Sanwa Supply 400-MA077) facilitated easy directional UI element selection (Figure 5). We developed *GlassMessaging* using Unity 3D (2021.3.6f1)[9], Mixed Reality Toolkit (MRTK 2.8)[10], leveraging MRTK's built-in functions for mid-air gestures, voice inputs, virtual keyboard[11], and content stabilization. To simulate a realistic messaging experience, we implemented a virtual chat server using Python, running on a tablet computer (Surface Pro 7+) connected to the HL2 via Wi-Fi, enabling bi-directional communication through a socket connection between the client and the server. For implementation details, see https://github.com/NUS-HCILab/GlassMessaging

4.2.2 Interface Design. To enhance learnability and maintain consistency [56] with familiar interfaces, we chose to modify the UIs of existing mobile *messaging* apps and tailor them to OHMDs, instead of entirely redeveloping them. The initial interface is shown in Figure 4a.

Main Interfaces. Our design focused on the three primary interfaces identified in *study 1* (sec 3.2.4): the chat screen, the contact list, and the notification. We positioned the chat screen at the middle-center, notifications at the top-center, and the contact list on the right side, improving their noticeability [14], based on importance (sec 3.2.4). Self-testing by four co-authors indicated that placing the chat list on the dominant hand side (contrary to the default left side in mobile/desktop *messaging* apps) minimized the effort and hand movements necessary for direct mid-air interactions.

---

[9]https://unity.com/
[10]https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/
[11]https://learn.microsoft.com/en-us/windows/mixed-reality/design/keyboard

UI Elements. Adhering to HL2's design guidelines [19] and existing findings [21], we employed a billboard style with a gray/dark background and white/green text to enhance legibility and distinguish UI elements visually. The background was semi-transparent (i.e., gray color) to see the environment while *messaging*. Font sizes ranged from 16-24 pt for comfortable legibility using the HL2's default Segoe UI font [18, 19]. As shown in Figure 4a, selected elements were highlighted by green or blue borders [11, 21].

Position of UI. In alignment with Microsoft's Mixed Reality guidelines [19, 27], we positioned the graphical interface 0.5m ahead and 10° below the eye level to facilitate comfortable direct manipulation via mid-air gestures. Furthermore, we implemented body-locking for a more comfortable reading experience during multitasking [27, 47].

(a) Initial interface design of *GlassMessaging*.   (b) Final interface of *GlassMessaging*.
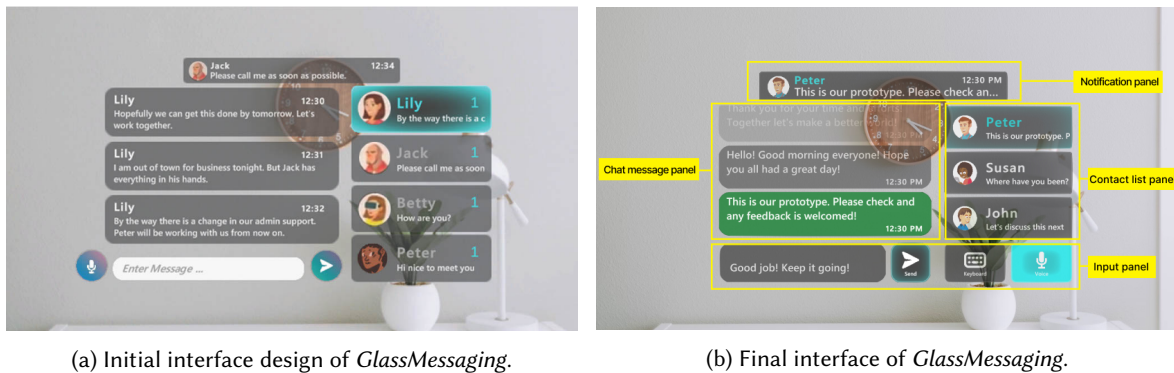
Fig. 4. (a) **Initial interface** includes a notification bar on top, a chat message box in the middle, a contact list on the right with numbers indicating unread messages, and a texting area at the bottom. (b) **Final interface** features a Notification Panel, Contact List Panel, Chat Messages Panel, and Input Panel (Voice/Keyboard, Send). **1) Notification panel**: temporarily displays new message notifications upon arrival. **2) Contact list**: shows displays all contacts available for communication. Selecting a contact (through voice or manual input) opens the respective chat conversation in the Chat Message panel. **3) Chat Message panel:** shows the chat history between the user and the selected contact. **4) Input panel**: contains a text entry field displaying the entered message, an input selection panel for toggling between voice dictation and virtual keyboard, and a 'send' button to transmit the entered text as a message. Note: For optimal visibility and detail, please view this content on a device with a color display.

4.2.3 Evaluation. Six volunteers (1 UI designer, 2 UX researchers, 1 HCI researcher, and 2 students) were invited to identify usability issues of the initial prototype while *messaging* during daily tasks, such as working on a computer while sitting, casually walking, and arranging items on the table while standing. Observations with the think-aloud protocol were utilized, similar to the previous iteration.

4.2.4 Insight for Iteration. Although participants found the initial *GlassMessaging* "interesting" and identified its potential advantages over existing mobile *messaging* platforms, they provided feedback for improvements on interface and input interactions, as shown in Table 1.

4.3 Final Version: *GlassMessaging*

The final version (see Figure 4b and 5) of *GlassMessaging* incorporates the feedback from the earlier iterations (e.g., Table 1).

Table 1. Problems identified in **Iteration 2** and their corresponding design solutions concluded for **Final Version**.

Problem	Design Solutions	Explanations
Blocked Vision	Adjust UI arrangement	• Limit the number of contacts and message history to around three to minimize clutter and visual interference with daily tasks, as messaging often occurs with only a few recent contacts. • Avoid virtual keyboard blocking the chat screen while typing.
	On-demand interface with alerts	• Allow fully hiding the interface when daily tasks need undivided attention, only allowing notifications to avoid delaying urgent messages.
	Increase transparency	• To increase the visibility of the environment, increase the transparency of messages from latest to oldest (i.e., the latest messages are more opaque and visible while old messages are gradually faded).
	Use non-visual feedback	• To increase visual attention on daily tasks, incorporate distinguishable auditory feedback into UI interactions and notifications. This strategy can reduce the need for visual attention on individual UI elements for feedback, such as clicking a button.
Inconvenient Interaction	Reduce interaction cost	• Position the most frequently used interactive UI elements closer to the user's dominant hand to enable easier mid-air manipulation. • Allow head-locking of the interface when the user needs to move their head for daily tasks, as body-locking requires additional head rotation to see the interface.
	Simplify voice commands	• Simplify voice commands for faster interaction (e.g., shorten 'SEND MESSAGE' to 'SEND').
	Voice command shortcuts	• Enable shortcuts for multi-step voice commands (e.g., use 'REPLY' to represent 'OPEN NOTIFICATION' and 'VOICE MESSAGE').
	Decrease learning curve	• Make ring interaction more intuitive to minimize confusion (e.g., due to hardware constraints, the ring mouse initially used the 'right' button counter-intuitively as a way to change selection 'leftward').
Errors	Reduce interfering feedback	• Disable hand-mesh visualization for mid-air gestures to avoid interfering with hand-eye coordination (e.g., although hand-mesh visualization provides feedback to users on whether mid-air interaction is working, it blocks users from seeing the fingertips, which are essential for fine-grain control of certain daily tasks)
	Increase error tolerance	• Increase the gap between interactable elements to minimize accidental triggers when using direct mid-aid manipulation

4.3.1 Visual Interface (Output). The visual interface of *GlassMessaging* (Figure 4b) consists of four main UI panels, namely, notifications, contacts, chat messages, and voice/keyboard input panels.

4.3.2 Audio Interface (Input-Output). As depicted in Figure 5, users can interact with *GlassMessaging* via voice commands (see Appendix B.1 for details) to navigate the UI (e.g., 'SCROLL UP', 'SCROLL TO TOP') and dictate text (using 'VOICE MESSAGE'). Audio feedback (e.g., beeps) accompanies some input interactions.

When the app is not in dictation mode, voice commands can directly activate various functionalities, such as opening notifications ('OPEN NOTIFICATION'), selecting contacts ('<NAME>'), sending the message ('SEND'), and hiding the interface ('HIDE CHAT'). Voice shortcuts such as 'TEXT <NAME>' are also available, which combine '<NAME>' and 'VOICE MESSAGE' for direct text entry. Similarly, the 'REPLY' command opens the notification and begins dictation for a reply immediately.

4.3.3 Manual-input Interface (Input). *GlassMessaging* supports two manual input methods: a wearable ring-mouse and mid-air hand gestures. See Appendix B.1 for details.

Ring Mouse. The user can scroll through the contact list using the ring mouse's 'up' and 'down' buttons. The 'right' button toggles between input modalities and selects the send button. The 'center' button activates the selected virtual button and serves as a long-press toggle to hide/reveal the entire interface.

Mid-air Interaction. The visual interface can be interacted with through mid-air gestures. The contact list can be scrolled by swiping, and a contact's chat can be opened by pressing their virtual icon. The input modality is chosen by selecting the corresponding virtual buttons (voice or keyboard). Pressing on a notification opens the chat with the sender.
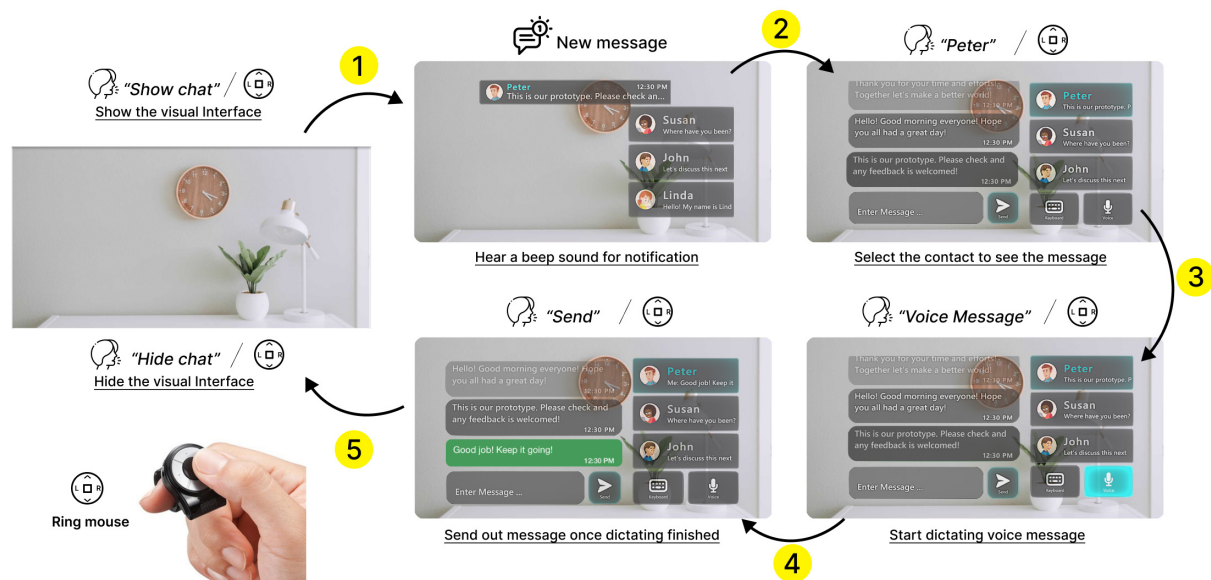

Fig. 5. Steps for sending a message after receiving a notification. The user wears the OHMD and a ring mouse and sees the environment. The user **(1) says 'SHOW CHAT'**, the interface is displayed, and a notification including the name of the sender and the sending time appears at the top of the view with a beep sound; **(2) says the name of the contact (e.g., 'PETER')**, and the system automatically navigates to the chat interface of the respective contact; **(3) says 'VOICE MESSAGE' and dictates the message via voice**. The system transcribes the user's utterances to text in real-time, displayed in the text entry box. Once the user stops speaking for a measured amount of time (silence gap), dictation turns off automatically; **(4) says 'SEND'**, and the system sends the message; **(5) says 'HIDE CHAT'**, and the full interface is hidden, restoring the full vision of the environment.

5 STUDY 3: EMPIRICAL COMPARISON OF *GLASSMESSAGING* WITH MOBILE PHONE *MESSAGING*

This study evaluates *GlassMessaging* (*Glass*, sec 4) by contrasting it with the current dominant solution, mobile phone (*Phone*)[12] messaging applications. The focus is on multitasking and messaging behaviors in mobile and hands-busy scenarios across various everyday tasks.

5.1 Participants

The study included 16 volunteers (P1-P16, 8M, 8F), aged 19-30 (mean age=23.2, SD=3.1), from the university community. All had normal or corrected vision and no reported color or visual impairments. To minimize bias, an equal number of native and non-native English speakers with professional working proficiency were included. Only 5 had previous experience (less than 2 hours) with OHMDs. All had used phones (10 Android, 6 iOS) for messaging for over six years, with Telegram and WhatsApp as the most utilized apps. All participants used typing on phones, and four occasionally used voice-to-text. Participants received USD14 compensation, and none had partaken in prior studies.

5.2 Apparatus

Participants employed *GlassMessaging* (final version) discussed in *study 2* (sec 4.3) for mobile *messaging*. For the mobile phone context, all participants used the Telegram mobile app[13] on a Google Pixel 6 Pro (Android)[14] phone to maintain consistency and avoid privacy concerns. We developed a Telegram Bot API[15] that autonomously sent and received messages, with a basic UI panel located above the main app for instructions (Figure 6A). The app delivered haptic and audio feedback to notify users about new messages.

To simulate a realistic experience, a scrolling contact list of 16 contacts with common English names was used for both platforms [6].

Both platforms were linked to the tablet computer running a Python program to display instructions, record participants' inputs, and compute messaging timing and accuracy. Moreover, participants' multitasking behavior and the screens of both platforms (Figure 6) were video-recorded.

5.3 Tasks

To evaluate the comparison between *GlassMessaging* (*Glass*) and mobile phones (*Phone*), a multitasking simulation involving multiple daily life activities (*study 1*, sec 3.1.1) was devised.

5.3.1 Primary Task: Walk to Eat Snacks in the Kitchen. The primary task involved a series of activities like walking, preparing food/drinks, washing dishes, and eating, representing mobility and hands/eyes/mouth busy situations (average duration: 8.1 min (SD=1.5), see Appendix C.1-Figure 12). This approach helped to comprehend the impact of realistic tasks on messaging behaviors using two *messaging* platforms.

5.3.2 Secondary Task: Messaging.

Task Focus. A *messaging task* (Figure 6) was designed for participants to perform concurrently with the primary task. The focus was on replying to messages, the most common and comprehensive behavior, which involves contact selection and message viewing (see sec 3.2.4, Figure 10, [6]).

---

[12]We do not consider other platforms, such as laptops, due to limited mobility support, or smart watches as they are not as commonly used and have limited text entry support [12, 45].

[13]Telegram for Android v9.1, https://telegram.org/android

[14]As anticipated, participants, especially iOS users, noted no differences in their use of the mobile app during the study, owing to the Telegram interface's uniformity across mobile operating systems, except for the differently located voice diction button.

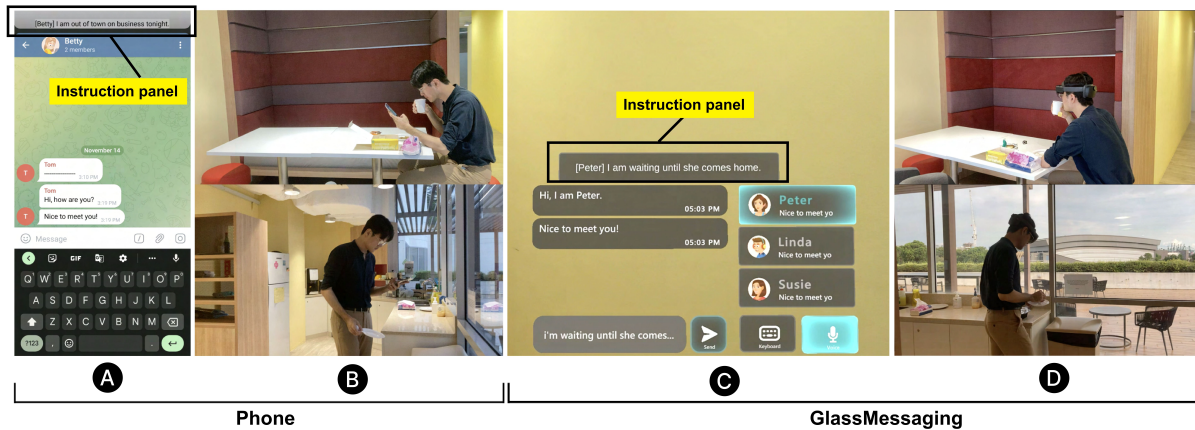[15]https://core.telegram.org/bots/api

Fig. 6. Apparatus including the Instruction Panel. (A) Telegram app with the Instruction Panel; (B) A participant uses *Phone* to perform *messaging* while eating (top) and cleaning (bottom); (C) *GlassMessaging* interface; (D) A participant uses *GlassMessaging* to perform *messaging task* while eating (top) and cleaning (bottom). Note: For optimal visibility and detail, please view this content on a device with a color display.

Message Content. In line with previous studies [66], the *messaging task* asked participants to re-type a preset message displayed in the instruction panel rather than drafting a reply independently. This design controls text length and minimizes the effects of confounding factors like thinking time and word choice. Participants were instructed to enter the text shown in the instruction panel and then send it as a message via typing or voice dictation.

These preset messages were drawn from a standard text entry phase set for mobile contexts, collected by Vertanen et al. [76]. Specifically, a set of *Long* messages with an average length of 13.9 (SD=1.6) words ($M =$ 71.4, SD = 3.2 characters) and *Short* messages with an average length of 6.6 (SD=1.2) words (M = 33.4, SD = 3.2 characters), totaling approximately 100 messages, was chosen based on previous findings[16] [53, 74].

Stimuli. Participants received a maximum of 10 instructions per session to simulate a realistic chat conversation. Each instruction contained a message the participant had to send to a contact, both selected randomly without repetitions. A minimum gap of 45 seconds was maintained between each instruction [4, 37, 55]. For example, if the instruction was "[John] Did you get this?", the participant had to send the message "Did you get this?" to "John". The current instruction was displayed continuously to facilitate easier typing and to eliminate the potential confounding factor of working memory capacity. New instructions were given one at a time and only after the previous instruction was addressed. The chat history was reset to a few initial messages before each session.

5.4 Study Design

The study was conducted using a within-subject, repeated-measures design. All the independent variables and measures are listed in Figure 7 and Table 2. Specifically, two independent variables, *Platform* (*Glass vs. Phone*) and *Message Length* (*Short vs. Long*), generated four conditions. A baseline condition, *no-messaging* (i.e., participants only performed the primary task), was included for comparison. The *Platform* variable was counterbalanced, and *Message Length* was administered in ascending lengths.

---

[16]The average *messaging* message length is around 14 words (70 characters)
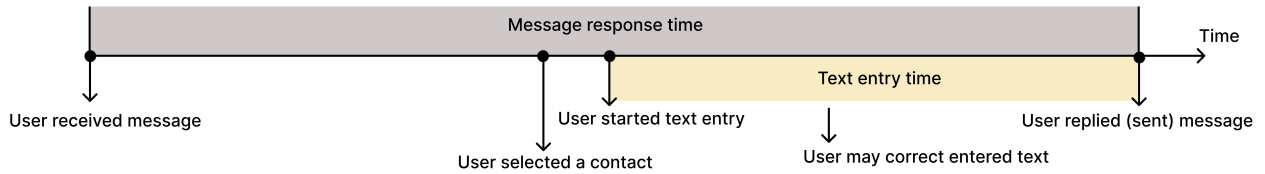
Fig. 7. Timing associated with receiving, composing, and sending messages.

Table 2. *Study 3* design table. See Figure 7 for details.

Independent Variables	**Platform**	• *Glass* Using *GlassMessaging* for *messaging*	• *Phone* Using Telegram for *messaging*
	Message Length	• *Short* The user receives/sends short messages	• *Long* The user receives/sends long messages
Measures (Dependent Variables)	**Primary task**	*Task Completion Increase* = $\frac{Task\ Completion\ Time}{Task\ Completion\ Time\ Without\ Texting}$, measuring time increase to complete the primary task	
	Secondary (*messaging*) task	• *Message Count* = The total number of messages the user sent • *Texting Speed* (WPM) = $\frac{Number\ of\ Words}{Text\ Entry\ Time}$, [3] where *text entry time* is calculated from the start of entering message to sending the message (Figure 7); • *Message Response Time* (seconds) = Average time gap between receiving and replying (Figure 7) • *Texting Accuracy* (%) = 1 − *Minimum String Distance* (*MSD*), where MSD [3, 70] means lowercase Levenshtein Distance between replied message and received message. • *Perceived Messaging Effectiveness* (1-7) = "It was effective to use the selected platform for texting while engaged with the primary task" • *User Experience Questionnaire - Short* (*UEQ-S* measures pragmatic and hedonic qualities using 8-item scale where each item is transformed to -3 to +3) [24, 40]	
	Overall	• *Raw NASA − Task Load Index* (*RTLX*, 0-100) [35] • *Preference* (Overall, Short, Long) and reasons	

5.5 Procedure

After obtaining consent, we conducted a 30-40 minute training session to familiarize the participant with the apparatus, primary, and secondary tasks. Participants were explicitly informed they could utilize existing typing/correction aids such as auto-correct, auto-complete, swipe typing, and voice input on the mobile app. Nevertheless, they were not briefed about the different text lengths before or during the study sessions. Once acclimated to the apparatus, they first experienced a *no-messaging* condition followed by four *messaging* conditions on two *Platforms*. Post each condition; participants completed questionnaires detailing their multitasking and *messaging* experiences with each *Platform* before taking a break. After the experiment, they filled out a subjective preference questionnaire and were interviewed for approximately 10-15 minutes. The whole experiment, including training, lasted for approximately 2 hours.

Participants were instructed to carry out primary tasks as naturally as possible and maintain consistency in food/drink preparation across conditions (e.g., snack quantity, water- and heat-levels of drinks). For the *messaging task*, participants were free to select their most comfortable modality (e.g., voice or keyboard input, ring mouse or mid-air gestures), as the study aimed to compare the natural usage of the two platforms rather than the different

modalities. Furthermore, they were guided to reply at their convenience, without the necessity for immediate responses.

5.6 Quantitative Results

Table 3 displays the participants' input interaction methods for each *Platform*. In most scenarios, participants utilized a blend of methods, particularly for *Glass*.

Table 3. Input interaction methods for *messaging* employed by participants during during *Study 3* (N=16). The count of participants is indicated in parentheses.

Platform	Contact selection	Text entry	Send message	Hide/Reveal UI
Glass	voice command (16) ring-mouse (5) mid-air gesture (4)	voice dictation (16)	voice command (16) ring-mouse (8) mid-air gesture (6)	voice command (10) ring-mouse (10)
Phone	touch scrolling (14) searching by typing (3)	typing (16): one-hand, two-hand, swipe voice dictation (2)	touch input (16)	Looking away from screen / Looking at screen (16)

Figures 8-9 present the mean performance of participants in terms of quantitative measures (see Appendix C.3-C.2 for detailed data and analysis). There were no significant differences between native and non-native speakers concerning the speed or accuracy of text entry.

(a) *Task Completion Increase*   (b) *Message Count*   (c) *Texting Speed* (WPM)   (d) *Message Response Time* (s)
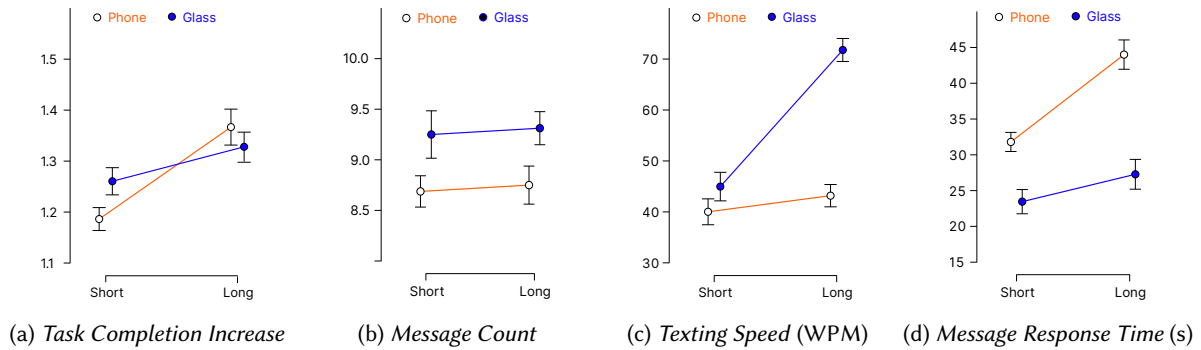
Fig. 8. Measures related to time and speed of primary and secondary tasks with 16 participants. Refer to Table 2 for measurement calculation details. The X-axis represents the *Message Length*, and the error bars represent the standard error. See Appendix C.3 for more details.

5.6.1 Primary Task Completion Increase. Figure 8a shows similar task completion times in both *Glass* and *Phone* for each *Message Length*. As expected, task completion time increases with *Message Length*. A repeated-measures ANOVA revealed a significant main effect of *Message Length* ($F_{1,45} = 15.68, p < 0.001, \eta_p^2 = 0.54$) and an interaction effect ($F_{1,45} = 4.62, p = 0.03, \eta_p^2 = 0.34$), but no significant main effect for *Platform*. Significant simple main effects were observed for *Phone* on *Message Length* ($p < 0.05$). Post-hoc analyses with Bonferroni correction showed that *Long messaging* ($M = 1.35, SD = 0.22$) significantly increased task completion time compared to *Short* ones ($M = 1.22, SD = 0.14, p_{bonf} < 0.001, r = 0.85$). Furthermore, *Short messaging* on *Phone* significantly decreased task completion time compared to *Long messaging* on both *Phone* and *Glass* ($p_{bonf} < 0.01$).

5.6.2 Message Count. Figure 8b demonstrates that *Glass* allows more *messaging* opportunities than *Phone*. A significant main effect was observed for *Platform* ($F_{1,45} = 7.43$, $p < 0.01$, $\eta_p^2 = 0.26$). The post-hoc analysis showed that the message count for *Glass* ($M = 9.41$, $SD = 1.34$) was significantly higher than *Phone* ($M = 8.81$, $SD = 1.51$, $p_{bonf} < 0.01$, $r = 0.71$). These results, along with the *Task Completion Increase* measure, suggest that *Glass* enables more *messaging* opportunities without extending task completion time.

5.6.3 Texting Speed. Figure 8c shows that *Glass* allows faster text entry than *Phone*. A repeated-measures ANOVA revealed a significant main effect of *Platform* ($F_{1,15} = 20.45$, $p < 0.001$, $\eta_p^2 = 0.58$) and *Message Length* ($F_{1,15} = 81.78$, $p < 0.001$, $\eta_p^2 = 0.85$), as well as their interaction ($F_{1,15} = 84.68$, $p < 0.001$, $\eta_p^2 = 0.85$). Additionally, significant simple main effects were found for *Platform* on *Long messaging* ($p < 0.001$). Post-hoc analyses showed that the texting speed of *Glass* ($M = 58.38$, $SD = 17.1$) was significantly higher than *Phone* ($M = 41.60$, $SD = 11.25$, $p_{bonf} < 0.001$, $d = 1.13$). Also, *Long messaging* ($M = 57.48$, $SD = 17.62$) had significantly higher speed than *Short messaging* ($M = 42.49$, $SD = 11.84$, $p_{bonf} < 0.001$, $d = 2.26$). Notably, the *Glass-Long* condition exhibited higher speed than others.

5.6.4 Message Response Time. Figure 8d demonstrates that *Glass* ensures shorter response time than *Phone*. Repeated measures ANOVA showed main effects of *Platform* ($F_{1,15} = 30.30$, $p < 0.001$, $\eta_p^2 = 0.67$) and *Message Length* ($F_{1,15} = 24.46$, $p < 0.001$, $\eta_p^2 = 0.62$), as well as their interaction ($F_{1,15} = 8.49$, $p < 0.05$, $\eta_p^2 = 0.36$). Moreover, significant simple main effects were observed for *Platform* on *Message Length* ($p < 0.001$). Post-hoc analysis showed that the response time of *Glass* ($M = 25.37$, $SD = 10.72$) was significantly shorter than *Phone* ($M = 37.90$, $SD = 12.94$, $p_{bonf} < 0.001$, $d = 1.13$). As expected, *Long messaging* ($M = 35.64$, $SD = 14.95$) required a longer response time than *Short messaging* ($M = 27.63$, $SD = 10.35$, $p_{bonf} < 0.001$, $d = 0.72$), particularly for *Phone*. Interestingly, for *Glass*, no significant difference was found between *Short* and *Long messaging*, suggesting that response times were similar for both *Message Lengths*.

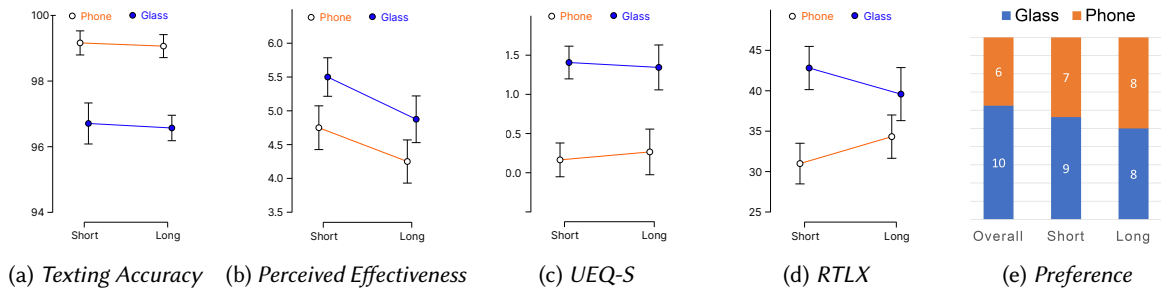(a) *Texting Accuracy* (b) *Perceived Effectiveness* (c) *UEQ-S* (d) *RTLX* (e) *Preference*

Fig. 9. Measures related to accuracy and perception of messaging task with 16 participants. Refer to Table 2 for measurement calculation details. The X-axis represents the *Message Length*, and the error bars represent the standard error. See Appendix C.3 for more details.

5.6.5 Texting Accuracy. Figure 9a illustrates that *Glass* yields less accurate *messaging* than *Phone*. Repeated measures ANOVAs showed a significant main effect of *Platform* for *Texting Accuracy* ($F_{1,15} = 17.16$, $p < 0.001$, $\eta_p^2 = 0.53$), with *Glass* ($M = 96.64$, $SD = 2.62$) showing significantly lower accuracy ($\approx 2.5\%$) than *Phone* ($M = 99.11$, $SD = 0.85$).

5.6.6 Perceived Messaging Effectiveness. Figure 9b shows that, after repeated-measures ANOVA, a significant main effect of *Platform* was observed ($F_{1,45} = 3.91$, $p = 0.05$, $\eta_p^2 = 0.15$). Here, *Glass* ($M = 5.19$, $SD = 1.40$) was perceived as more effective than *Phone* ($M = 4.50$, $SD = 1.70$) for *messaging* during multitasking.

5.6.7 UEQ-S. Figure 9c indicates that, after repeated-measures ANOVA, a significant main effect of *Platform* was observed ($F_{1,45} = 25.85$, $p < 0.001$, $\eta_p^2 = 0.36$). *Glass* ($M = 1.38$, $SD = 1.19$) provided a significantly better experience in pragmatic and hedonic qualities [40] than *Phone* ($M = 0.21$, $SD = 1.70$).

5.6.8 RTLX. A repeated-measures ANOVA showed only a significant main effect of *Platform* ($F_{1,15} = 5.17$, $p < 0.05$, $\eta_p^2 = 0.26$), where *Glass* ($M = 41.20$, $SD = 17.71$) incurred significantly higher task load than *Phone* ($M = 32.66$, $SD = 18.88$) during multitasking. Although there was no significant interaction effect ($p = 0.15$), as reflected in Figure 9d (see Appendix C.3-Figure 13 for individual scores) and participants' feedback, task load reduced for *Glass* with practice[17].

5.6.9 Preference. Figure 9e indicates a split preference between *Glass* and *Phone*, with a slight overall inclination towards *Glass* (*10/16*). The performance data was segregated based on overall preference and analyzed using paired-sample t-tests to explore the relationship between preference and performance. Results showed that participants favoring *Glass* exhibited significantly higher *Texting Speed* ($t(9) = 4.49$, $p = 0.004$, $d = 1.42$), *Perceived Messaging Effectiveness* ($t(9) = 3.54$, $p = 0.006$, $d = 1.12$), *UEQ-S* ($t(9) = 7.56$, $p < 0.001$, $d = 2.39$), and lower *Message Response Time* ($t(9) = -5.12$, $p < 0.001$, $d = -1.62$) and *Texting Accuracy* ($t(9) = -4.77$, $p = 0.001$, $d = -1.51$) compared to those who preferred *Phone*. This indicates a potential correlation of higher *Texting Speed*, *Perceived Messaging Effectiveness*, *UEQ-S*, and lower *Message Response Time* with *Glass* preference, but unexpectedly not lower *Texting Accuracy*. Conversely, participants preferring *Phone* demonstrated significantly lower *RTLX* ($t(5) = 4.50$, $p = 0.006$, $d = 1.84$) (and non significantly higher *Texting Accuracy*, $p = 0.112$) compared to *Glass* users, suggesting that lower *RTLX* might influence their preference.

Participants leaning towards *Glass* highlighted the advantages of voice input convenience, hands-free operation, and modality switching based on situational needs. On the other hand, *Phone* enthusiasts pointed out familiarity, easier message editing, and limitations of *GlassMessaging* (sec 5.7.3) as determining factors. Participants perceived short message entry to be more accurate using voice on *Glass*, but voice correction of long messages was deemed more "demanding" than typing on *Phone*, influencing their preference.

5.7 Qualitative Results

Overall, participants had **higher texting speed** and **quicker response times** with *Glass* than *Phone* due to its **hands-free nature**, **convenient access to voice inputs**, and **multimodal input support**. However, *Glass*'s lack of text editing support, occasional voice recognition errors, and platform unfamiliarity resulted in decreased accuracy and higher cognitive load.

5.7.1 GlassMessaging Promotes Quicker Responses by Providing Increased Messaging Opportunities. The hands-free, wearable, and multimodal aspects of *GlassMessaging* facilitated enhanced *messaging* opportunities and expedited responses compared to *Phone* (sec 5.6: *Message Count* and *Message Response Time*).

- **Hands-free operation.** *GlassMessaging* allowed hands-free UI navigation and dictation activation, while phone users had to manipulate the chat interface using their hands after accomplishing primary tasks. Specifically, *Glass* enabled *messaging* in *situations where hands were unoccupied but not accessible for phone interaction*, such as when they were wet, unclean, or unhygienic. With its voice commands and mid-air gestures, the wearable nature of *Glass* made the *messaging* application more easily accessible than the phone, eliminating the need to find a place to set it when hands were required for primary tasks.
- **Wearable, always-on nature.** The expedited response time of *GlassMessaging* was also a product of the *prompt reception of incoming messages*, which enabled quicker replies while multitasking. This was in contrast with *Phone* users who had to conclude primary tasks before accessing their phones to check

---

[17]Note: *Long* sessions were administered after *Short* sessions for both platforms (sec 5.4).

messages. P5 mentioned that "Smart glasses come in, so we can easily read messages and reply" during tasks like dishwashing, where holding a phone is inconvenient. Most participants (11/16) found *GlassMessaging* to be more convenient for staying connected and promptly processing information, thus accelerating the reading and response times for urgent messages.

- **Multimodal input.** *GlassMessaging* enabled improved modality coordination for multitasking between messaging and primary tasks. During activities like walking or stair climbing, users could utilize voice or the ring mouse to navigate the UI and input text, keeping their focus on the environment and the app. Likewise, while eating, participants could view messages with their eyes and respond using mid-air gestures or ring interaction without the need to look down at a phone. Interestingly, participants reported that this approach made messaging safer and more efficient as it negated the necessity to divert their visual and motor attention away from primary tasks. "I don't need to look down to see the message [compared to Phone], which is safer while walking and going upstairs (P1)" "[while eating] I have to shift my visual and hands attention to the Phone. With the Glasses, I don't have to do that. I can continue eating, and as long as I can speak and the Glasses can comprehend me, I can message while eating concurrently. (P9)"

5.7.2 GlassMessaging *Enhances Input Efficiency During Multitasking.* The hands-free voice input of *Glass* facilitated seamless interface navigation and efficient text entry during multitasking with *GlassMessaging* (sec 5.6: *Texting Speed*, *Perceived Messaging Effectiveness*).

- **Navigation.** Compared to phones, where users must search or scroll to locate a contact before they can compose a message, voice input with *Glass* reduced interaction steps, allowing users to select a chat and initiate dictation with a single command.
- **Text entry.** All participants in *GlassMessaging* utilized dictation, while only two did so for *Phone* (Table 3). As expected, in *Glass*, participants favored hands-free voice input to minimize interference with primary tasks requiring hands, "glass-based interaction is more convenient because voice input is generally faster than touch input for texting. Plus, I don't need to touch the [phone] screen while eating and texting (P10)". Most participants found dictation in *Glass* quicker than typing on *Phone*, particularly for long messages, and perceived it as less effortful. As anticipated in *Phone*, the majority (*15/16*) were used to typing as a habit, finding it easier to edit and generally more accurate, although "typing out long sentences can be a little tedious (P14)". One participant occasionally used voice dictation while eating, while another frequently used it because "voice is easier and faster [than typing] (P3)". In summary, the holdable nature of *Phone* and its superior accuracy facilitated typing, but it reduced input efficiency during multitasking, especially for lengthy messages.

5.7.3 GlassMessaging *Limitations.* Despite the potential benefits of *GlassMessaging* in multitasking scenarios, further enhancements are required to address its limitations.

- **Tradeoff: Efficiency vs. Accuracy.** While hands-free interaction and voice input improved texting speed and decreased response time, lower texting accuracy was a concern for participants, especially for long messages; "sending a wrong message can be worse than not sending anything (P9)". The accuracy worsened in unsuitable environments (e.g., speaking while chewing food) or with accents different from native speakers. In contrast, *Phone* has obvious advantages in accuracy, as auto-corrections and selective editing simplify error correction. "I can edit the problematic word [by typing on *Phone*] instead of re-speaking the whole sentence [on *Glass*]. But the typing act [on mid-air virtual keyboard] is obviously a no-go for the Glasses. (P4)" Yet, some participants still preferred the higher efficiency, tolerating the slight tradeoff in accuracy; "I sacrificed a little bit of accuracy to achieve efficiency. (P2)"
- **Tradeoff: Efficiency vs. Cognitive load.** Participants reported increased cognitive load when using *Glass* due to learning new technology, handling input errors, and overcoming technical obstacles. Learning

GlassMessaging involved getting used to mid-air gestures, remembering voice commands, and adapting to OHMDs. Participants also had to deal with input errors such as repeating voice commands to correct text and accidental mid-air gesture triggers, "I looked down at my food, and the glass would misclassify my eating hand as some gesture (P1)". However, participants found it easier to use *GlassMessaging* with practice.

- **Technical challenges of current OHMD development.** Participants found that content stabilization on the OHMD caused shaking during mobility tasks like walking or climbing stairs, making text reading difficult. This could be mitigated by combining multiple content stabilization methods based on mobility and primary task (e.g., body-locking while walking, head-locking while stationary). Furthermore, the accuracy of gesture detection decreased in mobile scenarios, and participants occasionally had to repeat voice commands, indicating the need for personalized customization of motion and dictation parameters (e.g., silence gap, dictation duration). As expected, the weight and size of the OHMD caused discomfort and obstructed movements, and external brightness reduced the interface's readability, "device takes up a lot of space on my head, and I need to tilt my head up more while drinking (P13)". These issues negatively impacted the evaluation measures of *RTLX* and *Preference* for *GlassMessaging*.

6 DISCUSSION

We identified mobile texting issues during daily activities (*Study 1*) and developed an OHMD app, *GlassMessaging*, to address some of them (*Study 2*). Comparing *GlassMessaging* to mobile phone texting showed how it could mitigate hands- and eyes-busy problems that occur during multitasking (*Study 3*) and proved its viability in everyday situations.

6.1 Is OHMD Better than Phone for *Messaging*?

With that said, we do not claim that our solution can completely replace current platforms like phones and smartwatches for *multitasking with messaging*. The phone is widely accepted, provides higher accuracy, and is applicable in almost all scenarios, while the potential of *GlassMessaging* mainly relies on its increased access to voice input and seamless hand interaction, which may not be suitable in all scenarios. Therefore, we view OHMDs as a promising complementary solution to the phone for multitasking with messaging, enabling heads-up usage [84] to reduce the "smartphone zombie" phenomenon [2].

6.2 Why is Voice Input More Natural on OHMDs than on Mobile Phones?

Our study shows that users prefer voice input with OHMDs but use virtual keyboards on phones. This is due to inconvenient mid-air keyboards on OHMDs [1, 80] and the affordances unique to each platform [54, 75, 86, 87] (see Table 4).

Table 4. A Summary of different affordances supported by OHMDs and Phones [44, 54, 75, 86, 87]. Note: This list is not exhaustive.

Features	Display	Access	Viewing	Wearing	Hand usage	Content privacy
OHMD	See-through, Near-eye	Always-on, On-demand	Heads-up	Wearable	Hands-free	Only seen by the wearer
Phone	Opaque, Distant	On-demand	Heads-down	Holdable	Hands-busy	Can be seen by bystanders

6.2.1 Wearable Nature of OHMD. The wearable nature of OHMDs may explain this preference for voice input. Its physical proximity allows effortless capture of verbal commands anytime, anywhere. The see-through near-eye display permits message viewing without impeding hands or platform access. In contrast, commonly used phone texting apps often require extra steps for voice commands, and their distant displays necessitate closer interaction, limiting their voice input support. Hence, unlike mobile devices, the proximity of OHMDs to the head as wearable devices naturally favors voice input.

6.2.2 Keyboard Use is Natural on Mobile Phones. Conversely, users favor phone keyboards, particularly when stationary. Touchscreen keyboards are intuitive, accurate, and provide reliable tactile feedback, an aspect that *GlassMessaging*'s virtual keyboard currently lacks due to OS limitations. Furthermore, keyboard usage also offers more privacy as voicing messages is not required.

6.2.3 Input Modality Switching on OHMDs. Multitasking restrictions can lead to situational disabilities and impairments [78]. When specific interactions are limited, switching to an alternative modality supported by the platform can overcome these limitations. For example, phone typing becomes impractical during hand-engaging tasks. Shifting to hands-free interactions can help, but mobile apps often require additional hand-based steps (e.g., opening the chat, pressing a button), limiting phone voice usage. In contrast, OHMDs enable seamless input modality switching to overcome such situational disabilities (sec 5.7.1).

6.2.4 Higher Perceived Trust of OHMDs. Some participants reported higher trust in OHMDs, enhancing their comfort with voicing messages. This trust stems from the device's exclusivity, ensuring that only the user views the content. The wearable nature of OHMDs also makes them less shareable compared to phones, enhancing perceived trustworthiness.

6.3 Comparison with Google Glass XE

Google Glass XE (2013-2017) [29, 30, 77], a discontinued product, supported heads-up messaging. Here, we distinguish between our application and Google Glass XE, showcasing our contributions from both practical and academic perspectives.

6.3.1 Google Glass XE (GG) Interface. GG incorporated a default set of voice action commands for messaging (Appendix D-Figure 15) [30]. Its lightweight and seamless design combined voice, head gestures, and touch gestures for inputs and an OHMD for output. To activate voice commands or send messages, users would utter "OK Glass" and "Send a message to", followed by the contact's name and message content. Users would respond to a message by saying "Reply" followed by their message content. Hence, *GG* provided an efficient method for sending and replying to **individual** messages.

6.3.2 Comparison. Table 5 compares *GlassMessaging* (*GM*) and *GG*, illustrating that both employ voice input for text entry and navigation. Our evaluation shows voice input as an efficient messaging method on OHMDs, indicated by quicker text entry and reduced response times, validating *GG*'s design choice. Speech recognition challenges do affect *GM*'s accuracy, which *GG* users likely also experienced[18]. Additionally, *GG* likely faced inconveniences editing messages, requiring input method switches for text correction, such as canceling incorrect messages using touchpad swipes and re-dictating the correct message. Below, we delve into their differences in messaging task handling, display mechanism, and interaction design.

One-off Messaging vs. Context-maintained Messaging. GG's voice commands catered to users' immediate messaging needs but struggled to facilitate multiple interleaved messages or reply to older messages. Hence, while it met basic messaging needs, *GG* fell short in handling complex messaging [6, 13, 31, 32] scenarios. Effective

---

[18]*GG* was using speech recognition technologies more than seven years ago.

Table 5. A comparison between *GlassMessaging* and Google Glass XE. LoS stands for Line of Sight, and FoV represents Field of View. Note: This list is not exhaustive and based on public online resources [29, 57, 77], as Google Glass XE has been discontinued since 2017.

Features	*GlassMessaging* with HL2	Google Glass XE (*GG*)
Interactions	Voice, Ring-mouse, Mid-air	Voice, Touchpad (on the right temple), Head gestures
Text entry	Voice, Mid-air keyboard	Voice
Display	Binocular, Higher-Resolution (2048x1080 px per eye) Larger-FoV (30° horizontal)	Monocular, Lower-resolution (640x360 px, right eye) Smaller-FoV (13° horizontal)
Chat history	Shows last 3 messages and 3 contacts	Shows last message and last contact
Chat position Notification position Contact position	LoS (middle-center) Above LoS (top-center) Right of LoS (middle-right)	Above LoS (top-center), Manual switching between each UI
UI opacity	Increased for new messages	Fixed
UI access	On-demand (using voice or ring-mouse)	On-demand (by looking up or using voice)

communication often necessitates retaining contextual information, such as message sequence, frequency, and priority [15, 58]. Modern messaging apps have more comprehensive features (e.g., full chat history, easy contact switching, unread indicators, etc.), a standard that *GM* follows as well.

Line-of-sight vs. Above Line-of-sight. The message display location is another key difference between *GG* and *GM*. *GG* positioned content above the line-of-sight (LoS) to maintain situational awareness, but this demanded user effort in switching away from the natural angle of forward vision. As higher-resolution displays with larger FoVs are now available, recent studies propose positioning some content within the LoS while leaving sufficient gaps to maintain environmental awareness [63, 85]. *GM* thus displays certain content within the LoS and uses gradual opacity adjustments to ensure situational awareness.

Interaction Support. In comparison to *GG*, *GM* offers a unique set of interaction possibilities for messaging while multitasking (sec 5.7.1-5.7.2). *GG* incorporated head-gesture functionality as a convenient way to initiate and respond to real-time messaging needs. Conversely, the integration of ring and mid-air interactions in *GM* enables users to alternate between input modalities based on situational constraints (sec 6.2.3).

In conclusion, both *GG* and *GM* facilitate heads-up messaging for daily usage, with different affordances and design emphases. *GG*, designed for an earlier generation of smart glasses with lower resolution and a smaller FoV, focused on convenient real-time one-off messaging needs. *GM* takes advantage of advanced OHMDs with higher resolution and larger FoVs and can handle immediate or delayed messaging needs involving multiple messages. Good synergies exist between the two designs, and combining them could offer a better messaging experience for users. For instance, *GM* could incorporate *GG*'s approach for handling one-off messaging needs by integrating head gestures for real-time messaging needs. Thus, the current *GM* design is more suitable for more complex messaging needs.

6.4 Can OHMD Improve People's Social Connections?

Unlike phones, the hands-free nature and wearable near-eye display of OHMDs permit timely message viewing. In our observational study, one participant (P10) checks messages promptly when eating alone, maintaining a connection with others. Yet, due to social etiquette, they tend not to do so when dining with others. Their message viewing and responding behaviors diverge: they view unread messages immediately but reply at their

discretion. Therefore, the wearable and hands-free capability of OHMDs may enhance users' social connections by facilitating swift message viewing.

6.5 Design Implications for *Messaging* on OHMDs

Besides the explored opportunities and constraints, we found more opportunities for *GlassMessaging* in the *multitasking with messaging* area, including text editing and multimedia support.

6.5.1 Supporting Text Editing. Voice input presents a trade-off between speed and accuracy. However, adding more text editing functions to user dictations could improve accuracy. For example, by incorporating *EYEditor* [28], a text editing system designed for OHMDs in on-the-go scenarios, *GlassMessaging* could achieve a similar accuracy level as phone messaging apps. Alternatively, AI capabilities, such as large language models, could predict text, correct writing errors, and enhance content [82].

6.5.2 Supporting Input Modalities Switching. Although *GlassMessaging* shows promise through easier voice input, it is not always suitable. Moreover, some input methods might lose precision due to situational impairments [78], like misrecognized hand gestures. In these cases, users should transition smoothly to another input mode. In public settings, users could switch from voice to socially acceptable alternatives such as ring interaction or finger taps (e.g., [83]).

6.5.3 Identifying Optimal Output Methods. *GlassMessaging* mainly uses text for messages, supplemented with auditory feedback for notifications and user actions. Participants appreciated the visual display for facilitating quick error identification, faster reading, and more efficient chat history scanning than temporal voice messages. Furthermore, two participants stated that the near-eye display enhances content privacy compared to phones, which bystanders can view. However, availability concerns arose for high-attention and eyes-busy scenarios. An alternative could be voice output, for example, text-to-speech or retaining the original audio alongside the text version. This solution could also address the accuracy issue by enabling text and audio version comparison if text inaccuracies impact readability. Further studies are needed to determine the most effective output and information display methods.

7 LIMITATIONS AND FUTURE WORK

7.1 Social Acceptability

Our *study 3* was conducted in private settings, bringing into question the applicability of *GlassMessaging* in public and social contexts. Participants expressed worries about voice input's potential awkwardness in public, disturbance to bystanders, privacy issues, and unexpected input/triggers in noisy environments. Moreover, voice input may not be suitable in quiet zones like libraries. Additional studies are required to comprehend the convenience and affordances of voice input on *Glass* in social settings.

Multiple text entry methods, including non-voice-based text entry [34, 50, 80, 83], can address this concern. These methods could allow for input method transitioning. Non-intrusive, concealable gestures [41] or interaction techniques [62] could be utilized for non-voice input. Alternatively, using *GlassMessaging* to view messages could complement phones, paralleling current smartwatch practices.

7.2 Long Time Usage

We examined *messaging* with tech-savvy participants for short periods in limited realistic scenarios. The novelty of *GlassMessaging* and the discrepancy in daily *Glass* and *Phone* usage could have influenced our findings. Moreover, our findings may not extend to other groups, like older adults, and do not consider potential long-term effects. Thus, more research involving diverse OHMD prototypes, user populations, and longitudinal field studies is warranted.

8   CONCLUSION

Multitasking with messaging is a common real-life occurrence, yet current mobile applications and platforms offer inadequate support. In this study, we initially analyzed people's messaging behaviors through a survey and observation study. We identified two primary situational impairments caused by existing mobile platforms (hands-busy and eyes-busy), leading us to iteratively develop *GlassMessaging* (https://github.com/NUS-HCILab/GlassMessaging), a *messaging* application designed for OHMDs to mitigate these limitations. A comparison with a phone-based *messaging* application in a realistic multitasking situation verified that *GlassMessaging* allows more *messaging* opportunities, faster responses, and increased texting speed due to its hands-free, wearable nature, and coordinated multimodal input. Consequently, we discussed the affordances of both devices and potential design implications. We anticipate that *messaging* on OHMD will represent the next communication frontier and serve as a beneficial complement to mobile phones during *multitasking with messaging*, propelled by technological advancements.

REFERENCES

[1] Jiban Adhikary and Keith Vertanen. 2021. Text Entry in Virtual Environments using Speech and a Midair Keyboard. *IEEE Transactions on Visualization and Computer Graphics* 27, 5 (May 2021), 2648–2658.   https://doi.org/10.1109/TVCG.2021.3067776 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[2] Markus Appel, Nina Krisch, Jan-Philipp Stein, and Silvana Weber. 2019. Smartphone zombies! Pedestrians' distracted walking as a function of their fear of missing out. *Journal of Environmental Psychology* 63 (June 2019), 130–133.   https://doi.org/10.1016/j.jenvp.2019.04.003

[3] Ahmed Sabbir Arif and Wolfgang Stuerzlinger. 2009. Analysis of text entry performance metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. 100–105.   https://doi.org/10.1109/TIC-STH.2009.5444533

[4] Daniel Avrahami, Susan R. Fussell, and Scott E. Hudson. 2008. IM waiting: timing and responsiveness in semi-synchronous communication. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08*. ACM Press, San Diego, CA, USA, 285.   https://doi.org/10.1145/1460563.1460610

[5] Shannon K.T. Bailey, Bradford L. Schroeder, Daphne E. Whitmer, and Valerie K. Sims. 2016. Perceptions of Mobile Instant Messaging Apps Are Comparable to Texting for Young Adults in the United States. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60, 1 (Sept. 2016), 1235–1239.   https://doi.org/10.1177/1541931213601288

[6] Agathe Battestini, Vidya Setlur, and Timothy Sohn. 2010. A large scale study of text-messaging use. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 229–238.   https://doi.org/10.1145/1851600.1851638

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

[8] Andreas Bulling, Raimund Dachselt, Andrew Duchowski, Robert Jacob, Sophie Stellmach, and Veronica Sundstedt. 2012. Gaze interaction in the post-WIMP world. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1221–1224.   https://doi.org/10.1145/2212776.2212428

[9] Runze Cai, Nuwan Nanayakkarawasam Peru Kandage Janaka, Shengdong Zhao, and Minghui Sun. 2023. ParaGlassMenu: Towards Social-Friendly Subtle Interactions in Conversations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–21.   https://doi.org/10.1145/3544548.3581065

[10] Jeff K. Caird, Kate A. Johnston, Chelsea R. Willness, Mark Asbridge, and Piers Steel. 2014. A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention* 71 (Oct. 2014), 311–318. https://doi.org/10.1016/j.aap.2014.06.005

[11] Isha Chaturvedi, Farshid Hassani Bijarbooneh, Tristan Braud, and Pan Hui. 2019. Peripheral vision: a new killer app for smart glasses. In *Proceedings of the 24th International Conference on Intelligent User Interfaces - IUI '19*. ACM Press, Marina del Ray, California, 625–636. https://doi.org/10.1145/3301275.3302263

[12] Xiao Chen, Wanli Chen, Kui Liu, Chunyang Chen, and Li Li. 2021. A comparative study of smartphone and smartwatch apps. In *Proceedings of the 36th Annual ACM Symposium on Applied Computing (SAC '21)*. Association for Computing Machinery, New York, NY, USA, 1484–1493. https://doi.org/10.1145/3412841.3442023

[13] Hyunsung Cho, Jinyoung Oh, Juho Kim, and Sung-Ju Lee. 2020. I Share, You Care: Private Status Sharing and Sender-Controlled Notifications in Mobile Instant Messaging. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 1–25. https://doi.org/10.1145/3392839

[14] Soon Hau Chua, Simon T. Perrault, Denys J. C. Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI - ChineseCHI2016*. ACM Press, San Jose, USA, 1–6. https://doi.org/10.1145/2948708.2948713

[15] Karen Church and Rodrigo de Oliveira. 2013. What's up with whatsapp?: comparing mobile instant messaging behaviors with traditional SMS. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services - MobileHCI '13*. ACM Press, Munich, Germany, 352. https://doi.org/10.1145/2493190.2493225

[16] P R Cohen and S L Oviatt. 1995. The role of voice input for human-machine communication. *Proceedings of the National Academy of Sciences* 92, 22 (Oct. 1995), 9921–9927. https://doi.org/10.1073/pnas.92.22.9921 Publisher: Proceedings of the National Academy of Sciences.

[17] Anna L. Cox, Paul A. Cairns, Alison Walton, and Sasha Lee. 2008. Tlk or txt? Using voice input for SMS composition. *Personal and Ubiquitous Computing* 12, 8 (Nov. 2008), 567–588. https://doi.org/10.1007/s00779-007-0178-8

[18] cre8ivepark. 2022. Text in Unity - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unity/text-in-unity

[19] cre8ivepark. 2022. Typography - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/typography

[20] David Curry. 2022. Most Popular Apps (2022). https://www.businessofapps.com/data/most-popular-apps/

[21] Saverio Debernardis, Michele Fiorentino, Michele Gattullo, Giuseppe Monno, and Antonio Emmanuele Uva. 2014. Text Readability in Head-Worn Displays: Color and Style Optimization in Video versus Optical See-Through Devices. *IEEE Transactions on Visualization and Computer Graphics* 20, 1 (Jan. 2014), 125–139. https://doi.org/10.1109/TVCG.2013.86

[22] Aaron M. Dollar. 2014. Classifying Human Hand Use and the Activities of Daily Living. In *The Human Hand as an Inspiration for Robot Hand Development*, Ravi Balasubramanian and Veronica J. Santos (Eds.). Springer International Publishing, Cham, 201–216. https://doi.org/10.1007/978-3-319-03017-3_10

[23] Tafadzwa Joseph Dube and Ahmed Sabbir Arif. 2019. Text Entry in Virtual Reality: A Comprehensive Review of the Literature. In *Human-Computer Interaction. Recognition and Interaction Technologies (Lecture Notes in Computer Science)*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 419–437. https://doi.org/10.1007/978-3-030-22643-5_33

[24] Díaz-Oreiro, López, Quesada, and Guerrero. 2019. Standardized Questionnaires for User Experience Evaluation: A Systematic Literature Review. *Proceedings* 31, 1 (Nov. 2019), 14. https://doi.org/10.3390/proceedings2019031014

[25] Peter F. Edemekong, Deb L. Bomgaars, Sukesh Sukumaran, and Shoshana B. Levy. 2022. Activities of Daily Living. In *StatPearls*. StatPearls Publishing, Treasure Island (FL). http://www.ncbi.nlm.nih.gov/books/NBK470404/

[26] Eric Enge. 2022. Mobile Voice Usage Trends in 2020 / Perficient, Inc. https://www.perficient.com/insights/research-hub/voice-usage-trends

[27] erickjpaul. 2022. Comfort - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/comfort

[28] Debjyoti Ghosh, Pin Sym Foong, Shengdong Zhao, Can Liu, Nuwan Janaka, and Vinitha Erusu. 2020. EYEditor: Towards On-the-Go Heads-Up Text Editing Using Voice and Manual Input. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376173

[29] Google Glass. 2013. Google Glass - YouTube. https://www.youtube.com/user/googleglass

[30] Google Glass. 2013. Google Glass: How to use voice actions. https://www.youtube.com/watch?v=rv3KU0Yo5ZM

[31] Rebecca Grinter and Margery Eldridge. 2003. Wan2tlk? everyday text messaging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. Association for Computing Machinery, New York, NY, USA, 441–448. https://doi.org/10.1145/642611.642688

[32] Rebecca E. Grinter and Leysia Palen. 2002. Instant messaging in teen life. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work (CSCW '02)*. Association for Computing Machinery, New York, NY, USA, 21–30. https://doi.org/10.1145/587078.587082

[33] Uwe Gruenefeld, Tim Claudius Stratmann, Jinki Jung, Hyeopwoo Lee, Jeehye Choi, Abhilasha Nanda, and Wilko Heuten. 2018. Guiding Smombies: Augmenting Peripheral Vision with Low-Cost Glasses to Shift the Attention of Smartphone Users. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Munich, Germany, 127–131. https://doi.org/10.1109/ISMAR-Adjunct.2018.00050

[34] Yizheng Gu, Chun Yu, Zhipeng Li, Zhaoheng Li, Xiaoying Wei, and Yuanchun Shi. 2020. QwertyRing: Text Entry on Physical Surfaces Using a Ring. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (Dec. 2020), 128:1–128:29. https://doi.org/10.1145/3432204

[35] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[36] Rami Hashish, Megan E. Toney-Bolger, Sarah S. Sharpe, Benjamin D. Lester, and Adam Mulliken. 2017. Texting during stair negotiation and implications for fall risk. *Gait & Posture* 58 (Oct. 2017), 409–414. https://doi.org/10.1016/j.gaitpost.2017.09.004

[37] Jibo He, Alex Chaparro, Bobby Nguyen, Rondell Burge, Joseph Crandall, Barbara Chaparro, Rui Ni, and Shi Cao. 2013. Texting while driving: is speech-based texting less risky than handheld texting?. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '13)*. Association for Computing Machinery, New York, NY, USA, 124–130. https://doi.org/10.1145/2516540.2516560

[38] Jibo He, William Choi, Jason S. McCarley, Barbara S. Chaparro, and Chun Wang. 2015. Texting while driving using Google Glass: Promising but not distraction-free. *Accident Analysis & Prevention* 81 (Aug. 2015), 218–229. https://doi.org/10.1016/j.aap.2015.03.033

[39] Jibo He, Jason S. McCarley, Kirsten Crager, Murtuza Jadliwala, Lesheng Hua, and Sheng Huang. 2018. Does wearable device bring distraction closer to drivers? Comparing smartphones and Google Glass. *Applied Ergonomics* 70 (July 2018), 156–166. https://doi.org/10.1016/j.apergo.2018.02.022

[40] Andreas Hinderks, Martin Schrepp, Francisco José Domínguez Mayo, María José Escalona, and Jörg Thomaschewski. 2019. Developing a UX KPI based on the user experience questionnaire. *Computer Standards & Interfaces* 65 (July 2019), 38–44. https://doi.org/10.1016/j.csi.2019.01.007

[41] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, Santa Clara, California, USA, 4203–4215. https://doi.org/10.1145/2858036.2858436

[42] Yuta Itoh, Tobias Langlotz, Jonathan Sutton, and Alexander Plopski. 2021. Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-mounted Displays. *Comput. Surveys* 54, 6 (July 2021), 120:1–120:36. https://doi.org/10.1145/3453157

[43] Nuwan Janaka, Chloe Haigh, Hyeongcheol Kim, Shan Zhang, and Shengdong Zhao. 2022. Paracentral and near-peripheral visualizations: Towards attention-maintaining secondary information presentation on OHMDs during in-person social interactions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3491102.3502127

[44] Nuwan Janaka, Xinke Wu, Shan Zhang, Shengdong Zhao, and Petr Slovak. 2022. Visual Behaviors and Mobile Information Acquisition. https://doi.org/10.48550/arXiv.2202.02748

[45] Sangeun Jin, Minsung Kim, Jihyeon Park, Minsung Jang, Kyuseok Chang, and Daemin Kim. 2019. A comparison of biomechanical workload between smartphone and smartwatch while sitting and standing. *Applied Ergonomics* 76 (April 2019), 105–112. https://doi.org/10.1016/j.apergo.2018.11.009

[46] Eleanor Jones, Jason Alexander, Andreas Andreou, Pourang Irani, and Sriram Subramanian. 2010. GesText: accelerometer-based gestural text-entry systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2173–2182. https://doi.org/10.1145/1753326.1753655

[47] Elisa Maria Klose, Nils Adrian Mack, Jens Hegenberg, and Ludger Schmidt. 2019. Text Presentation for Augmented Reality Applications in Dual-Task Situations. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Osaka, Japan, 636–644. https://doi.org/10.1109/VR.2019.8797992

[48] Barry Kollee, Sven Kratz, and Anthony Dunnigan. 2014. Exploring gestural interaction in smart spaces using head mounted devices with ego-centric sensing. In *Proceedings of the 2nd ACM symposium on Spatial user interaction (SUI '14)*. Association for Computing Machinery, New York, NY, USA, 40–49. https://doi.org/10.1145/2659766.2659781

[49] Tuomo Kujala and Hilkka Grahn. 2017. Visual Distraction Effects of In-Car Text Entry Methods: Comparing Keyboard, Handwriting and Voice Recognition. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3122986.3122987

[50] Lik-Hang Lee and Pan Hui. 2018. Interaction Methods for Smart Glasses: A Survey. *IEEE Access* 6 (2018), 28712–28732. https://doi.org/10.1109/ACCESS.2018.2831081

[51] Feiyu Lu, Shakiba Davari, Lee Lisle, Yuan Li, and Doug A. Bowman. 2020. Glanceable AR: Evaluating Information Access Methods for Head-Worn Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 930–939. https://doi.org/10.1109/VR46266.2020.00113 ISSN: 2642-5254.

[52] Andr\'{e}s Lucero and Akos Vetek. 2014. NotifEye: using interactive glasses to deal with notifications while walking in public. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology - ACE '14*. ACM Press, Funchal, Portugal, 1–10. https://doi.org/10.1145/2663806.2663824

[53] Fiona Lyddy, Francesca Farina, James Hanney, Lynn Farrell, and Niamh Kelly O'Neill. 2014. An Analysis of Language in University Students' Text Messages. *Journal of Computer-Mediated Communication* 19, 3 (April 2014), 546–561. https://doi.org/10.1111/jcc4.12045

[54] Denys J. C. Matthies, Marian Haescher, Rebekka Alm, and Bodo Urban. 2015. Properties of a Peripheral Head-Mounted Display (PHMD). In *HCI International 2015 - Posters' Extended Abstracts*, Constantine Stephanidis (Ed.). Vol. 528. Springer International Publishing, Cham, 208–213. https://doi.org/10.1007/978-3-319-21380-4_37

[55] Dana Monzer, Assem Abou Ali, Maya Abou-Zeid, and Nadine Marie Moacdieh. 2022. Voice messaging while driving: Effects on driving performance and attention. *Applied Ergonomics* 101 (May 2022), 11. https://doi.org/10.1016/j.apergo.2022.103692

[56] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. Association for Computing Machinery, New York, NY, USA, 152–158. https://doi.org/10.1145/191666.191729

[57] Niora. 2023. Google Glass 3.0 - Review. https://www.niora.net/en/p/google_glass_3

[58] Kenton P. O'Hara, Michael Massimi, Richard Harper, Simon Rubens, and Jessica Morris. 2014. Everyday dwelling with WhatsApp. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 1131–1143. https://doi.org/10.1145/2531602.2531679

[59] Jason Orlosky, Kiyoshi Kiyokawa, and Haruo Takemura. 2014. Managing mobile text in head mounted displays: studies on visual preference and text placement. *ACM SIGMOBILE Mobile Computing and Communications Review* 18, 2 (June 2014), 20–31. https://doi.org/10.1145/2636242.2636246

[60] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1 (Jan. 2012), 105–114. https://doi.org/10.1007/s00779-011-0412-2

[61] Justin M. Owens, Shane B. McLaughlin, and Jeremy Sudweeks. 2011. Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention* 43, 3 (May 2011), 939–947. https://doi.org/10.1016/j.aap.2010.11.019

[62] Henning Pohl, Andreea Muresan, and Kasper Hornbæk. 2019. Charting Subtle Interaction in the HCI Literature. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland Uk, 1–15. https://doi.org/10.1145/3290605.3300648

[63] Ashwin Ram and Shengdong Zhao. 2021. LSVP: Towards Effective On-the-go Video Learning Using Optical Head-Mounted Displays. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (March 2021), 30:1–30:27. https://doi.org/10.1145/3448118

[64] Melanie Revilla, Mick P. Couper, Oriol J. Bosch, and Marc Asensio. 2020. Testing the Use of Voice Input in a Smartphone Web Survey. *Social Science Computer Review* 38, 2 (April 2020), 207–224. https://doi.org/10.1177/0894439318810715 Publisher: SAGE Publications Inc.

[65] Mikko J. Rissanen, Samantha Vu, Owen Noel Fernando, Natalie Pang, and Schubert Foo. 2013. Subtle, Natural and Socially Acceptable Interaction Techniques for Ringterfaces – Finger-Ring Shaped User Interfaces. In *Proceedings of the First International Conference on Distributed, Ambient, and Pervasive Interactions - Volume 8028*. Springer-Verlag, Berlin, Heidelberg, 52–61. https://doi.org/10.1007/978-3-642-39351-8_6

[66] Sherry Ruan, Jacob O. Wobbrock, Kenny Liou, Andrew Ng, and James A. Landay. 2018. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (Jan. 2018), 159:1–159:23. https://doi.org/10.1145/3161187

[67] Shardul Sapkota, Ashwin Ram, and Shengdong Zhao. 2021. Ubiquitous Interactions for Heads-Up Computing: Understanding Users' Preferences for Subtle Interaction Techniques in Everyday Settings. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*. ACM, Toulouse & Virtual France, 1–15. https://doi.org/10.1145/3447526.3472035

[68] Ben D. Sawyer, Victor S. Finomore, Andres A. Calvo, and P. A. Hancock. 2014. Google Glass: A Driver Distraction Cause or Cure? *Human Factors* 56, 7 (Nov. 2014), 1307–1321. https://doi.org/10.1177/0018720814555723 Number: 7 52 citations (Crossref) [2022-06-09] Publisher: SAGE Publications Inc.

[69] Seiler Sj. 2015. Hand on the wheel, mind on the mobile: an analysis of social factors contributing to texting while driving. *Cyberpsychology, behavior and social networking* 18, 2 (Feb. 2015). https://doi.org/10.1089/cyber.2014.0535

[70] R. William Soukoreff and I. Scott MacKenzie. 2003. Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. Association for Computing Machinery, New York, NY, USA, 113–120. https://doi.org/10.1145/642611.642632

[71] Mark J. M. Sullman, Aneta M. Przepiorka, Agata P. Błachnio, and Tetiana Hill. 2021. Can't text, I'm driving – Factors influencing intentions to text while driving in the UK. *Accident Analysis & Prevention* 153 (April 2021), 106027. https://doi.org/10.1016/j.aap.2021.106027

[72] Ying Tang and Khe Foon Hew. 2022. Effects of using mobile instant messaging on student behavioral, emotional, and cognitive engagement: a quasi-experimental study. *International Journal of Educational Technology in Higher Education* 19, 1 (2022), 3. https://doi.org/10.1186/s41239-021-00306-6

[73] technavio. 2022. Smart Glass Market by Application and Geography - Forecast and Analysis 2022-2026. https://www.technavio.com/report/global-smart-glasses-market-industry-analysis

[74] Crispin Thurlow and A Brown. 2003. Generation Txt? The sociolinguistics of young people's text-messaging. *Discourse analysis online* 1 (Jan. 2003), 30.

[75] Kristin Vadas, Kent Lyons, Daniel Ashbrook, Ji Soo Yi, Thad Starner, and Julie Jacko. 2006. Reading on the Go: An Evaluation of Three Mobile Display Technologies. (2006), 16.

[76] Keith Vertanen and Per Ola Kristensson. 2011. A versatile dataset for text entry evaluations based on genuine mobile emails. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*. ACM Press, Stockholm, Sweden, 295. https://doi.org/10.1145/2037373.2037418

[77] Wikipedia. 2023. Google Glass. https://en.wikipedia.org/w/index.php?title=Google_Glass&oldid=1139958030 Page Version ID: 1139958030.

[78] Jacob O. Wobbrock. 2019. Situationally-Induced Impairments and Disabilities. In *Web Accessibility*, Yeliz Yesilada and Simon Harper (Eds.). Springer London, London, 59–92. https://doi.org/10.1007/978-1-4471-7440-0_5

[79] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[80] Zhican Yang, Chun Yu, Xin Yi, and Yuanchun Shi. 2019. Investigating Gesture Typing for Indirect Touch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (Sept. 2019), 117:1–117:22. https://doi.org/10.1145/3351275

[81] Shanhe Yi, Zhengrui Qin, Ed Novak, Yafeng Yin, and Qun Li. 2016. GlassGesture: Exploring head gesture interface of smart glasses. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*. 1–9. https://doi.org/10.1109/INFOCOM.2016.7524542

[82] Mingrui Ray Zhang, He Wen, and Jacob O. Wobbrock. 2019. Type, Then Correct: Intelligent Text Correction Techniques for Mobile Text Entry Using Neural Networks. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 843–855. https://doi.org/10.1145/3332165.3347924

[83] Mingrui Ray Zhang, Shumin Zhai, and Jacob O. Wobbrock. 2022. TypeAnywhere: A QWERTY-Based Text Entry Solution for Ubiquitous Computing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3491102.3517686

[84] Shengdong Zhao, Felicia Tan, and Katherine Fennedy. 2023. Heads-Up Computing: Moving Beyond the Device-Centered Paradigm. *arXiv:2202.02748 [cs]* (2023), 11 pages. arXiv:2305.05292 https://doi.org/10.48550/arXiv.2305.05292

[85] 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 (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3544548.3581430

[86] Qiushi Zhou, Joshua Newn, Benjamin Tag, Hao-Ping Lee, Chaofan Wang, and Eduardo Velloso. 2019. Ubiquitous smart eyewear interactions using implicit sensing and unobtrusive information output. In *Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers - UbiComp/ISWC '19*. ACM Press, London, United Kingdom, 661–666. https://doi.org/10.1145/3341162.3348392

[87] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376233

A STUDY 1

Figure 10 shows the participants' replying behavior, and Figure 11 depicts the themes and codes that emerged from the thematic analysis in *study 1*.

B *GLASSMESSAGING*

B.1 *GlassMessaging* Input Interactions

Table 6 shows the supported input interactions of *GlassMessaging*.

C STUDY 3

C.1 Primary Task

Figure 12 indicate the primary task details.

C.2 Analysis

Factorial repeated measures ANOVAs or factorial repeated measures ANOVAs after Aligned Rank Transform (ART) [79], in cases of violation in ANOVA assumptions, were applied to analyze the data. The Shapiro-Wilk
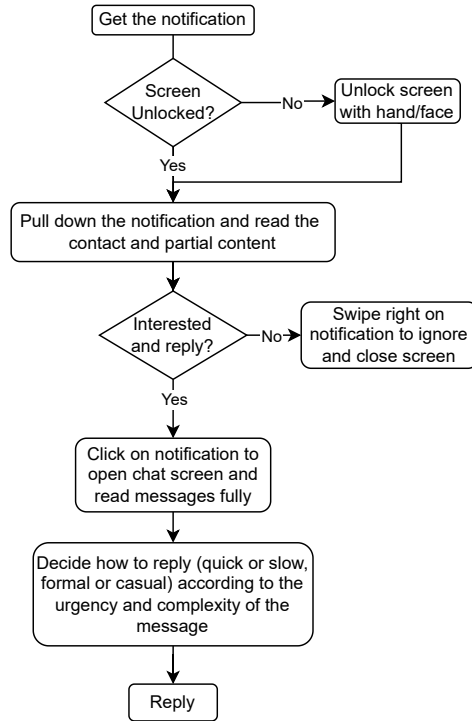
Fig. 10. Process of receiving and replying to a message on the phone observed for most participants in *study 1*. Note that the real situation may slightly differ from this flow due to phone configurations and user habits (e.g., reading notifications on a locked screen)

test and the Mauchly test were used to test normality and sphericity. Lastly, paired-sample t-test or Wilcoxon signed-rank test were used as post-hoc tests, and Bonferroni correction was applied for multiple comparisons. When non-parametric distributions take a large range of values (e.g., NASA-TLX, which ranged from 0-100) and follow parametric assumptions, parametric tests were used.

Transcribed interview recordings and open-ended answers in questionnaires were analyzed following the same procedure as in *study 1*, sec 3.2.3. Figure 14 depicts the themes and codes that emerged from the thematic analysis.

C.3 Quantitative Data

Table 7 indicates participants' mean performance ('mean (sd)') and Figure 13 shows the individual and overall *RTLX* scores.

D GOOGLE GLASS XE

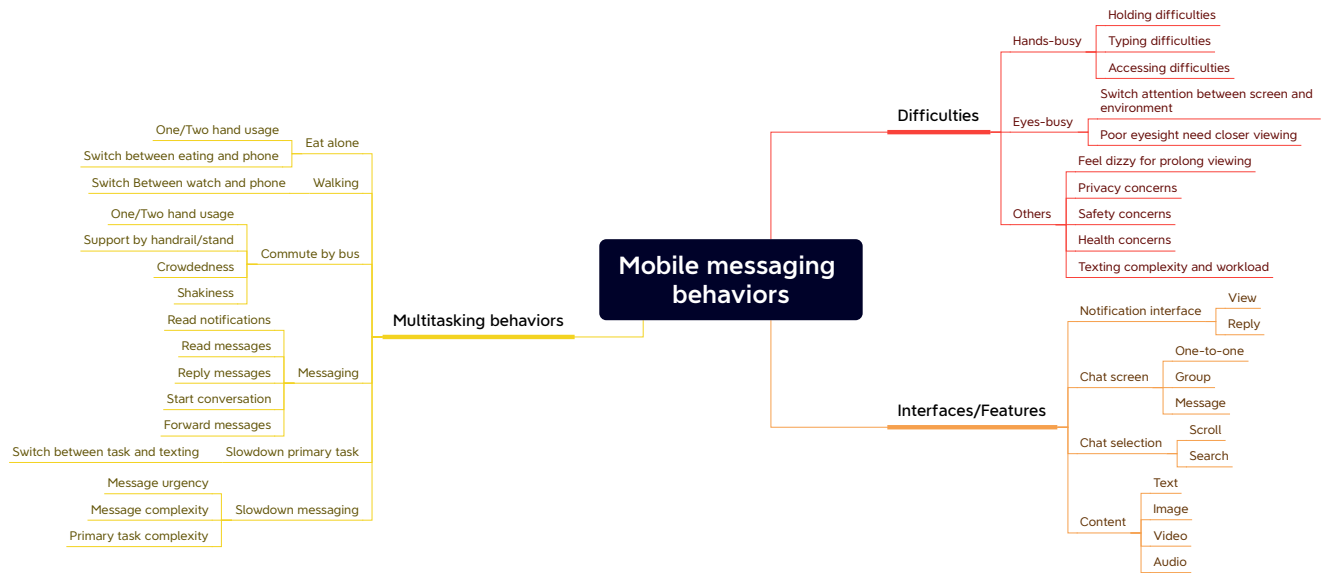Figure 15 shows the process of sending and replying to a message.

Fig. 11. Themes and codes emerged in *study 1*.

Table 6. Supported input interactions of *GlassMessaging* (Final Version)

| Function	Mid-air gesture	Ring interaction	Voice command
Reveal interface		Click 'center' button for 1 second	Show chat
Hide interface		Click 'center' button for 1 second	Hide chat
Open the chat related to notification	Press on notification		Open notification
Open the chat with the contact, <Name>	Press on contact <Name>	Click 'up'/'down' button to navigate	<Name>
Activate voice dictation	Press on 'voice' button	Click 'right' button to navigate + click 'center' button to activate	Voice message
Send the message	Press on 'send' button		Send
Open the virtual keyboard	Press on 'keyboard' button		Open keyboard
Close the virtual keyboard	Press anywhere on the interface	Click any button	Close keyboard
Go to the topmost contact	Scroll up using the finger	Click and hold the 'up' button	Scroll to the top
Go to closest top contact	Scroll up using the finger	Click 'up' button	Scroll up
Go to closest bottom contact	Scroll down using the finger	Click 'down' button	Scroll down
Start voice dictating to message received contact			Reply (Open notification + Voice message)
Start voice dictation to the contact, <Name>			Text <Name> (<Name>+ Voice message)
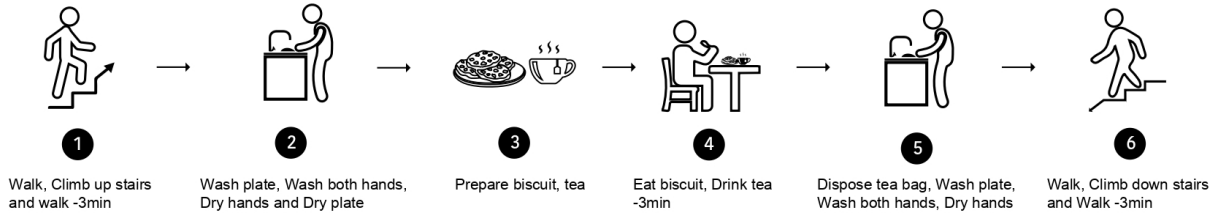
Fig. 12. Details of the primary task. The primary task required participants to walk back and forth for a total distance of around 43×2*m*. This included walking from the starting position to a circular stair (≈ 25*m*), climbing up a one-floor circular stair (≈ 4*m* height), walking (≈ 14m) to a kitchen (≈ 6×3m^2), performing washing and eating activities, and finally walking back to the starting position.

Table 7. Average performance (*'mean (sd)'*) in *study 3* with 16 participants. The first column represents the *Platform-Message Length* combination using the first letters of each (G = *Glass*, P = *Phone*; S = *Short*, L = *Long*).

|	Primary task	Secondary task					Overall
	Task Completion Increase	*Message Count*	*Texting Speed*	*Message Response Time*	*Texting Accuracy*	*Perceived Messaging Effectiveness*	*RTLX*
G-S	1.26 (0.14)	9.4 (1.4)	44.97 (11.37)	23.46 (9.94)	96.71 (3.05)	5.50 (1.27)	42.8 (16.9)
P-S	1.19 (0.13)	8.8 (1.4)	40.02 (12.14)	31.80 (9.23)	99.16 (0.96)	4.70 (1.61)	31.0 (17.2)
G-L	1.32 (0.21)	9.4 (1.4)	71.78 (9.83)	27.28 (11.44)	96.57 (2.20)	4.88 (1.50)	39.6 (18.9)
P-L	1.37 (0.24)	8.9 (1.7)	43.19 (10.43)	44.00 (13.48)	99.07 (0.75)	4.25 (1.81)	34.3 (20.8)
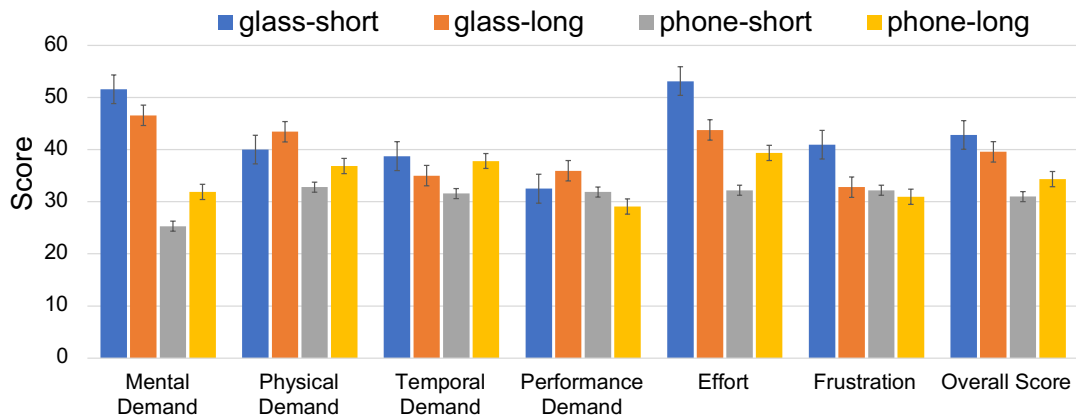
Fig. 13. NASA-TLX scores for the four conditions (*Glass-Short*, *Glass-Long*, *Phone-Short*, and *Phone-Long*) in *study 3* (N=16). Error bars represent standard errors. Note that the *Long* sessions were administered after *Short* sessions for both platforms.

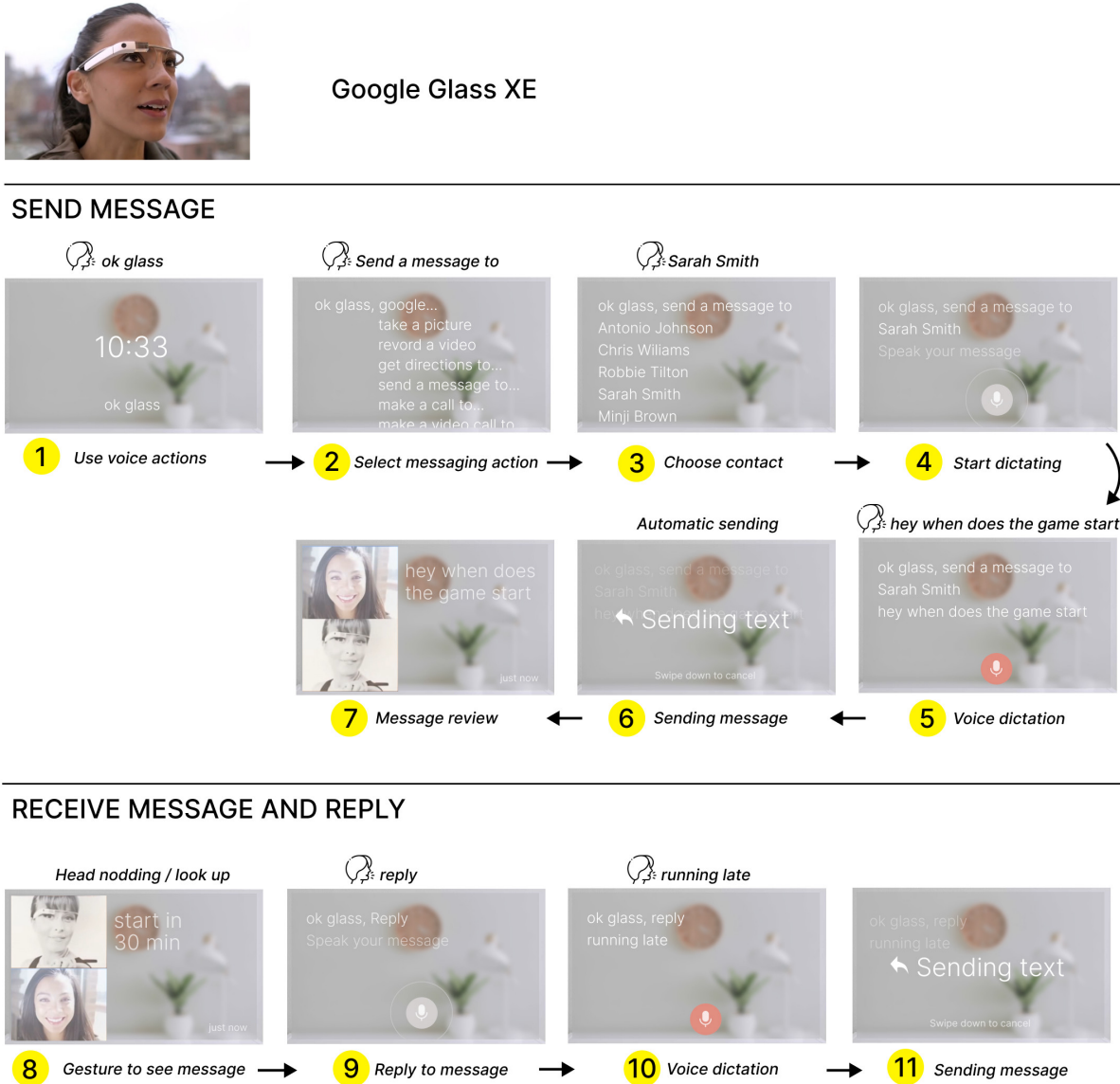Fig. 14. Themes and codes that emerged in *study 3*.

Fig. 15. Steps of sending a new message and Reply to a Message using Google Glass XE (Source: [30]). To **Send a New Message**, the user: (1) activates voice commands by saying 'OK GLASS'; (2) initiates the messaging process by saying 'SEND A MESSAGE TO'; (3) selects the recipient by saying their name (e.g., 'SARAH SMITH'); (4) starts voice dictation; (5) the system transcribes the user's utterances to text in real-time and stops transcription after a silence; (6) the system automatically sends the message unless the user explicitly cancels it using the touch interface; (7) the user sees the sent message as a notification while looking up, and it disappears when the user looks forward. To **Reply to a Message**, the user: (8) nods the head or looks up to see the received message; (9) initiates the replying process by saying 'REPLY'; (10) the system transcribes the user's utterances to text in real-time and stops transcription after a silence; (11) the system automatically sends the message unless the user explicitly cancels it using the touch interface.