

# Multimedia-Enabled 911: Exploring 911 Callers' Experience of Call Taker Controlled Video Calling in Simulated Emergencies

Punyashlok Dash  
School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, Canada  
punyashlok\_dash@sfu.ca

Benett Axtell  
School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, Canada  
benett\_axtell@sfu.ca

Denise Y. Geiskkovitch  
Department of Computing and  
Software, McMaster University,  
Hamilton, Canada  
geiskkod@mcmaster.ca

Carman Neustaedter  
School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, Canada  
carman@sfu.ca

Wolfgang Stuerzlinger  
School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, Canada  
wodom@sfu.ca

## ABSTRACT

Emergency response to large-scale disasters is often supported with multimedia from social media. However, while these features are common in everyday video calls, the complex needs of 911 and other systems make it difficult to directly incorporate these features. We assess an ME911 (Multimedia-Enabled 911) app to understand how the design will need to deviate from common norms and how callers will respond to those non-standard choices. We expand the role of 911 call taker control over emergency situations to the calling interface while incorporating key features like map-based location finding. Participants' experiences in mock emergencies show the non-standard design helps callers in the unfamiliar setting of emergency calling yet it also causes confusion and delays. We find the need for emergency-specific deviations from design norms is supported by participant feedback. We discuss how broader system changes will support callers to use these non-standard designs during emergencies.

## CCS CONCEPTS

• **Human-centered computing** → Interaction design; Interaction design process and methods; User interface design; Human computer interaction (HCI); HCI design and evaluation methods; User studies.

## KEYWORDS

Emergency response, 911 calling, Non-standard design, Video call, User assessment

## ACM Reference Format:

Punyashlok Dash, Benett Axtell, Denise Y. Geiskkovitch, Carman Neustaedter, and Wolfgang Stuerzlinger. 2024. Multimedia-Enabled 911:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3643055>

Exploring 911 Callers' Experience of Call Taker Controlled Video Calling in Simulated Emergencies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3613904.3643055>

## 1 INTRODUCTION

When someone witnesses or experiences an emergency, they are generally able to call an emergency number, for example 911 in North America, and reach a dispatch centre where a trained call taker will assess the situation and dispatch the appropriate responders. Call takers use a broad array of software to determine and dispatch services to support the given emergency [12, 36]. Recently, during large-scale emergencies like natural disasters, call takers have incorporated posts and multimedia from Twitter (now known as X) and other social media sites, which people are using to report and document their situation [21, 30, 39]. The multimedia included in these tools, including photos, videos, and location, provide valuable information to call takers to deliver timely emergency support [4, 35].

However, call takers and emergency services are not able to take advantage of such multimedia during small-scale emergencies, reported by one or a few people calling 911, because 911 still relies on audio-only telephone calls for information gathering. While video calling with smartphones is commonplace, existing options, like Skype or Facetime, are not designed with emergency calling in mind and lack key considerations for 911 calls, like enabling a call taker to control a caller's interface [35]. National organizations are working towards these needed changes to emergency calling systems, such as by introducing needed network changes to support "next generation 911" features in Canada [4, 5].

In this paper, we explore the potential of multimedia-enabled 911 (ME911) with a 911 calling app prototype, targeted at small-scale emergencies like car crashes and robberies. We present the design of this ME911 app based on the needs and expectations of ME911 calls from call takers and callers [35, 40]. Our design enables call taker control over interactive elements of the caller's app (e.g., making features visible only as needed) to fit the structured process call takers use to collect needed information and dispatch emergency

services efficiently [35]. This structured process provides call takers with a question tree to communicate standardized emergency response cards in sequence and walk callers through a very scripted conversation [17]. In parallel with this, call takers are trained to take control of the call to manage caller stress or panic as well as to determine what support is needed as quickly as possible [18]. Because of this, a ME911 interface will not necessarily follow common design norms, and we use our prototype to understand how potential 911 callers respond to these unfamiliar design elements in the context of ME911.

We presented participants with five scenarios in a controlled lab environment, each targeting specific aspects of ME911 interactions and how they differ from common experiences of video calling or multimedia sharing. Through these scenarios and semi-structured interviews, we observed potential 911 callers as they encountered how ME911 calling is different both from familiar video calling tools and from audio-only 911 calls. Our work builds on previous explorations into the needs of both callers and call takers for ME911 [35, 40]. We expand on this by reporting participants' direct experience with ME911 calling using realistic scenarios and actual emergency footage in a controlled setting. We discuss their experiences with call taker interface control, including both struggling with and relying on the lack of control, as well as using the app to follow directions for managing the camera or emergency care.

In this paper, we present two main contributions:

- First, design considerations for the non-standard needs of ME911, supporting key features determined by call takers, including the design of our ME911 app as one implementation of these.
- Second, observations of potential callers' use of this app and how they react to and interact with their lack of control during a (simulated) emergency.

With our contributions, future ME911 designs can be created to better support the sometimes-conflicting needs of call takers and callers while considering the privacy needs of both sides.

## 2 RELATED WORK

We ground our research in existing systems and procedures of 911 calling. Typically, call takers ask callers a series of standardized questions [17], recording the collected information to be sent to dispatchers who forward it to first responders [36]. The call taker relies heavily on the information provided by the caller [31], but this is sometimes inaccurate or difficult to gather, due to various challenges, such as panicked callers or language and communication barriers [18, 25, 31]. Call takers effectively manage these issues by taking control [18, 25]. Prior literature has found that video calls could benefit 911 call takers by allowing them to assess emergencies with their first-hand experience [35]. Live video has also been seen to benefit other members of the emergency team, like first responders [29].

In this section, we review research from HCI and related fields on known needs of ME911, how multimedia is already being used in large-scale disasters, and understandings of smartphone video calling in everyday (non-emergency) settings.

### 2.1 Expected Benefits and Needs of ME911

Currently 911 calls in Canada are audio-only, but the benefits of live video to emergency response are well known, as seen in the use of video from social media during large-scale disasters [21, 39]. A field study of video broadcasting by on-scene emergency responders in Sweden found video was primarily used for time sensitive information and provides more detail for the rest of the response team than verbal reports [2].

Most of what is known about 911 calling and multimedia comes from interviews with both call takers and callers. Call takers have shared that viewing caller video would clarify information when callers are unsure of what they are seeing, have difficulty speaking, or miss a potential hazard [35]. People with experience calling 911 struggle to explain their precise location and details of the emergency [31]. They also agree with call takers that including video will allow call takers to interpret their inaccurate or imprecise descriptions [40]. Both callers and call takers share similar privacy and safety concerns around video sharing, such as risks to bystanders and protecting call takers from disturbing or upsetting scenes [35, 40]. Call takers feel that they should not share their video with callers as they are often busy checking various resources and communicating with other team members, so would not be giving their visual attention to the caller [35].

ME911 will also require systemwide technical and infrastructure changes, such as to incorporate the diverse channels of a ME911 system [32], to share information between emergency response teams using phones [27] or novel technologies like tabletops [20], and to add the display of multimedia to the large, complex displays used in dispatch centres [35]. Other works investigate needed improvements to phone networks to improve accuracy of cell phone location data to support the desired standards for emergencies [11]. Supporting ME911 will require coordinated changes across the full network of 911 calling and response systems, alongside emergency-specific interfaces for 911 callers.

### 2.2 Multimedia in Disaster Response

Publicly available multimedia, from sites like Twitter (now X) and Flickr, have supported emergency responses to large-scale emergencies but call takers and their teams are best supported by decentralized systems to manage these media [28, 30]. Additionally, call takers require access to the broader context to identify relevant media [30]. Some emergency teams already use social media for outreach, but generally simply duplicate what they communicate through more official channels [39] or cannot use them without bending their policy rules to effectively communicate with civilians [21]. Disaster response can also use multimedia within local communities, such as a neighbourhood map enabling individuals to share their personal status with the community as well as how they can help [7]. While the public nature of major disasters allows for sharing and gathering multimedia from many civilians and large emergency teams across emergency services, this is not yet possible for small-scale emergencies where only one or a few people call 911 and a subset of emergency services respond.

## 2.3 Camerawork in Everyday Settings

Smartphones and other mobile devices are commonly used to support video calling between friends and family, and require careful camerawork to sufficiently show the remote viewer a scene [24, 33]. Camerawork includes changing camera orientation, position, and zoom level to show a useful view [22]. This can require holding phones at awkward angles or other challenges, such as during family activities [22, 23, 37]. Improvements to this have focused mainly on providing or combining multiple camera views [14, 15, 26, 38]. In an emergency setting, call takers have described a need to efficiently direct callers' camerawork to show relevant aspects of the scene as part of their standardized processes [35].

Thus, while there is some research on emergency response and the possibilities of multimedia-enabled 911 calls, the potential of such an interface and how callers respond to the novel experience have both not been investigated, as we do in this work. Our work also complements related literature on everyday camerawork to explore camerawork in small-scale emergencies.

## 3 DESIGN MOTIVATION AND REQUIREMENTS FOR ME911

Our design work is part of a longer series of studies of 911 callers and call centres. To drive the design of our ME911 app, we look to previous research that worked directly with call takers (in their dispatch centres) [35] and people with experience calling 911 [31, 40]. With these works we were able to consider both perspectives in how a ME911 app might be used. Building on the key takeaways of these works together, we identified four main design guidelines for ME911:

1. **Call Taker Direction of Camerawork:** Call takers want the benefits of seeing the scene first-hand, but are also aware that their need to control the call will expand to controlling what is seen on the video as well, including directing caller camerawork [35].
2. **Call Taker Control Over Caller Interface:** Callers have shared that the strong emotions that come with calling 911 meant they sometimes did not think to use even the limited features of an audio-only call, like speaker phone [40]. The added features of ME911 risk similarly overwhelming callers and so we explore mitigating this by allowing call takers to control and change what is shown on the user interface depending on the situation and which questions are being asked in the call [40].
3. **Two-Way Media Sharing:** In addition to viewing multimedia themselves, call takers also see the benefit of sharing videos or images themselves, such as passing on visual information to emergency responders and providing video instructions to callers [35].
4. **Caller Location Finding with Map Support:** Despite GPS-enabled phones, explaining the exact location, and possibly directions to that location, is a commonly experienced difficulty for 911 callers [35]. This issue can be addressed by providing a map with clear landmarks for callers to identify. Also, relying solely on technology to explain location can be limited when lacking strong cellular connections [35]. This is also complicated by situations like callers on an overpass

above an accident or in far rural areas where GPS-based instructions can lead emergency responders to the wrong street level or unmarked dead ends [40].

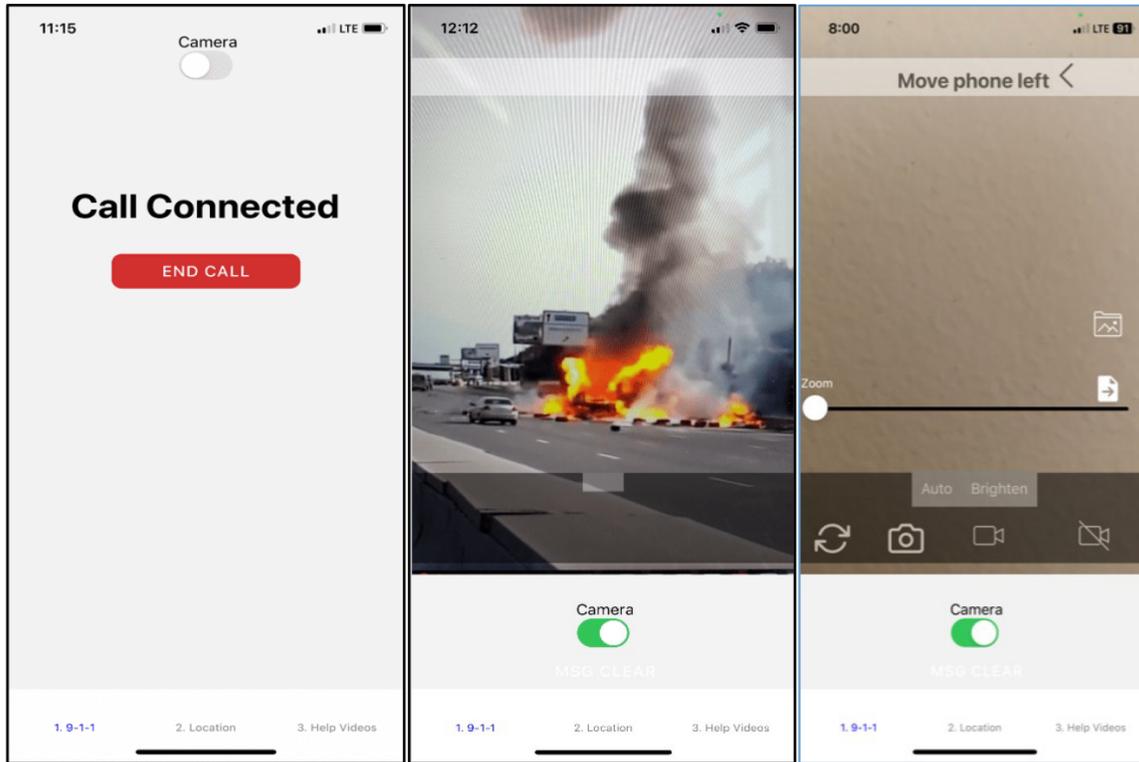
Here, we describe each of these themes within the context of our ME911 app. We emphasize where the design intentionally deviates from commonly accepted design norms to support the needs of emergency situations, primarily the need for the call taker to control the call. This is similar to how audio-only 911 calls deviate from standard phone calls, as the call taker directs the conversation and does not allow for the free back-and-forth conversation that is common in social settings [18]. In particular, the design needs to balance that callers want ME911 to be as easy as current audio-only 911 calls [40] and that call takers need a consistent level of control over all aspects of the call [35]. It is also important to recognize that use of ME911, like calling 911, is likely to be infrequent and possible have long gaps between uses [35]; thus, app features are unlikely to be familiar during each use by callers. Calling 911 is already an unfamiliar and non-standard experience and the call taker's scripted control over the call is used to help the caller to navigate the high stress situation without making it feel like a familiar social call. Using ME911 will require a similar balance between not being so unusual that it slows caller responses, but not so familiar that they do not stay aware of the urgency. Previous work into cognitive friction has shown how disrupting automatic and familiar patterns of interaction lead to more intentional interactions [9], so the unfamiliar interface can help callers to focus on the emergency at hand. Appendix A contains a descriptive example of using our ME911 app, and a video is included in the supplement.

### 3.1 Call Taker Direction of Camerawork

In our ME911 app, we enable video calling from the caller, but not from the call taker, in order to protect the privacy of the call taker [40], and incorporate how the call taker will direct the camerawork. The call is initially audio-only (Figure 1, left) and callers can enable their video with the camera toggle button (Figure 1, centre). While call taker control is key to an efficient 911 call, the call taker cannot perform camerawork (e.g., panning and zooming the camera) for the caller [35]. Our ME911 app allows call takers to guide callers to complete camerawork by overlaying on-screen text and/or visual instructions. For example, in the right of Figure 1, the caller's screen shows the text (with icon) "Move phone left <" to indicate that they need to pan the phone to the left. These visual camerawork instructions are shown on screen in addition to the call taker verbally describing them.

### 3.2 Call Taker Control Over Caller Interface

When considering future ME911 capabilities, call takers have acknowledged that emergency response cards and calling technology will need to be updated or replaced to support the inclusion of multimedia before ME911 can be used, particularly with regards to maintaining control [35]. Our ME911 app expands the call taker control to encompass the caller interface, giving call takers control over which app features are available to the caller. Specifically, callers have no access to camera features at the start of the call (Figure 1, centre) and the call taker can enable and disable features as needed. Table 1 lists all ME911 features with call taker control



**Figure 1:** Initial view of a call using ME911 showing camera toggle at top centre (left), ME911 call with video streaming (centre), ME911 interface showing all possible video features enabled and visual instructions (right).

**Table 1:** Call Taker Controlled Features in ME911

Feature	Task	Emergency Function and Rationale
Zoom video	Zoom video stream in and out	Call takers can better assess specific aspects of a scene with more detail and fewer distractions [40].
Flip video	Toggle between front and rear cameras	The default view is from the rear camera to show a scene in front of the caller. There are times that a caller may need to be visible, such as if they were injured and need to be assessed by the call taker [35].
Capture image	Capture a still image from the video stream	Image capture from mobile devices is generally higher quality than video screenshots that could be captured by the call taker. Call takers can review detailed still images more closely than moving video [35].
Record video	Start and stop recording of the video stream	Recorded video may be of higher quality than a live video stream, due to poor cell signal or limited capabilities of the caller's phone [35].
Brighten video	Auto enhance video brightness	When the scene is dark or at night, the video can be brightened to help call takers visually assess the scene [40].
Open gallery	View media stored on caller's phone	Caller can select media captured before or during a call to send to the call taker, as in the cases described above for capturing images or video [35].

and explains the motivation for each feature, as identified in previous work [35, 40], and how the feature could be needed in an emergency. Figure 1 (right) shows the ME911 interface with all possible video features enabled.

This method of revealing features only as needed is quite unusual to modern app designs and we expect callers to be unfamiliar with

buttons appearing on their screen. As discussed above, these departures from design norms are expected and even required to support the emergency setting [35]. To better understand how potential callers experience these design choices, we presented the app to our participants in scenarios with call taker control over the interface as well as with all features available throughout the mock 911 call. We discuss this in more detail in the Methods section below.

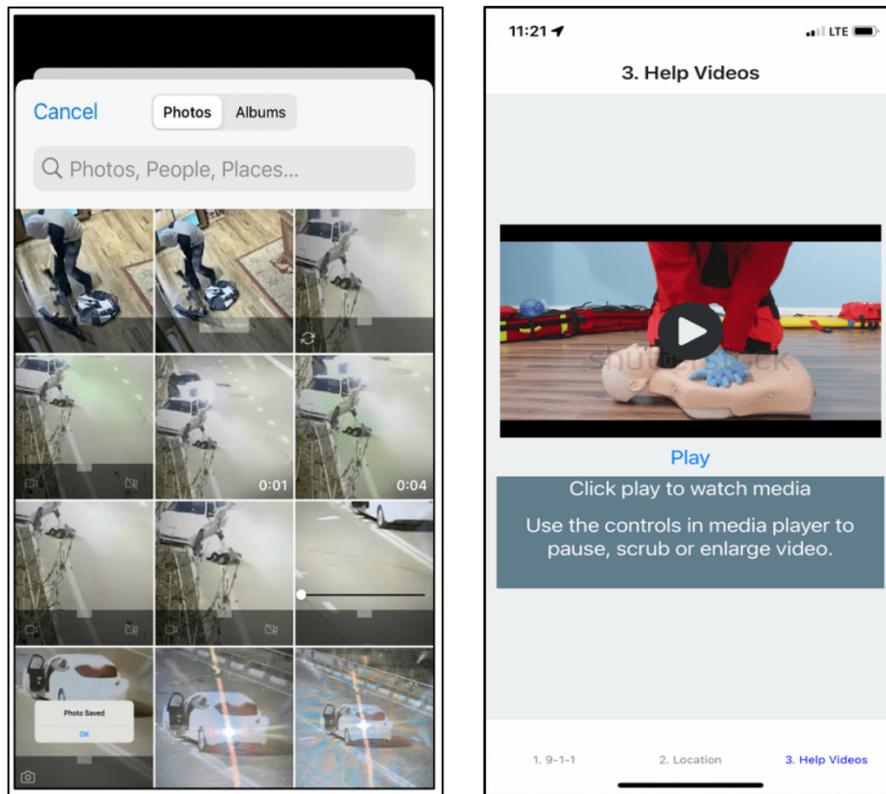


Figure 2: Sharing of recent video or images through ME911 (left), viewing an instructional video on CPR in ME911 (right).

### 3.3 Two-Way Media Sharing

Sharing multimedia is another opportunity to improve communication during 911 calls both for call takers and callers [31, 35, 40]. It is not always possible to stream live video (e.g., due to insufficient cell network bandwidth or phone limitations), and a moving video stream can be difficult for a call taker to assess, compared to a still image or a saved video that can be watched multiple times or paused [35]. Call takers also are called on to guide a caller through lifesaving manoeuvres, such as CPR (Cardio-Pulmonary Resuscitation), which could be facilitated by sending an instructional video file [35]. We consider both of these cases by introducing media sharing both by the caller and the call taker into ME911.

When technical issues limit a caller's ability to stream live video or when crucial information was captured before a 911 call (e.g., seeing a car crash in the background of an ongoing recording), callers can share videos or images that they captured before starting the call or capture and share media live during the call (Figure 2, left).

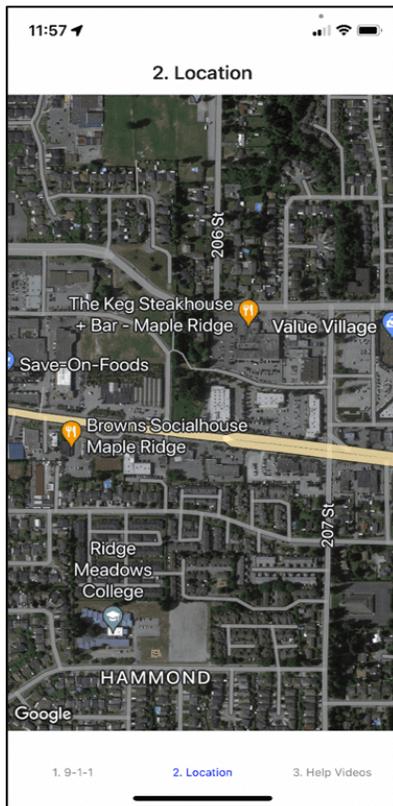
Instructional videos in the ME911 app are shared by call takers in a separate tab (Figure 2, right). While the caller is viewing this tab, the call taker continues to be able to see the video call from the caller's phone, enabling them to guide the caller in performing the given task by watching their technique. People with experience being guided verbally by a call taker to perform CPR have had difficulty with where to place or hold their phone so that they can

hear the instructions while also performing the tasks as directed [40]. Viewing a video could simplify this as a phone can be placed beside them where they can both see the screen and hear the call taker's directions.

### 3.4 Caller Location Finding with Map Support

A caller's location is a crucial aspect of 911 information gathering. While landlines provide call takers with a precise address, cell phone location is determined by cell towers and can be imprecise by up to half a kilometer [35]. Because of this, callers are currently required to describe their location based on major streets or other landmarks as part of the standard 911 call taker script. However, callers' descriptions can be inaccurate or imprecise and language or communication barriers can complicate this essential step [18, 25, 31]. Additionally, the specific location information needed in order to dispatch the needed services is often more complicated than the caller's coordinates [31, 40]. For example, if the caller is on an overpass looking at a traffic accident below them, sending an ambulance to the caller's location will not give responders access to the accident. In rural areas, where location services are often more limited, small roads or driveways may not be present on maps, so again the directions given to responders need to be aided by the caller's description.

To simplify the caller's role in describing the relevant location, our ME911 app includes a map view (Figure 3), using the Google



**Figure 3: Location finding in ME911 using a standard Google Maps view.**

Maps API (unchanged), showing street names and pins with location labels. Using this, callers can show call takers where they are visually and accurately name nearby streets and landmarks.

## 4 METHOD

In order to investigate how people experience calling 911 using our ME911 app, we selected five small-scale emergency scenarios, based on collaborations with 911 call takers and with access to emergency response cards, and presented these to participants using a counterbalanced within-participants study in a controlled lab setting. Our study focuses on the caller's experience of call taker control over their interface and how they respond to camerawork instruction. The within participants approach allowed the participants to experience scenarios twice: once with call taker control over the visibility of various features of the ME911 interface (the Call Taker Control condition) and once with all features fully available to the participants (the Full Functionality condition). We included these two conditions because we expected participants to be unfamiliar with the design of ME911. As discussed above, the needs of ME911 and of call takers require the design of our ME911 app to depart from accepted design norms, which might confuse or otherwise impact the use of the app by regular smartphone users and specifically their experience with video calling. Overall, we observe the potential impact of the ways our ME911 design deviates

from standard design for video calling, as is necessary to support the needs and expectations of emergency settings.

Ethical considerations informed all aspects of our research design. We chose a controlled lab setting because of the sensitive nature of this research. It is not technically feasible to perform this research in the wild during actual emergencies, as the infrastructure and other support tools (e.g., response cards) do not yet exist to support ME911 [35], and introducing untested communications tools into real emergencies puts all people involved at risk if any aspect should fail. We chose a fully simulated setting by showing videos of actual emergencies on a large immersive display over reenacting to ensure that participants felt safe but were aware of the seriousness of emergencies. We acknowledge that this is one approach to research in this area and thus expect future work to move towards more realistic evaluations of ME911, while continuing to prioritize the ethical concerns of both callers, call takers, and researchers. This protocol has been approved by our university's Research Ethics Board.

### 4.1 Participants

We conducted our study with 20 participants (11 men, 9 women). Their ages ranged from 19 to 61 ( $M=34$ ). We recruited our participants via university mailing lists and snowball sampling. Eight participants were between 18-25, seven between 26-35, and five over 35 years of age. All were regular smartphone users. Five had previously called emergency services.

### 4.2 Materials and Scenarios

We identified five scenarios for this study to include a range of common small-scale emergencies (requiring police, fire, and/or medical services), and also to require the use of the various ME911 features. We chose these scenarios in collaboration with call takers and by incorporating the emergency response cards into our study protocol. For this, we found YouTube videos showing actual footage of small-scale emergencies. The five scenarios are:

- **House fire** – The caller is on a walk and sees a house on fire, so they call 911. The call taker directs the caller's camerawork to show different parts of the house, to assess the fire's severity, as well as surrounding areas, to determine if there are additional risks.
- **Large truck fire on highway** – The caller is driving and sees a large truck carrying gasoline on fire, and pulls over to call 911. The call taker guides the caller to find and describe their precise location on the map and to use camerawork to show the whole scene.
- **Unconscious person** – The caller is walking and sees a passerby fall to the ground motionless. They go to the fallen person and call 911. The call taker instructs the caller to capture and share a still image of the person, due to limited connection, and then sends CPR instructional videos to guide the caller to complete the manoeuvre.
- **Multiple vehicle collision** – The caller is recording a video on a highway overpass, happens to record a crash between two vehicles, and calls 911. The call taker guides them to show the scene with camerawork, to share their video of the crash, and to identify the precise location of the crash on the



Figure 4: Immersive multi-screen display showing the shop robbery scenario.

map, rather than the caller's location, which is further down the road and above on an overpass.

- **Shop robbery** – The caller is in a side room of a jewelry shop and sees someone enter the main space and begin to rob the shop and the employee. The robber does not notice them, so they call 911. The call taker guides them through camerawork using on screen instructions to show the robber without revealing that they are filming.

Scenarios were presented to participants using a large, multi-screen display to create an immersive environment that helps participants imagine themselves in the given scenario [6] without risking their well-being. Each video was about two minutes in length and was displayed across the multiple screens to best fit the resolution of each video. The maximum resolution of the immersive screens was 15120x3840 (seven 4K displays rotated vertically), and our videos had between 480p and 4k resolution (Figure 4).

We chose to use immersive screens to simulate proximity to an emergency over reenactment (e.g., participants reacting to an actor pretending to collapse) and over a hybrid of reenactment and screens because not all scenarios are practical or safe to reenact (e.g., vehicle crash or performing CPR). Additionally, a hybrid presentation could distract participants as they would switch between looking at the screen and the actor(s), which could make the experience less realistic than an immersive video. None of

these study designs create a situation that is truly as stressful as calling 911, but, for practical and ethical reasons, we do not aim to explore the impact of stress on interaction with ME911 in this work. Our study design emphasizes the safety of both participants and researchers in order to explore the potential impact of unfamiliar design requirements on 911 callers.

We built our ME911 app with React Native, which supports both Android and iOS systems. For simplicity, we used the Google Sheets API to allow the call taker to dynamically enable and disable which features are available to the caller.

### 4.3 Procedure

The study was performed in a lab setting with the immersive screen setup described above. Participants interacted with the ME911 app for each scenario (counterbalanced within-participants for a total of ten scenarios per participant), and we conducted pre- and post-interviews for each scenario. Each study session lasted approximately 90 minutes and was audio recorded.

For all participants, the set of five scenarios in the order listed above was presented twice, once in each condition (Call Taker Control and Full Functionality). The order of these two conditions was counterbalanced across participants to control for learning effects in the repeated scenarios, while the order of scenarios within those conditions remained the same. That is, P1 experienced all

scenarios with Full Functionality and then again with Call Taker Control, while P2 began with Call Taker Control and then with Full Functionality, and so on. With these two conditions, participants were able to compare the less familiar experience of having a call taker control the conversation and interaction with something more familiar but within the same emergency setting. We repeated all scenarios with both conditions as each scenario targets different design needs of ME911, so participants would be able to comment on the likely unfamiliar experience of the Call Taker Control condition as it compares to Full Functionality across the different needs of the various scenarios.

In the Call Taker Control condition, we intentionally introduced varying delays of one to four seconds before buttons appear on the caller's phone. This simulates real world behavior of cellular networks, despite ongoing improvements [1]. Cellular networks already have known limitations for 911 call takers, like poor location accuracy [35], and we include these realistic delays to understand the impacts of these additional limitations.

The study began with a short interview collecting demographic information and prior experience calling 911. Participants were then asked to stand in front of the immersive display and we explained that they would watch each scenario and respond by (mock) calling 911 using a phone we provided. We clarified that the device would not actually call 911 and that the researcher present would be acting as call taker. We asked participants to imagine being present at the site of the emergency and that the researcher was an actual 911 call taker. We incorporated dialogue from the emergency response cards used by call takers into the protocol to more realistically match an actual 911 call. Call taker scripts are very consistent and controlled so that call takers can manage callers and quickly dispatch the needed help [17]. Because of this, we are confident that the instructions heard by participants were authentic to actual 911 calls. We consider how these scripts will have to be adapted to the needs of ME911 in the discussion section.

At the start of each scenario, the participant watched the video play once and the final image of the video remained on the screen while they completed the call. This provided a visual for the participant to interact with when using the ME911 app without overwhelming them with unrealistically looping videos of the emergency. For example, crashed vehicles are often stationary until they can be safely removed, so a still image of crashed vehicles is more realistic than repeatedly seeing a crash. After viewing the video, the participant (mock) called ME911. The researcher, as call taker, gathered the nature of the emergency and directed their use of ME911 relevant features.

Between scenarios, we asked questions to prompt participants to expand on their experiences. These were similar across scenarios and example questions included: "What worked well with the app in this scenario?", "What did not work well when the call taker had control?", and "Did you prefer having the call taker control or not? Why?". The researcher then reset the ME911 app and played the video for the next scenario on the immersive display.

After the participant completed all scenarios in both conditions, the study ended with a closing interview. We asked questions regarding the overall experience using the app as an emergency response tool and around their contrasting experiences with and

without call taker control over the interface. These questions included: "How did you feel losing autonomy over the user interface?" and "What did you think about being able to share media during a call to 911?". We also asked about other features that may be helpful for ME911 apps.

#### 4.4 Analysis

We audio-recorded and transcribed all sessions. We analyzed the transcripts using axial coding and thematic analysis [3] in NVivo. The lead researcher coded each transcript inductively, which was reviewed by a second researcher who proposed additional codes as relevant. Given the specific application area of this research, coding requires a level of expertise and familiarity with the topic. In these cases, it is common to use a single primary expert coder who can explain their choices to the rest of the research team [34]. Examples codes include "issues watching instructional videos", "legal concerns", and "delay in app interaction". These two researchers grouped codes into categories, forming axial codes, and then the primary coder determined themes from these axial codes, which were discussed and refined by the two researchers.

### 5 FINDINGS

In this section we structure our findings around the themes we identified from our participants' experiences using ME911: Camerawork and Call Taker Guidance, Caller Interactions with Call Taker Controlled Interface, How Callers Add Information Using ME911, and Privacy Concerns and the Need for Empathy.

#### 5.1 Camerawork and Call Taker Guidance

One aspect of ME911 that must be controlled by the caller rather than the call taker is camerawork. Both the call taker's guidance and the features supporting camerawork are novel to this app, which is reflected in our participants' reactions.

*5.1.1 Visual Instructions Support Safe Interactions.* First, we found that the written instructions on the video stream helped participants feel safe during 911 interactions. Some emergency scenarios put callers in noisy environments, like the vehicle explosion. Our participants described that the scene's ambient noise was too loud and distracting. Some felt it was difficult to hear the call taker, so they relied on visual instructions. Other scenarios risked the caller's safety, like the robbery. Here, many participants felt anxious even imagining being present at a robbery. Many felt speaking out loud would reveal their hiding spot, so spoke in whispers. Our participants found that visual instructions allowed them to read and follow instructions with as little audible interaction as possible.

"I do find [visual instructions] useful. They're better than a person telling you what to do. Cause sometimes it can be hard to hear. Also if someone is breaking in, not talking and interacting via visual instructions can be super handy" – P17

*5.1.2 Non-Normative Design Elements Need High Visibility.* Our participants easily understood the on-screen messages in ME911. However, in almost every scenario, participants understood in which direction to pan the camera from the on-screen message, but most moved the phone either too slowly or too quickly for the call

taker. The call taker then needed to verbally instruct them to repeat the action faster or slower. This embarrassed some callers who felt proficient with phone cameras but were not used to following camerawork directions. Other participants got frustrated when asked to do repeat camerawork.

“I was going a little slow and you noticed. So I thought, I’m not sure how you could tell someone to go in a mid-speed.” – P16

Icons indicating the panning direction alongside written message did not seem to improve participants’ understanding or performance. Most participants either understood the meaning from the text alone or did not understand the intended meaning of the icon.

“I did look at the icon beside the text but could not figure out how I was supposed to move the phone” – P18

Overlaid text and icons are an aspect of ME911 design that deviates from familiar video calling, and we see here how this is helpful to the specific needs of emergency situations. However, some participants struggled to read these instructions when they blended with the video background due to poor contrast. While some participants panned their phones to a higher contrast area to adjust for this, others missed the instructions entirely, in which case the call taker needed to verbally prompt the caller.

The standard use of video calling or recording primarily involves looking at the centre of the video, at the other caller or at what they are filming. Our participants continued to focus on the centre of the screen when using ME911, making it more difficult for them to notice instructions as they were added to the screen because the instruction bar was consistently present at the top of the screen and adding text did not create enough of a visual difference to draw attention. Placing a static element away from the centre of a video is common design practice, so the video is not obstructed. In our scenarios, the caller actually should be distracted by these instructions, as they need to quickly respond to that. This demonstrates the broader need for design considerations specific to ME911 (and distinct from non-emergency calling).

**5.1.3 Balancing Following Direction and Camerawork.** Participants experienced various issues with camerawork and phone placement before conducting CPR. The call taker asked them to find somewhere to place their phone upright so the call taker could see both the caller and the unconscious person. Participants struggled to place the phone upright in unexpected locations, and many did not want to put their phone on a cold or hard surface.

Some participants preferred to have a third person nearby hold the phone and do camerawork while they did CPR, whereas others said they would have preferred to hold the camera themselves while a third person performed CPR.

“I wouldn’t want to put my phone down on snow, in case it damaged my phone, I would prefer someone else hold the phone and take on the call, while I did the CPR.” – P4

Currently, with audio-only 911 calls, callers have similar difficulties in balancing the need to keep hearing and speaking to the call taker while performing hands-on tasks like CPR, and often have

a third person support them [35]. Most participants found that watching the short instructional CPR video, and then performing CPR with the call taker’s verbal instructions was enough to confidently complete the task. Others would have preferred seeing the CPR video while they performed CPR so they could follow along.

“If it’s something that I’m not familiar with or haven’t like done a CPR in a long time. It would probably be handy for me to be able to see the CPR video so I can continue to refer to it and make sure that I haven’t lost the beat or move of my hands” – P15.

As is common with 911 callers, most of our participants worried they could harm or injure the person while doing CPR and were concerned about legal liability. Participants were also unsure if they were exerting the right amount of force on the patient, which is another common concern of callers without medical training. The addition of instructional videos did not seem to change the issues that 911 callers commonly experience in managing both phone and task or their concerns that they might be hurting the person they are trying to help. There is more work to be done to see how multimedia can improve instruction for medical or other complex tasks like CPR.

**5.1.4 Emotional Responses to Delays and Stressful Situation Delay Interactions.** The unfamiliar design needs of ME911 often made participants curious. Some participants would interact with features without instruction from the call taker, particularly in the Call Taker Control condition as the appearance of new features and the delay in appearance piqued curiosity. Mimicking realistic delays in phone network performance by delaying button appearance raised anxiety and anticipation in our participants, leading them to over-interact. Some were irritated with the delay, so kept pressing the button until they got feedback, causing, for example, many photos to be captured instead of one. These participants recognized that this could be an issue during an emergency and apologized to the call taker for adding many pictures to the information they needed to process instead of only one. A few other participants felt the delay was a bug and decided that sending multiple images could be better than none.

“I was wondering if there is a bug in the app, I was thinking what’s happening? What should I do now? How to reset everything? I should keep pressing the screen till something happens” – P8

Our participants’ frustration and anxiety around the unexpected appearance of features, especially their delayed appearance, shows the importance of integrating an ME911 app tightly with the call taker’s work and scripted interactions with a caller. Call takers will need to manage callers’ emotional responses to looking for or waiting for features to appear.

## 5.2 Caller Interactions with Call Taker Controlled Interface

In the Call Taker Control condition, we observed reactions to the unfamiliar lack of control over and access to the UI. This is an extension of how call takers are currently trained to maintain control over a call, but it requires the design of ME911 to depart from common design norms. This theme highlights the ways participants

responded to this notable difference from common video calling platforms.

**5.2.1 Non-Standard Interface Confusing Without Call Taker Support.** In the Call Taker Control condition, participants initially saw their phone's video stream with no interactive features available. Most participants shared that this lack of clear task was frustrating and confusing, saying that they had no idea what to do after the video stream started. These participants needed verbal explanations from the call taker, in addition to the activation of individual features, before they understood how to perform each needed task. Many participants did not understand that the call taker had remote control until after the first features were activated. This shows the continuing importance of the clear and scripted role of a call taker to guide and reassure the caller.

The delays in the appearance of features, simulating realistic cellular network behaviour, caused some participants to feel nervous and anxious. They would receive verbal instruction to click a button before they could see it. As discussed above, some participants worried that this issue was with their phone, possibly because they trusted the call taker to provide accurate instructions. This worry led some to interact randomly with the blank screen, hoping that something would show up or they would find the button by chance.

“I felt a little bit anxious about that, partly because, I'm waiting for something to happen on the phone rather than I can interpret, you know, get used to the phone layout and see what I'm meant to be doing. So it looks unfamiliar, you know, from my normal phone layout.” – P9

Standard video calling apps have all features available by default and the unfamiliarity of ME911 led participants to feel surprised or confused. They typically required a few seconds to notice and understand each button after it appeared. In addition, participants felt that by adding features as needed, the call taker was judging their ability to carry out a video call.

Once a button was added in the Call Taker Control condition, it remained enabled for the rest of the call. Our participants found that this consistency helped them to get familiar with the app more broadly. They knew that they could fall back on those features without delay. We observed that after a few scenarios, they felt more familiar with both the app and the Call Taker Control. They felt less overwhelmed by the changing number of buttons and became more accustomed to their delayed appearance.

It is not unexpected that a non-standard design causes strong reactions in users. Most participants shared that the Call Taker Control condition was unusual and even bothered them. They could not understand the functionality of the ME911 app on their own, as it stands in contrast to the strong emphasis on intuitive learnability in modern app designs. In most scenarios, participants did not want someone else to be in control of their app. This unfamiliarity is felt in current audio-only calling too [40], though it is likely exacerbated here by the novelty of ME911.

In the Full Functionality condition, participants took little to no time to familiarize themselves with the interface. If they experienced any issues with the app or following instructions, they simply asked the call taker for help. Comparing the conditions, some participants found that seeing all features from the start helped them

understand the layout and general functionality of the app. They felt they were more easily able to carry out camerawork because they had already seen how the app would support that. A few participants expressed that having full functionality showed that the call taker trusted them to complete the needed tasks.

“All icons on helped me perform my task slightly faster as I did not wait for icons to pop up. All camera related icons were neatly present at the bottom bar, so it was easy to follow one instruction after another” – P4

Most participants felt that the Full Functionality condition aligned with their expectations for using a camera-based or video calling app and they preferred that. The unfamiliarity of the Call Taker Control condition was managed by the call taker's scripted instructions guiding users through tasks. Consequently, there is a conflict between participants' expectation for familiarity in interface design and the call taker's need to control the situation. Our next subtheme explores how participants experienced the other side of this conflict.

**5.2.2 Caller Comforted by Call Taker Control.** Participants who preferred the Call Taker Control condition felt that they would be panicked and stressed during a 911 call and would probably not think straight. We saw examples of this in participants' extra interactions we described above, like tapping randomly to find a missing button. These participants preferred the call taker to maintain control over the app to support their step-by-step guidance.

“I've been in situations where my brain kind of shuts down because there's a certain amount panic going on, if I'm in that situation, I really want to be talked through step one, step two, and have that somebody assuring me and telling me what to do rather than me figuring out things on my own.” – P7

Some participants were overwhelmed by the variety of features offered in the Full Functionality condition. Some felt they lacked experience with this app and the variety of features confused them as to what they were supposed to do. They already felt overwhelmed by the emergency scenario and many features on the video screen made them even more anxious.

“I think initially . . . there is so much information on screen. It feels a bit overwhelming. And if you're in a panicky situation, you might be a little bit nervous. Given that anxiety is high, it might be easier to be directed to one icon at a time and not necessarily have everything on screen” – P7

With control of the app, call takers can calm participants and clarify the needed task in the high stress environment, even when the unfamiliarity of the design is uncomfortable. When participants were already anxious, they were reassured by the simplicity of one feature appearing at a time. When the design does conform to norms (i.e., in the Full Functionality condition), stressed participants relied more on the call taker, and needed more verbal and visual guidance to find specific features.

### 5.3 How Callers Add Information Using ME911

Expanding 911 calls beyond audio-only increases the types of information a call taker has available. We designed the ME911 app to

expand their control of an audio call to video and media sharing features. With all this information available, our participants saw the role that callers might play to direct call taker attention to what they might have missed or what their script might not consider. Participants used available features in different ways to draw call taker attention and discussed additional features that could support more collaborative information sharing between caller and call taker.

**5.3.1 Directing Call Taker Attention with Camerawork.** In some cases, participants wanted to draw the attention of a call taker to a particular object visible in the video because they thought the call taker had neglected to ask about that specifically. In the truck fire scenario, participants were asked to pan across the entire scene, showing the extent of the fire but also that debris had scattered across all lanes of the highway. The call taker did not ask for any more information about the debris, but some participants decided to highlight it to the call taker by taking pictures or zooming in on it. Others waited until they had completed the camerawork as instructed and then described the debris verbally. The few participants who did not mention the debris explained afterwards that they felt that the call taker had probably seen it and would have said if they needed to do anything more.

These reactions led us to ask participants to consider other potential features to support drawing call taker attention, like on-screen annotations. Some participants were satisfied with the available ME911 features, like capturing images or zooming video and did not want the extra work of annotating on-screen.

“I wouldn’t have the patience to draw. It is like, I would just point something out, I would just tell them, oh, look here, I will zoom in and say it’s this object in the centre of the screen.” - P17

Some were concerned that it would be difficult to annotate on screen in winter, since gloves would make interactions difficult or impossible.

“I can imagine that being too clunky for example, if I’m wearing gloves in the winter and I don’t have smart fingers on, on the gloves, it might not work” - P15

Others said that annotation would be helpful to identify and label specific objects, but only on top of a captured image since the live video image would change as they annotated.

Most participants felt obligated to make sure the call taker did not miss important information. In an audio-only call, the call taker needs the caller to explain the entire scene, but the addition of video can reduce that work and allow the call taker to directly assess the scene. However, because visual media can contain so much detail, especially in the potential chaos of emergency settings, callers may need to direct attention to what may have been missed, making for a more collaborative exchange than in audio-only calls.

**5.3.2 Verbally Supporting Map-Based Location Finding.** Participants were able to use the map to identify their current address fairly well, but some struggled to identify relevant nearby landmarks. The standard Google Maps interface was familiar, but showed many diverse options that could be chosen as landmarks, make it difficult for participants to choose a single guiding landmark.

Most participants said they would have been better supported by a map specific to the emergency setting that highlighted three to five potential landmarks that would be relevant and helpful to the call taker.

“Finding where I am and the bridge and then the nearest location mark, I guess that can be a bit confusing for people because there’s like lots of places that are nearby. You don’t know which one to say.” - P17

While the integrated map is helpful in location finding, we find that participants need more support in efficiently describing location. From our participants’ experiences, the inclusion of a simple, familiar map interface was not sufficient as they still had to describe their location from all of the information available. An ME911 app could help to guide and control callers, such as by highlighting the most useful landmarks based on a caller’s approximate location.

## 5.4 Privacy Concerns

Previous work has found that both callers and call takers feel that adding multimedia to 911 calls, particularly video, will add new privacy and emotional concerns that will need to be considered in the design of these tools [35, 40]. After having used our ME911 app, we asked our participants to reflect on these issues. While most participants were comfortable sharing a pre-recorded video with a call taker, they were not willing to share videos that showed their family or friends.

“I could be streaming goofy stuff prior to the accident or might have a friend in the video and would not like to share these types of content with the call taker” - P4.

In response to this specific concern, we asked participants to consider how ME911 could support them to protect other people’s privacy, such as trimming a video. Most participants would trim a video themselves but were concerned call takers could be too busy to trim and might keep the full video for simplicity over protecting privacy. A few other participants thought that call takers would be better for the task as the stressful situation might cause callers to incorrectly remove a needed part of the video.

“I’m terrible at cropping videos. It takes me forever to figure it out, so I would not feel like I could do that fast enough in order to send a cropped video during a call like this” - P15

The addition of visual media to 911 calling may pose further privacy risks to non-callers than audio-only calls. Thus, designing for ME911 should consider including features, like trimming, that support callers to quickly and easily protect other’s privacy.

## 6 DISCUSSION

From our study, we see a range of experiences and reactions to the unfamiliar nature of ME911 calling. Here, we highlight personal, social, and structural challenges and considerations for ME911 calling systems. With our design and study, we observed how participants experience non-standard design elements of ME911 and we discuss how these important features might be incorporated into actual emergency response systems.

From our observations, we see that the design of ME911 needs to intentionally support the call taker to guide callers through a stressful and unfamiliar situation. Overall, we see that introducing ME911 features to smartphones cannot be done without parallel changes to the full 911 calling system. These multimedia features will require more collaboration between call taker and caller, which will also need to be integrated into the broader emergency response system. These considerations show the importance of building a complete picture of what ME911 systems will entail. Here we investigated the caller's experience and future work will expand on our findings, such as considering ME911 tools for dispatch centres and emergency responders, towards safe and functional ME911 systems.

## 6.1 ME911 Caller Features Tightly Integrated with Full 911 System

911 call-taking is largely about control. Call takers are trained to remain in command of a situation to efficiently gather the needed information from a caller [10, 17]. To support the emergency, this information flow and control needs to be asymmetric: callers and call takers cannot be equal conversational partners. However, with the addition of multimedia features, there is a need for collaboration between the caller and call taker [40], as we see in participants volunteering potentially missed details during camerawork.

Despite the non-emergency setting of our study, our participants experienced similar feelings to those experienced by actual 911 callers, like panic and stress [18, 25], not only in relation to the high stakes of successfully completing the given task but also from the non-standard design of ME911. Call takers currently use standard scripted questions and responses to directly and consistently collect the needed information. Video could simplify some of this information gathering by allowing the call taker to see the scene for themselves [35]. However, our observations showed that participants needed clearer camerawork directions to complete tasks efficiently. As we based the study design on current response cards for audio-only calling, we added both spoken and on-screen instructions for camerawork, neither of which were sufficient for participants to complete a task on the first try. The instructions were understood by participants, but they were missing additional relevant details, like panning speed.

Our study thus shows that of the various novel elements of ME911, **guiding camerawork within the interface design is central as it supports efficient camerawork by the caller**, and should be a key focus for future work. This is distinct from camerawork in non-emergency settings, which generally does not need to be guided for the caller [24]. The time-sensitive nature of emergency calling and the scripted interactions controlled by call takers are particular to this context and with this study we see a need for specific support for call takers to guide callers through various, likely unfamiliar, camerawork tasks. Emergency response cards and instructions need to be updated to incorporate the tasks callers will complete in ME911, with a particular focus on how to describe precise camerawork. Additionally, on-screen instructions should suffice when the call taker cannot speak or cannot be heard, such as in noisy or dangerous situations. The ME911 application will

need to clearly display those on-screen instructions to draw the caller's attention and be easily legible.

When the call taker added a feature, that change on the screen typically drew participants' eyes, as opposed to finding one button among all the static elements. This addition resembles an icon animation, which can help users learn novel layouts and draw their attention [8]. This helped callers to adjust to the unfamiliar design. However, the non-standard design also confused participants, who generally wanted control over the app's video features and felt anxious seeing the video screen lacking common features. This was expected, and we use our observations of this confusion to suggest ways of preparing for it. Our findings are also in contrast to past research in which 911 callers felt they would prefer to give up control over multimedia features in a 911 call to manage anxiety [40].

Call takers can prepare callers for this unusual app design by explaining ahead of time that features are hidden until needed and can also include directions of where to look for features before they appear, to help callers understand instructions quickly. As an alternative to hiding all features, the various features could initially be visible but clearly disabled, so callers are not distracted by their appearance or delayed by searching for a button. Whether initially hidden or visibly disabled, the design should clearly emphasize where an icon is enabled or added with a strong visual cue, like a simple animation or other highlighting, to draw the caller's attention that a feature is now enabled [8]. In our study, the dynamic appearance of an icon where there was none resembled an animation, in some cases pulling caller attention to that new feature. We suggest that the design of ME911 should explore further emphasizing that effect with intentional animation. In combination with call taker direction, **animation within the ME911 interface can ease caller stress and help them to quickly understand what they need to do**. Generally, we find that callers can struggle to find and use features both when all are available and when individually added because of the high stakes during a 911 call. The design of ME911 apps needs to balance an easily understandable system and clear visual cues through clear visualization of which features are currently needed and intentionally directing attention towards those. This balance will support call takers to guide callers through each needed step.

Though initially unfamiliar, participants generally became more comfortable with the non-standard design aspects of ME911 across the different scenarios. Calling 911 is not a common or familiar activity compared to everyday phone calls. Similarly, the irregular interactions of ME911 reminded participants of the urgency of the situation, so the call taker's guidance was key to directing the caller's actions. From this, we propose that a combination of carefully scripted call taker guidance to use the interface, clear visual cues directing callers towards each feature, and growing access to actual ME911 calls will collectively familiarize most callers to the non-standard nature of ME911.

## 6.2 ME911 Needs to Balance Caller as Collaborative Participant with Call Taker Control

In some cases, our participants felt the need to add information or draw the call taker's attention to aspects of the scene. This is not a standard part of a call taker's script, so there was not a clear way for callers to do this. We saw various reactions and adaptations that used the available features. Drawing attention in non-emergency video calling has largely explored on-screen annotations [16]. However, our participants felt annotations were too much work during an emergency or would distract from their task. Alternatively, callers may also wait until the call taker finishes the current task before they go back to point something out. They may also avoid mentioning it altogether when there is no clear way to do so. When caller input may be helpful, the ME911 app can encourage that input by using image recognition to highlight elements of the scene or by visually indicating when the call taker's speech and/or script has space for their suggestions. Rather than a static video feed, **suggesting ways to collaborate with the call taker could help hesitant callers to speak up**. These features can also be disabled when their input would slow crucial tasks, like performing CPR. Adding video to 911 calling increases the amount of detailed information available to call takers, and callers will have more of a role in highlighting what might have been missed.

Another aspect of collaboration comes from the non-standard interface. As mentioned above, the unfamiliar design caused participants to struggle to recognize icons or messages. In our observations, these struggles were exasperated by the high stress situation of calling 911, and drawing the caller's attention to part of the screen will likely need call taker support. We found that the call taker's control, only adding features as needed, combined with their verbal instructions helped callers locate buttons despite the stressful setting. Though our participants were less comfortable with the unfamiliar interface and lack of control, limiting callers' knowledge of features might allow them to act more quickly when the call taker introduces each needed task.

**Design of ME911 apps needs to provide opportunities for callers to quickly contribute their observations as part of the call taker's process.** This would place the caller in a more collaborative role than has been the case with audio-only 911 calls, and the broader system of 911 call-taking will need to consider how to support this without disrupting call taker control.

## 6.3 ME911 Features Pose Similar Privacy Concerns as Audio-Only 911

Privacy is an important aspect in an emergency call taking situation and designers need to design ME911 to consider the privacy of both callers and call takers [13, 18, 19, 25]. While ME911 features may improve the quality and quantity of data for call takers, the addition of a caller's video to a 911 call also adds another dimension of privacy concerns. **The design of ME911 apps must consider the privacy of any bystanders to the call** (including friends and family) who may be filmed as well. Our participants felt that video trimming or other privacy protection features could allow callers to comfortably share previously recorded videos with call takers, so ME911 apps can actively prompt them to use these features after

media is selected, but before it is sent to the call taker and stored in their dispatch centre. Additionally, future work should consider how videos uploaded to social media (whether by the caller or other bystanders) should be used by call takers. While videos shared by callers can provide further important details, such as for court proceedings as a result of a 911 call [2], there is also a need for further consideration of safeguarding the privacy of the people filmed, which is beyond the scope of this current work. Research into how streamers consider privacy concerns of bystanders highlights their different approaches, including disabling video as needed and announcing when the camera will be on. This might be able to be adapted to the ME911 context, such by reminding and guiding the caller to inform bystanders who may not expect to be in a video stream [41].

The privacy and safety concerns shared by participants using our mock ME911 app were often similar to what has been shared by previous 911 callers, such as calling during a dangerous situation and the difference between their active consent as a caller and a passerby's lack of consent [40]. Future work should focus on how visual media introduces new privacy concerns and consider how ME911 apps and systems of gathering videos from other online sources can proactively protect privacy of both callers and others.

## 7 LIMITATIONS AND FUTURE WORK

The safety of our participants and of us as researchers were key in designing and running this study. This decision has necessarily introduced limitations in the study design. We designed our study to mimic a 911 call as realistically as possible by basing the scripts used by the researcher (acting as call taker) on actual emergency response cards used across Canada and on our past work with call takers. Thus, our findings may differ from how emergency calling is handled in other areas of the world, and future work should explore emergency calling and the associated needs for incorporating multimedia in other world regions.

We tested our prototype in a lab setting and used very large screens to immerse participants in the context of emergency scenarios. This was, again, done to prioritize the safety of participants and researchers. We acknowledge that such a setting may not create the same level of stress and other emotions as real emergencies, but we see it as a necessary compromise between realism and safety, given that ME911 features are completely novel to 911 calling. Safely and effectively changing actual 911 call systems is incredibly challenging. Before an evaluation can happen with real emergency situations, the impact of ME911 across the larger system needs to be assessed, including updated response cards, as we discuss, as well as detailed input from call takers and necessary changes to dispatch centre software and hardware (e.g., to receive and process multimedia). This is a potential direction for future work.

Our controlled study is an important early step in safely understanding the design needs of ME911 calling applications. This work will guide researchers and designers of 911 systems to continue the work of safely designing for ME911, and eventually can lead to the introduction of these features to actual emergency calling. Future work on the ME911 calling interface will incorporate tools that guide the caller as a support for the call taker. Future research

should also turn to the call taker's perspective of this same application. Specifically, designs will need to incorporate the caller's video feed into the already busy displays of dispatch centres [35] and interactions to enable or disable features and deliver instructions for camerawork or otherwise communicate with the caller. Future work should continue to incorporate novel and improving features as phones and networks advance, such as combining Bluetooth and/or Wi-Fi with GPS to provide more accurate location information to support the caller's description or using live video stabilization. Finally, future work can explore ways to safely and ethically assess the non-standard design and interactions of ME911 in emergency settings, such as by starting with minor or non-urgent emergencies like a superficial vehicle collision.

## 8 CONCLUSION

Our work considers how potential 911 callers react to the non-standard design needs of multimedia-enabled 911 calling (ME911). Our ME911 application allows callers to complete camerawork and media sharing to the call taker. We consider how standardized call taker scripting will need to be adapted to the complex needs of ME911, including how to direct caller camerawork. From these findings, we consider what is still needed to bring access to multimedia into 911 calling. We hope that our work helps designers to explore the space of 911 systems and to look more closely at integrated systems enabling multimedia-enabled emergency calls.

## ACKNOWLEDGMENTS

We wish to acknowledge the Musqueam, Squamish, Tsleil-Waututh, Katzie, Kwikwetlem, Qayqayt, Kwantlen, Semiahmoo, and Tsawwassen peoples on whose unceded traditional territories our three campuses reside. We thank our reviewers for their valuable feedback.

## REFERENCES

- [1] Mainak Adhikari and Abhishek Hazra. 2022. 6G-Enabled Ultra-Reliable Low-Latency Communication in Edge Networks. *IEEE Communications Standards Magazine* 6, 1: 67–74. <https://doi.org/10.1109/MCOMSTD.0001.2100098>
- [2] Fredrik Bergstrand and Jonas Landgren. 2011. Visual reporting in time-critical work: exploring video use in emergency response. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 415–424. <https://doi.org/10.1145/2037373.2037436>
- [3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2: 77–101. <https://doi.org/10.1191/1478088706QP063OA>
- [4] Canadian Radio-television and Telecommunications Commission (CRTC) Government of Canada. 2017. Next-generation 9-1-1. Retrieved July 24, 2023 from <https://crtc.gc.ca/eng/phone/911/gen.htm>
- [5] Canadian Radio-television and Telecommunications Commission (CRTC) Government of Canada. 2021. Establishment of new deadlines for Canada's transition to next-generation 9-1-1. Retrieved July 25, 2023 from <https://crtc.gc.ca/eng/archive/2021/2021-199.htm>
- [6] Gokhan Cetin, Wolfgang Stuerzlinger, and John Dill. 2018. Visual analytics on large displays: exploring user spatialization and how size and resolution affect task performance. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, 1–10.
- [7] Choyce A. Chappell. 2022. Design of a mobile crisis weather app for residents of Prince Edward Island. University of Prince Edward Island.
- [8] Fanny Chevalier, Nathalie Henry Riche, Catherine Plaisant, Amira Chalbi, and Christophe Hurter. 2016. Animations 25 Years Later: New Roles and Opportunities. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, 280–287. <https://doi.org/10.1145/2909132.2909255>
- [9] Anna L. Cox, Sandy J.J. Gould, Marta E. Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*, 1389–1397. <https://doi.org/10.1145/2851581.2892410>
- [10] Punyashlok Dash, Carman Neustaedter, Brennan Jones, and Carolyn Yip. 2022. The Design and Evaluation of Emergency Call Taking User Interfaces for Next Generation 9-1-1. *Frontiers in Human Dynamics* 4: 670647. <https://doi.org/10.3389/fhumd.2022.670647>
- [11] Guido De Angelis, Giuseppe Baruffa, and Saverio Cacopardi. 2013. GNSS/Cellular Hybrid Positioning System for Mobile Users in Urban Scenarios. *IEEE Transactions on Intelligent Transportation Systems* 14, 1: 313–321. <https://doi.org/10.1109/ITITS.2012.2215855>
- [12] Timothy Denton. 2013. *A Report on Matters Related to Emergency 9-1-1 - 5 July 2013*. Canadian Radio-television and Telecommunications Commission. Retrieved July 24, 2023 from <https://crtc.gc.ca/eng/publications/reports/rp130705.htm>
- [13] Rae-Lynne Heather Dicks. 2014. Prevalence of PTSD symptoms in Canadian 911 operators. University of the Fraser Valley. Retrieved July 24, 2023 from <https://ufv.arcabc.ca/islandora/object/ufv:334>
- [14] Mica R. Endsley, Betty Bolte, and Debra G. Jones. 2003. *Designing for Situation Awareness*. CRC Press. <https://doi.org/10.1201/9780203485088>
- [15] Arvid Engström, Goranka Zoric, Oskar Juhlin, and Ramin Toussi. 2012. The mobile vision mixer: a mobile network based live video broadcasting system in your mobile phone. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 1–4. <https://doi.org/10.1145/2406367.2406390>
- [16] Omid Fakourfar, Kevin Ta, Richard Tang, Scott Bateman, and Anthony Tang. 2016. Stabilized Annotations for Mobile Remote Assistance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1548–1560. <https://doi.org/10.1145/2858036.2858171>
- [17] Giolo Fele. 2023. The Organization of Emergency Calls. In *Emergency Communication*. Springer International Publishing, Cham, 11–42. [https://doi.org/10.1007/978-3-031-26239-5\\_2](https://doi.org/10.1007/978-3-031-26239-5_2)
- [18] Kerstin Forslund, Annica Kihlgren, and Mona Kihlgren. 2004. Operators' experiences of emergency calls. *Journal of Telemedicine and Telecare* 10, 5: 290–297. <https://doi.org/10.1258/1357633042026323>
- [19] Elizabeth Hayes. 2017. Commonly Identified Symptoms of Stress Among Dispatchers: A Descriptive Assessment of Emotional, Mental, and Physical Health Consequences. St. Cloud State University. Retrieved from [https://repository.stcloudstate.edu/cjs\\_etds/7](https://repository.stcloudstate.edu/cjs_etds/7)
- [20] Harmen Hofstra, Henk Scholten, Sisi Zlatanova, and Alessandra Scotta. 2008. Multi-user tangible interfaces for effective decision-making in disaster management. In *Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters*, Shailesh Nayak and Sisi Zlatanova (eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 243–266. [https://doi.org/10.1007/978-3-540-79259-8\\_14](https://doi.org/10.1007/978-3-540-79259-8_14)
- [21] Amanda L. Hughes, Lise A. A. St. Denis, Leysia Palen, and Kenneth M. Anderson. 2014. Online public communications by police & fire services during the 2012 Hurricane Sandy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1505–1514. <https://doi.org/10.1145/2556288.2557227>
- [22] Kori Inkpen, Brett Taylor, Sasa Junuzovic, John Tang, and Gina Venolia. 2013. Experiences2Go: sharing kids' activities outside the home with remote family members. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 1329–1340. <https://doi.org/10.1145/2441776.2441926>
- [23] Brennan Jones, Kody Dillman, Richard Tang, Anthony Tang, Ehud Sharlin, Lora Oehlberg, Carman Neustaedter, and Scott Bateman. 2016. Elevating Communication, Collaboration, and Shared Experiences in Mobile Video through Drones. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 1123–1135. <https://doi.org/10.1145/2901790.2901847>
- [24] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. 2015. Mechanics of Camera Work in Mobile Video Collaboration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 957–966. <https://doi.org/10.1145/2702123.2702345>
- [25] Heidi Kevoe Feldman. 2021. Calming emotional 911 callers: Using redirection as a patient-focused directive in emergency medical calls. *Language & Communication* 81: 81–92. <https://doi.org/10.1016/j.langcom.2021.09.001>
- [26] David S. Kirk, Abigail Sellen, and Xiang Cao. 2010. Home video communication: mediating "closeness." In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 135–144. <https://doi.org/10.1145/1718918.1718945>
- [27] Jonas Landgren and Urban Nulden. 2007. A study of emergency response work: patterns of mobile phone interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, 1323–1332. <https://doi.org/10.1145/1240624.1240824>
- [28] Sophia B. Liu, Leysia Palen, Jeannette Sutton, Amanda L. Hughes, and Sarah Vieweg. 2008. In search of the bigger picture: The emergent role of on-line photo sharing in times of disaster. In *Proceedings of the information systems for crisis response and management conference (ISCRAM)*, 4–7.
- [29] Thomas Ludwig, Christian Reuter, and Volkmar Pipek. 2013. What You See is What I Need: Mobile Reporting Practices in Emergencies. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, Olav W. Bertelsen, Luigina Ciolfi, Maria Antonietta Grasso and George Angelos Papadopoulos (eds.). Springer London, London, 181–206. [https://doi.org/10.1007/978-1-4471-5346-7\\_10](https://doi.org/10.1007/978-1-4471-5346-7_10)

- [30] Thomas Ludwig, Christian Reuter, and Volkmar Pipek. 2015. Social Haystack: Dynamic Quality Assessment of Citizen-Generated Content during Emergencies. *ACM Transactions on Computer-Human Interaction* 22, 4: 1–27. <https://doi.org/10.1145/2749461>
- [31] Leanna Machado. 2022. Investigating 9-1-1 Call Experience for Medical Emergencies for Future Design. Pace University. Retrieved from [https://digitalcommons.pace.edu/honorscollege\\_theses/366](https://digitalcommons.pace.edu/honorscollege_theses/366)
- [32] Evangelos K. Markakis, Asimakis Lykourgiotis, Ilias Politis, Anastasios Dagjuklas, Yacine Rebahi, and Evangelos Pallis. 2017. EMYNOS: Next Generation Emergency Communication. *IEEE Communications Magazine* 55, 1: 139–145. <https://doi.org/10.1109/MCOM.2017.1600284CM>
- [33] Michael Massimi and Carman Neustaedter. 2014. Moving from talking heads to newlyweds: exploring video chat use during major life events. In *Proceedings of the 2014 conference on Designing interactive systems*, 43–52. <https://doi.org/10.1145/2598510.2598570>
- [34] Nora Mcdonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction*: 23. <https://doi.org/10.1145/3359174>
- [35] Carman Neustaedter, Brennan Jones, Kenton O'Hara, and Abigail Sellen. 2018. The Benefits and Challenges of Video Calling for Emergency Situations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3173574.3174231>
- [36] S. Rebecca Neusteter, Maris Mapolski, Mawia Khogali, and Megan O'Toole. 2019. *The 911 call processing system: A review of the literature as it relates to policing*. Vera Institute of Justice. Retrieved from <https://www.vera.org/downloads/publications/911-call-processing-system-review-of-policing-literature.pdf>
- [37] Irene Rae, Gina Venolia, John C. Tang, and David Molnar. 2015. A Framework for Understanding and Designing Telepresence. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1552–1566. <https://doi.org/10.1145/2675133.2675141>
- [38] Stuart Reeves, Christian Greiffenhagen, Martin Flintham, Steve Benford, Matt Adams, Ju Row Farr, and Nicholas Tandavantij. 2015. I'd Hide You: Performing Live Broadcasting in Public. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2573–2582. <https://doi.org/10.1145/2702123.2702257>
- [39] Aleksandra Sarcevic, Leysia Palen, Joanne White, Kate Starbird, Mossaab Bagdouri, and Kenneth Anderson. 2012. "Beacons of hope" in decentralized coordination: learning from on-the-ground medical twitterers during the 2010 Haiti earthquake. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 47–56. <https://doi.org/10.1145/2145204.2145217>
- [40] Samarth Singhal and Carman Neustaedter. 2018. Caller Needs and Reactions to 9-1-1 Video Calling for Emergencies. In *Proceedings of the 2018 Designing Interactive Systems Conference*, 985–997. <https://doi.org/10.1145/3196709.3196742>
- [41] Yanlai Wu, Xinning Gui, Pamela J. Wisniewski, and Yao Li. 2023. Do Streamers Care about Bystanders' Privacy? An Examination of Live Streamers' Considerations and Strategies for Bystanders' Privacy Management. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1: 1–29. <https://doi.org/10.1145/3579603>

## APPENDICES

### A APPENDIX A - ME911 CALLING SCENARIO

To situate the ME911 app and features in the expected setting, we present a scenario of use. We use one of our study scenarios

(multiple vehicle car crash) to show how we combine standard call taker emergency response cards with our design.

- Two friends are recording a video from a bridge over a multi-lane highway when they see two cars crash in the background of their camera view. One friend stops the recording and calls 911 from their ME911 app.
- A call taker initially answers the multimedia-enabled call in audio-only mode and goes through the standard initial questions to collect the caller's contact information, in case the call is disconnected and the call taker has to call back.
- To get a precise location, the call taker directs the caller to open the Maps tab. They find and describe their location by naming a few nearby streets and landmarks.
- The call taker then asks the caller to start the video feed from the rear camera by pressing the video button. The video feed shows on both the caller's and call taker's screens, without any features or buttons on the caller's side.
- From the highway overpass, the caller is too far from the crash to show detail and cars continue to pass by in the open lanes, adding distractions to the video. The call taker explains that they are enabling a zoom slider and asks the caller to zoom in on the crash to provide a clearer view. The call taker can now identify one crashed vehicle off the road, a person sitting nearby, and a second crashed vehicle in the oncoming lane.
- The call taker then enables image capturing and asks the caller to take photos of the target at the current zoom level. To observe the scene more comprehensively, the call taker then sends visual instructions to pan the phone left and right to take additional pictures.
- The call taker then asks the caller if they have any recorded media or photos of the crash. The caller explains that they were recording video and captured the crash in the background, so the call taker enables the media sharing button and instructs the caller to share the prerecorded crash video.
- The call taker then informs the caller that they have dispatched emergency services to the scene and that they can end the call.