

# PokerFace Mask: Exploring Augmenting Masks with Captions through an Interactive, Mixed-Reality Prototype

Josh Urban Davis  
Dartmouth  
[josh.u.davis.gr@dartmouth.edu](mailto:josh.u.davis.gr@dartmouth.edu)

John Tang  
Microsoft Research  
[johntang@microsoft.com](mailto:johntang@microsoft.com)

Edward Cutrell  
Microsoft Research  
[cutrell@microsoft.com](mailto:cutrell@microsoft.com)

Teddy Seyed  
Microsoft Research  
[Teddy.Seyed@gmail.com](mailto:Teddy.Seyed@gmail.com)

## Abstract

The 2020 COVID-19 pandemic made masks a daily wearable for personal protective equipment as a public health precaution. Traditional mask designs obscure communication by obstructing the face and muffling the voice which can make communication especially difficult for users who are deaf or hard of hearing (DHH). PokerFace uses a commodity smartphone and recycled materials to display a live-stream of a user’s mouth and nose on the mask surface. This maintains the safety precautions afforded by the mask, while mitigating the obfuscation of traditional mask designs. To compare PokerFace’s ability to facilitate communication with traditional masks, we conducted a user study with 18 participants, who played a collaborative communication game similar to charades. Participants performed better at this collaborative communication task with our prototype than with traditional masks, and even non-DHH users became aware of the importance of lip-reading and facial cues in communication due to study participation.

## 1. Introduction

During the COVID-19 pandemic response, masks for personal protective equipment (PPE) became a daily wearable as a public health precaution in 2020. As a result, communication and socialization was affected by regulations of social distancing and had to adjust to having the face partially obscured by masks. Besides visually obscuring a large portion of the face, traditional masks also muffle the voice of the wearer by covering their mouth with multiple layers of fabric or protective material. These difficulties are exacerbated for users who are deaf or hard of hearing (DHH) since they may rely on lip reading or facial cues to comprehend communication. Subsequently, many innovations began emerging to ameliorate the communication limitations imposed by daily mask usage. This included focusing mask design on personal expression [1], using variable

materials to allow the facepiece to be transparent, or even replacing the entire cloth mask material with hard plastic to mitigate visual obstruction to the mouth and nose area. Innovation to mask design was not strictly relegated to design of the mask itself, but also included embedding masks with computation capabilities [2]. This process of imbuing computational capabilities into daily wearables proceeds from a history of ubiquitous computation including smart fabrics, e-textiles [3], and smart jewelry [4][5]. Thus, we explored the potential for the mask as a face-worn wearable embedded with computation and interactivity, as well as the capability of this design space to mitigate limitations in communication resulting from the mask form-factor.

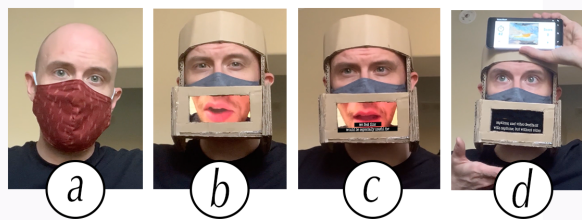


Figure 1. Four User-Study Conditions: a. traditional mask; b. PokerFace mask with video but without captions; c. PokerFace mask with video and captions; d. PokerFace mask with captions but without video;

In this paper, we present PokerFace, an interactive, mixed-reality mask which allows for face-worn interaction and expression using a commodity capacitive touch-screen device and cost-effective recyclable materials (Figure 1b, 1c, and 1d). The mask situates a smartphone in a cradle over the user’s mouth and nose, similar to a traditional mask. PokerFace uses the display of the device to live-stream a video of the wearer’s mouth and nose in real time, as well as present captions of user’s speech using a real-time speech-to-text function. Our goal with this prototype was to explore the following research questions:

- R1. How can embedding computation into

*face-worn wearables support interaction?*

- R2. *Can augmented masks facilitate better communication for users than traditional masks?*

To explore our research questions, we iteratively designed a prototype mixed-reality mask, and conducted a counterbalanced user study with 18 participants. During the user study, participants were asked to play a series of games of Guess What, a charades-like game, where users try to communicate concepts to each other using gestures and pantomime. We compared across 4 conditions: traditional masks, PokerFace prototype with video but not captions, PokerFace prototype with video and captions, and PokerFace prototype with captions but without video. Our study suggests that users are able to perform better playing this collaborative communication game with the PokerFace prototype than with traditional masks. The principal contributions of this work include:

- A prototype mixed-reality mask, called PokerFace, made using iterative design process using a commodity smartphone (Section 3)
- Evaluation of prototype PokerFace mask in a counterbalanced study examining its ability to facilitate communication (Section 4)
- Suggesting prototype applications which explore the design space of mixed-reality masks and face-worn wearables (Section 7)

While the prototype studied in this paper is not itself a practical implementation, it demonstrates the potential of exploring this design space while technologies that could make it viable are still being developed.

## 2. Related Work

We first examine the history of embedding computation in wearables and computationally augmenting the body (Section 2.1), before providing an overview of existing research regarding masks for PPE (Section 2.2). Finally, we examine prior work in augmenting the face with computation (Section 2.3).

### 2.1. Computationally Augmenting the Body

Computationally augmenting the body is a popular interdisciplinary domain of research within the field of human computer interaction because on-body embedded systems provide immediate, convenient access to information and services. Early successes of the ubiquitous computing movement imbued daily wearable garments with computation such as smartwatches and

headphones [6]. Research interest in developing the capabilities of these common wearables continues through innovation in interaction [7], biosensing [8], tangibles [9] and other domains. Computational augmentation of the body also takes a wide variety of form-factors from e-textiles [3], and robotic jewelry [10]. Research into virtual, augmented, and mixed reality (XR) augmentation of the body is limited. Prior approaches incorporate displays [4] or projection [11] onto the body to enable interaction or data visualization [12]. XR work in this domain often blurs boundaries between virtual reality and augmented reality. For example, Kim, Lee and Koh incorporated a display onto the back of a virtual reality headset to provide user's not wearing the VR headset a view of what the user in VR was seeing [13]. Previous work examined how assistive wearables [14] could support end users with hearing impairments (deaf, deafened, and hard of hearing) [15] [16] [17], cognitive impairments [18], visual impairments (such as: blindness [19] and color-blindness [20]), as well as motor impairments [21]. Wearables also present a promising venue for augmenting, restoring, or replacing human skills, abilities, and senses [22][23][24]. While this prior work is promising, little investigation has examined the potential for mixed-reality imbued masks as an interactive face-worn wearable.

### 2.2. Masks, PPE, and Facial Cues

Due to the spread of the COVID-19 pandemic in early 2020, the U.S. Centers for Disease Control and Prevention (CDC) implored individuals to regularly use disposable medical or reusable cloth masks when interacting in social public environments to mitigate pandemic contagion [25]. This sudden increase in mask usage also stimulated creativity and innovation and making and wearing masks for PPE [26]. Research into mask safety is ongoing with novel developments and innovations emerging at the time of this research [1]. Previous work on masks focused primarily on health considerations, material affordances, and design considerations. Given that daily garments often become the subject of embedded wearable innovation such as smartwatches and e-textiles, masks present a promising candidate as computational augmentation.

One consideration within this domain of mask design affordances is the obfuscation of the face incurred through mask usage. Prior work in the literature demonstrates that facial cues and non-verbal communication are vital to understanding verbal communication in social interactions [27]. Non-verbal communication is especially important for populations

with dementia, who are deaf or hard of hearing (DHH), or have cognitive impairments [28]. Proliferation of mask usage due to the COVID-19 pandemic has made interpersonal communication increasingly difficult, due in large part to the obstruction of facial cues by mask usage [29]. While some work has explored social and behavioral means of mitigating these obstacles to interpersonal communication, there has been little investigation of innovating mask design to better facilitate non-verbal communication [30]. Bowman et al. demonstrated through their Looking Glass prototype the potential for enriching interpersonal communication through computationally augmented wearables [31]. Computational augmentation accounting for non-verbal communication is exhibited elsewhere in computer mediated communication. Emoticons and pictographs, for example, have been demonstrated to services as non-verbal communication modalities in computer-mediated technologies, as well as other textual means of nonverbal communication [32] [33]. Given this prior literature, it is evident that further investigation into the potential for computationally augmented wearables as a means of enriching interpersonal and nonverbal communication is needed.

### **2.3. Augmenting the Face with Computation**

Work within the realm of computationally augmenting the face is limited, and reflects many of the use cases evident in the greater field of wearables. Many face-worn computational devices provide medical services to users [34]. The face has also been investigated as a source of novel interaction capabilities, such as using the nose [35], ear [36], or tongue [37]. More closely related to our avenue of inquiry are systems which enable mixed-reality capability for face-worn wearables. Chameleon, for example, uses a large display worn over a person's face to provide embodied telepresence interaction for a remote user [38]. The potential of this paradigm has also been studied within the realm of accessibility. Augmenting and alternative communication (AAC) systems provide services for facilitating communication interaction for users who might have difficulty producing speech or language. Feuston and Jackson, for example, used a projection mapping system to display facial expressions directly on user's faces whom had facial paralysis, thus enabling a display of the user's emotional affect [39]. Developments in computational masks have emerged within the research literature as well [40]. Recently, commercial products such as the C-mask, and research prototypes such as MAScreen have begun integrating speakers and LED arrays into masks to provide

facilitate communication [41]. While MAScreen is able to provide a semblance of facial expression, they are abstracted and displayed using a grid of LEDs. In addition, unlike PokerFace, prior literature is unable to display captions and facial expression at the same time.

### **2.4. Literature Gap**

To summarize, prior research highlights the potential of imbuing daily face-worn wearables with computation. We explore mixed-reality affordances to facilitate communication and mitigate the interaction barriers endemic to the mask form-factor. Building a working prototype enables investigating how it could potentially enhance communication, interaction, and collaboration and end-user reactions and feedback to the affordances presented by this design space.

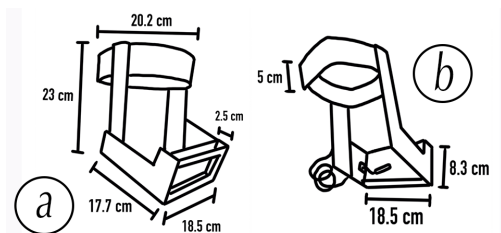
## **3. PokerFace**

To explore the potential for computationally augmented masks to better facilitate communication, we designed and implemented a mixed reality prototype mask called PokerFace. Informed by prior literature, our design accounts for the loss of facial-cue and other non-verbal communication through textual and visual mixed-reality augmentation [27]. Captions and future augmentations, in particular, present a promising medium for facilitating the non-verbal interpersonal communication obfuscated by mask form-factor [32]. Our prototype PokerFace mask uses recyclable cardboard material and a commodity smartphone to embed computation into a helmet-like mask form-factor. This design allows a live video stream of the user's mouth to be displayed on the mask's surface, rendering the user's face visible while keeping it covered and sealed by the mask's facepiece. Additionally, captioning of the user's speech can be displayed at the bottom of the video feed in real-time.

### **3.1. Implementation**

PokerFace supports mixed reality interactions by embedding a smartphone in the mask's facepiece. The infrastructure of the mask was created using recycled cardboard to architect sufficient support for the smartphone, while cradling the facepiece a sufficient distance from the face (Figure 2a). Double-ply cotton cloth was then used to seal edges of the mask around the nose and chin to prevent potential leakage of particles from the mouth cavity. To accommodate the offset position of the built-in camera on the smartphone, we also used a commodity endoscopic camera which attached to the USB-C port on the smartphone (Figure

2b). To display captions, either with or without video, we used the Microsoft Teams video calling platform with its real-time speech-to-text captioning service.



**Figure 2. PokerFace Prototype Sketches with Dimensions: a. front piece of mask with slot for smartphone; b. back of mask with endoscopic camera**

## 4. User Evaluation

To understand how the PokerFace prototype facilitates communication compared with traditional masks, we recruited 18 participants for a user study. The study comprised an entrance questionnaire, a 4-condition counterbalanced evaluation of our prototype, and an exit questionnaire and interview. Our study design was reviewed and approved by our institution's Ethics Review Board, and took place over the duration of 1 week.

### 4.1. Method

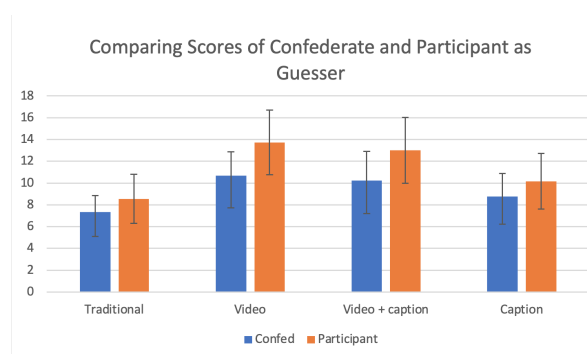
The 13 question online entrance questionnaire collected demographic information including age, occupation, self-described disabilities (if any), and gender, as well as surveying the user's previous experience wearing masks for PPE. Participants were then given a virtual tutorial via video conferencing software on playing Guess What, a charades-like game for mobile devices. Guess What is played by one player, called the guesser, placing a mobile phone on their forehead while it displays a picture of a noun which has not been seen by the guesser. The other player, called the clue giver, gives clues without saying the actual word displayed to get the guesser to say the noun. Upon correctly guessing the noun, the guesser tilts the phone down to display a new word on the mobile device. Players can "pass" at any time by tilting up if the word proves to be too difficult. The number of nouns the clue giver is able to communicate to the guesser within 90 seconds is tabulated as the score. The tutorial detailed how to play the game, and a practice round where both participant and research team member were the clue giver. The study involved a series of 8 games of Guess What under 4 game conditions:

traditional mask, PokerFace prototype mask with video, PokerFace prototype mask with video and captions, and PokerFace prototype mask with captions and without video. To protect the health of our participants under pandemic conditions, it was imperative that we limited the amount of mask removal and mask-changing which they performed during the study. Thus, we employed the use of a researcher as a confederate, who served as the playing partner for each participant, and was the only player to change their masks under each condition. For each of the 4 conditions, the participant's mask did not change, wearing their personal cloth or disposable traditional mask for PPE throughout the duration of the study, while the confederate changed their masks for each of the 4 conditions. Once the confederate had changed into 1 of the 4 masks as dictated by the condition, they played 2 games of Guess What with the participant, one game where the confederate served as the guesser and the participant served as the clue giver and a second game with these roles reversed. We only allowed the participant to "pass" on any noun in a game, since the confederate became quite proficient with the game. Between conditions, the confederate changed their mask (a safe distance away) according to the next condition, and again played 2 games as before. After completing the 8 games of Guess What under 4 conditions between the participant and the confederate, participants completed an online exit questionnaire followed by an online exit interview. The exit questionnaire comprised 12 short-answer and 20 Likert Scale questions asking participants to rate their experience with various aspects of interacting with and wearing traditional masks, as well as experiences interacting with the prototype mask during the study. When possible, questions involving specific aspects of interacting with people wearing masks were repeated using the same language for both traditional masks, as well as the PokerFace prototype mask. This was done to afford a Likert scale basis of directly comparing participant experiences with these mask features. The exit questionnaire also asked 16 short-answer questions (total of 36 questions) reflecting on their experiences with the 4 mask conditions. Semi-structured exit interviews directly followed completion of the exit questionnaire, giving our researcher an opportunity to review participant's responses to the exit questionnaire and clarify any vagueness or inquire about inconsistencies.

### 4.2. Participants and Data Collection

All 18 participants ranged in age from 18 to 84 and had previous experience using masks for PPE.

Study participation was strictly voluntary, taking place over a duration of approximately 90 minutes, and all participants received a \$50 gratuity for their time. Eight participants (44%) identified as “woman”, eight (44%) as “man”, one (6%) as “genderqueer”, and one (6%) preferred not to say, and 2 participants identified that they are DHH. All of our participants indicated that they had previously played Guess What or similar Charades-like game. We collected quantitative data in the form of the number nouns guessed correctly per game (score). During the exit questionnaire, Likert scores on a scale of 1 to 5 were collected with 1 meaning “not at all” and 5 meaning “very much so”. Qualitative data were accumulated during both the entrance and exit questionnaire, as well as the exit interviews, and recordings of the game sessions. These participant responses were selectively reviewed by 2 members of the research team who devised an overall coding approach using a grounded theory methodology. One of these two researchers then applied this agreed-upon coding approach to the remainder of the participant responses. Qualitative responses were coded for recurring themes in participant experience interacting with the confederate under the 4 study conditions. This analysis, as well as descriptive statistical treatment of the quantitative and Likert responses are detailed below.



**Figure 3. Results of user study comparing confederate and participants scores, reflecting the same trend across both datasets.**

## 5. Results

We first analyze the results of our quantitative data reflecting participant overall performance interacting with various features of the PokerFace prototype versus traditional masks. Then we look at the qualitative data collected in the form of Likert results and open-ended questions from our entrance and exit questionnaires. After examining the limitations of our design approach, we proceed to a discussion of these

**Table 1. Descriptive statistics of performance by condition including Confederate (C), Participant (P) and Combined Confederate and Participant (CP)**

Condition	Mean	Standard Dev.	Range
Traditional (C)	7.33	1.53	[4-11]
Traditional (P)	8.55	2.24	[5-12]
Traditional (CP)	7.94	1.53	[4-12]
Video (C)	10.67	2.21	[7-15]
Video (P)	13.72	2.96	[8-20]
Video (CP)	12.19	1.03	[7-20]
Video Captions (C)	10.22	2.70	[7-17]
Video Captions (P)	13	3.04	[9-18]
Video Captions (CP)	11.61	1.00	[7-18]
Captions Only (C)	8.78	2.10	[6-13]
Captions only (P)	10.17	2.54	[6-15]
Captions (CP)	9.47	0.94	[6-15]

results contextualized by the feedback provided by our participants in the exit interviews and questionnaires.

### 5.1. Participant Task Performance

Overall, our participants performed better at correctly guessing clues with any version of the prototype PokerFace mask compared to traditional masks (Table 1). The performance score was similar under the video without captions (mean 12.2, SD = 3.0) and video with captions condition (mean 11.6, SD = 3.2). Interestingly, participants performed better using the prototype mask with captions but without video (mean 9.5, SD = 2.4), than with traditional masks (mean 7.9, SD = 2.0). These findings suggest that the video of the mouth contributed more toward increased performance than the captions. Due to the small dataset of users (n=18), standard deviation for the reported results was rather wide (Figure 3).

### 5.2. Qualitative Likert Feedback

We also collected Likert scale results of participant experiences as part of our exit questionnaire. Participants were asked to rate various aspects of their experience interacting with the prototype, as well as traditional masks on a scale of 1 to 5 where 1 indicated “not at all” and 5 indicated “very much so”. Figure 4 shows a selected subset of Likert scale questions that compare reactions to traditional masks with the prototype. Overall, participants reported that communication was easier when wearing the PokerFace mask than traditional masks, with 90% of participants reporting a score of 4 or 5. When asked this same question in regard to traditional masks, 22% of

participants reported a score of 4, with none reporting 5. Other general trends evident in the Likert results indicate that participants generally do not enjoy wearing traditional masks, with 65% reporting 2 or below when asked if they “liked wearing a traditional mask” and none reported a score of 4 or 5. This is echoed by the 78% of participants who reported that traditional masks were not fun to use (rating them a 1 or 2). In comparison, most participants also found the PokerFace mask “fun” to interact with, reporting a score of 4 or 5 for 75% of participants, and none indicating a score of 1 or 2.

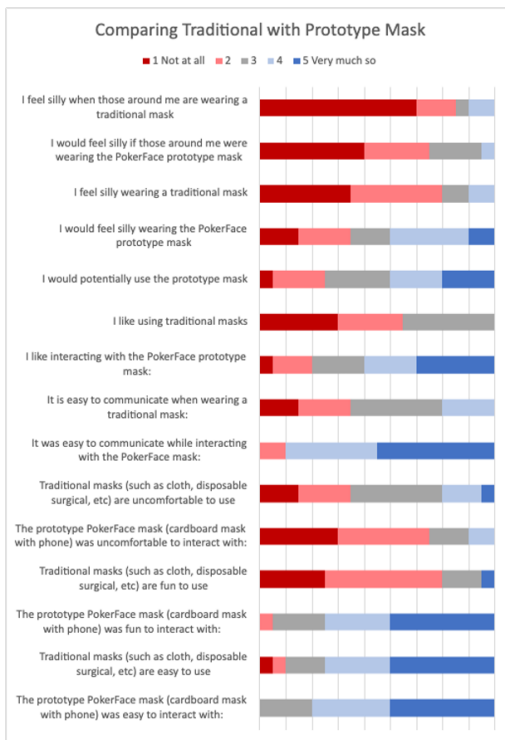


Figure 4. Results of Likert Scale questions from exit questionnaire and interviews.

### 5.3. Limitations of Our Results

Our 4-condition study design did not achieve the 24 participants needed for a full Latin Square counterbalanced study accounting for all possible permutations of condition ordering. Our principal concern in orchestrating our study was on the safety and health of our participants and researcher. Using a confederate as partner for all games played during the study introduces limitations to the data we collected as well as potential biases in our reported results. Since we used the same confederate across all conditions and orders, there is the potential that our confederate

could improve task performance over time due to learning effects. To account for this potential bias, we conducted an analysis averaging confederate and participant performance over time per trial (Figure 5).

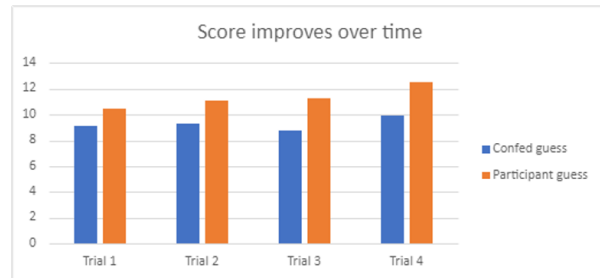


Figure 5. Averaged results by trial reflecting comparable performance by participants and confederate over time.

Figure 5 shows that both confederate and participant guesses remained comparable across all trials, with a slight improvement for both confederate and participant guessing over time. Note that the data also reflect that participants consistently performed better at guessing (on average) than the confederate. A similar per-condition analysis above that compared performance of participant and confederate for each condition (Figure 3) also found that the participant consistently performed better at guessing than the confederate across all conditions. These data confirm that the task remained meaningfully challenging over time, perhaps in part due to the mitigation of not allowing the confederate to pass on difficult clues. This indicates to us that influences of learning and order effects are fairly minimal. While acknowledging the limitations of our quantitative data, we emphasize that we relied on our qualitative and Likert results to understand the interaction experience with PokerFace. Furthermore, our participants were only able to reflect upon experiences interacting with the confederate wearing the prototype, and not wearing it themselves. However, given that the main benefits of the mask are experienced by those interacting with the mask wearer, and not in wearing the mask itself, reported participant experiences interacting with the mask are valuable to understand. Analysis of this is crucial to inform the design of future face-worn wearables and understand communication while wearing these devices.

## 6. Discussion

While general trends indicate that participants preferred the prototype mask because of its ability to better facilitate communication, these results were not unanimous. Participants indicated several social

stressors which should be considered for future XR mask designs. Users also provided feedback specifying use-cases where they thought the prototype would be useful in their work or social interactions. We discuss our prototype's perceived effect on communication, social awkwardness and wearability, and participant-reported use case scenarios.

## 6.1. XR Masks and Communication

Participants overall perceived communication to be easier when interacting with the PokerFace mask than when interacting with a person wearing a traditional mask, as also corroborated by the quantitative performance results. This perceptual and quantifiable improvement to communication can be attributed to two principal affordances enabled by the prototype: facial visibility and speech captioning.

**6.1.1. Face Visibility** Quantitative analysis of participant performance indicated that using the PokerFace prototype mask with video tended toward better performance at the collaborative communication game. Similarly, most participants indicated that the principal benefit to communication afforded by the prototype is the ability to see their communication partner's face. Participants reported that being able to read facial cues from the PokerFace video display was extremely helpful in understanding the wearer's intent. *"The interaction was easier and more engaging when wearing the prototype. Normally with a traditional mask I have to get closer than the 6ft minimum to hear what the other person is saying if they are a senior citizen or someone who is just soft spoken. I didn't realize how helpful facial cues are when talking. The prototype mask was a fun way to engage with the other person and see their mouth ...without removing the mask and feeling at risk of a virus."* (p9) Face visibility was almost unanimously indicated by our participants as being incredibly helpful in communicating, with many echoing p9's sentiment about reading facial cues. Including p9, fourteen participants (78%) volunteered that they noticed themselves reading their game partner's lips when visible through the live video feed. While two (p17 and p18) of these participants disclosed they were DHH, the others did not, suggesting that lip and facial cue reading is a useful affordance of the mask for all users. Even when not addressing the mask directly, participants indicated desires to see other's faces. *"I'll be glad when things go back to normal, if they ever do. I miss seeing peoples smiles without the masks."* (p12) The ability to perceive emotional cues

and facial expressions was a priority among numerous participants, especially those with highly social occupations. Participants who worked as educators (p8, p9, p11, p15) expressed a need to have their and their students' mouths be visible when teaching. Other participants who held social occupations indicated similar priorities with being able to see other people's faces when performing their work. For example, p1 is a beauty consultant, reflected that being able to see their customer's face was imperative for making successful suggestions, as well as their customers being able to see their smiling face to establish rapport. Furthermore, 89% of participants indicated in the exit questionnaire that the live-video feed was helpful in facilitating communication, and thus was a promising feature of our mixed-reality prototype.

**6.1.2. Captions** The speech-to-text captions displayed in real-time on the PokerFace screen was the second integral element in facilitating communication and improving performance in our study task. The better performance of the PokerFace mask with captions but without video condition seems to indicate an intrinsic benefit of captions independent of a live video display. All saw the potential for captions to be useful, and expressed preferences for how the captions should be displayed. Interestingly, some participants (p4, p6, p11, p12) indicated they preferred the captions without the video feed because the captions with the video were overwhelming. *"I'd actually prefer [captions without video]. Only because reading body language, listening to my partners voice, reading captions as well as looking at the screen of the other phone was more than enough to play charades. Introducing video was information overload."* (p12) Other participants indicated that they found captions with video more useful because they could use the captions to confirm what was being said. Two participants (p13, p15) compared this experience to "watching movies with subtitles". Our participants who are DHH reported the most enthusiastic support for captions in the qualitative feedback, indicating that the captions often filled-in words that the person may have missed when someone is speaking. *"I sometimes don't hear what said because the mask covers people's mouths. So, I just nod my head like I hear them...I don't want to be rude and say I missed something. So [seeing the captions displayed] was nice because I could check the screen if I didn't hear [the speaker]."* (p18) Both of our participants who are DHH preferred video with captions over video without captions. One DHH participant reported a reduction in stress when they stopped trying to hear the

speaker, and instead focused on reading the captions and lips of the speaker wearing PokerFace. These results indicate that captions are a promising accessibility feature for a mixed reality face-worn display.

## 6.2. Mask Form Factor and Social Interaction

Participants reflected a notable period of adjustment when interacting with a PokerFace mask wearer, citing a “social awkwardness” (p15) or an “uncanny valley effect” (p11) that subsided after continued interaction with the device. As mentioned above in Section 5.2, participants indicated an incongruity between their feelings of silliness when wearing a mask in public, and their feelings of silliness when others wear masks in public. When participants were asked if they would use PokerFace if available, 50% said they would. Participants who were skeptical about adopting PokerFace often held highly active, or less social positions, and their feedback reflected that the device would be less useful to them in these circumstances. When asked if they would feel silly wearing the PokerFace mask in public 50% of participants indicated that they would feel silly, while 73% of participants indicated that they would not feel silly if someone around them was wearing a PokerFace prototype mask. This reaction was contextually dependent, and even participants who anticipated feeling silly while wearing PokerFace in public, still delineated several scenarios where using the prototype in public could be beneficial. For example, several participants felt the novelty of the prototype was “fun” and would be jovial to wear to a party (p3, p4, p6, p8, p11, p12). Others identified use case scenarios where communicating may be difficult, such as trying to speak with friends at a location with substantial background noise, and thus the prototype could be beneficial for enabling communication (p1, p16, p17). These user-reported use cases, in combination with the feedback from participants who were enthusiastic about the potential of PokerFace, seem to indicate that all participants acknowledge that the device would be beneficial in facilitating communication in social circumstances.

The mask form factor became a topic of hesitation for participants as well, several of whom noted that the current prototype was “too bulky” (p8) for regular usage or “wouldn’t be able to fit in my purse” (p1). While future iterations of PokerFace could substantially reduce the overall bulkiness of its form factor, this concern from participants reflects a design priority for portability among participants. Subsequent designs of mixed reality face-worn wearables could make use of flexible displays to substantially reduce the overall size

and weight of the device. Weighing the trade-off cost of portability and durability of material is a design consideration that should be incorporated into the future creation of mixed-reality masks.

## 7. Mixed-Reality Mask Applications

Motivated by our study results and participant feedback, we describe potential applications and interactions enabled by mixed-reality masks such as PokerFace. Many of these interactions were provided by our participants during the user study, or were indicated as potentially useful applications extracted from qualitative user feedback analysis. While some of these use cases could be enabled by a simpler, transparent mask that enabled viewing the mask-wearer’s mouth, most rely on features that could only be enabled by digitally augmenting the mask. We acknowledge that the size and form factor of the current PokerFace prototype make it impractically complex, we believe that our study demonstrates the opportunities in this design space. In light of the lightweight, bendable displays and advances in generating avatar re-enactments of bodily expressions obscured by wearable devices [[42]], we believe that a practical prototype that affords the features of PokerFace could be implemented in the near future.

A common application suggested by our participants was communication with people who are DHH. This suggestion was particularly emphatic from our two participants who are DHH. Affording lip reading and presenting live captioning could enable more fluid communication for users who may be DHH. The built-in microphone and speaker on most commodity smartphones could be used to amplify the voice of the person speaking while wearing the PokerFace prototype, overcoming the muffling effect of traditional masks. Participants who were not DHH also identified additional communication use cases with these features, indicating that widespread adoption of an accessibility device with these considerations could be possible. Several participants worked in environments which were heavily bilingual, but were not bilingual themselves, often relying on translation services. Overwhelmingly, integration of live translation into future iterations of the PokerFace prototype or other mixed-reality masks was indicated as a valuable feature by these participants working in bilingual professions. Many real-time translation services currently exist, and integration of this feature into PokerFace, combined with the live video feed, could enable a user to communicate in a foreign language while the expression of their mouth remained visible. XR Filtering such



as those used in Snapchat could enable a variety of augmented communication capabilities, from personal expression and creative pursuits, to serving as an integrated element of a costume. This use case seemed particularly popular with users who often offered whimsical applications of the prototype in their feedback. For example, p10 asked “Is there any way to make a mask with the face of Mickey Mouse smoking a fake cigarette attached to it?” Similarly, users with facial paralysis [39] could potentially benefit from having augmentation filters which facilitate exhibiting desired facial expressions or restoring facial symmetry.

## 8. Conclusion

The PokerFace prototype explored how a mixed reality mask could overcome communication limitations of traditional face masks that became widely used during the pandemic. Our user study explored the various features of our prototype PokerFace mask, and compared the effectiveness of these features at allowing players to communicate while playing a collaborative communication game. On average, participants performed better using any of the 3 PokerFace conditions than with traditional masks. Feedback from participants indicated preferences for various features of PokerFace, as well as suggested potential use cases. Some participants preferred to use PokerFace with captions and without live video because it mitigated communication difficulties arising from voice muffling, without inducing information overload. Participants who worked in social disciplines such as education were more enthusiastic about the potential of PokerFace, while all participants, even those skeptical of our mixed reality mask, suggested social use cases where PokerFace could be beneficial. Social situations such as communicating with a person who is DHH, in environments which are loud, or in a foreign language were all indicated as potential communication use cases which could benefit from mixed reality masks. PokerFace is a preliminary work that highlights a promising new design space that could provide practical solutions to enable and facilitate better collaboration, as well as prove fruitful ground for novel XR interactions. We believe the potential posed by integrating mixed reality into masks merits future research even beyond the use of masks during a pandemic.

## References

[1] K. O’Dowd, K. M. Nair, P. Forouzandeh, S. Mathew, J. Grant, R. Moran, J. Bartlett, J. Bird, and S. C. Pillai, “Face Masks and Respirators in the Fight Against the COVID-19 Pandemic: A Review of Current Materials, Advances and Future Perspectives,”

*Materials*, vol. 13, p. 3363, Jan. 2020. Number: 15  
 Publisher: Multidisciplinary Digital Publishing Institute.

[2] K. Misawa and J. Rekimoto, “ChameleonMask: a human-surrogate system with a telepresence face,” in *SIGGRAPH Asia 2015 Emerging Technologies*, SA ’15, (New York, NY, USA), pp. 1–3, Association for Computing Machinery, Nov. 2015.

[3] J. Davis, “IllumiWear: A Fiber-Optic eTextile for MultiMedia Interactions,” in *NIME*, 2019.

[4] J. Lee, B. Kim, B. Suh, and E. Koh, “Exploring the Front Touch Interface for Virtual Reality Headsets,” in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’16, (New York, NY, USA), pp. 2585–2591, Association for Computing Machinery, May 2016.

[5] V. G. Motti, “Wearable Health: Opportunities and Challenges,” in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth’19, (New York, NY, USA), pp. 356–359, Association for Computing Machinery, May 2019.

[6] T. A. M. Spil, B. Kijl, and V. Romijnders, “The Adoption and Diffusion of Wearables,” in *ICT Unbounded, Social Impact of Bright ICT Adoption: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2019 Accra, Ghana, June 21-22, 2019. Proceedings*, pp. 31–47, Springer, May 2019.

[7] S. Eriksson, F. Gustafsson, G. Larsson, and P. Hansen, “Takt: The Wearable Timepiece That Enables Sensory Perception of Time,” in *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*, DIS ’17 Companion, (New York, NY, USA), pp. 223–227, Association for Computing Machinery, June 2017.

[8] B. MIJOVIC, I. GLIGORIJEVIC, D. Popovic, and M. D. VOS, “Wireless eeg headphones for cognitive tracking and neurofeedback,” Feb. 2019.

[9] J. Du, P. Markopoulos, Q. Wang, M. Toeters, and T. Gong, “ShapeTex: Implementing Shape-Changing Structures in Fabric for Wearable Actuation,” in *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI ’18, (New York, NY, USA), pp. 166–176, Association for Computing Machinery, Mar. 2018.

[10] C. H.-L. Kao, “Exploring Interactions and Perceptions of Kinetic Wearables.”

[11] C. Harrison, H. Benko, and A. D. Wilson, “OmniTouch: wearable multitouch interaction everywhere,” in *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST ’11*, (Santa Barbara, California, USA), p. 441, ACM Press, 2011.

[12] R. Xiao, T. Cao, N. Guo, J. Zhuo, Y. Zhang, and C. Harrison, “LumiWatch: On-Arm Projected Graphics and Touch Input,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*, (Montreal QC, Canada), pp. 1–11, ACM Press, 2018.

[13] B. Kim, J. Lee, and E. Koh, “Input techniques for virtual reality headset devices with front touch screens,” Nov. 2018.

[14] S. K. Kane, “Wearables,” in *Web Accessibility: A Foundation for Research* (Y. Yesilada and S. Harper, eds.), Human-Computer Interaction Series, pp. 701–714, London: Springer, 2019.

- [15] D. Jain, L. Findlater, J. Gilkeson, B. Holland, R. Duraiswami, D. Zotkin, C. Vogler, and J. E. Froehlich, "Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, (New York, NY, USA), pp. 241–250, Association for Computing Machinery, Apr. 2015.
- [16] M. Mielke and R. Brück, "A Pilot Study about the Smartwatch as Assistive Device for Deaf People," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, ASSETS '15*, (New York, NY, USA), pp. 301–302, Association for Computing Machinery, Oct. 2015.
- [17] M. Mott, J. Tang, S. Kane, E. Cutrell, and M. R. Morris, "'I just went into it assuming that I wouldn't be able to have the full experience': Understanding the Accessibility of Virtual Reality for People with Limited Mobility," July 2020.
- [18] "Smartwatch as an Assistive Technology: Tracking System for Detecting Irregular User Movement | Request PDF."
- [19] A. Fiannaca, I. Apostolopoulous, and E. Folmer, "Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility, ASSETS '14*, (New York, NY, USA), pp. 19–26, Association for Computing Machinery, Oct. 2014.
- [20] E. Tanuwidjaja, D. Huynh, K. Koa, C. Nguyen, C. Shao, P. Torbett, C. Emmenegger, and N. Weibel, "Chroma: a wearable augmented-reality solution for color blindness," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14*, (New York, NY, USA), pp. 799–810, Association for Computing Machinery, Sept. 2014.
- [21] P. Carrington, A. Hurst, and S. K. Kane, "Wearables and chairables: inclusive design of mobile input and output techniques for power wheelchair users," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, (New York, NY, USA), pp. 3103–3112, Association for Computing Machinery, Apr. 2014.
- [22] J. Gomez, J. C. Torrado, and G. Montoro, "Using smartwatches for behavioral issues in ASD," in *Proceedings of the XVII International Conference on Human Computer Interaction, Interacci&#xf3;n '16*, (New York, NY, USA), pp. 1–2, Association for Computing Machinery, Sept. 2016.
- [23] J. C. Torrado, J. Gomez, and G. Montoro, "Emotional Self-Regulation of Individuals with Autism Spectrum Disorders: Smartwatches for Monitoring and Interaction," *Sensors (Basel, Switzerland)*, vol. 17, June 2017.
- [24] V. G. Motti, "Assistive wearables: opportunities and challenges," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, UbiComp/ISWC '19 Adjunct*, (New York, NY, USA), pp. 1040–1043, Association for Computing Machinery, Sept. 2019.
- [25] CDC, "Coronavirus Disease 2019 (COVID-19)," Feb. 2020.
- [26] W. Kearns, J. M. Jasiewicz, J. L. Fozard, P. Webster, S. Scott, J. Craighead, M. E. Bowen, and J. McCarthy, "Temporo-spacial prompting for persons with cognitive impairment using smart wrist-worn interface," *Journal of Rehabilitation Research and Development*, vol. 50, no. 10, pp. vii–xiv, 2013.
- [27] C. S. Fichten, V. Tagalakakis, D. Judd, J. Wright, and R. Amsel, "Verbal and Nonverbal Communication Cues in Daily Conversations and Dating," *The Journal of Social Psychology*, vol. 132, pp. 751–769, Dec. 1992. Publisher: Routledge \_eprint: <https://doi.org/10.1080/00224545.1992.9712105>.
- [28] A. Mehrabian, *Nonverbal Communication*. Routledge, July 2017. Google-Books-ID: KIAPEAAAQBAJ.
- [29] N. Mheidly, M. Y. Fares, H. Zalzale, and J. Fares, "Effect of Face Masks on Interpersonal Communication During the COVID-19 Pandemic," *Frontiers in Public Health*, vol. 8, p. 898, 2020.
- [30] M. Schlögl and C. A. Jones, "Maintaining Our Humanity Through the Mask: Mindful Communication During COVID-19," *Journal of the American Geriatrics Society*, vol. 68, pp. E12–E13, May 2020.
- [31] "Through the looking glass (self): The impact of wearable technology on perceptions of face-to-face interaction."
- [32] S.-K. Lo, "The Nonverbal Communication Functions of Emoticons in Computer-Mediated Communication," *CyberPsychology & Behavior*, vol. 11, pp. 595–597, Oct. 2008. Publisher: Mary Ann Liebert, Inc., publishers.
- [33] S. Suntwal, S. Brown, and L. Brandimarte, *Pictographs, Ideograms, and Emojis (PIE): A Framework for Empirical Research Using Non-verbal Cues*. Jan. 2021. Accepted: 2020-12-24T20:20:00Z Pages: 6400.
- [34] Y. Wang, J. Amores, and P. Maes, "On-Face Olfactory Interfaces," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, (Honolulu HI USA), pp. 1–9, ACM, Apr. 2020.
- [35] T. R. Henry, A. K. Yeatts, S. E. Hudson, B. A. Myers, and S. Feiner, "A Nose Gesture Interface Device: Extending Virtual Realities," *Presence: Teleoperators and Virtual Environments*, vol. 1, pp. 258–261, Jan. 1992.
- [36] R. Lissermann, J. Huber, A. Hadjakos, S. Nanayakkara, and M. Mühlhäuser, "EarPut: augmenting ear-worn devices for ear-based interaction," in *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures the Future of Design - OzCHI '14*, (Sydney, New South Wales, Australia), pp. 300–307, ACM Press, 2014.
- [37] P. Nguyen, N. Bui, A. Nguyen, H. Truong, A. Suresh, M. Whitlock, D. Pham, T. Dinh, and T. Vu, "TYTH-Typing On Your Teeth: Tongue-Teeth Localization for Human-Computer Interface," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, (Munich Germany), pp. 269–282, ACM, June 2018.
- [38] K. Misawa and J. Rekimoto, "Wearing another's personality: a human-surrogate system with a telepresence face," in *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*, (Osaka, Japan), pp. 125–132, ACM Press, 2015.
- [39] J. L. Feuston, "Expressiv: An Expressive System for Individuals with Facial Paralysis," Apr. 2015. Accepted: 2016-06-07T17:56:14Z Publisher: Georgia Institute of Technology.

- [40] H. Y. Nam, I. Hernandez, and B. Harmon, “Unmasked,” in *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20 Adjunct, (New York, NY, USA), pp. 111–113, Association for Computing Machinery, Oct. 2020.
- [41] “UX research | Hyein Lee.”
- [42] E. Christinaki, N. Vidakis, and G. Triantafyllidis, “Facial expression recognition teaching to preschoolers with autism: a natural user interface approach,” in *Proceedings of the 6th Balkan Conference in Informatics*, BCI '13, (New York, NY, USA), pp. 141–148, Association for Computing Machinery, Sept. 2013.