
International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

# Touchless Interaction for Smart Glasses in Emergency Medical Services: User Needs and Experiences

**Zhan Zhang, Enze Bai, Aram Stepanian, Swathi Jagannath & Sun Young Park**

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2024.2328917



# Touchless Interaction for Smart Glasses in Emergency Medical Services: User Needs and Experiences

Zhan Zhang[a], Enze Bai[a], Aram Stepanian[a], Swathi Jagannath[b], and Sun Young Park[c,d]

[a]School of Computer Science and Information Systems, Pace University, New York City, NY, USA; [b]Microsoft, Redmond, WA, USA; [c]School of Information, University of Michigan, Ann Arbor, MI, USA; [d]School of Art and Design, University of Michigan, Ann Arbor, MI, USA

**ABSTRACT**

The application of smart glasses in healthcare, particularly for providers engaged in hands-on patient care tasks, presents unique design challenges. This study combines participatory design and usability testing to assess the user experiences of touchless interaction methods for smart glasses in the context of Emergency Medical Services (EMS). The participatory design workshops with 16 EMS providers reveal a preference for touchless interaction methods such as voice commands and pinch hand gestures, driven by the need to keep hands free and minimize cross-contamination risk. Despite this preference, the laboratory-based usability testing with 16 EMS providers indicates that both voice commands and hand gestures fall short in task performance compared to the default tangible buttons on smart glasses, primarily due to software limitations and EMS providers' unfamiliarity with touchless techniques. Our findings reveal specific issues associated with using different interaction methods when operating smart glasses. Building on these insights, we discuss design implications for smart glasses to better align with the dynamic and unique characteristics of fast-paced medical work.

**KEYWORDS**

Smart glasses; touchless interaction; usability; user experience; emergency medical services; healthcare

## 1. Introduction

The launch of Google Glasses in 2013 started a new wave of enthusiasm for smart glasses. This technology typically pertains to augmented reality (AR)-powered, head-mounted devices capable of displaying virtual information on near-eye displays (Zuidhof et al., 2021). Although they are still in the early stages of development, smart glasses have seen significant advancements over the past decade, particularly in computing power and battery life. These improvements have greatly enhanced the adaptability of smart glasses for various tasks and work settings, including fast-paced medical work such as Emergency Medical Services (EMS) or pre-hospital care (Schaer et al., 2016; Schlosser et al., 2021; Zhang, Joy, et al., 2022).

EMS represents a specialized medical domain where emergency care providers such as paramedics or emergency medical technicians (EMTs) deliver urgent medical assistance outside of hospital settings and transport critically ill or injured patients to the most appropriate point of definitive care. Several technologies, such as electronic health record (EHR) (Pilerot & Maurin Söderholm, 2019) and telemedicine systems in ambulances (Chapman Smith et al., 2019; Cho et al., 2015; Geisler et al., 2019; Yperzeele et al., 2014), have been developed to facilitate data collection, integration, and sharing in the field. However, due to the demanding nature of treating high-acuity patients, EMS providers often have limited capacity to use handheld computing devices in real time (Pilerot & Maurin Söderholm, 2019; Zhang, Joy, et al., 2021). It is therefore understandable that many technologies created for EMS providers have a low adoption rate, primarily because of their heavy reliance on manual input and control (Chapman Smith et al., 2019; Cho et al., 2015; Hertzum et al., 2019; Rogers et al., 2017). Therefore, there is a pressing need for innovative technologies and interaction methods that can better accommodate the hands-busy nature of EMS work. Smart glasses are considered a promising solution because of their potential to free up providers' hands from the use of computing devices (Schaer et al., 2016).

The dynamic nature of EMS work requires that smart glasses be user-friendly and support "touchless" operation to mitigate possible cross-contamination and interference with EMS providers' manual tasks (Schlosser et al., 2021; Zhang, Joy, et al., 2022). However, the majority of current smart glasses rely on tangible buttons or touchpads for interaction (Lee & Hui, 2018; Tung et al., 2015). Researchers in the Human-Computer Interaction (HCI) field thus have sought to develop various touchless interaction methods, such as voice commands, hand gestures, head movements, facial expressions, and gaze inputs (Aigner et al., 2012; Gjoreski et al., 2023; Ha et al., 2014; Jones et al., 2010; Lee & Hui, 2018; Toyama et al., 2014). Advanced sensing technology within smart glasses can process these input methods, enabling touchless interactions with the virtual content displayed. Despite this body of research, there has been

**CONTACT** Zhan Zhang zzhang@pace.edu School of Computer Science and Information Systems, Pace University, New York City, NY, USA

relatively limited research focused on examining two critical aspects: (1) EMS providers' inputs on the preferred touchless interaction methods viable for a fast-paced work environment, and (2) the user experience and usability of the preferred touchless interaction methods when using smart glasses to assist EMS work.

To address these research gaps, we adopted a user-centered design and evaluation approach, consisting of two studies. In the first study, we organized participatory design workshops to gather insights from EMS providers about their preferred touchless interaction methods for smart glasses. We found that voice commands and pinch hand gestures were their most preferred touchless interaction methods among various methods synthesized in prior work (Lee & Hui, 2018). Building on our workshop findings, we iteratively designed and developed a system prototype that incorporated providers' preferred interaction methods. We then conducted a follow-up usability study in a controlled setting to compare the user experiences with voice commands and pinch gestures against those with tangible buttons, which are the default interaction method on most commercial smart glasses. Our findings suggest that tangible buttons yielded the most efficient task performance, followed by voice commands and then pinch gestures. Many of the errors participants made with voice commands and pinch gestures could be attributed to software limitations and unfamiliarity with certain commands and hand postures. Despite the less-than-optimal performance, EMS providers still favored touchless interaction methods over tangible buttons, particularly given the unique challenges of their work, such as hands-on tasks and the risk of cross-contamination. Finally, we detail the EMS providers' perceived advantages and challenges of both touchless and tangible interaction methods. We conclude the paper by discussing design implications and research opportunities for making smart glasses easy-to-use in dynamic and time-critical medical domains.

Our contributions to the field of HCI are three-fold. First, given that there has been limited exploration into the interaction between medical providers and smart glass devices, our study contributes new insights into designing appropriate interaction methods of smart glasses for fast-paced medical settings where the use of handheld computing devices is often challenging. Second, while other hospital settings have embraced technological advancements, EMS practice still lacks adequate technological support (Gausche-Hill et al., 2021). Our work uncovers new technology opportunities that can greatly benefit this critical medical domain. Lastly, our study reveals the perspectives of EMS providers on the benefits and challenges of different smart glass interaction approaches, providing actionable insights for designers, researchers, and smart glass manufacturers to enhance the usability and adaptability of this technology in dynamic, hands-busy settings.

In the remaining sections of the paper, we present a literature review on smart glass usage and interaction (Section 2) and provide an outline of our study design (Section 3). The findings from our first study—participatory design workshops with EMS providers—are reported in Section 4. The results of our second study, which focused on the usability evaluation of touchless interaction methods, are detailed in Section 5. Finally, we discuss the implications of our findings (Section 6) and conclude the paper with key take-away messages (Section 7).

## 2. Related work

### *2.1. Applications of smart glasses in medical settings*

In recent years, smart glasses have been explored and used across various medical settings. A predominant application area for smart glasses is facilitating communication and care coordination among distributed care teams. For instance, they have been used for broadcasting surgeries to remote consultants (Weibel et al., 2020), facilitating the remote evaluation of critical patients (Broach et al., 2018; Cicero et al., 2015; Follmann et al., 2019; Noorian et al., 2019), and enabling virtual ward rounds during the COVID-19 pandemic (Munusamy et al., 2021). Beyond serving as a telemedicine tool, smart glasses have also been used for recording patient encounters (Aldaz et al., 2015; Klinker et al., 2020a; Odenheimer et al., 2018), enhancing awareness of patient statuses in critical care (Romare & Skär, 2020), and aiding decision-making processes (Faiola et al., 2019). These studies have demonstrated the usefulness of smart glasses in patient care, including their ability to integrate with existing medical devices (e.g., vital signs monitor) to offer a holistic view of the patient's condition (Chai et al., 2014; Mitrasinovic et al., 2015; Yu et al., 2016).

Of particular relevance to this study is the growing interest in utilizing smart glasses to support EMS work. For example, a few studies have evaluated the use of off-the-shelf smart glasses in assisting patient triage during mass casualty incidents in the field (Berndt et al., 2016; Broach et al., 2018; Cicero et al., 2015; Follmann et al., 2019; Schaer et al., 2016). These studies have underscored the advantages of employing smart glasses to improve prioritization of casualties and real-time sharing of visual medical information directly from the scene. Additionally, a few recent studies, including our own (Schlosser et al., 2021; Zhang, Joy, et al., 2022; Zhang, Ramiya Ramesh Babu, et al., 2022), have explored the perceptions and opinions of EMS providers regarding the adoption of smart glasses. These studies revealed several potential application areas for smart glasses in the EMS context, such as enhanced communication between distributed care providers, augmented decision-making, and hands-free data collection and retrieval.

Despite these efforts, a research gap persists—researchers have primarily focused on evaluating or exploring off-the-shelf smart glass devices in EMS work, with little attention paid to the interaction between EMS providers and the smart glasses themselves. Furthermore, limited work has taken a user-centered approach to investigate appropriate interaction methods for fast-paced medical settings, such as EMS. For example, in the study conducted by Zhang, Joy, et al. (2022), EMS providers' opinions about voice commands and hand gestures were sought, but their investigations were restricted to these two specific methods and did not evaluate user preferences across a range of interaction methods. Therefore, our study aims to bridge this research gap by systematically investigating favorable and feasible interaction methods for smart glasses from EMS providers' perspectives while considering a range of options that have been developed and explored in prior work (Lee & Hui, 2018). In the next section, we review existing work related to smart glass interaction methods.

### 2.2. Touchless interaction methods for smart glasses

Compared to smartphones, smart glasses have a more limited display size and viewing angle (Syberfeldt et al., 2017). These unique characteristics of smart glasses, combined with their default interaction methods like tangible button, can complicate the user interactions with the device (Grossman et al., 2015; Ok et al., 2015). Even more concerning is that tangible buttons require physical touch or pressing, which increases the risk of cross-contamination in medical settings. To address these limitations, HCI researchers have developed and evaluated various touchless interaction methods. In this section, we review several such methods that are particularly relevant to environments where hand control is either impractical or inconvenient [a comprehensive review of interaction methods for smart glasses can be found in Lee and Hui (2018)].

Voice recognition is perhaps the most well-known and established touchless interaction methods for smart glasses. This technology, widely applied in smartphones and conversational agents, allows computing devices to be controlled through a set of pre-programmed voice commands, which users must learn before operating the device. Although intuitive, voice commands can be problematic in noisy and dynamic settings (Lee & Hui, 2018). To mitigate this challenge, recent research has attempted to leverage advanced speech recognition and natural language processing (NLP) techniques to enhance voice-based interactions with smart glasses (Almutairi et al., 2020; Zhang, Luo, et al., 2022). For example, Almutairi et al. (2020) developed a smart glass application with integrated speech recognition to assist visually impaired individuals in navigating from one location to another.

Hand gestures represent another touchless input method for smart glasses that has been widely explored. This interaction method enables users to perform 2D-based tasks (e.g., selection or pointing) or 3D-based tasks in an AR environment (e.g., 3D object orientation and manipulation) (Aigner et al., 2012; Ha et al., 2014). Various hand gesture-based interactions have been developed to facilitate the use of smart glasses across various tasks (Lee & Hui, 2018). For example, Lee et al. (2019) introduced a freehand mid-air interaction technique that captures user hand movements (e.g., pointing). In a similar vein, another study by Lee et al. (2019) implemented a text input system for smart glasses that allows users to enter text using hand gestures in a keyboard-less interface.

In addition to voice-based and hand gesture-based controls, previous research has also explored other forms of touchless input. For example, by utilizing the built-in accelerometers and gyroscopes in smart glasses, head movement has been proposed as a method to control text input (Jones et al., 2010), user authentication (Yi et al., 2016), and picture-taking (Klinker et al., 2020a). However, due to ergonomic concerns, head-tilt gestures are rarely considered as a major input method for smart glasses, especially when users are doing excessive activities. For example, a study noted that certain head gestures (e.g., tilting) were difficult to apply in practice when medical providers engaged in care activities (Prilla et al., 2019). Another non-obtrusive interaction method for smart glasses is gaze input, which controls cursor movement on near-eye displays. For instance, Toyama et al. (2014) explored the utilization of gaze movement for tasks such as text reading and translation. While gaze-based interaction is compelling and involves minimal muscle movement, it has major drawbacks such as being error-prone and requiring excessive calibration (Bulling et al., 2012; Lee & Hui, 2018). Lastly, a few recent studies have explored the use of users' facial gestures and expressions (e.g., eye-wink, frown, smile, etc.) to realize touchless interaction with smart glasses or mobile devices (e.g., tablets) (Gjoreski et al., 2023; Goel et al., 2015; Matthies et al., 2021). This innovative interaction method uses sensors on smart glasses to monitor the user's face and leverages advanced artificial intelligence (AI) techniques to recognize facial expressions. While promising, calibration tailored to each wearer and user training remain essential. For example, users might need to memorize the correlation between facial gestures and specific actions, such as using smile to confirm a selection.

In summary, HCI researchers have proposed, implemented, and evaluated various touchless interaction approaches for smart glasses to overcome the limitations of tangible input controls. However, most of these approaches have been tailored for non-critical environments, such as office settings (Huang et al., 2015) and gaming (Tung et al., 2015). There is limited research exploring their potential in supporting time-critical tasks like those in EMS. To address this research gap, we conducted two consecutive studies to (1) *identify touchless interaction methods preferred by EMS providers, taking into account their unique work practices*, and (2) *compare the user experience of touchless interaction methods with that of tangible inputs*. Our findings provide valuable insights for designing intuitive smart glass interactions to meet the needs of dynamic and fast-paced medical teams.

## 3. Overview of the study design

The work presented in this paper are part of a larger research project that aims to design and develop user-friendly smart glass applications to support EMS work. As an initial phase of the larger project, our focus was on exploring suitable interaction methods for the use of smart glasses in the EMS context and evaluating the usability of those interaction methods. To achieve this goal, we conducted participatory design workshops and usability evaluations with EMS providers.

This research was approved by the first author's university Institutional Review Board (IRB).

The participatory design workshop served as a platform to gather insights into the needs of EMS providers regarding smart glass design, including their preferences for touchless interaction methods. These user inputs informed our iterative design and development of a prototype smart glass system. Following this, we conducted a comparative evaluation study with the smart glass prototype to assess the usability of three specific interaction methods: the two most favored touchless methods (voice commands and hand gestures) and the default interaction method (tangible buttons). Using a within-subject, repeated-measure design, participants engaged with each interaction method in a randomized order to complete an identical set of tasks. When presenting qualitative results in the subsequent sections, we adopt the notation PD# to refer to a design workshop participant and T# for a usability testing participant.

We recruited study participants from four EMS agencies, including a fire-based agency located in the rural area of the mountain region of the U.S., and three hospital-based EMS agencies located in an urban area in the east coast region of the U.S. The various characteristics of these participating teams (e.g., fire vs. hospital-based agency, urban vs. rural area) could help improve the generalizability of this research. The director at each EMS agency sent out a recruitment email to the whole team and instructed them to contact the researcher, if interested. The recruited participants have different occupations (e.g., paramedic vs. EMT) and varying experiences (e.g., less than a year vs. more than 40 years). The details of participant demographics are presented in Table 1. All participants were compensated for their participation with a $60/hour rate.

Table 1. Participant demographics.

| Study | ID | Gender | Years of experience | Occupation |
|---|---|---|---|---|
| PD Workshop | PD1 | Male | 42 years | Paramedic & Director |
| | PD2 | Male | 3 years | EMT |
| | PD3 | Male | 10 years | Paramedic |
| | PD4 | Male | 5 years | EMT |
| | PD5 | Female | 10 years | Paramedic |
| | PD6 | Male | 20 years | Paramedic |
| | PD7 | Male | 27 years | Paramedic |
| | PD8 | Male | 6 years | Paramedic |
| | PD9 | Female | 18 years | Paramedic |
| | PD10 | Female | 7 years | Paramedic |
| | PD11 | Male | 9 years | Paramedic |
| | PD12 | Male | 16 years | Paramedic |
| | PD13 | Male | 6 years | EMT |
| | PD14 | Male | 2 years | EMT |
| | PD15 | Male | 5 years | EMT |
| | PD16 | Male | 20 years | Paramedic |
| Usability Testing | T1 | Male | 5 years | EMT |
| | T2 | Male | 2.5 years | EMT |
| | T3 | Male | 40 years | Paramedic & Director |
| | T4 | Male | 13 years | Paramedic |
| | T5 | Female | 4 years | Paramedic |
| | T6 | Male | 4 years | Paramedic |
| | T7 | Male | 3 years | EMT |
| | T8 | Male | 6 years | EMT |
| | T9 | Female | 10 years | Paramedic |
| | T10 | Female | 16 years | Paramedic |
| | T11 | Male | 25 years | Paramedic |
| | T12 | Male | 6 years | EMT |
| | T13 | Male | 8 years | EMT |
| | T14 | Male | 25 years | Paramedic |
| | T15 | Male | 5 years | EMT |
| | T16 | Male | 1 year | EMT |

## 4. Study 1: Exploring EMS providers' preferred touchless interaction methods of smart glasses

### 4.1. Method

#### 4.1.1. Study procedure

We conducted four participatory design workshops with 16 EMS providers. Each workshop session consisted of four participants and lasted for up to two hours. The workshops focused on discussing the necessary system features for supporting EMS work and identifying the preferred touchless interaction methods among EMS providers. Participants were encouraged to share the challenges they encountered in their work, express their needs and suggestions for system features, and discuss their preferred interaction methods for smart glasses. Additionally, participants were asked to create sketches to illustrate their technology needs and preferences and collaborate as a group to develop design concepts.

We selected six touchless interaction methods from a range of options described in Lee and Hui (2018). These options were evaluated based on two criteria in our study: (1) whether the interaction method is suitable for the dynamic and mobile work practices of EMS, and (2) the feasibility of its technical implementation on most commercial smart glass devices. For the first criterion, two project consultants—senior EMS providers with over 20 years of experience and some familiarity with using wearable technologies in medical training—assessed each option and compiled a shortlist of interaction methods potentially suitable for EMS work. Subsequently, our research team investigated the feasibility of implementing these shortlisted methods on commercial smart glass devices. For example, while head-tilt controls can technically be implemented on most smart glasses, our consultants noted that EMS providers are constantly on the move and engage in physically demanding activities such as lifting patients, which makes head-tilt controls impractical for this context. Similarly, since EMS care involves direct patient interaction, using facial expressions as a control method was considered inappropriate for this setting. In contrast, eye-tracking offers a more subtle way of interacting with smart glasses (e.g., compared to voice commands); however, due to cost and hardware limitations, this method has not yet been implemented on commercial smart glasses (Haque et al., 2015; Prilla et al., 2019). More details about the pre-selection process for the touchless interaction methods are presented in Table 2.

After an iterative process of review and selection, six interaction methods were included in our final list to be ranked and discussed by EMS participants in the workshops: (1) voice commands (Firouzian et al., 2017), (2) pinch gestures (performing a pinch hand gesture in front of the device camera to invoke click interaction on the screen) (Guimbretière & Nguyen, 2012), (3) ring input (a finger-worn device that detects touch on the ring surface as input for smart glasses) (Ens et al., 2016), (4) haptic glove (a

**Table 2.** A Summary of the pre-selection process of touchless interaction methods described in Lee and Hui (2018).

| Method name (reference) | Included/excluded | Rationale |
|---|---|---|
| Finger-Worn Ring (Ens et al., 2016) | Included | Deemed suitable for EMS workflow as it is unobtrusive and socially acceptable (as some EMS providers already wear rings). |
| Wristband (Ham et al., 2014) | Included | Considered potentially suitable for EMS workflow due to its non-obtrusive nature (some providers already wear wristband). |
| Touch-Belt Device (Dobbelstein et al., 2015) | Excluded | Deemed inconvenient and impractical to wear. |
| Facial Expression (Serrano et al., 2014) | Excluded | Not practical for EMS due to extensive patient interaction requirements; Cognitive burden from associating expressions with commands. |
| Touch on Palm (Wang et al., 2015) | Included | Using augmented reality to interact with smart glasses via palm was not seen as obtrusive. |
| Forearm as Surface (Azai et al., 2017) | Excluded | Wearing additional equipment on the forearm was considered inconvenient and socially awkward. |
| Pinch Gesture (Guimbretière & Nguyen, 2012) | Included | No external device needed; Convenient and intuitive to use as it is similar to mouse clicking. |
| Voice Recognition (Firouzian et al., 2017) | Included | Preferred due to familiarity compared to other touchless methods. |
| Head Gesture (Yi et al., 2016) | Excluded | Deemed impractical as EMS providers need to move frequently and excessively. |
| Gaze Movement (Toyama et al., 2014) | Excluded | Initially included by consultants but later excluded by researchers due to technical constraints due to its technical constraints, for example, eye-tracking hardware and software are not yet fully developed and implemented in current smart glasses. |
| Tongue Gesture (Saponas et al., 2009) | Excluded | Inapplicable in EMS contexts where communication with patients and partners is essential. |
| Haptic Glove (Hsieh et al., 2014) | Included | Acceptance was tentative; use of palm/hand as a surface is feasible, but compatibility with gloves remains uncertain. |

tactile glove equipped with touch-sensible textile for pointing and selecting actions on the screen of smart glasses) (Hsieh et al., 2014), (5) wristband (tracking the movement of a users' wrist as gestural inputs for smart glasses) (Ham et al., 2014), and (6) touch on palm (using the palm as a surface to interact with the projected virtual content of smart glasses on the user's palm) (Wang et al., 2015). We used relevant videos, pictures, and online resources to illustrate these touchless interaction methods and facilitate discussion on their preferred interaction methods. After presenting and describing the interaction methods, participants were asked to discuss and rank them from most to least preferred, taking into account the feasibility and ease of use of each interaction method in their workflow. The rankings were anonymously submitted via Google Form. Out of 16 participants, 15 rankings were successfully recorded.

#### 4.1.2. Data analysis

We transcribed the discussions from the design workshop verbatim and imported the transcripts into NVivo (version 12) for further analysis. The transcribed data were then analyzed using an open coding technique (Blair, 2015; Holton, 2007). Two researchers first reviewed all the transcripts to familiarize themselves with the data, and subsequently, they independently analyzed one transcript to develop a codebook. The analysis focused on user requirements for smart glass design, user preferences and perceived ease-of-use for different interaction methods, and perceptions of using smart glasses in the field. All codes were discussed among the researchers to decide if a code needed revision, and to determine which codes to keep, merge, or discard. The researchers then used the finalized codebook to standardize the coding process for the remaining transcripts. Any new codes that emerged during this process were thoroughly discussed and added to the codebook. Finally, all researchers employed the thematic analysis approach (Braun & Clarke, 2012) to collaboratively construct the affinity diagram—an inductive method often used to organize low-level codes into high-level categories (Creswell & Poth, 2016)—to identify overarching themes.

The participants' preferences for touchless interaction methods were analyzed by examining their "first choice" (the most preferred method), "second choice" (the second most preferred method), and "third choice" (the third most preferred method).

### 4.2. Results

Out of the 16 workshop participants, 15 successfully submitted their rankings for the touchless interaction methods, which are depicted in Figure 1. Voice command emerged as the most preferred method, with 10 participants choosing it as their "first choice" and the remaining 5 selecting it as their "second choice." Participants explained that voice command was more straightforward and easier to learn compared to the other methods: "*With things like Alexa and stuff like that, I would say the voice commands would be my first choice in terms of familiarity.*" (PD 11)

Pinch gestures were deemed the second most preferred method, with 5 participants selecting it as their "first choice" and 10 participants choosing it as their "second choice." Participants found using pinch gestures to be easier and less cumbersome compared to other gestural input methods that require an additional device (e.g., wristband): "*I would rather do the pinching, even if it's socially less acceptable than use any of these other methods.*" (PD 6)

Among the remaining four methods, wristband-based gestural input received the highest ranking, with nine participants selecting it as their "third choice." Participants cited comfort and social acceptability as the primary reasons for

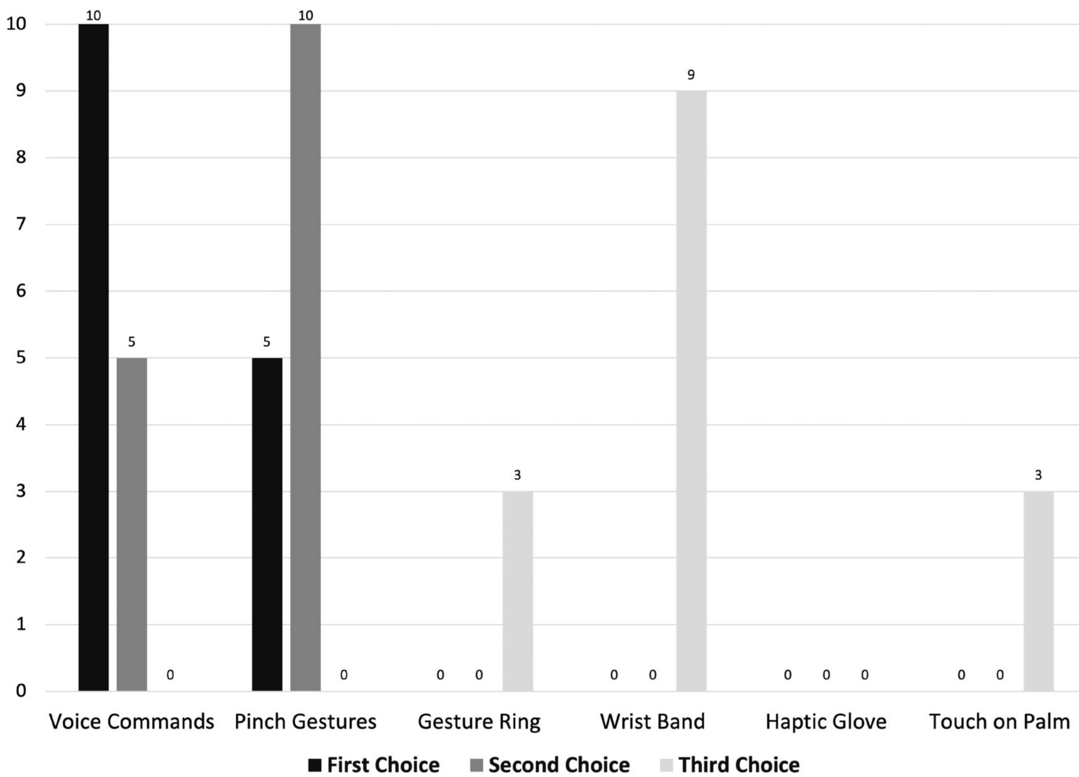


**Figure 1.** Participants' ranking of preference for six illustrated touchless interaction methods in the participatory design workshops.

preferring the wristband over other options such as the ring input or touch on the user's palm for smart glass interaction. One participant explained: "*I'm already wearing a smartwatch, which is about the same size. So, wearing a wristband is not a big deal to me. Also, I think that would be somewhat more socially acceptable.*" (PD 11)

The other methods (ring, touch on the palm, and haptic glove) were perceived as impractical to use (e.g., "*I don't see it (haptic glove) working well underneath our medical gloves.*"), prone to contamination (e.g., "*If you touch any sort of blood, urine, body fluids, we're constantly getting stuff all over our gloves and the rings going to be on top of that.*"), easy to lose (e.g., "*They're gonna lose those rings.*"), or too complex to use (e.g., "*The touch on the palm just from the video itself seemed more complex than gestures and voice recognition. You're adding another step because you have to put on an extra piece of equipment rather than just the gloves.*").

## 5. Study 2: Usability evaluation of EMS providers' preferred touchless interaction methods for smart glasses

### 5.1. Method

#### 5.1.1. Prototype development

We developed our smart glass application on the Vuzix M400 platform (Figure 2), which operates on Android. The device's camera allows for capturing still images and recording videos. This device is water-resistant and has up to 12 hours of battery life, while also being capable of operating across a wide temperature range, i.e., from −20 °C to 45 °C. Additionally, the device is compatible with wearers' prescription glasses and protective gear (e.g., helmets, hats, etc.). Compared to the first generation of smart glasses released almost a decade ago, these hardware advancements have the potential to address longstanding concerns related to battery life and ergonomic issues when using smart glasses in daily work or life (Zuidhof et al., 2021).

The primary user interface (UI) of the M400 device is the near-eye display, where users interact with virtual content and various elements such as virtual buttons, lists, and menus (Figure 2). The default interaction method for this device is tangible buttons (buttons 1–3 in Figure 2): Button 1 (rearmost) enables forward navigation, button 2 (middle) enables backward navigation, and button 3 (foremost) is used for selection.

Based on the findings from our design workshops, which identified voice commands and pinch gestures as the most preferred touchless interaction methods, we integrated these two methods into our prototype for subsequent usability evaluations. Specifically, we utilized the Vuzix software development kit (SDK) along with its Speech Command engine[1] to develop the voice command functionality. The Vuzix Speech Command engine is a phrase-matching recognition system designed to interpret and respond to voice commands. It comes with a base vocabulary (e.g., saying "hello Vuzix" to activate the system) and supports the addition of custom vocabulary to carry out application-specific actions, such as taking a photograph in response to the command "take a picture." In our system prototype, we programmed a set of simple voice commands that correspond to the text labels on the UI elements. We also implemented

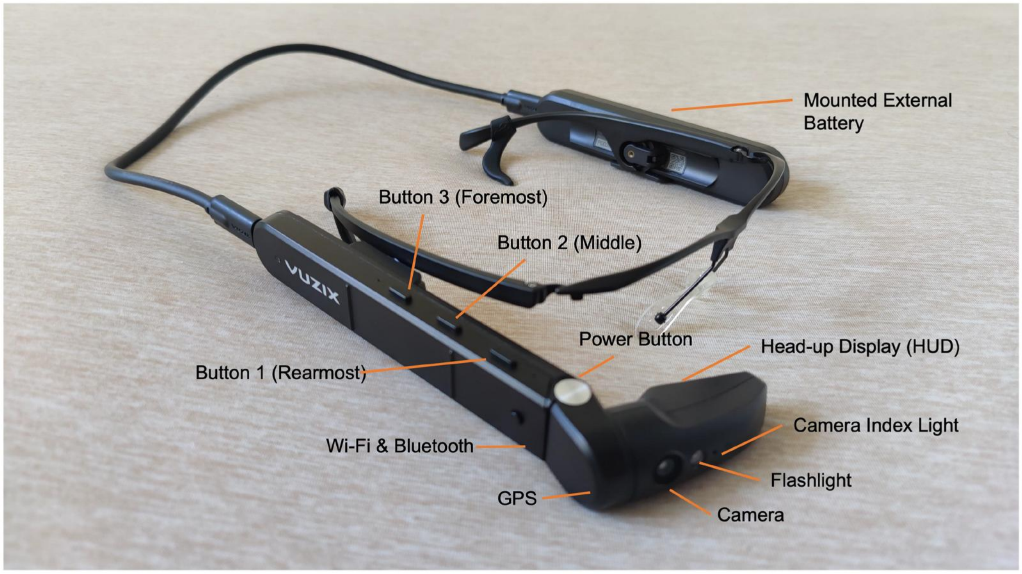


**Figure 2.** Vuzix M400 smart glass used in the study.

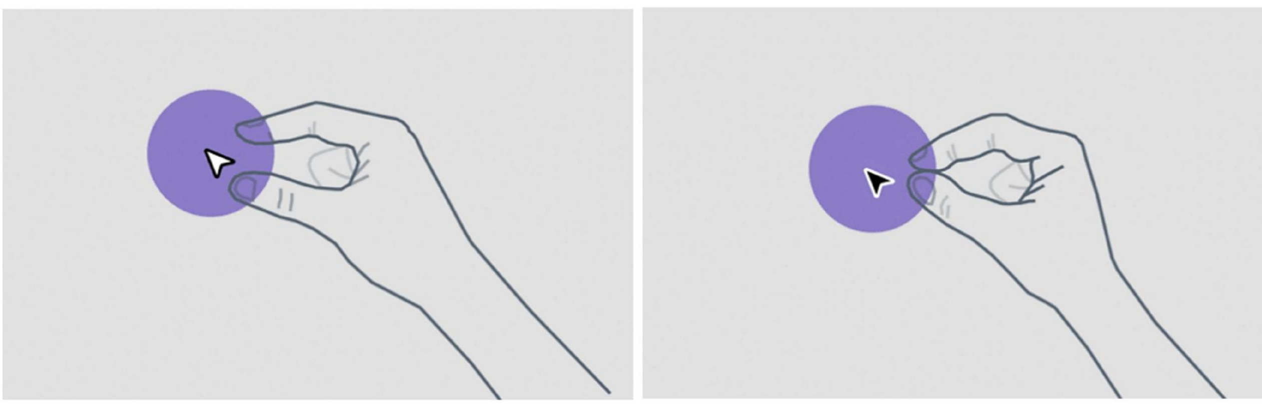

**Figure 3.** The pinch hand gesture allows users to interact with different UI elements on the near-eye display. To execute this gesture, the user begins by performing an open pinch to activate the cursor (left), followed by a pinch click achieved by tapping the thumb and index finger together (right).

multiple voice commands for a single action to provide users with greater flexibility. For example, users can activate the system's listener for voice input by saying "hey glass," "hey Vuzix," "hello glass," or "hello Vuzix." To implement the pinch gesture-based interaction method, we used a commercial gesture control software (CrunchFish[2]). This software is powered by hand-tracking algorithms, allowing a user to perform a pinch gesture to select and activate a virtual button on the screen in a touchless manner. Operating the pinch gesture involves two steps: first, the user performs an open pinch to summon the cursor (Figure 3, Left), and then, to complete an interaction such as clicking or dragging, the user performs a pinch click by tapping the thumb and index finger together (Figure 3, Right).

To prepare tasks for the usability testing phase, we iteratively designed and implemented five features in our prototype, drawing on insights from the participatory design workshops. These features are intended to support key EMS tasks that are currently challenging or impractical to perform using existing tools. Below, we provide a brief description of each feature.

*Feature 1: Video-based consultation with a remote physician.* EMS providers often require consultations with remote physicians for decision support or medical guidance, such as obtaining medication authorization or receiving expert advice on treatment and patient destination. By using smart glasses, the remote physician can observe and hear what the EMS providers are experiencing in the field, thus gaining a better understanding of the patient's condition. In our current prototype design, the smart glass wearer initiates the call by indicating the reason (Figure 4, Left). During the call, the wearer can control the camera and audio settings, and end the call as needed (Figure 4, Right).

*Feature 2: Hospital notification via video call.* Due to technical limitations, EMS providers are currently restricted to using radio communication inside the ambulance to notify the receiving hospital. This work practice causes inconvenience for EMS providers; for example, one provider needs to leave the patient and get into the ambulance to notify the receiving care team (Zhang, Sarcevic, et al., 2021). Our system allows EMS providers to notify the receiving care team anytime and anywhere.

*Feature 3: Medication scanning.* Accurate collection and documentation of medication information in the field can be both time-consuming and challenging. To address this issue, we have developed a feature that allows EMS providers to scan the

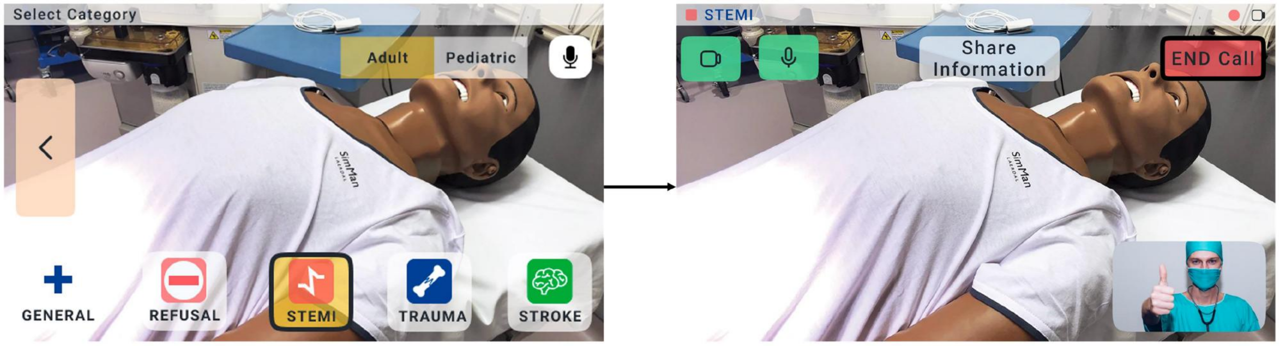


**Figure 4.** The specific steps for conducting a teleconsultation with an online medical control doctor (system feature 1) using our system.

**Table 3.** An overview of the five testing tasks.

| Task ID | Task content | Corresponding system feature |
|---|---|---|
| Task 1 | Call a remote doctor for online medical control | Feature 1 |
| Task 2 | Notify a hospital about the patient's arrival via a video call | Feature 2 |
| Task 3 | Scan the medication barcode | Feature 3 |
| Task 4 | Take a photo of the surrounding environment | Feature 4 |
| Task 5 | Dictate patient information for documentation | Feature 5 |

barcode on administered medications, which automatically records detailed information such as the medication's name and administered time.

*Feature 4: Visual information capture.* It is not uncommon for EMS providers to take pictures of the scene, such as a crashed vehicle, using personal cellphones (Zhang et al., 2017). These visual records provide additional context (e.g., patient injuries) for the hospital care team, which facilitates discussions between EMS and hospital teams during patient hand-off. To enhance this practice, our prototype enables picture-taking and video recording.

*Feature 5: Dictation of patient information for documenting on EHR.* Documenting patient information in the field using an EHR can be time-consuming to the providers and often incomplete in time-critical settings (Jagannath et al., 2019). To streamline EMS data collection and documentation, we aim to integrate our smart glass prototype with EHR systems. This will allow EMS providers to dictate patient information directly to the smart glasses. Leveraging NLP techniques, the smart glasses automatically transcribe and process the dictation in real-time to record essential medical information into the EHR, which supports continuity of patient care.

### 5.1.2. Usability evaluation

*5.1.2.1. Study procedure.* We first conducted a power analysis using G*Power (version 3.1.9.7) to determine the requisite number of participants to meet the minimum sample size needed for our study (Faul et al., 2007). While conducting the power analysis, we considered an expected effect size of 0.5, an alpha level of 0.05, and a power of 0.80. Based on the results of the power analysis, we recruited 16 participants to participate in individual usability testing. The tests took place in a controlled environment within the EMS agency (e.g., an office or a simulation lab), with each session lasting about 60 min.

At the beginning of each testing session, we explained the study's purpose and obtained consent from the participants. We then demonstrated the key system features and interaction methods of our prototype. Additionally, we provided a reference sheet with the necessary voice commands to facilitate task completion. Since none of the participants had prior experience with smart glasses, we provided training to help them become acquainted with the device's operation. Each participant was given four sample tasks to practice operating the prototype and to become familiar with the three interaction methods (voice commands, pinch gestures, and tangible buttons). The training typically lasted between 15 and 20 min, though some participants required more time to become comfortable with the device and different interaction methods.

The participants were then asked to use the three interaction methods in a randomized order to complete an identical set of tasks, each corresponding to one of the five major features of our prototype. A summary of the tasks is presented in Table 3. To mitigate the implications of learning effects (e.g., participants might become more familiar with the tasks as the testing progresses, potentially resulting in better performance with the interaction method tested later), we randomized the testing order of the interaction methods for each participant. For example, participant#1 started with hand gestures, followed by voice commands and tangible buttons, while participant#2 tested the interaction methods in a different order (voice commands→hand gestures→tangible buttons).

After completing the testing phase, we administered a survey that utilized a Likert scale ranging from 1 to 5, where "1" represented strong disagreement and "5" indicated strong agreement with the given statements. The survey included statements such as "I think the hand gesture interaction method is easy to use," "I think the voice commands are easy to use," and "I would prefer using tangible buttons over voice commands and hand gestures." Additionally, we conducted semi-structured interviews to gather the participants' perceptions, expectations, and experiences regarding the system features and interaction methods.

All testing sessions were audio-recorded and videotaped. We also monitored and recorded the users' interactions with the application UI by connecting the smart glasses to the researcher's laptop via Bluetooth. This setup allowed us to observe participants' actions and movements within the interface, address any questions they had, and provide necessary support when needed. More importantly, the recordings of users' operations within the application enabled us to perform an in-depth analysis of which UI elements participants selected and whether an error occurred.

*5.1.2.1. Data analysis.* In our analysis, we examined both objective measurements (e.g., task performance) and subjective measurements (e.g., user experiences collected through surveys and interviews). Below, we provide a detailed description of how we analyzed these data.

*Task Performance Analysis:* We analyzed the video recordings of the smart glass screen frame-by-frame to measure three key aspects: task completion time, errors encountered, and the time taken to recover from errors. Task completion time represents the duration (in seconds) participants spent on completing each task using a specific interaction method. Errors encountered refer to any failed, incorrect, or unintended operations that occurred during task performance. The time taken to recover from errors indicates the amount of time (in seconds) that participants needed to resolve any encountered errors.

Two researchers first reviewed two video recordings together to get familiar with the dataset and to determine what constituted an error. The task completion time was manually measured by calculating the elapsed time from the start of a task to its successful completion. Similarly, error recovery time was measured by determining the elapsed time from the occurrence of an error to its resolution. An iterative process was used to develop a coding schema for errors, which included a variety of error types encountered during task performance, along with their definitions and examples. Subsequently, the same researchers randomly selected and independently reviewed an additional four video recordings. To ensure inter-rater reliability, we used Cohen's Kappa coefficient to compare the error codes assigned by the researchers. The resulting kappa values were interpreted according to the scale proposed by Landis and Koch (1977). The coders achieved an "Almost Perfect" level of agreement for the error codes, with a kappa value of 0.85. Any disagreements were resolved through discussions involving all researchers. Following this step, the researchers refined the coding schema for errors and completed the analysis of the remaining video recordings.

To assess differences in task performance among different interaction methods, we conducted a set of statistical analyses. We began with Shapiro-Wilk tests to check the normality of participants' performance data for each measurement (e.g., errors, task completion time, and error recovery time). Since all measurements deviated from normal distribution ($W > 0.55$, $p < 0.01$), we proceeded with non-parametric statistical methods to analyze the task performance data. Specifically, we applied the Friedman test to determine if there were any significant differences in measurements across the three interaction methods. When a significant difference was detected, we performed post hoc pairwise comparisons using the Wilcoxon Signed-Rank test with Bonferroni corrected p-values.

*Questionnaire Data Analysis*: For the analysis of questionnaire data, which comprised responses to survey questions, we employed descriptive statistical analysis (e.g., calculating the average rating of all responses for each question). In addition, to assess the significance of differences in user ratings, we used the Kruskal-Wallis test to examine group differences, followed by post hoc Mann-Whitney U test with Bonferroni-corrected p-values for pairwise comparisons.

Regarding the qualitative data collected through interviews, we adopted the same inductive qualitative analysis approach specified in Section 4.1.2. This analysis focused specifically on EMS providers' experiences and perceptions of using smart glasses and different interaction methods in their work.

### 5.2. Results

#### 5.2.1. Usability comparison of touchless and tangible interaction methods

*5.2.1.1. Task completion time.* We observed a significant difference among the three interaction methods for this measurement ($\chi^2(2) = 20.14$, $p < 0.01$). Post hoc pairwise analyses, as illustrated in Figure 5, revealed significant differences in task completion time between pinch hand gestures (Median = 22.5) and voice commands ($Z = 1023$, $p = 0.015$), between voice commands (Median = 18) and buttons ($Z = 919.5$, $p = 0.012$), as well as between buttons (Median = 16) and pinch gestures ($Z = 604.5$, $p < 0.01$). These results indicate that participants could complete tasks significantly faster when using tangible buttons. While voice commands took longer than tangible buttons, they were still significantly faster than hand gestures in terms of task completion.

*5.2.1.2. Errors.* We observed a significant difference in the occurrence of errors among the three interaction methods ($\chi^2(2) = 17.59$, $p < 0.01$). As shown in Figure 5, participants encountered significantly more errors when using pinch gestures (Median = 1) compared to using voice commands (Median = 1) and buttons (Median = 0) ($Z = 361$ and 232.5, respectively; both $p$ $values < 0.01$). However, the difference in the number of errors between buttons and voice commands was not found to be significant ($Z = 263$, $p = 0.10$). These findings suggest that pinch hand gestures were more prone to errors compared to the other two methods.

Upon further examination, we categorized the types of errors encountered by participants for each interaction method and examined whether participants could recover from these errors independently or required assistance from the researchers. A summary of this analysis is presented in Table 4, with more details described below.

*Pinch hand gestures* accounted for the majority of errors (90/160, 56.25%). However, participants were able to resolve

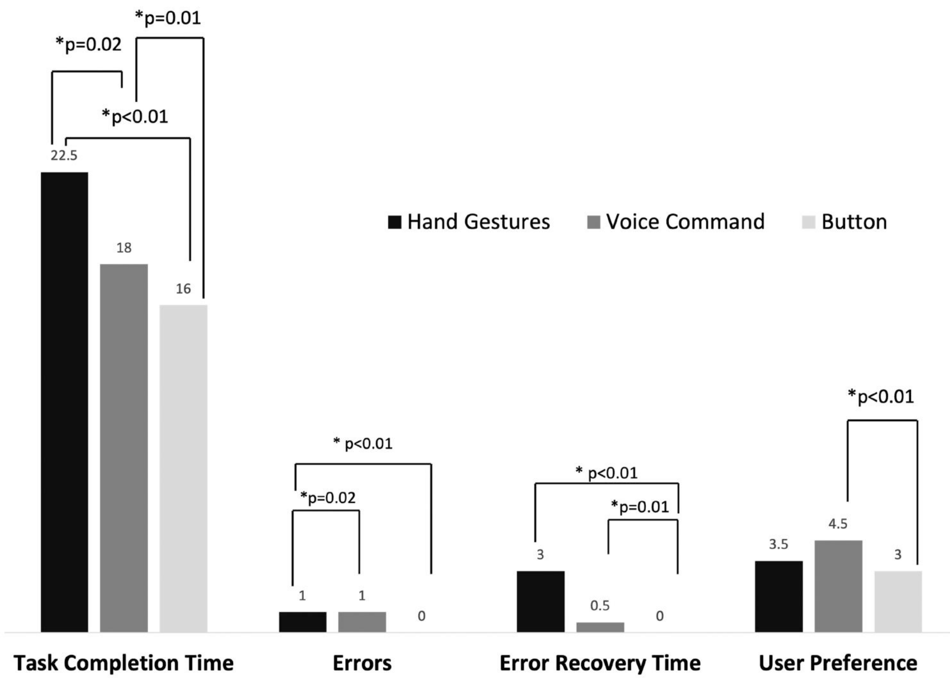


**Figure 5.** Results of usability comparison of pinch hand gestures, voice commands, and tangible buttons across four measurements: task completion time, encountered errors, error recovery time, and user preference. *Note*: *Median values are used as data labels for each bar.*

**Table 4.** An overview of the analysis of errors encountered by our participants when using each interaction method.

| Interaction method | Error type | Number of total occurrences | Number of instances (W/O H) | Number of instances (WH) |
|---|---|---|---|---|
| Pinch Gestures | -The system failed to recognize the close pinch gesture | 75 | 74 | 1 |
| | -The user used a wrong gesture | 5 | 4 | 1 |
| | -The system failed to recognize the entire hand posture | 10 | 9 | 1 |
| | **Total** | **90** | **87** | **3** |
| Voice Commands | - The system did not respond to voice commands | 22 | 20 | 2 |
| | -The user forgot to activate or reactivate the device's listener | 20 | 17 | 3 |
| | -The user used an incorrect command | 7 | 5 | 2 |
| | **Total** | **49** | **42** | **7** |
| Tangible Buttons | -The user pressed a wrong button | 21 | 21 | 0 |
| | **Total** | **21** | **21** | **0** |

*Note*: The term "Number of Total Occurrences" refers to the overall count of errors that occurred. "Number of Instances (W/O H)" represents the number of instances the instances where participants were able to identify and resolve the error independently, without requiring assistance or guidance from the research team ("W/O H" stands for "without help"). On the other hand, "Number of Instances (WH)" indicates the number of instances where participants sought help from the research team to rectify the error ("WH" stands for "with help").

these errors without assistance from the researchers in most cases (87/90, 96.67%). The most common error was the system's failure to recognize the pinch gesture (75/90, 83.33%). This could be attributed to the specific requirements of the gesture-sensing software used in the study: (1) the hand must be positioned at a minimum distance of 20 cm from the device's camera (Figure 6, Left), and (2) the plane of both fingers (index and thumb) needs to be parallel to the camera when performing a close pinch gesture (Figure 6, Right). Due to participants' unfamiliarity with this new method, they faced challenges in performing the pinch gestures correctly, especially at the beginning of the task.

*Voice commands* had the second highest number of errors among the three methods (49/160, 30.63%), but participants were able to resolve most of these errors on their own (42/49, 85.71%). The most common error was the voice-recognition software not responding to participants' commands ($n = 22$). Typically, participants had to attempt their command a second time to get a response from the system. The second most frequent error occurred when users forgot to activate or reactivate the device's listener (n = 20), likely due to their unfamiliarity with the activation process.

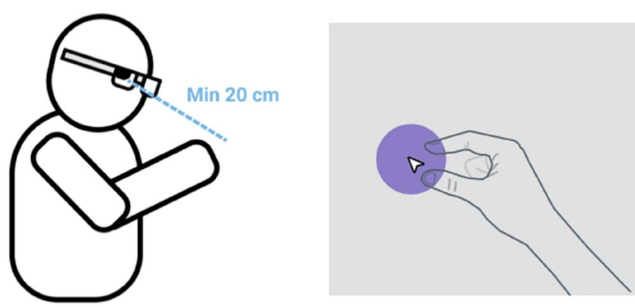


**Figure 6.** Required hand posture by the gesture-sensing software.

*Tangible buttons* had the fewest errors (21/160, 13.13%), and all participants were able to rectify these errors independently without seeking assistance (21/21, 100%). The only error encountered by participants was pressing the

**Table 5.** A summary of participants' perceived advantages and disadvantages of each interaction method.

| Interaction method | Perceived advantages | Perceived disadvantages |
|---|---|---|
| Voice Commands | • Enabling touchless operation<br>• Voice commands are easy to use and learn as they correspond to the text label displayed on the UI elements | • Challenging to use in noisy and chaotic environments<br>• Prone to mis-triggers due to cross-talks<br>• Uncomfortable to speak commands in public spaces |
| Pinch Gestures | • Facilitating touchless operation<br>• Minimal impact on communication with the patient and partner | • Lengthy learning curve led to incorrect hand posture, resulting in unrecognized hand gestures by the system<br>• Challenging to use while inside a moving ambulance |
| Tangible Buttons | • Already familiar with tangible buttons<br>• Exhibits fewer errors compared to using voice commands or hand gestures | • Not feasible to use when attending to patients<br>• Raise concerns about potential cross-contamination |

wrong button, which could be attributed to the counterintuitive design of the buttons' primary functions. As depicted in Figure 1, button 1 (rearmost) and button 2 (middle) are used for navigating forward and backward, respectively, while button 3 (foremost) is for selection. During our follow-up discussions with participants (see Section 5.2.2.3), it emerged that most had presumed the middle button (button 2) was for selection and the foremost button (button 3) was for backward navigation. This confusion of the buttons' functions accounted for the majority of the errors.

*5.2.1.3. Time taken to recover from errors.* Overall, there were significant differences in the time taken to recover from errors among the three interaction methods ($\chi^2(2)$ = 16.28, $p < 0.01$). As illustrated in Figure 5, post hoc pairwise analyses revealed that both voice commands (Median = 0.5) and pinch gestures (Median = 3) required significantly more time to recover from errors compared to buttons (Median = 0) ($Z = 297.5$ and 301.5, respectively; both *p values* $< 0.01$). However, no significant difference was found between voice commands and pinch gestures in this measure ($Z = 0.677$, $p = 0.12$). In summary, tangible buttons performed the best in terms of error recovery, while voice commands and pinch gestures showed comparable performance in error correction, despite voice commands requiring less time for error recovery.

### 5.2.2. User experience of touchless interaction methods

The analysis of the survey questionnaire shows that buttons ($M = 3.0/5$, $SD = 1.01$) and pinch gestures ($M = 3.47/5$, $SD = 1.24$) received lower user ratings compared to voice commands ($M = 4.38/5.0$, $SD = 0.72$). Higher ratings indicate a more positive attitude towards an interaction method. The Kruskal-Wallis test shows a significant difference in user ratings across the three methods (H ($df = 2$) = 12.36, $p < 0.01$). As illustrated in Figure 5, pairwise comparisons show no significant difference between pinch gestures (Median = 3.5) and voice commands (Median = 4.5) ($U = 70.5$, $n1 = n2 = 16$, $p = 0.03$). However, there is a significant difference in user ratings between voice commands (Median = 4.5) and buttons (Median = 3) ($U = 37$, $n1 = n2 = 16$, $p < 0.01$). These findings are interesting because user preference for the interaction methods, after testing them, does not entirely align with their task performance. These results indicate that despite the suboptimal task performance of touchless methods, EMS providers still preferred using them over tangible inputs. Participants also elaborated on their perceived benefits and barriers for each interaction method, which are further described below. A summary of participants' perceptions is presented in Table 5.

*5.2.2.1. Voice commands.* Participants favored this method the most ($n = 8$), because of its ease of use and intuitiveness: "*I think the voice commands were nice because they match what you saw on the screen. I didn't have to remember anything super special. It was pretty user-friendly and pretty straightforward. Like if you use a smartphone, you can probably use this.*" [T14] Additionally, voice commands allowed for hands-free operation, reducing the risk of cross-contamination: "*They're very helpful and beneficial because, you know, I do have a lot of blood and stuff on me all the time.*" [T8]

Regarding perceived disadvantages, a primary concern is that using voice commands in the noisy and chaotic prehospital environment could be very challenging:

> "*The voice commands work, but they will not work in all scenarios. There's a lot of noise, chaos, other people talking, all those different things that can make it difficult.*" [T13] In such environments, voice-recognition software might be easily mis-triggered by crosstalk at the scene: "*The big thing is the trigger phrases that we talked about. They might be mis-triggered from just the regular conversation. So, let's say, you know, I talk to my patients like, 'Hey, what are those glasses you're wearing?' Now, the trigger phrase is included in what we're talking about. So, you need to find a way to make the trigger phrase unique enough and not common enough so that the glasses won't be easily triggered from just normal conversation.*" [T15] Additionally, using voice commands could confuse others at the scene, as one participant explained: "*I'm not a huge voice control fan, especially when you're trying to have a conversation with a patient or your crew. I think that might be confusing or hard to work when you're talking to your system.*" [T12] Lastly, participants raised concerns about the social implications of using voice commands, with one noting: "*If there was someone else next to me, I might feel weird using the voice command.*" [T16].

*5.2.2.2. Pinch hand gestures.* Participants acknowledged the usefulness of pinch gestures because this interaction method may be more suitable for noisy environments than voice-based interaction. They appreciated its ability to interact seamlessly with the smart glasses without interrupting their conversation with the patient or their partner: "*I could continue to talk and evaluate and do some things while I'm using those gestures.*" [T1]

Although many participants liked and were amazed by what pinch hand gestures could accomplish, they still expressed a few concerns about this method. First, unfamiliarity with the required hand position resulted in difficulties using hand gestures effectively to interact with the device. Some participants found this method a bit unnatural and required some time to become accustomed to it: "*I found this method a little bit of unnatural, and then it also took a little bit of time getting used to.*" [T15] Relatedly, the learning curve associated with using hand gestures led to several errors. For example, incorrect positioning of the user's hand could make it unrecognizable by the system: "*I was having a hard time trying to use the hand gestures, because sometimes my hand would be out of the screen, or it's not even recognizing my hand on the screen.*" [T10] Finally, using hand gestures might pose challenges for providers in a moving ambulance: "*Is it going to work reliably when you're bouncing around in the back of an ambulance?*" [T4]

Our participants provided some suggestions to improve the usability and acceptability of pinch gesture interaction. They emphasized the need for training to ensure that EMS providers are familiar with performing pinch gestures correctly. As one participant mentioned, they found the gestural input much easier to use after more practice: "*Once I kind of understood the basic controls from it, I think that was the easiest to do.*" [T6] Additionally, they expressed a desire for improved gesture-sensing software that could more easily and accurately recognize their hand movements. Finally, they recommended enlarging the virtual buttons on the interface and positioning them more appropriately for easy interaction with gestures. For example, it was noted that participants often encountered difficulty when interacting with the back button located in the corner. This observation implies that to prevent usability issues during gestural input, UI elements should be positioned away from the edges or corners of the screen.

***5.2.2.3. Tangible buttons.*** A few participants ($n = 5$) stated that the tangible buttons on the smart glasses were easy to use because they were already familiar with such input: "*I think that button was the easiest, just because it was more intuitive than the other two. I think it is probably just because that's what I'm used to using.*" [T1] Moreover, unlike using voice commands, providers can use buttons without disturbing others, as one participant stated: "*I prefer the buttons because you know it's quieter without stating what you're doing.*" [T7]

However, our participants highlighted several concerns about this method. First, the default setup of the buttons' primary functions was not very intuitive for some participants. Despite the initial training, at least five participants found it difficult to remember the function of each button: "*Because you might get confused. My intuitive thought of how the buttons should work is that the foremost button should be going backward, and then the middle one would be select.*" [T16] Second, using the buttons could interfere with manual tasks and increase the risk of cross-contamination: "*So during patient care, you're always wearing gloves. But as soon as those are contaminated, I don't want to be touching anything near my face. And so, it's not likely for me to be using those buttons.*" [T6] Another participant shared a similar opinion and added that using buttons could eliminate the hands-free advantage of smart glasses: "*If part of the goal of this is to offload your hands, then buttons are probably eliminating a little bit of that advantage.*" [T6]

Nevertheless, EMS providers noted that the tangible buttons could serve as a backup approach to interact with the smart glasses if voice commands or hand gestures became ineffective, or when their hands are not occupied: "*I would say the buttons could be our back-up because if the voice commands or hand gestures are not working, you're going to revert to manually doing it.*" [T11]

## 6. Discussion

In this section, we will first reflect on the user-centered approach we employed to examine smart glass interaction methods for EMS teams. We will then discuss our study's implications for enhancing smart glass interaction in dynamic medical settings. Lastly, we will address the study's limitations and suggest directions for future work.

### 6.1. User-centered approach for determining appropriate smart glass interaction methods for fast-paced medical teams

Prior work on the application of smart glasses in healthcare has primarily focused on using off-the-shelf smart glass devices to support various medical tasks (e.g., documentation, teleconsultation, etc.) and evaluating their efficacy (Mitrasinovic et al., 2015). However, there has been minimal focus on exploring the interaction between medical providers and smart glass devices. To the best of our knowledge, no previous research has worked closely with care providers to determine the most suitable way of interacting with smart glass devices in dynamic medical settings. Therefore, our study represents the first to employ a participatory design approach to examine the preferred interaction methods for smart glasses by fast-paced medical teams such as EMS. This approach allowed us to integrate interaction methods into smart glasses that align more closely with the end-users' needs.

Through the design workshops, we identified voice commands and pinch gestures as the two most preferred touchless interaction methods. This user preference might be attributed to two primary factors. First, EMS providers may be more familiar with these two touchless methods, compared to others. Voice commands is widely deployed in many smart devices so users have already become aquatinted with this interaction method. Pinch gestures, on the other hand, mimic the operation of mouse clicks. Therefore, these two methods may seem more natural and convenient to the participants. This is understandable because in high-risk medical settings, providers tend to rely on technologies they are more familiar with (Safi et al., 2018). Second, some other options, such as ring-based input or haptic gloves, require additional devices to operate. Among all participants, it became evident that EMS providers preferred not to use additional devices to interact with smart glasses, as it would add an extra burden. This finding implies that alignment with clinical workflow and specific work conditions is also an important determinant of user preference and the adoption of touchless interaction methods.

Given the great number of touchless interaction methods available, we chose to initially collaborate closely with senior EMS consultants, who have prior experience with wearable head-mounted display devices, to identify a shortlist of options for EMS providers to evaluate during workshops. This iterative pre-selection process considered not only the applicability of each option to the EMS context but also the technical feasibility of implementing each method on commercial smart glasses. This strategy allowed the researchers to familiarize themselves with the types of touchless interaction techniques proposed or examined in previous research, and to narrow down a long list of options to a shorter list that was more practical for further investigation.

Another key consideration in our study design was that EMS providers might have varied opinions regarding touchless interaction methods and the smart glasses themselves. Therefore, we intentionally recruited EMS participants with diverse experiences, ages, and backgrounds from different agencies and locations. This diversity aimed to enhance the generalizability of our findings and minimize potential sampling biases. Our inclusive approach ensured that care providers with a range of perspectives had equal opportunities to express their preferences and concerns.

Finally, when designing the touchless interaction methods during prototype development, we sought to adhere to established design guidelines. However, we discovered that design guidelines specific to gesture-based interactions are limited. Moreover, voice interaction design largely follows the set of guidelines for graphical interfaces (Murad et al., 2018). In line with arguments from previous literature (Suhm, 2003), we recognize an emerging need for design principles tailored to touchless interactions with computing devices, especially wearable head-mounted displays. Nevertheless, we did follow several general design guidelines proposed by Nielsen (1994) and Shneiderman et al. (2016). For example, we mapped most voice commands to text labels on virtual buttons to reduce users' cognitive load associated with recalling an extensive list of voice commands. This design strategy resonated with design principles such as "recognition rather than recall" and "minimalism in design and dialogue." Guided by the "flexibility and efficiency of use" principle, we implemented multiple voice commands for the same action to offer users flexibility. For example, users could activate the system's listener for voice input using phrases like "hey glass," "hey Vuzix," "hello glass," or "hello Vuzix." Despite our best efforts to create an intuitive voice-based interaction, we noted that EMS providers sometimes faced challenges using voice commands and were uncertain why the system was not responding to their voice input. This observation underscores the potential value of providing users with feedback regarding system status to enhance the system's transparency—a design approach that aligns with the principle of "visibility/feedback of system status." Previous research has explored the utility of visual feedback in helping users understand how their voice input was processed—some studies indicated that users relied on visual feedback to gauge whether their voice input was correctly understood (Begany et al., 2015; Cowan et al., 2017), while others found that users relied on "guessing" tactics and that providing visual feedback offered no discernible advantage over scenarios where no visual support was given (Myers et al., 2018; Prilla & Mantel, 2021). Given these varied findings, our future work will explore the most effective methods of delivering feedback on the voice interface's actions and status to keep cognitively overwhelmed care providers informed and help them more easily recover from errors.

### 6.2. Implications for enhancing smart glass interaction in time-critical medical settings

Our usability testing revealed that tangible buttons achieved the best task performance, while voice commands and pinch gestures performed less effectively than buttons. These findings are not surprising, considering that users are already familiar with tangible interaction methods (Kyung et al., 2008). However, most errors related to voice commands and pinch gestures could be primarily attributed to software limitations as well as users' unfamiliarity with the specific commands and required hand posture for performing pinch gestures. Additionally, despite the sub-optimal task performance, participants still favored touchless interaction methods over tangible buttons since the latter method could interfere with their manual tasks and increase the risk of cross-contamination, eliminating the benefits of using smart glasses.

To increase the user adoption of smart glasses in in time- and safety-critical medical settings like EMS, it is crucial to make the interaction between care providers and smart glasses as seamlessly as possible. Below, we discuss several ideas to enhance smart glass interaction for fast-paced medical teams, including improving the technical performance of hand or voice recognition, accounting for domain-specific constraints, considering the social implications of smart glass usage, offering flexible and hybrid user interactions, and providing necessary training.

#### 6.2.1. Improving the technical performance of voice-recognition and gesture-sensing software

With the increased prevalence of voice-based interaction in daily life and work (e.g., voice assistants), it has become one of the dominant input sources for many mobile applications (Syberfeldt et al., 2017). In our study, we found that although tasks took longer to complete with voice commands compared to buttons, this method did not result in significantly more errors. Consistent with prior work (Lee & Hui, 2018), our study also found that the primary challenge for participants using voice commands was the voice-recognition software's limitation in recognizing the voice inputs. This issue often arose due to the software's failure to accurately capture the voice input or participants' unfamiliarity with the pre-defined voice commands. These findings suggest that with improvements in voice-recognition software and additional training, voice commands could become a reliable input method for smart glasses. Recent research is increasingly focused on enhancing the technical performance of voice-recognition software; for example, studies have explored utilizing sensing and signal processing techniques to achieve high performance automatic speech recognition for smart glasses (Maruri et al., 2018; Zhang et al., 2023). Additionally, some speech recognition software tailored to the medical domain, such as Amazon Transcribe Medical and Google Speech-to-Text Clinical Conversation, have been designed for more accurate capture and transcription of medical terms and clinical conversations (Tran et al., 2023). Future studies could test the use of advanced voice-recognition tools with noise cancellation techniques in simulated or real EMS contexts to measure their effectiveness in processing voice inputs in noisy environments.

Our study revealed that pinch gestures were the least efficient in terms of task completion. This finding aligns with previous research (Sambrooks & Wilkinson, 2013), which reports that gestural interaction is less effective than touch interaction and is prone to inaccurate recognition and muscle fatigue. In our study, we found that most errors in using the pinch gesture method were due to incorrect hand positioning, as the gesture-sensing software has strict requirements for hand postures. However, it is encouraging to see that participants were able to independently resolve nearly all these errors and issues associated with performing pinch gestures. Therefore, we believe that with improvements in software that enhance hand recognition and intuitiveness, the pinch gesture-based method has the potential to gain user acceptance. Recent work using deep learning methods to improve the accuracy of skeleton-based hand gesture recognition has shown impressive results (Mohammed et al., 2023). In our future work, we plan to evaluate more advanced gesture-sensing software with higher accuracy in hand posture recognition (e.g., CrunchFish XR Skeleton) and compare its task performance with the three methods assessed in the current study.

### 6.2.2. Accounting for domain-specific constraints

Although usability testing was conducted in controlled settings, participants envisioned the use of smart glasses in the real world and identified several potential barriers associated with using voice commands and pinch gestures in the dynamic pre-hospital environment. It is imperative for the design of smart glass interactions to consider and address these issues.

First, unlike other hospital domains with relatively quiet workspaces, EMS providers operate in noisy and chaotic environments. Consequently, voice commands may be challenging to capture and recognize, or may be accidentally triggered by ambient noise (Lee & Hui, 2018). This issue is exacerbated by the current lack of efficient voice-recognition software for smart glasses that can accurately recognize medical terminology and syntax (Muensterer et al., 2014). Second, the nature of EMS work involves frequent movement or operating within a moving ambulance, which makes the use of gesture-based interaction difficult. For example, hand gestures often require a relatively long dwelling time compared to touch or click interactions, as users need to maintain a posture long enough for the system to recognize the start and end of the gesture (Istance et al., 2008). Given this limitation, using hand gestures in a turbulent environment, can become cumbersome. Future research should focus on testing the effectiveness of gestural input in a moving ambulance.

In brief, the EMS context has unique characteristics that distinguish it from hospital-based care environments. As a result, using voice or gesture interaction methods can be challenging under certain conditions. Future research should aim to rigorously test touchless interaction methods in real settings to assess their suitability and user-friendliness.

### 6.2.3. Offering hybrid, multi-model user interaction

Our study revealed that user preference for interaction methods varied among providers. While some participants found voice commands easy to use, others appreciated the usefulness of pinch gestures. Additionally, although participants considered tangible buttons intuitive to use, they expressed hesitance to adopt this method due to concerns such as interference with manual tasks and the risk of patient contamination. Additionally, participants emphasized the importance of multimodal user interactions to provide alternative options when one method becomes unresponsive or impractical. For example, several participants mentioned that tangible buttons could serve as a backup approach for smart glass interactions.

Our findings are in line with prior research that suggests enabling hybrid user interaction to mitigate the limitations inherent in each individual method. For instance, studies found that single- or multi-finger touch gestures are more effective for interacting with virtual 2D content (e.g., selection and dragging), while direct manipulation gestures are better suited for 3D content (e.g., rotating) (Kim et al., 2019; Lee & Hui, 2018). From these insights, we propose offering hybrid, multimodal interaction methods for EMS providers to flexibly interact with different UI elements and perform various tasks with varying levels of complexity.

However, adopting a multimodal approach doesn't necessarily entail offering an abundance of interaction methods. As the number of interaction methods on smart glasses increases, additional problems may arise. For example, users might experience increased cognitive load if required to remember specific ways to use different interaction methods. Moreover, the computing power and battery life of the smart glass device could be quickly drained by running multiple advanced sensing software simultaneously. We also observed that the gesture-sensing software could conflict with other features that require the smart glasses' camera resources, such as video-based teleconsultation with remote experts. Given these considerations, future work should explore the optimal balance between the number of interaction methods to be integrated into smart glasses and the physical, cognitive, and computational constraints in using and implementing these methods.

#### 6.2.4. Considering social implications of smart glass usage

Our participants raised concerns about the social implications of using voice commands in their work environment. They noted that using voice controls could attract a lot of attention in public spaces and potentially disturb others' work (Kollee et al., 2014). This finding aligns with previous studies that have found interactions with smart glasses can be obtrusive when not thoughtfully designed, leading to reluctance among users to use smart glasses at their workplace (Hsieh et al., 2016; Koelle et al., 2015; Lucero & Vetek, 2014; Serrano et al., 2014).

This user concern highlights the need to consider the social implications of using smart glasses in medical settings, where vulnerable patients and their family members may be affected by the presence of smart glasses. Social acceptability plays a critical role in the adoption and use of new technologies (Rico & Brewster, 2010). It encompasses aspects such as the device's appearance, its perceived social status, and its conformity with cultural norms, all of which can impact both users and spectators (e.g., bystanders) (Montero et al., 2010). For example, Denning et al. (2014) emphasized the importance of bystanders' assumptions about the purpose of using smart glasses. In a similar vein, Koelle et al. (2015) examined the social acceptability of smart glasses from the viewpoint of both users and spectators and found that the usage was perceived more positively by users, while the acceptability by spectators depended on understanding the purpose of use.

These previous studies clearly indicate that the use of smart glasses can be more acceptable when the purpose is known to others (Koelle et al., 2015; Montero et al., 2010). From this perspective, EMS providers may need to explain the purpose of using smart glasses if time permits and the patient is conscious (similar to how law enforcement officers explain the use of body-worn cameras to individuals after arriving at an incident scene). Additionally, the use of smart glasses affects not only EMS providers but also patients, their families, and bystanders (Klinker et al., 2020b). Therefore, it is necessary to engage different stakeholders involved in EMS care (e.g., providers and patients) to investigate the social implications of using voice commands, hand gestures, or other input methods when using smart glasses in patient care. These insights can inform the exploration of new ideas to make the use and interaction of smart glasses less obtrusive and more socially acceptable.

Finally, given that the widespread use of smart glasses often raises privacy and ethical concerns (Denning et al., 2014), it is imperative for researchers to seek ways to enhance the protection of privacy for bystanders and patients. For example, certain studies have explored consent mechanisms, such as opt-in and opt-out gestures, which enable bystanders to express their recording preferences to camera-equipped devices (Denning et al., 2014; Koelle et al., 2018). Additionally, research has suggested using facial recognition techniques to automatically blur or obscure faces of individuals (e.g., bystanders) in video recording, thus circumventing potential ethical or medico-legal issues (Senior & Pankanti, 2011). Such approaches can be considered when deploying smart glass applications in sensitive contexts like emergency care.

#### 6.2.5. Providing training to increase user adoption of smart glasses

We identified that participants' lack of familiarity with the voice command library and correct hand posture were key factors contributing to the ineffective task performance when using touchless interaction methods. This finding is consistent with previous studies that reported similar issues (Pino et al., 2013; Sambrooks & Wilkinson, 2013). Therefore, it is crucial to provide the necessary instruction and comprehensive training on the proper use of and interaction with the device to enhance the uptake and effective utilization of smart glasses by medical providers. With training, EMS providers can improve their recall of voice commands and develop muscle memory for performing hand gestures. In line with research on health information technology implementation (Sittig & Singh, 2015), we believe that offering not only initial onboarding training but also regular training sessions to EMS users is essential. This approach promotes user adaptation and facilitates the integration of smart glasses into their actual work practices (Park & Chen, 2012). Moreover, the level of technological proficiency among users plays a significant role in the adoption of touchless interaction methods. Some tech-savvy medical providers may find it easier to learn and use these methods, while others may not. Therefore, training programs should be tailored to the needs of individuals who require additional support and more time to become familiar with operating the device. Lastly, after the system deployment, it is beneficial to regularly assess the actual usage of the device in real-world scenarios and collect user feedback. Such data can provide insights for improving the design of smart glasses and addressing any unexpected system issues or unintended consequences. By actively monitoring user experiences and committing to continuous improvement, we can ensure that smart glasses meet the evolving needs and expectations of EMS providers.

### 6.3. Limitations and future work

Our study has several limitations that should be acknowledged. First, the training provided to participants on how to interact with smart glasses was not extensive enough. A more comprehensive training could have improved users' familiarity with touchless interaction methods and potentially led to better task performance. Second, our evaluation of the interaction methods was conducted in controlled environments. To address this study limitation, our next study phase will involve conducting evaluations in simulated events, which will allow us to uncover additional usability issues and gain a deeper understanding of how each interaction method performs in more realistic scenarios. Third, the number of participants in the usability testing was not large, even though the power analysis indicated that 16 participants would meet the minimum requirements for our study's sample size. Access to EMS providers is challenging due to their exhaustive and high-stress work schedules (e.g., 8-h or 12-h shifts). In future work, we plan to recruit more EMS providers by collaborating with EMS agencies beyond the four that participated in the current study. Lastly, it is important to note that the usability and user experiences of voice commands and pinch gestures in our study could be influenced by the quality and performance of the software used. For instance, using more sophisticated gesture-sensing software could potentially enhance the task performance of the gesture-based method. Therefore, some of the issues identified with using touchless interaction methods, such as difficulties in recognizing voice commands or pinch gestures, may not be applicable to other devices and applications with different software implementations.

## 7. Conclusions

In this paper, we present an exploratory investigation into the user needs and experiences of touchless interaction methods for smart glasses in the context of EMS work. Despite recent advancements, our study reveals that touchless interaction methods still have limitations in facilitating seamless interactions with smart glasses compared to tangible methods. However, the less effective task performance of touchless interaction methods could be attributed to participants' lack of familiarity with these new methods and to limitations of the software. More importantly, it is noteworthy that voice commands and pinch gestures were rated higher in user preferences than buttons by our study participants. These findings highlight the importance of improving the technical performance and usability of touchless interaction methods to better meet user needs and expectations. Furthermore, it is crucial to account for the distinct characteristics of the EMS environment when designing smart glass interactions for dynamic settings. Beyond technical constraints, smart glass designers and manufacturers must make trade-off decisions by balancing factors such as technology maturity and performance, user preferences, social implications, and other contextual limitations. Our research uncovers many research opportunities for designing intuitive interaction experiences for smart glasses that can enhance the efficiency and effectiveness of fast-paced medical work and similar contexts. The insights from this study can inform future design efforts to create seamless and intuitive interactions that empower professionals in time-critical situations.

## Notes

1. https://intercom.help/vuzix/en/articles/5954802-overview.
2. https://www.crunchfish.com/gestures/.


## Acknowledgements

We extend our heartfelt gratitude to all the EMS providers who participated in our research. We would also like to express our sincere appreciation to Jack Finkelstein at Interfaith Medical Center (New York City, New York), Dr. Jared Kutzin at Mout Sinai Health System (New York City, New York), Humaira Ali at Maimonides Medical Center (New York City, New York), Dr. Kathleen Adelgais at the University of Colorado and Children's Hospital Colorado (Aurora, Colorado), and Brittany Olvera at Colorado Springs Fire Department (Colorado Springs, Colorado) for their support in coordinating the studies and recruiting participants.


## Disclosure statement

No potential conflict of interest was reported by the author(s).


## Funding

This study was supported by the National Science Foundation [grant# 1948292] and the Agency for Healthcare Research and Quality [grant# 1R21HS028104-01A1].


## References


Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S., & Koh, J. (2012). *Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI*. (Microsoft Research TechReport MSR-TR-2012-111, 2, 30).

Aldaz, G., Shluzas, L. A., Pickham, D., Eris, O., Sadler, J., Joshi, S., & Leifer, L. (2015). Hands-free image capture, data tagging and transfer using Google Glass: A pilot study for improved wound care management. *PloS One*, *10*(4), e0121179. https://doi.org/10.1371/journal.pone.0121179

Almutairi, K., Ismail, A., Abdlerazek, S., & Elbakry, H. (2020). Development of smart healthcare system for visually impaired using speech recognition. *International Journal of Advanced Computer Science and Applications*, *11*(12), 647–654. https://doi.org/10.14569/IJACSA.2020.0111275

Azai, T., Ogawa, S., Otsuki, M., Shibata, F., & Kimura, A. (2017). *Selection and manipulation methods for a menu widget on the human forearm* [Paper presentation]. Proceedings of the ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems, Denver, Colorado. https://doi.org/10.1145/3027063.3052959

Begany, G. M., Sa, N., & Yuan, X. (2015). Factors affecting user perception of a spoken language vs. textual search interface: A content analysis. *Interacting with Computers*, *28*(2), iwv029. https://doi.org/10.1093/iwc/iwv029

Berndt, H., Mentler, T., & Herczeg, M. (2016). Smartglasses for the triage of casualties and the identification of hazardous materials: How smartglasses can help emergency medical services managing challenging rescue missions. *i-com*, *15*(2), 145–153. https://doi.org/10.1515/icom-2016-0024

Blair, E. (2015). A reflexive exploration of two qualitative data coding techniques. *Journal of Methods and Measurement in the Social Sciences*, *6*(1), 14–29. https://doi.org/10.2458/v6i1.18772

Braun, V., & Clarke, V. (2012). *Thematic analysis*. American Psychological Association.

Broach, J., Hart, A., Griswold, M., Lai, J., Boyer, E. W., Skolnik, A. B., & Chai, P. R. (2018). *Usability and reliability of smart glasses for secondary triage during mass casualty incidents* [Paper presentation]. Proceedings of the Annual Hawaii International Conference on System Sciences, Hawaii, USA. https://doi.org/10.24251/HICSS.2018.175

Bulling, A., Dachselt, R., Duchowski, A., Jacob, R., Stellmach, S., & Sundstedt, V. (2012). *Gaze interaction in the post-WIMP world* [Paper presentation]. Extended Abstracts of the ACM CHI Conference on Human Factors in Computing Systems, Austin, TX. https://doi.org/10.1145/2212776.2212428

Chai, P. R., Wu, R. Y., Ranney, M. L., Porter, P. S., Babu, K. M., & Boyer, E. W. (2014). The virtual toxicology service: Wearable head-mounted devices for medical toxicology. *Journal of Medical Toxicology:*

*Official Journal of the American College of Medical Toxicology*, *10*(4), 382–387. https://doi.org/10.1007/s13181-014-0420-5

Chapman Smith, S. N., Brown, P. C., Waits, K. H., Wong, J. S., Bhatti, M. S., Toqeer, Q., Ricks, J. V., Stockner, M. L., Habtamu, T., Seelam, J., Britt, R. C., Giovia, J. M., Blankson, B. K., Bennam, P., Gormley, M. A., Lu, J., & Ornato, J. P. (2019). Development and evaluation of a user-centered mobile telestroke platform. *Telemedicine Journal and e-Health: The Official Journal of the American Telemedicine Association*, *25*(7), 638–648. https://doi.org/10.1089/tmj.2018.0044

Cho, S. J., Kwon, I. H., & Jeong, J. (2015). Application of telemedicine system to prehospital medical control. *Healthcare Informatics Research*, *21*(3), 196–200. https://doi.org/10.4258/hir.2015.21.3.196

Cicero, M. X., Walsh, B., Solad, Y., Whitfill, T., Paesano, G., Kim, K., Baum, C. R., & Cone, D. C. (2015). Do you see what I see? Insights from using google glass for disaster telemedicine triage. *Prehospital and Disaster Medicine*, *30*(1), 4–8. https://doi.org/10.1017/S1049023X1400140X

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D., & Bandeira, N. (2017). *"What can I help you with?" Infrequent users' experiences of intelligent personal assistants* [Paper presentation]. Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, Vienna, Austria.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.

Denning, T., Dehlawi, Z., & Kohno, T. (2014). *In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Toronto, Canada.

Dobbelstein, D., Hock, P., & Rukzio, E. (2015). *Belt: An unobtrusive touch input device for head-worn displays* [Paper presentation]. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea.

Ens, B., Byagowi, A., Han, T., Hincapié-Ramos, J. D., & Irani, P. (2016). *Combining ring input with hand tracking for precise, natural interaction with spatial analytic interfaces* [Paper presentation]. Proceedings of the 2016 Symposium on Spatial User Interaction, Tokyo, Japan. https://doi.org/10.1145/2983310.2985757

Faiola, A., Belkacem, I., Bergey, D., Pecci, I., & Martin, B. (2019). *Towards the design of a smart glasses application for MICU decision-support: Assessing the human factors impact of data portability & accessibility* [Paper presentation]. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care, Chicago, IL. https://doi.org/10.1177/2327857919081012

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Firouzian, A., Pulli, P., Pleva, M., Juhar, J., & Ondas, S. (2017). *Speech interface dialog with smart glasses* [Paper presentation]. 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia.

Follmann, A., Ohligs, M., Hochhausen, N., Beckers, S. K., Rossaint, R., & Czaplik, M. (2019). Technical support by smart glasses during a mass casualty incident: A randomized controlled simulation trial on technically assisted triage and telemedical app use in disaster medicine. *Journal of Medical Internet Research*, *21*(1), e11939. https://doi.org/10.2196/11939

Gausche-Hill, M., Krug, S., & Wright, J. (2021). Emergency medical services (EMS) 2050: A vision for the future of pediatric prehospital care. *Prehospital Emergency Care*, *25*(1), 91–94. https://doi.org/10.1080/10903127.2020.1734123

Geisler, F., Kunz, A., Winter, B., Rozanski, M., Waldschmidt, C., Weber, J. E., Wendt, M., Zieschang, K., Ebinger, M., & Audebert, H. J. (2019). Telemedicine in prehospital acute stroke care. *Journal of the American Heart Association*, *8*(6), e011729. https://doi.org/10.1161/JAHA.118.011729

Gjoreski, H., Mavridou, I., Archer, J. A. W., Cleal, A., Stankoski, S., Kiprijanovska, I., Fatoorechi, M., Walas, P., Broulidakis, J., & Gjoreski, M. (2023). *OCOsense glasses–monitoring facial gestures and expressions for augmented human-computer interaction: OCOsense glasses for monitoring facial gestures and expressions* [Paper presentation]. Extended Abstracts of the ACM CHI Conference on Human Factors in Computing Systems, Hamburg, Germany.

Goel, M., Zhao, C., Vinisha, R., & Patel, S. N. (2015). *Tongue-in-cheek: Using wireless signals to enable non-intrusive and flexible facial gestures detection* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Seoul, South Korea.

Grossman, T., Chen, X. A., & Fitzmaurice, G. (2015). *Typing on glasses: Adapting text entry to smart eyewear* [Paper presentation]. Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, Copenhagen, Denmark.

Guimbretière, F., & Nguyen, C. (2012). *Bimanual marking menu for near surface interactions* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Austin, TX. https://doi.org/10.1145/2207676.2208521

Ha, T., Feiner, S., & Woo, W. (2014). *WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception* [Paper presentation]. IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany.

Ham, J., Hong, J., Jang, Y., Ko, S. H., & Woo, W. (2014, June 22–27). *Smart wristband: Touch-and-motion–tracking wearable 3D input device for smart glasses* [Paper presentation]. Distributed, Ambient, and Pervasive Interactions: Second International Conference, DAPI 2014, Heraklion, Crete, Greece.

Haque, F., Nancel, M., & Vogel, D. (2015). *Myopoint: Pointing and clicking using forearm mounted electromyography and inertial motion sensors* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Seoul, South Korea.

Hertzum, M., Manikas, M. I., & á Torkilsheyggi, A. (2019). Grappling with the future: The messiness of pilot implementation in information systems design. *Health Informatics Journal*, *25*(2), 372–388. https://doi.org/10.1177/1460458217712058

Holton, J. A. (2007). The coding process and its challenges. In J. Holton (Ed.), The Sage *handbook of grounded theory* (Vol. 3, pp. 265–289). Sociology Press.

Hsieh, Y.-T., Jylhä, A., & Jacucci, G. (2014, October 30–31). *Pointing and selecting with tactile glove in 3D environment* [Paper presentation]. Symbiotic Interaction: Third International Workshop, Helsinki, Finland.

Hsieh, Y.-T., Jylhä, A., Orso, V., Gamberini, L., & Jacucci, G. (2016). *Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, San Jose, CA. https://doi.org/10.1145/2858036.2858436

Huang, Z., Li, W., Hui, P. (2015). Ubii: Towards seamless interaction between digital and physical worlds. *Proceedings of the 23rd ACM International Conference on Multimedia*, Montreal, Canada.

Istance, H., Bates, R., Hyrskykari, A., & Vickers, S. (2008). *Snap clutch, a moded approach to solving the Midas touch problem* [Paper presentation]. Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, Savannah, TX. https://doi.org/10.1145/1344471.1344523

Jagannath, S., Sarcevic, A., Young, V., & Myers, S. (2019). *Temporal rhythms and patterns of electronic documentation in time-critical medical work* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Glasgow, UK. https://doi.org/10.1145/3290605.3300564

Jones, E., Alexander, J., Andreou, A., Irani, P., & Subramanian, S. (2010). *GesText: Accelerometer-based gestural text-entry systems* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Atlanta, GA.

Kim, M., Choi, S. H., Park, K.-B., & Lee, J. Y. (2019). User Interactions for Augmented Reality Smart Glasses: A comparative evaluation of visual contexts and interaction gestures. *Applied Sciences*, *9*(15), 3171. https://doi.org/10.3390/app9153171

Klinker, K., Wiesche, M., & Krcmar, H. (2020a). Digital transformation in health care: Augmented reality for hands-free service innovation. *Information Systems Frontiers*, *22*(6), 1419–1431. https://doi.org/10.1007/s10796-019-09937-7

Klinker, K., Wiesche, M., & Krcmar, H. (2020b). *Smart glasses in health care: A patient trust perspective* [Paper presentation]. Proceedings of the 53rd Hawaii International Conference on System Sciences, Hawaii, USA. https://doi.org/10.24251/HICSS.2020.435

Koelle, M., Ananthanarayan, S., Czupalla, S., Heuten, W., & Boll, S. (2018). *Your smart glasses' camera bothers me! Exploring opt-in and opt-out gestures for privacy mediation* [Paper presentation]. Proceedings of the 10th Nordic Conference on Human-Computer Interaction, Oslo, Norway.

Koelle, M., Kranz, M., & Möller, A. (2015). *Don't look at me that way! Understanding user attitudes towards data glasses usage* [Paper presentation]. Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, Copenhagen, Denmark.

Kollee, B., Kratz, S., & Dunnigan, A. (2014). *Exploring gestural interaction in smart spaces using head mounted devices with ego-centric sensing* [Paper presentation]. Proceedings of the 2nd ACM Symposium on Spatial User Interaction, Honolulu, HI. https://doi.org/10.1145/2659766.2659781

Kyung, K.-U., Lee, J.-Y., & Park, J. (2008). Haptic stylus and empirical studies on braille, button, and texture display. *Journal of Biomedicine & Biotechnology*, *2008*, 369651–369611. https://doi.org/10.1155/2008/369651

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, *33*(2), 363–374. https://doi.org/10.2307/2529786

Lee, L. H., Braud, T., Bijarbooneh, F. H., & Hui, P. (2019). *Tipoint: Detecting fingertip for mid-air interaction on computational resource constrained smartglasses* [Paper presentation]. Proceedings of the 23rd International Symposium on Wearable Computers, London, UK.

Lee, L.-H., & Hui, P. (2018). Interaction methods for smart glasses: A survey. In *IEEE Access* (Vol. 6, pp. 28712–28732). https://doi.org/10.1109/ACCESS.2018.2831081

Lee, L. H., Lam, K. Y., Yau, Y. P., Braud, T., & Hui, P. (2019). *Hibey: Hide the keyboard in augmented reality* [Paper presentation]. IEEE International Conference on Pervasive Computing and Communications, Kyoto, Japan.

Lucero, A., & Vetek, A. (2014). *NotifEye: Using interactive glasses to deal with notifications while walking in public* [Paper presentation]. Proceedings of the 11th Conference on Advances in Computer Entertainment Technology, Funchal, Madeira, Portugal.

Maruri, H. A. C., Lopez-Meyer, P., Huang, J., Beltman, W. M., Nachman, L., & Lu, H. (2018). V-Speech: Noise-robust speech capturing glasses using vibration sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(4), 1–23. https://doi.org/10.1145/3287058

Matthies, D. J., Weerasinghe, C., Urban, B., & Nanayakkara, S. (2021). *Capglasses: Untethered capacitive sensing with smart glasses* [Paper presentation]. Proceedings of the Augmented Humans International Conference, Rovaniemi, Finland. https://doi.org/10.1145/3458709.3458945

Mitrasinovic, S., Camacho, E., Trivedi, N., Logan, J., Campbell, C., Zilinyi, R., Lieber, B., Bruce, E., Taylor, B., Martineau, D., Dumont, E. L. P., Appelboom, G., & Connolly, E. S. (2015). Clinical and surgical applications of smart glasses. *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine*, *23*(4), 381–401. https://doi.org/10.3233/THC-150910

Mohammed, A. A., Lv, J., Islam, M. S., & Sang, Y. (2023). Multi-model ensemble gesture recognition network for high-accuracy dynamic hand gesture recognition. *Journal of Ambient Intelligence and Humanized Computing*, *14*(6), 6829–6842. https://doi.org/10.1007/s12652-021-03546-6

Montero, C. S., Alexander, J., Marshall, M. T., & Subramanian, S. (2010). *Would you do that? Understanding social acceptance of gestural interfaces* [Paper presentation]. Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services, Lisbon, Portugal.

Muensterer, O. J., Lacher, M., Zoeller, C., Bronstein, M., & Kübler, J. (2014). Google Glass in pediatric surgery: An exploratory study. *International Journal of Surgery (London, England)*, *12*(4), 281–289. https://doi.org/10.1016/j.ijsu.2014.02.003

Munusamy, T., Karuppiah, R., Bahuri, N. F. A., Sockalingam, S., Cham, C. Y., & Waran, V. (2021). Telemedicine via smart glasses in critical care of the neurosurgical patient—COVID-19 pandemic preparedness and response in neurosurgery. *World Neurosurgery*, *145*, e53–e60. https://doi.org/10.1016/j.wneu.2020.09.076

Murad, C., Munteanu, C., Clark, L., & Cowan, B. R. (2018). *Design guidelines for hands-free speech interaction* [Paper presentation]. Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, Barcelona Spain. https://doi.org/10.1145/3236112.3236149

Myers, C., Furqan, A., Nebolsky, J., Caro, K., & Zhu, J. (2018). *Patterns for how users overcome obstacles in voice user interfaces* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Montreal, Canada. https://doi.org/10.1145/3173574.3173580

Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Boston, MA. https://doi.org/10.1145/259963.260333

Noorian, A. R., Hosseini, M. B., Avila, G., Gerardi, R., Andrle, A.-F., Su, M., Starkman, S., Saver, J. L., & Sharma, L. K. (2019). Use of wearable technology in remote evaluation of acute stroke patients: Feasibility and reliability of a Google Glass-based device. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, *28*(10), 104258. https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.06.016

Odenheimer, S., Goyal, D., Jones, V. G., Rosenblum, R., Ho, L., & Chan, A. S. (2018). Patient acceptance of remote scribing powered by Google glass in outpatient dermatology: Cross-sectional study. *Journal of Medical Internet Research*, *20*(6), e10762. https://doi.org/10.2196/10762

Ok, A. E., Basoglu, N. A., & Daim, T. (2015). *Exploring the design factors of smart glasses* [Paper presentation]. Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR.

Park, S. Y., & Chen, Y. (2012). *Adaptation as design: Learning from an EMR deployment study* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Austin, TX.

Pilerot, O., & Maurin Söderholm, H. (2019). *A conceptual framework for investigating documentary practices in prehospital emergency care* [Paper presentation]. Proceedings of the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia.

Pino, A., Tzemis, E., Ioannou, N., & Kouroupetroglou, G. (2013). *Using kinect for 2D and 3D pointing tasks: Performance evaluation* [Paper presentation]. International Conference on Human-Computer Interaction, Las Vegas, NV.

Prilla, M., Janßen, M., & Kunzendorff, T. (2019). How to interact with augmented reality head mounted devices in care work? A study comparing handheld touch (hands-on) and gesture (hands-free) interaction. *AIS Transactions on Human-Computer Interaction*, *11*(3), 157–178. https://doi.org/10.17705/1thci.00118

Prilla, M., & Mantel, A. M. (2021). Analysing a UI's impact on the usability of hands-free interaction on smart glasses. *IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Virtual.

Rico, J., & Brewster, S. (2010). *Usable gestures for mobile interfaces: Evaluating social acceptability* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA.

Rogers, H., Madathil, K. C., Agnisarman, S., Narasimha, S., Ashok, A., Nair, A., Welch, B. M., & McElligott, J. T. (2017). A systematic review of the implementation challenges of telemedicine systems in ambulances. *Telemedicine Journal and e-Health: The Official Journal of the American Telemedicine Association*, *23*(9), 707–717. https://doi.org/10.1089/tmj.2016.0248

Romare, C., & Skär, L. (2020). Smart glasses for caring situations in complex care environments: Scoping review. *JMIR mHealth and uHealth*, *8*(4), e16055. https://doi.org/10.2196/16055

Safi, S., Thiessen, T., & Schmailzl, K. J. (2018). Acceptance and resistance of new digital technologies in medicine: Qualitative study. *JMIR Research Protocols*, *7*(12), e11072. https://doi.org/10.2196/11072

Sambrooks, L., & Wilkinson, B. (2013). *Comparison of gestural, touch, and mouse interaction with Fitts' law* [Paper presentation]. Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, Adelaide Australia. https://doi.org/10.1145/2541016.2541066

Saponas, T. S., Kelly, D., Parviz, B. A., & Tan, D. S. (2009). *Optically sensing tongue gestures for computer input* [Paper presentation]. Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, Victoria, Canada. https://doi.org/10.1145/1622176.1622209

Schaer, R., Melly, T., Muller, H., & Widmer, A. (2016). *Using smart glasses in medical emergency situations, a qualitative pilot study* [Paper presentation]. IEEE Wireless Health, Bethesda, MD.

Schlosser, P., Matthews, B., Salisbury, I., Sanderson, P., & Hayes, S. (2021). *Head-worn displays for emergency medical services staff: Properties of prehospital work, use cases, and design considerations* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Hamburg, Germany.

Senior, A. W., & Pankanti, S. (2011). Privacy protection and face recognition. In *Handbook of face recognition* (pp. 671–691). Springer.

Serrano, M., Ens, B. M., & Irani, P. P. (2014). *Exploring the use of hand-to-face input for interacting with head-worn displays* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Toronto, Canada. https://doi.org/10.1145/2556288.2556984

Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). *Designing the user interface: Strategies for effective human-computer interaction*. Pearson.

Sittig, D. F., & Singh, H. (2015). A new socio-technical model for studying health information technology in complex adaptive healthcare systems. In *Cognitive informatics for biomedicine* (pp. 59–80). Springer.

Suhm, B. (2003). *Towards best practices for speech user interface design* [Paper presentation]. Eighth European Conference on Speech Communication and Technology, Geneva, Switzerland.

Syberfeldt, A., Danielsson, O., & Gustavsson, P. (2017). Augmented reality smart glasses in the smart factory: Product evaluation guidelines and review of available products. *IEEE Access*, *5*, 9118–9130. https://doi.org/10.1109/ACCESS.2017.2703952

Toyama, T., Sonntag, D., Dengel, A., Matsuda, T., Iwamura, M., & Kise, K. (2014). *A mixed reality head-mounted text translation system using eye gaze input* [Paper presentation]. Proceedings of the 19th International Conference on Intelligent User Interfaces, Los Angeles, CA. https://doi.org/10.1145/2557500.2557528

Tran, B. D., Latif, K., Reynolds, T. L., Park, J., Elston Lafata, J., Tai-Seale, M., & Zheng, K. (2023). "Mm-hm,""Uh-uh": Are non-lexical conversational sounds deal breakers for the ambient clinical documentation technology? *Journal of the American Medical Informatics Association*, *30*(4), 703–711. https://doi.org/10.1093/jamia/ocad001

Tung, Y.-C., Hsu, C.-Y., Wang, H.-Y., Chyou, S., Lin, J.-W., Wu, P.-J., Valstar, A., & Chen, M. Y. (2015). *User-defined game input for smart glasses in public space* [Paper presentation]. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Seoul, South Korea. https://doi.org/10.1145/2702123.2702214

Wang, C.-Y., Chu, W.-C., Chiu, P.-T., Hsiu, M.-C., Chiang, Y.-H., & Chen, M. Y. (2015). *PalmType: Using palms as keyboards for smart glasses* [Paper presentation]. Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, Copenhagen, Denmark.

Weibel, N., Gasques, D., Johnson, J., Sharkey, T., Xu, Z. R., Zhang, X., Zavala, E., Yip, M., & Davis, K. (2020). *Artemis: Mixed-reality environment for immersive surgical telementoring* [Paper presentation]. Extended Abstracts of the ACM CHI Conference on Human Factors in Computing Systems, Virtual. https://doi.org/10.1145/3334480.3383169

Yi, S., Qin, Z., Novak, E., Yin, Y., & Li, Q. (2016). *Glassgesture: Exploring head gesture interface of smart glasses* [Paper presentation]. The 35th Annual IEEE International Conference on Computer Communications, San Francisco, CA.

Yperzeele, L., Van Hooff, R.-J., De Smedt, A., Valenzuela Espinoza, A., Van Dyck, R., Van de Casseye, R., Convents, A., Hubloue, I., Lauwaert, D., De Keyser, J., & Brouns, R. (2014). Feasibility of AmbulanCe-Based Telemedicine (FACT) study: Safety, feasibility and reliability of third generation in-ambulance telemedicine. *PLoS One*, *9*(10), e110043. https://doi.org/10.1371/journal.pone.0110043

Yu, J., Ferniany, W., Guthrie, B., Parekh, S. G., & Ponce, B. (2016). Lessons learned from Google Glass: Telemedical spark or unfulfilled promise? *Surgical Innovation*, *23*(2), 156–165. https://doi.org/10.1177/1553350615597085

Zhang, J., Wu, J., Qiu, Y., Song, A., Li, W., Li, X., & Liu, Y. (2023). Intelligent speech technologies for transcription, disease diagnosis, and medical equipment interactive control in smart hospitals: A review. *Computers in Biology and Medicine*, *153*, 106517. https://doi.org/10.1016/j.compbiomed.2022.106517

Zhang, Z., Joy, K., Harris, R., Ozkaynak, M., Adelgais, K., & Munjal, K. (2022). Applications and user perceptions of smart glasses in emergency medical services: Semistructured interview study. *JMIR Human Factors*, *9*(1), e30883. https://doi.org/10.2196/30883

Zhang, Z., Joy, K., Upadhyayula, P., Ozkaynak, M., Harris, R., & Adelgais, K. (2021). Data work and decision making in emergency medical services: A distributed cognition perspective. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–32. https://doi.org/10.1145/3479500

Zhang, Z., Luo, X., Harris, R., George, S., & Finkelstein, J. (2022). *Hands-free electronic documentation in emergency care work through smart glasses* [Paper presentation]. International Conference on Information, Virtual.

Zhang, Z., Ramiya Ramesh Babu, N. A., Adelgais, K., & Ozkaynak, M. (2022). Designing and implementing smart glass technology for emergency medical services: A sociotechnical perspective. *JAMIA Open*, *5*(4), ooac113. https://doi.org/10.1093/jamiaopen/ooac113

Zhang, Z., Sarcevic, A., & Bossen, C. (2017). *Constructing common information spaces across distributed emergency medical teams* [Paper presentation]. Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, OR. https://doi.org/10.1145/2998181.2998328

Zhang, Z., Sarcevic, A., Joy, K., Ozkaynak, M., & Adelgais, K. (2021). *User needs and challenges in information sharing between pre-hospital and hospital emergency care providers* [Paper presentation]. AMIA Annual Symposium Proceedings, San Diego, CA.

Zuidhof, N., Ben Allouch, S., Peters, O., & Verbeek, P.-P. (2021). Defining smart glasses: A rapid review of state-of-the-art perspectives and future challenges from a social sciences' perspective. *Augmented Human Research*, *6*(1), 1–18. https://doi.org/10.1007/s41133-021-00053-3

## About the authors

**Zhan Zhang** is an Associate Professor in the School of Computer Science and Information Systems at Pace University. His research spans the fields of Human-Computer Interaction and Health Informatics, with a particular focus on the design and development of innovative technologies that facilitate the work of fast-paced medical teams.

**Enze Bai** is a PhD candidate in Computer Science in the School of Computer Science and Information Systems at Pace University. Supervised by Dr. Zhan Zhang, his research focuses on developing smart glass applications to support patient information sharing and collection in pre-hospital care.

**Aram Stepanian** graduated from the School of Computer Science and Information Systems at Pace University with a concentration on user experience design. He was supervised by Dr. Zhan Zhang and is now a Senior User Experience Designer at Walmart Global Tech.

**Swathi Jagannath** is a User Experience Researcher at Microsoft currently working on Azure Cloud products. Her prior research focus was on healthcare information systems, specifically understanding activity delays in time-critical settings to inform the design of clinical decision-support systems.

**Sun Young Park** is an Associate Professor in Stamps School of Art and Design and the School of Information at the University of Michigan. Her research focuses on evaluating and designing health information systems and technologies in clinical and non-clinical settings.