



# Design and Evaluation of a Hands-Free Video Game Controller for Individuals With Motor Impairments

Atieh Taheri<sup>1\*</sup>, Ziv Weissman<sup>2</sup> and Misha Sra<sup>3</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, CA, United States, <sup>2</sup>Palo Alto Senior High School, Palo Alto, CA, United States, <sup>3</sup>Computer Science Department, University of California, Santa Barbara, Santa Barbara, CA, United States

## OPEN ACCESS

### Edited by:

Thomas Kosch,  
Darmstadt University of Technology,  
Germany

### Reviewed by:

Jun Shen,  
University of Wollongong, Australia  
Danny Hughes,  
KU Leuven, Belgium

### \*Correspondence:

Atieh Taheri  
atieh@ece.ucsb.edu

### Specialty section:

This article was submitted to  
Mobile and Ubiquitous Computing,  
a section of the journal  
Frontiers in Computer Science

**Received:** 01 August 2021

**Accepted:** 22 October 2021

**Published:** 10 December 2021

### Citation:

Taheri A, Weissman Z and Sra M  
(2021) Design and Evaluation of a  
Hands-Free Video Game Controller for  
Individuals With Motor Impairments.  
Front. Comput. Sci. 3:751455.  
doi: 10.3389/fcomp.2021.751455

Over the past few decades, video gaming has evolved at a tremendous rate although game input methods have been slower to change. Game input methods continue to rely on two-handed control of the joystick and D-pad or the keyboard and mouse for simultaneously controlling player movement and camera actions. Bi-manual input poses a significant play impediment to those with severe motor impairments. In this work, we propose and evaluate a hands-free game input control method that uses real-time facial expression recognition. Through our novel input method, our goal is to enable and empower individuals with neurological and neuromuscular diseases, who may lack hand muscle control, to be able to independently play video games. To evaluate the usability and acceptance of our system, we conducted a remote user study with eight severely motor-impaired individuals. Our results indicate high user satisfaction and greater preference for our input system with participants rating the input system as easy to learn. With this work, we aim to highlight that facial expression recognition can be a valuable input method.

**Keywords:** hands-free input, video game controller, quadriplegia, motor impairment, accessibility, video gaming, facial expression recognition

## 1 INTRODUCTION

For many people, video games are about experiencing great adventures and visiting new places that are often not possible in real life. People also build social and emotional connections through gaming (Granic et al., 2014). Yet, as prolific as gaming is, it is largely inaccessible to a significant number of people with disabilities (one of four American adults have a disability<sup>1</sup>). Video games are increasingly being used for purposes other than entertainment, such as education (Gee, 2003), rehabilitation (Lange et al., 2009; Howcroft et al., 2012), or health (Warburton et al., 2007; Kato, 2010). These new uses make game accessibility increasingly critical and even more so for players with disabilities who stand to benefit greatly from the new opportunities that video games offer.

Gaming is usually far more demanding than other entertainment media “in terms of motor and sensory skills needed for interaction control, due to special purpose input devices, complicated interaction techniques, and the primary emphasis on visual control and attention” (Grammenos et al., 2009). For individuals with degenerative neurological diseases such as muscular dystrophy

<sup>1</sup><https://www.cdc.gov/media/releases/2018/p0816-disability.html>.

(MD) or spinal muscular atrophy (SMA) grasping, holding, moving, clicking, or doing pushing and pulling actions, often needed for using console game controllers, is challenging and may present an insurmountable hurdle to playing. PC mouse and keyboard input is also not suitable for many of these users due to the need for bi-manual control necessary to control game camera and movement (Cecilio et al., 2016). To improve access, a number of input devices such as mechanical switches by Perkins and Stenning (1986), mouth and tongue controllers (Krishnamurthy and Ghovanloo, 2006; Peng and Budinger, 2007), mouth joysticks (Quadstick, 2020), brain-computer interfaces (BCIs) (Pires et al., 2012), and eye-gaze controllers (Gips and Olivieri, 1996; Smith and Graham, 2006) have been explored. The suitability of a device depends on the requirements of an individual as determined by the degree and type of muscle function available and targeted by that device.

Despite the wide variety, most of these devices are quite constrained with regard to the input that they can provide when compared with the variety and complexity of input required in conventional games. As a result, players may be restricted to playing greatly simplified games compared to games created for those with full muscle control. Although gaming software has started to include more options for different types of disabled players, there is still a great need for the design of new gaming input methods. Newer input types can give disabled players similar amounts of agency and control as non-disabled players, especially when software accessibility features are not helpful, as is the case for players with severe motor disabilities where using hands to control an input device is not an option.

To facilitate this, in this paper, we propose a novel hands-free input system, which translates facial expressions (FEs), recognized in a webcam video stream, into game input controls. The system is designed in collaboration with a quadriplegic player and includes speech recognition to serve as a secondary hands-free input modality. Our system contribution specifically pertains to the design of a hands-free interaction system. Although built using the known technique of FE recognition (FER), put together with keyboard mapping, speech input, and custom test games, the system holistically accomplishes novel functionality, which has not been explored in prior work. Specifically, our system provides a new hands-free method of playing video games that quadriplegics are otherwise unable to play with traditional input methods like a keyboard plus mouse, or a joystick, or a gamepad. Unlike BCIs and other input technologies designed for motor-impaired users, our system is inexpensive, easy to learn, and flexible and works without encumbering the user with sensors and devices.

The main contributions of this work are as follows:

- A fully functional prototype of FER-based video game control.
- Design of two games that demonstrate the mapping of FEs to game actions with focus on user agency, user comfort, ease of use, memorability, and reliability of recognition along with some design reflections.
- Results of an evaluation with quadriplegic individuals with neuromuscular diseases.

## 2 RELATED WORK

Over the past 20 years, researchers have been exploring the development of assistive technologies (ATs) to increase independence in individuals with motor impairments. We present here a subset of the work that investigates interaction with video games for a broader audience of disabled players as well as specifically for those with motor disabilities.

### 2.1 Accessibility in Consumer Video Games

Consumer games are increasingly incorporating accessibility options. For example, the accessibility options in Naughty Dog's 2016 release *Uncharted 4: A Thief's End*,<sup>2</sup> supported features like auto-locking the aiming reticle onto enemies, changing colors for colorblind users, and adding help to highlight enemies. Sony has also included a number of accessibility functions in their PS4 console,<sup>3</sup> including text-to-speech, button remapping, and larger font for players with visual and auditory impairments. In *The Last of Us: Part II*, released in June 2020, players can choose from approximately 60 accessibility features, such as directional subtitles and awareness indicators for deaf players, or auto-targeting and auto-pickup for those with motor disabilities.<sup>4</sup>

Players of a recently released game, *Animal Crossing: New Horizons*, are using the customization options of the game to make the game more accessible. For example, a blind player demonstrated how they modified the game in ways that do not rely on sight, whereas another player, a low-vision player, covered their island in grass and flowers to force fossils and rocks to spawn in familiar spots.<sup>5</sup> Not all commercial games are customizable, which leaves some players with disabilities to either ignore those games or seek the help of a friend or assistant to "play" the game. We designed our test games to match the visuals, difficulty, and gameplay of equivalent consumer games to evaluate playability with our proposed FE-based input method. In addition to adding accessibility options in commercial games, many special purpose games have been developed especially for blind players (Friberg and Gårdenfors, 2004; Yuan and Folmer, 2008; Morelli et al., 2010), with a large list of games available on the audio game website.<sup>6</sup>

### 2.2 Accessible Input Devices

Research has explored the design of interaction devices like the Canetroller by Zhao et al. (2018), which enables visually impaired individuals to navigate a virtual reality environment with haptic feedback through a programmable braking mechanism and vibrations supported by three-dimensional auditory feedback. *Virtual Showdown* by Wedoff et al. (2019) is a virtual reality

<sup>2</sup><https://daggersystem.com/disability-review-uncharted-4/>.

<sup>3</sup>[https://support.playstation.com/s/article/PS4-Accessibility-Settings?language=en\\_US](https://support.playstation.com/s/article/PS4-Accessibility-Settings?language=en_US).

<sup>4</sup><https://www.playstation.com/en-us/games/the-last-of-us-part-ii-ps4/accessibility/>.

<sup>5</sup><https://kotaku.com/how-animal-crossing-new-horizons-players-use-the-game-1844843087>.

<sup>6</sup><https://audiogames.net/>.

game designed for youth with visual impairments that teaches them to play the game using verbal and vibrotactile feedback.

The leading example of an accessible game controller is the Xbox Adaptive game controller by Microsoft (2020) that allows people with physical disabilities who retain hand/finger movement and control, to be able to interact and play games. By connecting the adaptive controller to external buttons, joysticks, switches, and mounts, gamers with a broad range of disabilities can customize their setup. The device can be used to play Xbox One and Windows 10 PC games and support Xbox Wireless Controller features such as button remapping (Bach, 2018).

The solutions presented here, although accessible, are not usable by those with severe motor disabilities as most of these solutions rely on hand-based control. In addition, although both software and hardware input controllers can make gaming accessible, we point out that the software can provide a more economical and customizable solution. Thus, in this work, we explore the design of an input system that works with any webcam or a mobile device camera without the need for any other hardware.

## 2.3 Game Input Methods for Quadriplegic Players

Assistive devices for users with quadriplegia include trackballs and joysticks (Fuhrer and Fridie, 2001), speech and sound recognition systems (Igarashi and Hughes, 2001; Hawley, 2002; Sears et al. 2003), camera-based interaction systems (Betke et al. 2002; Krapic et al., 2015), gaze-based interaction using EOG (Bulling et al., 2009) or eye tracking systems (Smith and Graham, 2006), switch-based systems (Switch, 2020), tongue- and breath-based systems (Patel and Abowd, 2007; Kim et al., 2013), mouth controllers (Quadstick, 2020), and electroencephalogram (EEG)/EMG-based systems (Keirn and Aunon, 1990; Williams and Kirsch, 2008). Of these ATs, we discuss below specific systems that have been used for controlling video games.

Most devices for gameplay collect signals from the tongue, brain, or muscles that the individual may have voluntary control over. There are several tongue machine interfaces (TMIs) such as tongue-operated switch arrays (Struijk, 2006) or permanent magnet tongue piercings that are detected by magnetic field sensors (Krishnamurthy and Ghovanloo, 2006; Huo et al., 2008) to enable interaction with a computer. Lau and O'Leary (1993) created a radio frequency transmitting device shaped like an orthodontic retainer containing Braille keys that could be activated by raising the tongue tip to the mouth superior palate. Leung and Chau (2008) presented a theoretical framework for using a multi-camera system for facial gesture recognition for children with severe spastic quadriplegic cerebral palsy. Chen and Chen (2003) mapped eye and lip movements to a computer mouse for a face-based input method.

Assistive devices based on BCI directly tap into the source of volitional control, the central nervous system. BCIs can use non-invasive or invasive techniques for recording the brain signals that convey the commands of the user. BCIs can

provide non-muscular control to people with severe motor impairments. Although non-invasive BCIs are based on scalp-recorded EEGs created using adaptive algorithms have been researched since the early 1970s (Vidal, 1973; Birbaumer et al., 1999; Wolpaw et al., 2002; McFarland et al., 2008), they have not yet become popular among users due to limitations, such as bandwidth and noise (Huo and Ghovanloo, 2010).

Motor-impaired users can play video games with a few consumer products. Switch (2020) is a non-profit dedicated to arcade style games that can be played with one switch. Quadstick (2020) enables three-way communication with computers and video games using a mouth-controlled device and to engage in social interaction through streaming on Twitch.<sup>7</sup> It includes sip/puff pressure sensors, a lip position sensor, and a joystick with customizable input and output mapping.

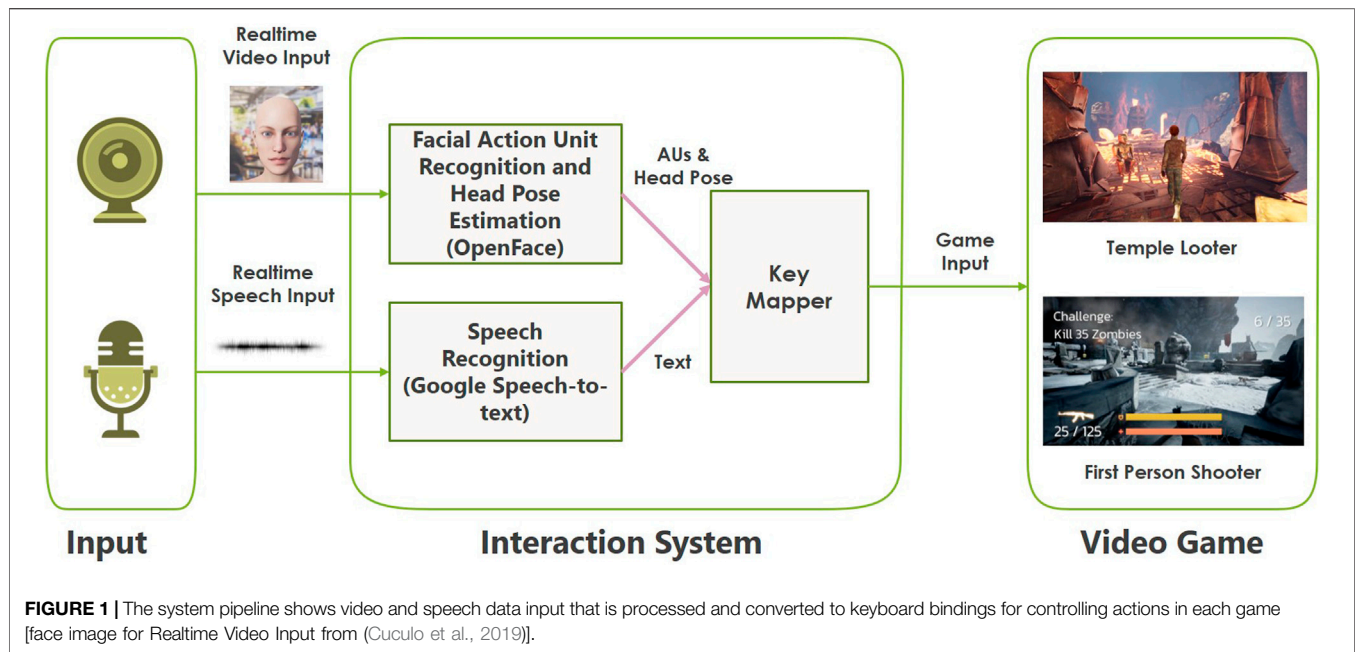
All these systems and devices have their unique affordances and limitations. For example, Quadstick (2020) is the most popular video game controller for quadriplegics, although it is expensive and needs updating with each new console release. There are several games where it is not possible to map a physical option on the Quadstick to a game action because of the large number of game actions possible. Our proposed software-based input method overcomes some of the limitations of prior devices by enabling fast and easy gameplay, the design of macros for complex input (e.g., jump + turn left) that can be mapped to a single FE, and customizable mapping of expressions to game actions. Hands-free interfaces like BCI require the user to wear a headset that may be difficult to wear and use for extended periods of time for playing games (Šumak et al., 2019). Our system is webcam-based and does not require the user to wear any sensors, trackers, or devices. To our knowledge, FER has not yet been investigated and evaluated on the basis of quantitative and subjective data in the context of game interaction for quadriplegic individuals.

## 3 SYSTEM DESIGN

Interaction design strives to create solutions that are generalizable to a large group of people. By contrast, ATs are usually tailored to the individual. In prior research, it has been shown that the best effects of an AT are seen when it is developed with and tested by potential end users (Šumak et al., 2019). The work that we present here uses the AT design method to develop a camera-based game input system and test games with the help of Aloy (real name withheld for anonymity), a quadriplegic engineering graduate student in our lab. Our co-design process is similar to that of Lin et al. (2014) who designed a game controller and a mouse for a quadriplegic teen.

Our design goal was to make use any small muscle movements available to people with severe mobility impairments to the fullest extent possible. The prior experience of Aloy with mouth-based and gaze-based systems was not so positive, so those input

<sup>7</sup><https://www.washingtonpost.com/video-games/2019/10/14/its-my-escape-how-video-games-help-people-cope-with-disabilities/>.



modalities were discarded. Because Aloy had voluntary control over only one finger, hand-based systems were also impractical. In contrast with other methods that require users to wear external hardware such as Earfieldsensing (Matthies et al., 2017) or Interferi (Iravantchi et al., 2019), we converged on a camera-based system that could use facial muscle control, which Aloy possessed, as input and support functionality using webcams or other camera devices that most users already own or can afford.

### 3.1 Pipeline

There are four main components to our system: 1) FER or facial action unit (AU) recognition along with head pose estimation, 2) speech recognition, 3) interaction design (AU and head movement to keyboard mapping), and 4) game design and gameplay. **Figure 1** illustrates the system pipeline. Through the two recognition systems (one for FEs and head pose and the other for voice), webcam and microphone input are sent to the keyboard mapper (**Section 3.4**), which converts them into keyboard input for each game. For creating the Temple Looter and First-Person Shooter (FPS) games, we used Unreal Engine (UE) version 4.25.1.

### 3.2 Video Input

Ekman and Friesen (1971) categorized facial muscle movements into facial AUs (FAUs) to develop the facial action coding system. There have been two major types of methods used for recognizing FAU over the years—those that use texture information and those that use geometrical information (Kotsia et al., 2008). Our system uses the OpenFace 2.0 toolkit developed by Baltrusaitis et al. (2018) that is based on capturing facial texture information, for FER and head pose estimation, hereafter referred to as OpenFace.

#### 3.2.1 Facial Input

AUs in our pipeline are detected in two ways: 1) AU presence—a binary value that shows whether an AU is present in the captured frame and 2) AU intensity—a real value between 0 and 5 that shows the intensity of the extracted AUs in the frame. OpenFace can detect AUs 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 25, 26, 28, and 45. In testing, we eliminated AUs 14, 17, and 20, because they were similar to other AUs. In addition, AU 45, which corresponds to blinking, could not serve as an input.

As soon as the player starts, the system begins to estimate AU presence and intensity values for all 18 AUs within each frame in the input video stream. The keyboard mapper then converts these values into game input. **Figure 2** (left) shows the six FEs that the player makes for taking actions in the games. These FEs are obtained by combining two or three FAUs (**Table 1**). **Figure 2** (right) shows Aloy playtesting the FPS game at home. The AU combinations of the game were determined experimentally by Aloy favoring the AUs with higher detection reliability than others. **Table 1** shows the AU combinations used in the games.

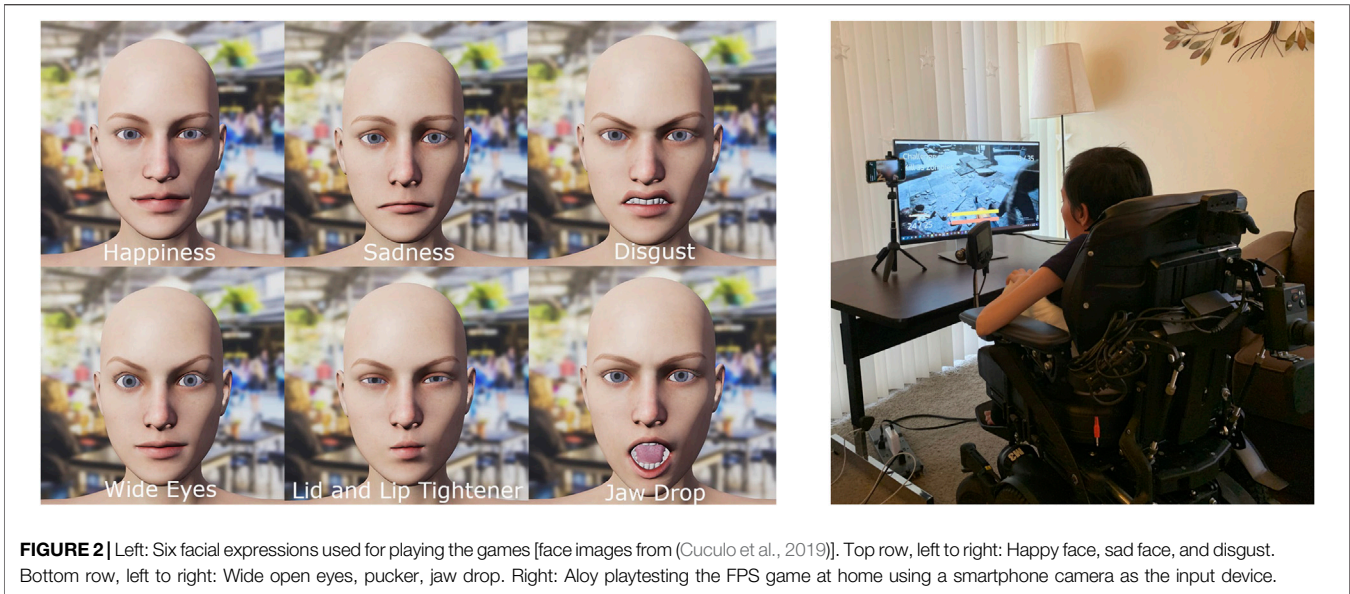
#### 3.2.2 Head Pose Input

Head gestures were included to augment the system because not all AU combinations are expected to work equally well for all users. We use head nodding instead of turning the head sideways that has greater potential for being falsely detected as input. In each frame, the six-dimensional head pose is estimated. If the player chooses to use head nodding as input in the customization interface, then we track the vertical movement through rotation angle around the x-axis to detect a nod.

### 3.3 Speech Input

We used a Python speech recognition library (Uberi, 2018) to communicate with Google Cloud Speech API Google (2020) for





**FIGURE 2** | Left: Six facial expressions used for playing the games [face images from (Cuculo et al., 2019)]. Top row, left to right: Happy face, sad face, and disgust. Bottom row, left to right: Wide open eyes, pucker, jaw drop. Right: Aloy playtesting the FPS game at home using a smartphone camera as the input device.

**TABLE 1** | Facial AU combinations with their descriptions and their approximate equivalent facial expressions mapped to keyboard keys.

Facial AUs combinations or facial expression			Keyboard key
AU6 + AU12	Cheek Raiser + Lip Corner Puller	Happiness	1
AU1 + AU4 + AU15	Inner Brow Raiser + Brow Lowerer + Lip Corner Depressor	Sadness	2
AU9 + AU10	Nose Wrinkler + Upper Lip Raiser	Disgust	3
AU1 + AU2 + AU5	Inner Brow Raiser + Outer Brow Raiser + Upper Lid Raiser	Wide Eyes	4
AU7 + AU23	Lid Tightener + Lip Tightener	Contempt	5
AU4 + AU25 + AU26	Brow Lowerer + Lips Part + Jaw Drop	Jaw Drop	6

converting spoken commands to text. The text data was scanned for specific keywords like “Walk” or “Yes” and converted into keyboard input using Pynput (Palmér, 2020) and mapped to keys previously programmed in UE for each action in each game. Speech interaction served as a backup modality to AU recognition and for interactions with the system such as pausing a game or choosing a game to play.

### 3.4 Key Mapping

FEs, head nods, and text keywords are mapped to keyboard input through the keyboard mapper. During testing, it became evident that AU recognition and mapping per input frame was frustrating to the user due to the system making multiple keyboard mappings per second leading to the Midas-touch problem. To resolve the issue, we set a threshold for the number of consecutive frames an AU combination needed to be visible in before getting mapped to the keyboard. This helped provide more control to the user and improved reliability. After testing with Aloy, all AU combinations were set to a five frame threshold.

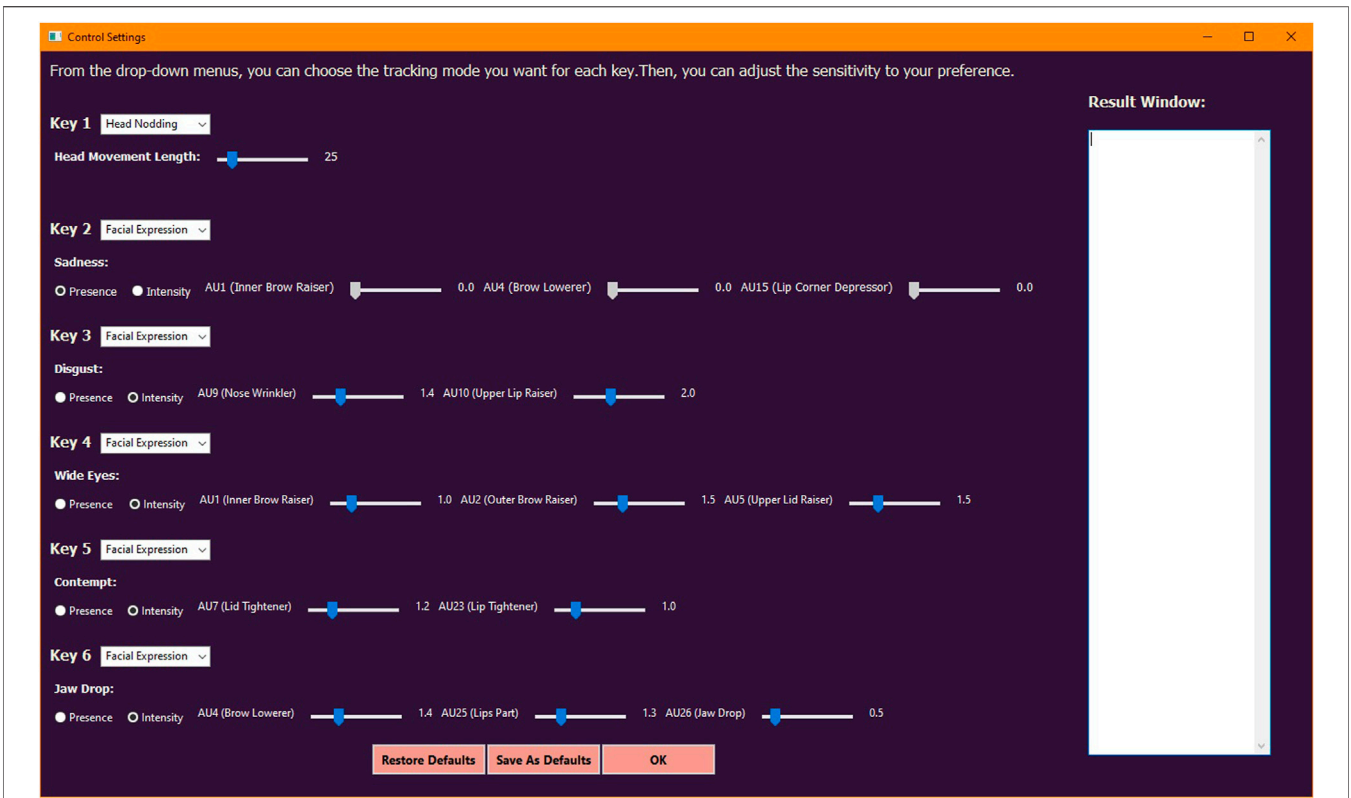
### 3.5 Customization

Although the setup and all adjustments presented in this work are best suited for Aloy and the thresholds configured to the

specific FE abilities of Aloy, we created an interface (Figure 3) to support personalization. Users can also choose a game action to be activated by head nodding, to replace the default input of a FE. When picking head nodding, the vertical range of movement is configurable and the user is encouraged to test and determine the values that work best for them. When choosing FEs, users can set the type and threshold of the AUs, again based on testing during the setup process. Customization gives the user the flexibility necessary for personalizing the system to their own needs. Figure 3 depicts the customization interface. As seen, head nodding has replaced the default FE of happiness for key 1.

## 4 GAME DESIGN AND GAMEPLAY

Over 91 different video game genres are available Wikipedia (2020). We implemented two games for the user evaluation that differ from one another. Temple Looter was used as a tutorial and the FPS was used for the study task. Table 1 shows the mapping of AU combinations to keyboard keys. The privacy of Aloy is maintained by replacing their face in all the figures with a virtual character making the same expression.



**FIGURE 3 |** We built a custom interface to enable users to change AU detection thresholds and head nod distance.

**TABLE 2 |** Left: Mappings of the keyboard keys to the actions defined in the Temple Looter game. Right: Mappings of keyboard keys to the actions defined in the FPS game.

Keyboard key	Game action	Input type	Keyboard key	Game action	Input type
1	Start Walking	AU	1	Start/Stop Walking Fwd	AU
2	Stop Walking	AU	2	Aim and Shoot	AU
3	Pick up	AU	3	Start/Stop Turning Left	AU
4	Sprint	AU	4	Start/Stop Turning Right	AU
5	Turn Yes	AU or Speech	5	Jump	AU
6	Turn No	AU or Speech	6	Pause	AU or Speech

### 4.1 Temple Looter

To introduce a player to the FER-based game input mechanism, we designed a simple game where the player moves through a cave environment and interacts with it using FEs as shown in Table 2 (left).

Adobe Fuse<sup>8</sup> was used to create the game character, which then was rigged and animated with Mixamo.<sup>9</sup> Ambient as well as task-related sounds are used in all the games. A map in Infinity Blade: Fire Lands<sup>10</sup> was modified to create the game level. An ancient temple scene similar to one seen in Indiana Jones movies

was created. Looting treasures in temples and running away is the objective of the game. A stamina bar was added, which the player needs to fill up by not spending too much energy before sprinting out of the temple (Figure 4, left).

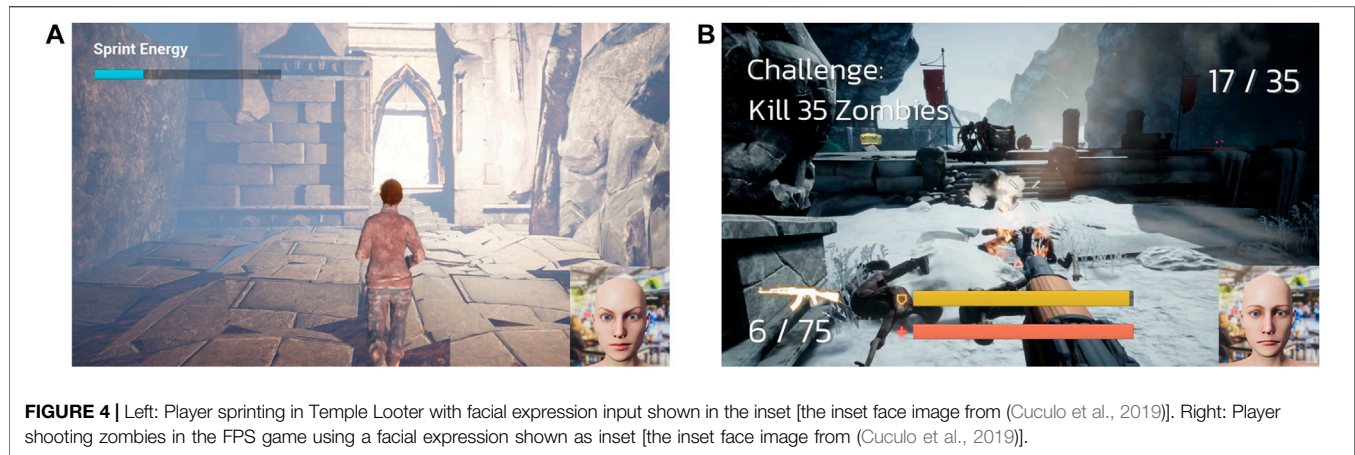
### 4.2 First-Person Shooter

The user study test game was an FPS, a popular video game genre. The FPS character and 40 animation sequences for covering typical movements in an FPS game were downloaded from Mixamo (Stefano Corazza, 2020). A blendspace was created to manage the animation logic playback with actions like walking, turning, jumping, aiming and shooting, reloading, and crouching. A single weapon option was added with sound effects and gunfire animation that showed as a flash at the tip of the gun (Figure ??, left). Zombies with different animations for walking, attacking, and dying were used as the enemy. Pathfinding logic was created

<sup>8</sup><https://www.adobe.com/products/fuse.html>.

<sup>9</sup><https://www.mixamo.com>.

<sup>10</sup><https://www.unrealengine.com/marketplace/en-US/product/infinity-blade-fire-lands?lang=en-US>.



for the zombies to move toward the player, when the player was within a certain distance range. A horde system was implemented to spawn new zombies on the basis of the heading of the player. In initial playtest of Aloy, the FPS proved difficult due to fast pacing for FEs and speech input. To improve gameplay and reduce frustration, we added an auto-aim feature and limited the number of zombies to 15. With auto-aim, the player character is turned by a defined amount per frame before the scoped gun is pointed at the nearest zombie that is within a predetermined range. Toggling a key for character movement replaced holding down a key continuously as in traditional FPS games. This way, the player could control the character by using their FEs, such as to walk forward or turn left or right. Because of the latency in cloud speech processing, speech input did not work as well for FPS gameplay as it did for Temple Looter. Thus, it was only used for pausing the game and not for the main actions. The mapping of FEs to FPS actions is presented in **Table 2** (right).

## 5 EVALUATION

To evaluate whether a FE-based input system is usable for playing video games by other individuals with neuromuscular diseases, especially those who have challenges playing with conventional PC game input (keyboard + mouse), we conducted a study with eight remotely located participants (in-person study was not allowed due to COVID-19 restrictions).

### 5.1 Method

#### 5.1.1 Participants

Twelve individuals were recruited from relevant Facebook groups created by and for people with MD and SMA. Eight of the 12 participants (five females, age range 18–45, two with MD, and six with SMA) were able to participate in the remote study. This sample size falls within the range of most prior research with quadriplegics (e.g., Corbett and Weber, 2016; Ammar and Taileb, 2017; Sliman, 2018), and in many cases, it is higher (e.g., Lyons et al., 2015; Soekadar et al., 2016; Nann et al., 2020). For instance, Ammar and Taileb (2017) explored EEG-based mobile phone control, and while they conducted an HCI requirement study

with 11 quadriplegic participants based on the work by Dias et al. (2012), they conducted their final usability study with five healthy participants. In contrast, our system was co-designed iteratively with participation of a quadriplegic individual (Aloy) and the full system usability and experience study was conducted with eight quadriplegic participants. It is worth noting that we recruited 12 participants. Of those 12, we conducted a pilot study with one individual. From the remaining 11 participants, we had to discard study data from three of them because of slow computers and non-working webcams that made it difficult to complete the study.

#### 5.1.2 Procedure

We conducted the study over Zoom videoconferencing and split the procedure into two separate sessions to minimize user fatigue, a regular response to physical exertion resulting from prolonged sitting, using the computer, or head nodding (Kizina et al., 2020). The study took each participant about 3 h. A 1.5-h first session began with participants providing informed consent (study approved by *anonymous for review*), completing a pre-study questionnaire with demographic questions and information about their background playing video games. Following the questionnaire, participants were walked through the installation of our system. Depending on their ability and hand muscle control as well as their computer setup (e.g., virtual keyboard, placement of webcam, and number of applications running on the computer), this step took the longest time, especially for those who did not have assistance or had assistants with little or no experience working with computers. After installation, participants were shown how to customize the system and tailor the settings to their facial muscle movement abilities. A tutorial game (Temple Looter; **Section 4.1**) was used to familiarize participants with FER-based gameplay.

Although the gameplay was different from the FPS (study task) game, the tutorial allowed participants to get comfortable with making FEs in front of their webcam, understand how long each expression needs to last, and control their expression speed when playing. To avoid exhausting the participants after 1.5 h of setup time, they played the FPS game in the second session,



**TABLE 3 |** The first part of the post-study questionnaire. The first two questions were rated on a seven-point Likert scale and the second two questions are open-ended questions.

Component	Question item	Minimum	Maximum
Efficiency	Overall, playing the FPS game was	Very Difficult	Very Easy
Satisfaction	How satisfied are you with using facial expressions as input for playing video games?	Very Unsatisfied	Very Satisfied
<b>Open-ended questions</b>			
What did you not like about the system?			
What did you like most about the system?			

**TABLE 4 |** The second part of the post-study questionnaire. The questions here were rated on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree). For each component, we also asked the participants to provide open-ended feedback on that component.

Category	Component	Question item
FER-based input-related	Sensitivity	Overall, the system is sensitive enough to detect my facial expressions while playing
	Learnable	I could learn how the system works in a short time
Game-related	Ease of Use	I found it easy to take actions in the FPS game using my facial expressions
	Comfort	Overall, using facial expressions to interact in the FPS game did not tire me

scheduled for a different day. Following the FPS gameplay, participants were asked to fill out a post-study questionnaire that was split into three parts: 1) system usability, 2) user experience, and 3) game experience. We also included an open-ended feedback question at the end of the post-study questionnaire asking about their overall experience.

### 5.2 Interaction Framework

For the first two parts of our post-study questionnaire, we used the theoretical framework of interaction by McNamara and Kirakowski (2006) to assess our usage of the input system. The framework focuses on functionality, usability, and user experience. It explores functionality by investigating how the controller supports the available game commands. The interaction method and how input is translated into game actions is presented in Section 3.

Iso (1998) describes usability as having three components: 1) *Efficiency*, 2) *Effectiveness*, and 3) *Satisfaction*. The purpose of our study is not to compare our input system with other input methods but to determine if a system such as ours can offer a viable option to those with limited choices in gaming input. We measured input *Effectiveness* on the basis of game completion set to a maximum of 25 min on the basis of a pilot study with one participant (not Aloy). *Efficiency* is measured by mental effort expended by the participants and *Satisfaction* by fulfillment of a mental desire. The first part of our post-study questionnaire measured *Efficiency* and *Satisfaction* as part of the overall usability measurement. There were four related questions: two evaluated with a seven-point Likert scale and two other open-ended questions asking participants what they liked and disliked most about the input system (shown in Table 3). The Likert scale for *Efficiency* is (1 = very difficult, 7 = very easy) and for *satisfaction* is (1 = very unsatisfied, 7 = very satisfied).

The last element of interaction design is user experience. Experience is the psychological and social impact of

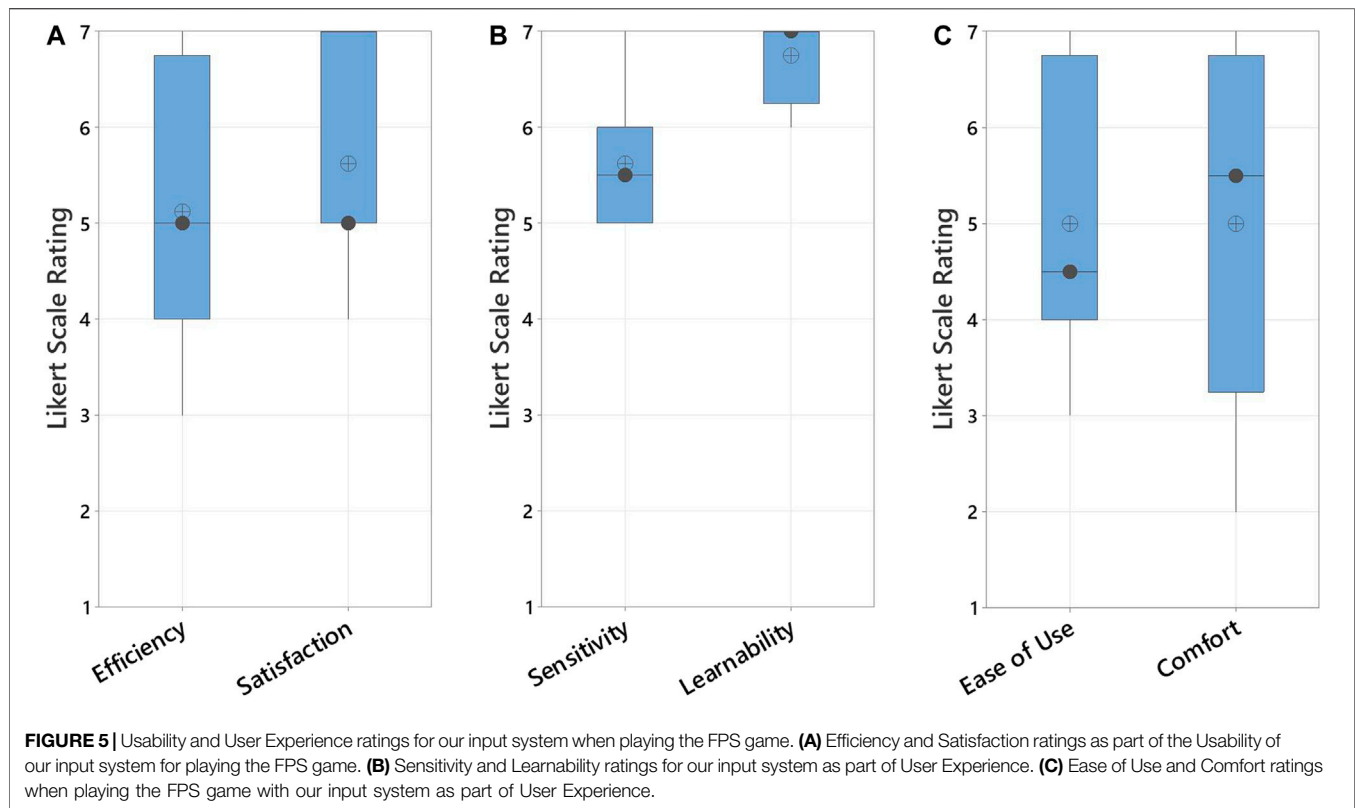
**TABLE 5 |** The third part of post-study questionnaire (GEQ) (Jsselsteijn et al., 2013). All questions were rated on a Likert scale (1 = strongly disagree, 7 = strongly agree).

Component	Statement
Sensory and Imaginative Immersion	S1: I was interested in the game’s story S2: It felt like a rich experience
Tension	S1: I felt frustrated playing the FPS game S2: I felt irritable
Competence	S1: I felt successful S2: I felt skillful
Flow	S1: I forgot everything around me S2: I was fully occupied with the game
Negative Effect	S1: I felt bored S2: I found it tiresome
Positive Effect	S1: I enjoyed it S2: I felt good
Challenges	S1: I felt challenged S2: I had to put a lot of effort into it

technology on users. This means impact beyond completing game tasks and is affected by external factors like design, marketing, social influence, and mood (Brown et al., 2010). The second part of the post-study questionnaire asked participants for feedback on their experience (Table 4). We collected data for this element using the Critical Incident Technique (CIT) (Flanagan, 1954). The questions fell into two categories:

- FER-based input-related questions: We asked participants four questions (two evaluated on a seven-point Likert scale with 1 = strongly disagree and 7 = strongly agree, and two open-ended) regarding how sensitive the input system was and how quickly they learned to use it.
- Game-related questions: These centered around the FPS video game played with our input system. Four questions





(two evaluated on a seven-point Likert scale with 1 = strongly disagree and 7 = strongly agree, and two open-ended) asked about the ease of use and how comfortable it was to use FER for playing the FPS game.

The third part of the questionnaire used questions from the Game Experience Questionnaire (GEQ) by IJsselstein *et al.* (2013). As many of the questions were not relevant to our task and game, we chose 14 of the 33 questions in the original questionnaire. GEQ categorizes all questions into seven factors. We picked two questions from each factor that were most relevant to our study (Table 5).

## 5.3 Results

### 5.3.1 Findings From the Usability Questionnaire

Results of *Efficiency* and *Satisfaction* as part of the usability of FER-based input system for playing the FPS game are shown in Figure 5A. As can be seen, the ratings are positive. Responses to open-ended questions about *Efficiency* and *Satisfaction* were also very positive.

The facts that participants were able to play without using their hands and that the system offers an easy to use alternative were two of the most appealing features of the system. P2 said, “[I]t was very intuitive and easy to learn how to use. Just the fact that I can have potentially one extra mode of input would be huge.” P3 agreed saying, “[B]eing able to toggle certain movements with just a FE was an interesting idea.” P4, P5, P6, and P7 also said the feature they liked most was being able to play

without using their hands. The favorite feature of P1 was that it is “Easy to navigate.”

In response to what they did not like about the FER-based input, participants expressed a desire to change all the mappings, although they appreciated the FER sensitivity customization that the system already provides. P4 said, “[I] wish you could swap FEs as inputs.” For some players, as also revealed in our testing, making some FEs was not easy. We expected this because each individual with neuromuscular diseases has different levels of facial muscle control. The comment of P5 reflected that “[I]t can be difficult to get the expressions right.” Interestingly, the *sadness* expression (AU1 + AU4 + AU15 or Inner Brow Raiser + Brow Lowerer + Lip Corner Depressor) was the most challenging for almost everyone). P7 said that there was “Nothing” they did not like, and P6 said that “I like it a lot”, which helps validate our fundamental design idea of using FER as a hands-free input method.

### 5.3.2 Findings From the User Experience Questionnaire

We collected data for “user experience” using the CIT (Flanagan, 1954). Table 4 summarizes the categorization and the components for each category that we developed: FER-based input-related = {*Sensitivity*, *Learnable*}, and Game-related = {*Ease of Use*, *Comfort*}. Each component has one question reporting the answer on a seven-point scale, depicted in Table 4, and one open-ended question asking the participants for additional feedback.

**TABLE 6** | Report of mean and standard deviation for each question item in GEQ along with Cronbach's alpha for each component.

Component	Statement numb	Mean	Standard deviation	Cronbach's alpha
Sensory and Imaginative Immersion	S1	4.75	2.49	0.95
	S2	5.25	2.28	
Tension	S1	3	1.66	0.77
	S2	1.75	1.09	
Competence	S1	4.75	1.36	0.9
	S2	5.88	1.66	
Flow	S1	4.5	2.12	0.52
	S2	5.38	1.49	
Negative Effect	S1	1.63	0.87	0.77
	S2	2.38	1.93	
Positive Effect	S1	6.13	0.93	1
	S2	6.13	0.93	
Challenges	S1	5.63	1.87	0.48
	S2	3.5	2	

As shown in **Figure 5B**, we see high ratings for both the *Sensitivity* and *Learnability* components of the first category—FER-based input. Several user comments for *Sensitivity* also provided positive feedback. Although P7 found the system to be fairly sensitive, P3 found it much too sensitive, and P8 was mixed. The high sensitivity can lead to the Midas-touch problem that we mitigated for Aloy (see **Section 3.4**) by increasing the number of frames in between detections. However, this element was not customizable for the user study experience. P7 offered design feedback on the customization interface: “I would like there to be some tooltips when you hover over the sensitivity sliders that tell you exactly what they govern and what they do. Some are obvious but others not much.” There were also comments regarding *Learnability* such as “As someone who uses computers a lot, it did not take me long to figure out how it all works and what the sliders did.”, P2 said. P7 remarked: “Learning process was pretty straightforward. I quickly figured out how to use it.” P6 even found the learning process entertaining and said, “It was fun.” Although many of the comments indicated that the learning process was short, P3 said, “It was usable but learning how to adjust the settings could be difficult for some users. It would certainly take some time.”

The results for the *Ease of Use* and *Comfort*, the two components of the game-related category, are shown in **Figure 5C**. In these two categories, the distribution of response ratings is wider, but it still resides on the higher side of the scale. The responses to open-ended questions were very varied. For the *Ease of Use* category, for example, P3 commented: “The overall system did take a lot of adjustments to get working with my FEs but worked decently when it was calibrated.” P7 said: “Once I learned which facial expression is connected to a specific action, it became very exciting.” On the other hand, P2 said: “Once I was able to figure out exactly how to do action number two, it became fairly easy. The facial expression required was not what I imagined when I was told to make a sad face. It required a lot more tension than I initially thought, but eventually it worked.”

Responses to the *Comfort* category were also diverse. “It could become tiring if playing for extended period.” said P5, and “Since I cannot play this type of game anymore, it was quite rewarding to be able to play without any difficulty in game control.” stated P7. With a fast-paced game like FPS, we expected some exhaustion, similar to

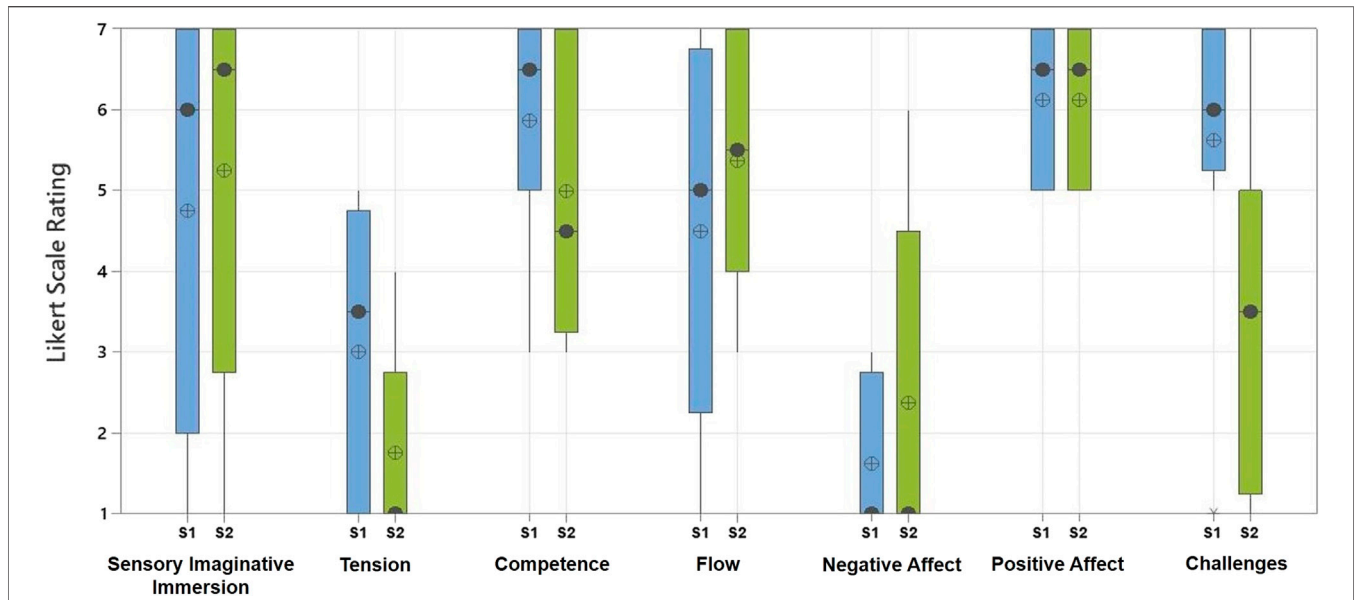
traditional input systems. Consequently, we put these last two categories under the game-related because gameplay affects the overall experience of the player, which does not depend solely on the controller. A criticism of the sad face expression from P2 showed the mapping of a FE to frequent game actions should be customizable along with game settings themselves (e.g., number and speed of zombies in the FPS game). P2 said, “Making a sad face to shoot zombies made my cheeks get a bit tired.”

### 5.3.3 Findings From the Game Experience Questionnaire

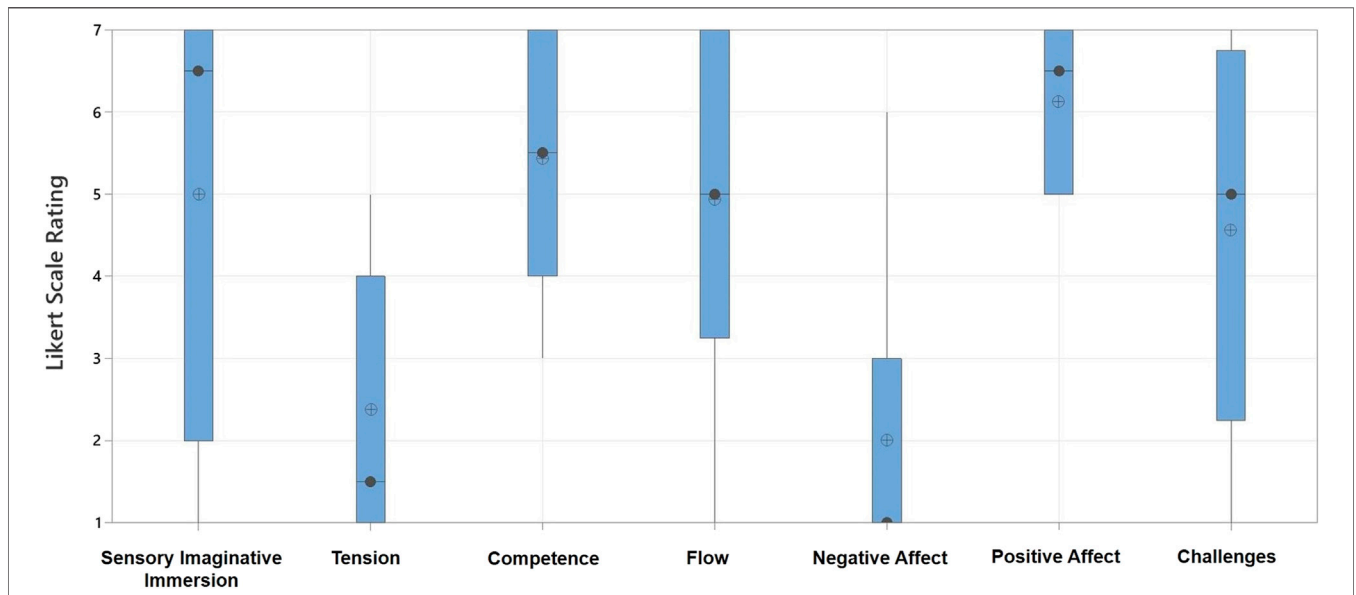
**Table 6** reports these values, along with the mean and standard deviation for each item. To measure the internal consistency between the two items in each component, Cronbach's alpha was calculated. Except for *Flow* and *Challenges*, all components have satisfactory internal consistency. These results indicate that the responses of the participants to the first question did not match their responses to the second question in these categories. For example, for *Flow*, although most of the participants were fully occupied in the game, it did not make them forget everything around them. On the basis of their ratings for *Challenges*, it appears that although they felt challenged, the effort required to play was variable across participants, leading to a wider distribution with a low mean (**Figure 6**). The visualization of the scores of the participants for each category is shown in **Figure 7**.

### 5.3.4 Overall Opinions About the System

At the end of the study, we asked participants for their overall feedback on the whole experience, which was very positive. A good number of comments mentioned that participants “loved it,” or that “it was fantastic” or “it was way too fun,” or “Overall, my experience today was easy and very straightforward. I had no issues getting anything to work the way it should.” One participant said they are “eagerly awaiting its availability” so that they can play games again and use it for other input. One participant mentioned that they would like to reduce the number of markers (AUs) required for FEs to one. As much as we value user input, we have to point out that limiting the number of AUs to one may potentially cause FEs to overlap, leading to difficulty in accurate detection. When we were iteratively developing our



**FIGURE 6 |** The ratings of the participants of each statements in every component of the GEQ post-study as part of post-study questionnaire. The statements for each component are shown in **Table 5**.



**FIGURE 7 |** participants ratings of each component of the GEQ questionnaire as part of the post-study questionnaire.

system, we realized that facial muscles are unconsciously linked together, and when you move one muscle on your face, it will inadvertently move one or two other muscles also, making them unusable for other controls. Another participant found the input system intriguing and said they would likely use it in combination with other input devices validating our inclusion of speech-based input as an added modality. Not only could additional modalities help with potential fatigue, they could also make the mappings more natural (e.g., saying “pause” vs. making a FE for pausing the game as used in the FPS game).

## 6 DISCUSSION

### 6.1 System Usability

As presented in **Section 5.3.1**, *Efficiency* and *Satisfaction* represent two components of system usability. The two related questions were rated moderately high and open-ended questions received very positive responses. Almost all of the participants were amazed that our system did not require hands to play video games and yet offered the ability to play a fast-paced and popular game like an FPS. Their feedback also indicates that changing the game mapping to suit different needs is

strongly appealing. For example, making a sad face was very easy for Aloy so they mapped it to a repetitive and perhaps critical action in the game. However, after conducting the study, we found that most participants found it difficult or tedious to make a sad face expression, which suggests that providing the option to change the mappings might be a way to accommodate the variable user abilities. Because most neuromuscular diseases affect facial muscle control over time, making all expressions might not be possible for everyone and even for one user that ability may evolve depending on physical therapy and disease progression.

## 6.2 User Experience

We evaluated user experience on the basis of *Sensitivity*, *Learnability*, *Ease of Use*, and *Comfort*. Participants rated the *Sensitivity* and *Learnability* of the system very high. This was also reflected in their answers to the open-ended questions. Although most participants felt the system was sufficiently sensitive, two said it was too sensitive. As discussed in **Section 3.4**, Midas-touch is a problem that we resolved for Aloy by setting a higher threshold for the number of consecutive frames that an AU combination must be visible in before registering it as detected input for mapping to the keyboard. However, on the basis of the study and feedback, this threshold may vary across users and thus needs to be customizable for each input expression during an initial system setup phase.

Participants rated *Ease of Use* and *Comfort* moderately high. Their comments also implied that the system is learnable and easy to use. Despite the fact that we had tried to map positive expressions to positive game actions, the comment about sad face being hard to make and the desire to change the mappings demonstrates the personal nature of expression-to-action associations and the ability to customize the input for each individual would be ideal.

## 6.3 Game Experience

Although our aim was to study and explore how users experienced our input system for video game playing, we had to find out how the players felt about the design of the game itself after they tested the input system. To do so, we utilized all seven factors from the GEQ that cover multiple aspects of game design. Four factors—*Tension*, *Competence*, *Negative Effect*, and *Positive Effect*—scored satisfactory.

We received a wide range of scores for *Sensory and Imaginative Immersion*. Cronbach's alpha and some analysis revealed why the distributions of scores for *Flow* and *Challenges* were wide—a low alpha value indicates inconsistency among questions in these factors, and for these two factors, we got low alpha values (**Table 6**). Although the majority of the participants did not strongly agree that they forgot everything around them, they did agree that they were fully occupied in the game which resulted in inconsistency in the *Flow* factor. Or, for the *Challenges*, the participants all strongly agreed they were challenged by the game. However, not all of them agreed that the challenge required a lot of effort. This is another affirmation that the game mechanism and input system were not overly complex and cumbersome.

## 7 DESIGN CONSIDERATIONS

Here, we articulate three design considerations that future developers of hands-free input systems for individuals with

neuromuscular diseases may want to consider. These are based on our experience of co-designing our system, feedback from the users, and the process of conducting the remote user study.

### 7.1 Design Input to Support Player Ability

Although it is obvious, understanding the capabilities of the target users, especially in this community where the situation and needs of each user are unique based on their disease progression, is the first step in determining the appropriate input method for them. In our design, we predominantly rely on facial muscle movements because those were the muscles Aloy had most voluntary control over. During the co-design process, we discovered that the number of muscles employed in each FE was an additional factor that needed to be considered for 1) the comfort level of the player, 2) the ability of the system to detect the expression reliably, and 3) the potential mapping to a game action. As Aloy tried several FEs, we found out that those with two or three AUs were most effective because they were easy to make repeatedly when needed and were most reliably detected without false positives.

### 7.2 Design for Personalization and Flexibility

The frequency of taking certain actions varies across games. For example, start/stop walking in *Temple Looter* or turning and shooting in the FPS is most frequently used. The FEs selected for these actions need to be fast and easy to make, whereas actions less often used can be relegated to either a secondary input method (e.g., speech input) or a more complex expression. During the study, user feedback pointed to a greater need for customization than our system currently supports, from mapping expressions to game actions to choosing whether or not to use expressions at all. The four FEs (happiness, sadness, disgust, and contempt) were relatively easy for Aloy to make, but the study showed that making the sadness expression was particularly challenging for some users. In addition, we had attempted to map positive expressions to positive game actions (e.g., smiling to moving forward) and negative expressions to negative actions (e.g., sadness to shooting) to assist the user in remembering the mapping. However, the study showed that mappings are more personal. Thus, to accommodate each player, enabling a change in mappings is another type of customization that should be supported. Last, given the limited set of FEs that are easy to make and detect reliably and the need to map them to a much larger set of game actions, combining game action sequences into “macros” is another customization possibility that would make the system broadly usable in a large variety of games.

### 7.3 Consider Fatigue and Disease Progression

The target audience is particularly prone to fatigue from muscle use and thus repeated actions like FEs to play a game can be exhausting, especially if a large number of muscles are involved in making those expressions. Reducing the number of muscles can lead to false positives. Hence, it is a fine balance between the number of AUs, the type of expression, and the game action, and this balance is best achieved by involving the individual in the design process. A characteristic of neuromuscular diseases is the progression and



change in voluntary muscle control over time. A system that integrates multimodal input (e.g., FEs and speech in our case) can help provide the player access for a longer duration as their disease progresses. Similarly, providing multimodal output through visuals, text, and sound data can help the player stay in control by letting them know that their FE or speech input was detected by the system and by allowing them to make informed decisions about next steps.

## 8 ASSUMPTIONS AND CHALLENGES

COVID-19–induced restrictions prevented in-person evaluation, and therefore, we conducted a remote study. This was incredibly challenging considering the degenerative neurological conditions of our participants. The ability to participate remotely required the ability to participate independently, even if assistance was available to install our application and setup the webcam. Despite being easy to install on any Windows 10 × 64 machine (a prerequisite for participating), the respective unique situations of our participants brought new challenges to each study session. The participant was presumed to be sitting in a wheelchair like Aloy, facing a monitor and webcam. However, one participant was unable to sit up and carried out the study lying down. Assuming everyone could hear over Zoom was another assumption (although we also shared instructions for installing and using Google Docs). Our best efforts failed to continue with one deaf participant. Furthermore, we assumed that having a PC meant having a functional GPU. The system of one participant had so many applications running that our application could not manage the real-time frame rate. Our goal was not to make our participants change their PC environment because they might have spent hours setting it up exactly how they needed it. As a result, this study had to be terminated prematurely. Our study attracted many people, but attempting to set up the input system on a PC with different specs can be challenging.

## 9 LIMITATIONS AND FUTURE WORK

Our work contributes to a body of research and design of hands-free gaming input for users with severe motor disabilities who need innovative solutions that can enable them to play independently. Our current system works for individuals who can voluntarily control their neck muscles (for the head nod gesture input) and facial muscles, which may exclude some users. In addition, it uses speech input that introduces latency and is not a viable option for some times of faster paced games and requires clarity in speech for reliable detection, which again may depend on muscle control for the user. The system also requires a front facing webcam that is pointed directly at the face of the user, which may not match the setup for some users.

Although our study focused on one of the most popular game genres (action games, specifically FPSs<sup>11</sup>), we did develop games from two other popular genres, sports, and adventure that we plan to conduct studies with in the future. On the basis of our pilot study, we determined that installing

and playing multiple games would be very time consuming and exhausting for our participant group, although actual gameplay may not be very long. Therefore, we decided to focus on one game in the study and seeing the results, we are encouraged to test our other games in the future. The main challenge with testing other games is the possibility of mapping game actions to FEs in a reasonable manner. This does rule out high speed games although our system enables creation of macros to successfully play games that require multiple keys to be pressed simultaneously or on quick succession (e.g., RPGs) to accomplish a task.

In addition, for future work, a greater degree of customization would be helpful in assisting a larger number of users with degenerative diseases. In addition, to make the system useful for a longer period for each user, it would be helpful if the system evolved as the disease progressed. Last, more work is needed to explore the use of FER for gameplay with commercially available games such that game developers can include FER as an accessibility feature to add to their growing list of features.

## 10 CONCLUSION

We proposed a hands-free game input system designed in collaboration with a quadriplegic student. We conducted a user study with eight participants with neuromuscular diseases to evaluate the usability of our system and the gameplay experience. In light of the unique needs of every motor-impaired person, we point out that our software solution can be easily customized to suit their abilities and needs, assuming that the individual is able to control their facial muscles. Because more and more game developers and companies are including accessibility features in their games, we are hopeful FER will be available soon, opening up a whole new world of gaming possibilities for people with severe mobility issues.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The UCSB Human Subjects Committee. The remote study participants provided informed verbal consent before the start of the study.

## AUTHOR CONTRIBUTIONS

All authors contributed to writing, editing, organizing, and reviewing the paper. The ideas presented in the paper were

<sup>11</sup><https://straitresearch.com/blog/top-10-most-popular-gaming-genres-in-2020/>.

formulated through a series of meetings to which all authors contributed.

## FUNDING

The user study was funded by the Google Research Scholar Award.

## REFERENCES

- Ammar, H., and Taileb, M. (2017). Smiletophone: A mobile Phone System for Quadriplegic Users Controlled by Eeg Signals. *INTERNATIONAL JOURNAL ADVANCED COMPUTER SCIENCE APPLICATIONS* 8, 537–541. doi:10.14569/ijacsa.2017.080566
- Bach, D. (2018). Xbox Adaptive Controller — the Next Level in Accessible Gaming.
- Baltrusaitis, T., Zadeh, A., Lim, Y. C., and Morency, L.-P. (2018). “Openface 2.0: Facial Behavior Analysis Toolkit,” in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi’an, China, 15–19 May 2018 (IEEE), 59–66. doi:10.1109/fg.2018.00019
- Betke, M., Gips, J., and Fleming, P. (2002). The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People with Severe Disabilities. *IEEE Trans. Neural Syst. Rehabil. Eng.* 10, 1–10. doi:10.1109/tnsre.2002.1021581
- Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., et al. (1999). A Spelling Device for the Paralyzed. *Nature* 398, 297–298. doi:10.1038/18581
- Brown, M., Kehoe, A., Kirakowski, J., and Pitt, I. (2010). “Beyond the Gamepad: Hci and Game Controller Design and Evaluation,” in *Evaluating User Experience in Games* (Springer), 209–219. doi:10.1007/978-1-84882-963-3\_12
- Bulling, A., Roggen, D., and Tröster, G. (2009). Wearable Eeg Goggles: Seamless Sensing and Context-Awareness in Everyday Environments. *J. Ambient Intelligence Smart Environments* 1, 157–171. doi:10.3233/ais-2009-0020
- Cecilio, J., Andrade, J., Martins, P., Castelo-Branco, M., and Furtado, P. (2016). Bci Framework Based on Games to Teach People with Cognitive and Motor Limitations. *Proced. Comp. Sci.* 83, 74–81. doi:10.1016/j.procs.2016.04.101
- Chen, C.-Y., and Chen, J.-H. (2003). “A Computer Interface for the Disabled by Using Real-Time Face Recognition,” in Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439), Cancun, Mexico, 17–21 Sept. 2003 (IEEE), 1644–1646. doi:10.1109/IEMBS.2003.1279691
- Corbett, E., and Weber, A. (2016). “What Can I Say? Addressing User Experience Challenges of a mobile Voice User Interface for Accessibility,” in Proceedings of the 18th international conference on human-computer interaction with mobile devices and services, Florence Italy, September 6 - 9, 2016 (New York City, NY: Association for Computing Machinery), 72–82. doi:10.1145/2935334.2935386
- Cuculo, V., and D’Amelio, A. (2019). OpenFACS: An Open Source FACS-Based 3D Face Animation System. *International Conference on Image and Graphics*. Springer, 232–242.
- Dias, M. S., Pires, C. G., Pinto, F. M., Teixeira, V. D., and Freitas, J. (2012). Multimodal User Interfaces to Improve Social Integration of Elderly and Mobility Impaired. *Stud. Health Technol. Inform.* 177, 14–25.
- Ekman, P., and Friesen, W. V. (1971). Constants across Cultures in the Face and Emotion. *J. Personal. Soc. Psychol.* 17, 124–129. doi:10.1037/h0030377
- Flanagan, J. C. (1954). The Critical Incident Technique. *Psychol. Bull.* 51, 327–358. doi:10.1037/h0061470
- Friberg, J., and Gärdenfors, D. (2004). “Audio Games: New Perspectives on Game Audio,” in Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology, Singapore, June 3–5, 2004, 148–154.
- Fuhrer, C., and Fridie, S. (2001). “There’s a Mouse Out There for Everyone,” in Calif. State Univ., Northridge’s Annual International Conf. on Technology and Persons with Disabilities.
- Gee, J. P. (2003). What Video Games Have to Teach Us about Learning and Literacy. *Comput. Entertain.* 1, 20. doi:10.1145/950566.950595
- Gips, J., and Olivieri, P. (1996). “Eagleeyes: An Eye Control System for Persons with Disabilities,” in The eleventh international conference on technology and persons with disabilities, 13.
- Google (2020). Speech-to-text: Automatic Speech Recognition | Google Cloud
- Grammenos, D., Savidis, A., and Stephanidis, C. (2009). Designing Universally Accessible Games. *Comput. Entertain.* 7, 1–29. doi:10.1145/1486508.1486516
- Granic, I., Lobel, A., and Engels, R. C. M. E. (2014). The Benefits of Playing Video Games. *Am. Psychol.* 69, 66–78. doi:10.1037/a0034857
- Hawley, M. S. (2002). Speech Recognition as an Input to Electronic Assistive Technology. *Br. J. Occup. Ther.* 65, 15–20. doi:10.1177/030802260206500104
- Howcroft, J., Klejman, S., Fehlings, D., Wright, V., Zabjek, K., Andrysek, J., et al. (2012). Active Video Game Play in Children with Cerebral Palsy: Potential for Physical Activity Promotion and Rehabilitation Therapies. *Arch. Phys. Med. Rehabil.* 93, 1448–1456. doi:10.1016/j.apmr.2012.02.033
- Huo, X., and Ghovanloo, M. (2010). Evaluation of a Wireless Wearable Tongue-Computer Interface by Individuals with High-Level Spinal Cord Injuries. *J. Neural Eng.* 7, 026008. doi:10.1088/1741-2560/7/2/026008
- Huo, X., Wang, J., and Ghovanloo, M. (2008). A Magneto-Inductive Sensor Based Wireless Tongue-Computer Interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 16, 497–504. doi:10.1109/tnsre.2008.2003375
- Igarashi, T., and Hughes, J. F. (2001). “Voice as Sound: Using Non-verbal Voice Input for Interactive Control,” in Proceedings of the 14th annual ACM symposium on User interface software and technology, Orlando Florida, November 11 - 14, 2001 (New York City, NY: Association for Computing Machinery), 155–156. doi:10.1145/502348.502372
- IJsselsteijn, W. A., de Kort, Y. A., and Poels, K. (2013). *The Game Experience Questionnaire*. Eindhoven: Technische Universiteit Eindhoven.
- Irvantchi, Y., Zhang, Y., Bernitsas, E., Goel, M., and Harrison, C. (2019). “Interfer: Gesture Sensing Using On-Body Acoustic Interferometry,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland Uk, May 4 - 9, 2019 (New York City, NY: Association for Computing Machinery), 1–13. doi:10.1145/3290605.3300506
- Iso (1998). *9241-11: 1998 Ergonomic Requirements for Office Work with Visual Display Terminals (Vdts)-part 11: Guidance on Usability*. Geneva, CH: ISO.
- Kato, P. M. (2010). Video Games in Health Care: Closing the gap. *Rev. Gen. Psychol.* 14, 113–121. doi:10.1037/a0019441
- Keirn, Z. A., and Aunon, J. I. (1990). Man-machine Communications through Brain-Wave Processing. *IEEE Eng. Med. Biol. Mag.* 9, 55–57. doi:10.1109/51.62907
- Kim, J., Park, H., Bruce, J., Sutton, E., Rowles, D., Pucci, D., et al. (2013). The Tongue Enables Computer and Wheelchair Control for People with Spinal Cord Injury. *Sci. Transl. Med.* 5, 213ra166. doi:10.1126/scitranslmed.3006296
- Kizina, K., Stolte, B., Totzeck, A., Bolz, S., Schlag, M., Ose, C., et al. (2020). Fatigue in Adults with Spinal Muscular Atrophy under Treatment with Nusinersen. *Sci. Rep.* 10, 1–11. doi:10.1038/s41598-020-68051-w
- Kotsia, I., Zafeiriou, S., and Pitas, I. (2008). Texture and Shape Information Fusion for Facial Expression and Facial Action Unit Recognition. *Pattern Recognition* 41, 833–851. doi:10.1016/j.patcog.2007.06.026
- Krapic, L., Lenac, K., and Ljubic, S. (2015). Integrating Blink Click Interaction into a Head Tracking System: Implementation and Usability Issues. *Univ. Access Inf. Soc.* 14, 247–264. doi:10.1007/s10209-013-0343-y
- Krishnamurthy, G., and Ghovanloo, M. (2006). “Tongue Drive: A Tongue Operated Magnetic Sensor Based Wireless Assistive Technology for People with Severe Disabilities,” in 2006 IEEE international symposium on circuits and systems, Kos, Greece, 21–24 May 2006 (IEEE), 4. doi:10.1109/ISCAS.2006.1693892
- Lange, B., Flynn, S., and Rizzo, A. (2009). Initial Usability Assessment of Off-The-Shelf Video Game Consoles for Clinical Game-Based Motor Rehabilitation. *Phys. Ther. Rev.* 14, 355–363. doi:10.1179/108331909x12488667117258

## ACKNOWLEDGMENTS

We would like to thank Mengyu Chen for assistance in testing our prototype. In addition, we would like to thank members of the Perceptual Engineering Lab in the Computer Science department for their insightful comments, questions and feedback.

- Lau, C., and O'Leary, S. (1993). Comparison of Computer Interface Devices for Persons with Severe Physical Disabilities. *Am. J. Occup. Ther.* 47, 1022–1030. doi:10.5014/ajot.47.11.1022
- Leung, B., and Chau, T. (2008). A Multiple Camera Approach to Facial Gesture Recognition for Children with Severe Spastic Quadriplegia. *CMBES Proc.* 31.
- Lin, H. W., Aflatoony, L., and Wakkary, R. (2014). "Design for One: a Game Controller for a Quadriplegic Gamer," in CHI'14 Extended Abstracts on Human Factors in Computing Systems, Toronto Ontario Canada, 26 April 2014 - 1 May 2014 (New York City, NY: Association for Computing Machinery), 1243–1248. doi:10.1145/2559206.2581334
- Lyons, F., Bridges, B., and McCloskey, B. (2015). "Accessibility and Dimensionality: Enhanced Real Time Creative independence for Digital Musicians with Quadriplegic Cerebral Palsy," in Proceedings of the International Conference on NIME, Baton Rouge, LA, June 2015 (NIME), 24–27.
- Matthies, D. J., Strecker, B. A., and Urban, B. (2017). "Earfieldsensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input through Facial Expressions," in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver Colorado USA, May 6 - 11, 2017 (New York City, NY: Association for Computing Machinery), 1911–1922. doi:10.1145/3025453.3025692
- McFarland, D. J., Krusienski, D. J., Sarnacki, W. A., and Wolpaw, J. R. (2008). Emulation of Computer Mouse Control with a Noninvasive Brain-Computer Interface. *J. Neural Eng.* 5, 101–110. doi:10.1088/1741-2560/5/2/001
- McNamara, N., and Kirakowski, J. (2006). Functionality, Usability, and User Experience. *interactions* 13, 26–28. doi:10.1145/1167948.1167972
- Microsoft (2020). Xbox Adaptive Controller. Available at: <https://www.xbox.com/en-US/accessories/controllers/xbox-adaptive-controller> (accessed on 0805, 2020).
- Morelli, T., Foley, J., Columna, L., Lieberman, L., and Folmer, E. (2010). "Vi-tennis: a Vibrotactile/audio Exergame for Players Who Are Visually Impaired," in Proceedings of the Fifth International Conference on the Foundations of Digital Games, Monterey California, June 19 - 21, 2010 (New York City, NY: Association for Computing Machinery), 147–154. doi:10.1145/1822348.1822368
- Nann, M., Peekhaus, N., Angerhöfer, C., and Soekadar, S. R. (2020). Feasibility and Safety of Bilateral Hybrid Eeg/eog Brain/neural-Machine Interaction. *Front. Hum. Neurosci.* 14, 521. doi:10.3389/fnhum.2020.580105
- Palmér, M. (2020). Pynput
- Patel, S. N., and Abowd, G. D. (2007). "Blui: Low-Cost Localized Blowable User Interfaces," in Proceedings of the 20th annual ACM symposium on User interface software and technology, Newport Rhode Island USA, October 7 - 10, 2007 (New York City, NY: Association for Computing Machinery), 217–220. doi:10.1145/1294211.1294250
- Peng, Q., and Budinger, T. F. (2007). "Zigbee-based Wireless Intra-oral Control System for Quadriplegic Patients," in 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE), 1647–1650. doi:10.1109/iembs.2007.4352623
- Perkins, W. J., and Stenning, B. F. (1986). Control Units for Operation of Computers by. *J. Med. Eng. Tech.* 10, 21–23. doi:10.3109/03091908609044332
- Pires, G., Nunes, U., and Castelo-Branco, M. (2012). Evaluation of Brain-Computer Interfaces in Accessing Computer and Other Devices by People with Severe Motor Impairments. *Proced. Comp. Sci.* 14, 283–292. doi:10.1016/j.procs.2012.10.032
- Quadstick (2020). Quadstick: A Game Controller for Quadriplegics. Available at: <https://www.quadstick.com/shop/quadstick-fps-game-controller> (accessed on 0805, 2020).
- Sears, A., Feng, J., Oseitutu, K., and Karat, C.-M. (2003). Hands-free, Speech-Based Navigation during Dictation: Difficulties, Consequences, and Solutions. *Human-comp. Interaction* 18, 229–257. doi:10.1207/s15327051hci1803\_2
- Sliman, L. (2018). "Ocular Guided Robotized Wheelchair for Quadriplegics Users," in 2018 TRON Symposium (TRONSHOW), Minato, Japan, 12-13 Dec. 2018 (IEEE), 1–7.
- Smith, J. D., and Graham, T. N. (2006). "Use of Eye Movements for Video Game Control," in Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, Hollywood California USA, June 14 - 16, 2006 (New York City, NY: Association for Computing Machinery), 20. doi:10.1145/1178823.1178847
- Soekadar, S. R., Witkowski, M., Gómez, C., Opisso, E., Medina, J., Cortese, M., et al. (2016). Hybrid Eeg/eog-Based Brain/neural Hand Exoskeleton Restores Fully Independent Daily Living Activities after Quadriplegia. *Sci. Robot* 1, eaag3296. doi:10.1126/scirobotics.aag3296
- Stefano Corazza, N. K. (2020). Mixamo - Get Animated.
- Struijk, L. N. S. A. (2006). An Inductive Tongue Computer Interface for Control of Computers and Assistive Devices. *IEEE Trans. Biomed. Eng.* 53, 2594–2597. doi:10.1109/tbme.2006.880871
- Šumak, B., Špindler, M., Debeljak, M., Heričko, M., and Pušnik, M. (2019). An Empirical Evaluation of a Hands-free Computer Interaction for Users with Motor Disabilities. *J. Biomed. Inform.* 96, 103249.
- Switch, O. (2020). OneSwitch.org.uk. Available at: <http://www.oneswitch.org.uk/> (accessed on 0805, 2020).
- Uberi, A. Z. (2018). Speechrecognition python Api
- Vidal, J. J. (1973). Toward Direct Brain-Computer Communication. *Annu. Rev. Biophys. Bioeng.* 2, 157–180. doi:10.1146/annurev.bb.02.060173.001105
- Warburton, D. E. R., Bredin, S. S. D., Horita, L. T. L., Zbogar, D., Scott, J. M., Esch, B. T. A., et al. (2007). The Health Benefits of Interactive Video Game Exercise. *Appl. Physiol. Nutr. Metab.* 32, 655–663. doi:10.1139/h07-038
- Wedoff, R., Ball, L., Wang, A., Khoo, Y. X., Lieberman, L., and Rector, K. (2019). "Virtual Showdown: An Accessible Virtual Reality Game with Scaffolds for Youth with Visual Impairments," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland Uk, May 4 - 9, 2019 (New York City, NY: Association for Computing Machinery), 1–15. doi:10.1145/3290605.3300371
- Wikipedia (2020). Video Game Genres. Available at: [https://en.wikipedia.org/wiki/Category:Video\\_game\\_genres](https://en.wikipedia.org/wiki/Category:Video_game_genres) (accessed on 0805, 2020).
- Williams, M. R., and Kirsch, R. F. (2008). Evaluation of Head Orientation and Neck Muscle EMG Signals as Command Inputs to a Human-Computer Interface for Individuals with High Tetraplegia. *IEEE Trans. Neural Syst. Rehabil. Eng.* 16, 485–496. doi:10.1109/tnsre.2008.2006216
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer Interfaces for Communication and Control. *Clin. Neurophysiol.* 113, 767–791. doi:10.1016/s1388-2457(02)00057-3
- Yuan, B., and Folmer, E. (2008). "Blind Hero: Enabling Guitar Hero for the Visually Impaired," in Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, Halifax Nova Scotia Canada, October 13 - 15, 2008 (New York City, NY: Association for Computing Machinery), 169–176. doi:10.1145/1414471.1414503
- Zhao, Y., Bennett, C. L., Benko, H., Cutrell, E., Holz, C., Morris, M. R., et al. (2018). "Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation," in Proceedings of the 2018 CHI conference on human factors in computing systems, Montreal QC Canada, April 21 - 26, 2018 (New York City, NY: Association for Computing Machinery), 1–14. doi:10.1145/3173574.3173690

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Taheri, Weissman and Sra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.