LOUVAIN RESEARCH INSTITUTE IN MANAGEMENT AND ORGANIZATIONS

WORKING PAPER

2019/10

A Gesture Elicitation Study for an Armband interacting with Netflix and Spotify

Robin Guerit,
Alessandro Cierro,

Jorge Perez-Medina
Louvain Research Institute in Management and Organizations

Louvain Research Institute in Management and Organizations
Working Paper Series
Editor: Prof. Valérie Swaen
(president-lourim@uclouvain.be)

A Gesture Elicitation Study for an Armband interacting with Netflix and Spotify

Robin Guerit, Alessandro Cierro, Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM)
Jorge Perez-Medina, Louvain Research Institute in Management and Organizations

Summary

Assessing the overall usability of mid-air and micro-gestures issued from an armband to control video, music, or an extended entertaining environment is still a research ques-tion today. To address this problem, this paper reports on results gained by jointly conducting and comparing two user studies. On one hand, a gesture elicitation study is conducted to measure the effectiveness and the user sub-jective satisfaction of gesture interaction by participants using a Thalmic Myo armband for interacting with NetFlix and Spotify, two plug-ins developed for this purpose. This study suggests new gestures which were not identified be-fore from the initial set of the 5 Myo standard gestures, based on agreement rates. On the other hand, a System Usability Scale (SUS) focusing on the usability and the satisfaction enables to identify which gestures are consid-ered usable and desirable in this context..

Keywords : Mid-air gestures; Gesture Elicitation Study; Wearable Computing; TV interaction.

JEL Classification: C81, C88.

.

Corresponding author :
Prof. Jorge Perez-Medina
CEMIS
Louvain Research Institute in Management and Organizations / Campus LLN
Université catholique de Louvain
Place des Doyens, 1
B-1348 Louvain-la-Neuve, BELGIUM
Email : jorge.perezmedina@uclouvain.be

A Gesture Elicitation Study for an Armband interacting with Netflix and Spotify

Robin Guerit, Alessandro Cierro
Université catholique de Louvain, ICTeam
B-1348 Louvain-la-Neuve, Belgium
{robin.guerit,alessandro.cierro}@uclouvain.be

Jorge Perez-Medina
Université catholique de Louvain, LouRIM
B-1348 Louvain-la-Neuve, Belgium
jorge.perezmedina@uclouvain.be

ABSTRACT
Assessing the overall usability of mid-air and micro-gestures issued from an armband to control video, music, or an extended entertaining environment is still a research question today. To address this problem, this paper reports on results gained by jointly conducting and comparing two user studies. On one hand, a gesture elicitation study is conducted to measure the effectiveness and the user subjective satisfaction of gesture interaction by participants using a Thalmic Myo armband for interacting with NetFlix and Spotify, two plug-ins developed for this purpose. This study suggests new gestures which were not identified before from the initial set of the 5 Myo standard gestures, based on agreement rates. On the other hand, a System Usability Scale (SUS) focusing on the usability and the satisfaction enables to identify which gestures are considered usable and desirable in this context.

Author Keywords
Myo Armband; Mid-air gestures; Gesture Elicitation Study; Wearable Computing; Netflix; TV interaction.

ACM Classification Keywords
• **Human-centered computing → Interaction devices; Gestural input**; *User studies; Mobile devices.*

INTRODUCTION
Fantasy and science fiction are ways to make humans dream about new kinds of interactions. As pointless as they may seem, Sci-Fi Fantasy stories are able to give insights on Human-Computer Interaction (HCI) design [21] while feeding the human sense with wonders [19]. For example, in George Lucas' Star wars movies [11], Luke Skywalker as a Jedi uses the Force to control and interact with his environment.

Figure 1. Luke Skywalker® uses hand-gestures to control the Force in Star Wars®.

With modern technologies, new devices came up and are used in many daily life contexts of use, such as domestic ones for controlling Internet of Things (IoT) devices, such as TV, recorder, communicating objects. There has been much research on how to control a TV [14,25] using different modalities of interaction, ranging from the traditional physical remote control, its substitute via a mobile application running on a tablet or a smartphone, to full-body gesture interaction. The main problem with traditional remote controls and their physical counterparts is that the end user is forced to always manipulate a physical device that may not be understandable, which is often inconsistent with another one, and which can be easily lost. When a new function is added to the TV, the remote control cannot be updated accordingly. Many types of full-body or hand gestures have been investigated to replace these physical devices in order to free the end user from using any physical device. However, these gestures need to be trained by the system and are sometimes difficult to learn, to produce, and to remember. This is especially the case when gestures are *system-defined*, i.e. gestures are standardized by the operating system and cannot be changed, no customization is possible. As opposed to system-defined gestures, *user-defined gestures* can be proposed by the end users themselves in order to come up with a consensus set of gestures [15]. It is not because these gestures emerge based on some consensus that they are simultaneously usable.

In this paper, we would like to investigate whether another interaction modality, i.e. an armband based on electromyography (EMG) [28], could become a realistic substitute for interacting with common tasks based on a TV by combining a usability testing and a gesture elicitation study.

The Myo Armband [28], built by Thalmic Labs is an off-the-shelf motion control device. It includes 8 EMG (electroMyographic sensors) able to measure electrical activity in the forearm muscles, thanks to which it identifies the user's arm gestures. It includes various sensors that enable the detection of arm movements. The sensors are a three-axis accelerometer, a three-axis gyroscope and a three-axis magnetometer. The Myo transmits a feedback to the user as the move he performs is recognized or when he unlocks the device. This feedback takes the form of vibrations (in a short interval of about 23.5ms, a medium interval of 45.5ms and a long interval of 100.5ms). Before using the device, the user can calibrate it (not mandatory but recommended). The calibration improves the gesture recognition as the muscles [23] in the arm are as unique as a fingerprint. If no calibration step is performed, then the default profile is applied. The device uses Bluetooth to be connected to the controlled tool (by using a smartphone or a computer).

This study used the control of Netflix thanks to the Netflix Connector: a free additional app built into Myo Connect (standalone Myo software) provided by Thalmic Labs and available in the official Myo Market. The Netflix Connector enables the control of the Netflix interface into the web browser. The control is set up through 7 standardized and non-changeable system-defined gestures (Fig. 2), which will be used for the navigation through both Netflix and Spotify.

Figure 2. Myo Amrband system-defined gestures.

Our practical results consist in: (1) an analysis of the agreement rates among the gesture proposals of the participants, (2) gestures comparison between TV interaction and music control with different positions, (3) usability testing scores about the predefined gestures of the Myo armband, (4) a comparison between the agreement rate pre and post experiment with the Myo armband, and (5) limits and discussion.

RELATED WORK

This section discusses prior work on the Myo armband technology and reviews the principles of the gesture elicitation methodology that we employ in this work to collect and understand the users preferences for mid-air gestures.

Technology and limitations of the Myo Armband

Previous studies show several technological aspects of the Myo armband in several domains of applications, such as, but not limited to [2,3]: healthcare [24], sports [13], virtual reality [30], and music [7]. For example, MyoGym [13] consists of an interactive application capturing up to 30 different physical exercises, monitor the movements of the end user, and reports on the progression over time. Another system [23] also relies on EMG to train the patient's muscles when they received a prosthetics. Squeeze gestures can be captured by the Myo in [6] by children suffering from cerebral palsy when they manipulate objects in virtual or mixed reality.

The aim of this section is to explore some of the limits or possibilities of the EMG and the other sensors included in the armband. In the field of bio-Mechatronics A. Khan and Z. Kausar [4] explained how they resolve multiple troubles thanks to the Myo armband in comparison with the Muscle Kit v3 (used with an Arduino Mega 2560). The purpose of their study was to analyze intermittent positions of the elbow anywhere in between the extreme positions using the EMG sensor. The research focused on the flexion-extension motions with a range of 0 to 80 deg. The problems they faced with the Muscle Kit v3 were unwanted noise caused by a little disturbance or the fact that only one intermittent position between the two extreme positions was captured. The EMG signal quality of the Muscle Kit v3 depended on the skin preparations of the participant. The EMG signal from the Myo armband was able to capture five angular positions: 0°, 20°, 40°, 60° and 80° [4].

Another research compared the use of the Nova biomedical base sensor (a non-invasive wearable system which allows the measure of complex physiological phenomena) [23] and the Myo armband for muscle fatigue detection. By applying the unidirectional ANOVA test, the results demonstrate the feasibility of using the Myo armband in HCI and limited medical scenarios where high accuracy levels of EMG signal are not required. Researchers used their own mid-air gesture recognition algorithm and showed an enhancement of the gesture recognition accuracy [10]: their approach outperforms the original Myo algorithm with an overall accuracy of 95% compared to 68%.

Gesture elicitation

Understand, collect and analyze users preferences with an interactive technology is the purpose of gesture elicitation. The end-users are often not fully consulted by designers and developers when they create gesture interfaces [8]. Many studies use gesture elicitation, also known as "participatory design" [9,18] or "guessability studies" [15] for a large variety of sensing devices and domains [26]. The outcome of a gesture elicitation study consists in a characterization of users' gesture input behavior with valuable information for designers, practitioners, and end users regarding the consensus levels between participants (computed as agreement or co-agreement scores and rates [15,25,26]), the most frequent (thus, generalizable across users) gesture proposals for a given task, and insights into users' conceptual models for performing tasks. The most recent formalization of the elicitation methodology proposed both repeated measures and between-subjects [25] experimental designs. Existing research worked on gesture elicitation in the domain of television interaction and showed that the gestures are not carried out on a 2D surface, but in mid-air [8]. Another gesture elicitation study worked in a smart home as a bigger environment [26].

RESEARCH QUESTIONS

First, we divided our experiment into four phases in order to understand and test our research questions and hypothesis defined below:

RQ1. What are the users' preferred gestures with an interactive armband for the TV interaction in their comfortable viewing position? Which gestures created by the participants are the most adapted to the requested action (referent)? Analysis and evaluation of gestures proposed under objective indicators (agreement rates) and a subjective questionnaire.

RQ2. What are the users' preferred gestures for the use of an interactive armband to control music in a different active position? How many changes? Hypothesis 1: Users use the same gestures to control music as for the TV interaction.

RQ3. What is the Myo usability? Learning and testing. What is the usability of the Myo for the Netflix interaction in a comfortable position?

RQ4. What are the actions ultimately taken after the experiments? Compromise and justification of the use of the Myo bracelet.

EXPERIMENT AND METHODOLOGY

We conducted a gesture elicitation experiment by following the methodology from the literature [8,13,15,25,26] and a usability test was also conducted [25, 3, 22].

Participants

We interviewed 21 participants (10 female), aged between 15 and 71 years old (M = 35.86, SD = 18.3), who volunteered for our experiment. Two pilots and 19 for the main study. We had to exclude 2 participants who did not succeed the pretest. The majority of the participants (16) were right-handed.

Apparatus

The experiments took place in two rooms set up to make the participant feel almost at home. In each setup there were a two-seat sofa and some blankets and cushions were available to allow a usual comfortable viewing position. We interviewed each participant one at a time and the experiment lasted about 1 hour. We used *Okja*, the South Korean-American movie, for all the Netflix phases and Michael Jackson's *Don't Stop 'Til You Get Enough* for the Spotify phases in order to stay neutral and avoid both choking scene and explicit lyrics. The participant wore the Myo armband (dominant arm) during all the experiments. We introduced the armband without showing any gesture. We informed the participant about the sensors included in the armband and about the possibility to perform mid-air gestures and finger gestures. The whole experiment was recorded on camera.

Figure 3. Participant performs Myo gestures to interact with Netflix with the support of gestures cards and the experimenter.

Procedure

Phase 1 – After introducing the Myo armband, we asked the participant to take a comfortable viewing position on the sofa with all the cushions needed and the first phase began: we presented the 7 referents they were going to elicit (unlock, play, pause, rewind, go forward, volume up, volume down). We allowed the participant to view all referents at once (after presenting the "unlock" referent alone) and they could revisit their gestures at any time. In doing so, we encouraged the participant to try out several possibilities for each referent. The participant had 5 minutes to think about his gestures. When he was ready, we asked him to perform each referent by using his own suggested gestures and in the same time we controlled the computer to allow him to foresee the results. We ended the first phase by asking the participant to complete a gesture questionnaire for each of his 7 suggested gestures and several subjective questions were asked.

No.	Referent	Description
1	Unlock	Turn on the TV set
2	Play	Turn off the TV set
3	Pause	Increase the volume
4	Forward	Decrease the volume
5	Rewind	Turn off the volume
6	Volume Up	Go to the next channel
7	Volume Down	Go to the previous channel

Table 1. Set of referents used for the elicitation experiment

Phase 2– The second phase took place in the same environment, but the participant had to quit their comfortable position and stand up. They were asked during this phase to use the Myo armband to navigate in Spotify instead of Netflix. The same 7 referents were used, and they had 5 minutes to find out if they were going to use the same gestures as in the phase 1, change all the gestures or change some of them. Once again, we simulated the Spotify system to allow the participant to imagine the interactions related to their gestures. The aim of this phase is to observe the transferability of the suggested gestures during the first phase from one context of use to another.

Phase 3– During the third phase we introduced the standardized Myo gestures for each referent and the armband manager gestures interfaces (Figure 4). To allow the best recognition with the armband, we created a new profile for each participant. We used the calibration step (needed to create a new profile) to teach the standardized gestures. Once the calibration set up, the participant was asked to lay back in their previous comfortable viewing position and to perform all the gestures for training. During the practice time we asked the participant to perform each action (randomly chosen) when asked, this phase lasted 5 minutes. We finally asked the participant to perform the standardized gestures following their own desires for 5 minutes. Then the participant completed a gesture questionnaire for each referent and a System Usability Scale test. During this phase we used the Myo armband instead of only simulate the results as in the two previous phases.

Figure 4. Myo information interface (visual feedback).

Phase 4– During the last phase, the participant was asked to stand up as we asked them to use Spotify one last time. We asked the participant to figure out (after suggesting their own gestures, during the phase 1 and 2, and learning the standardized Myo gestures) what gestures they wanted to keep to perform the referent actions for Spotify. They also were aware they could suggest completely new gestures. Having taken 5 minutes to consider the matter, we asked the participant to perform each referent while we simulated the results (as it was the case during the phase 2). An IBM CSUQ questionnaire was completed by the participant, and the final open questions were asked through a discussion. This 19-question questionnaire is preferred because it is empirically validated with a large number of participants on a significant set of stimuli [17], it is widely applicable for any user interface, and it benefits from a proved α=0.89 reliability coefficient between its results and the perceived system usability [16].

Measures

We employed measures to evaluate and understand the users preferred gestures and the users' performances in using the Myo armband:

1. We asked the participants to assume a VIEWING POSITION that they usually employed when watching TV at home. According to the results of the [7], we included the influence of a natural, comfortable position during the interactive test with Netflix. To assure a great ecological validity in our research, we will describe all different participants positions and include their remarks if they had difficulties related to their positions.
2. A CREATIVITY test was asked before the beginning of the experiment. It is made of 40 questions and gives a free assessment of one's level of creativity through 8 different metrics (abstraction, connection, perspective, curiosity, boldness, paradox, complexity and persistence)[1]. The participants needed a score higher than 55 to be part of the experiment. We rejected 2 participants because of a too low score.
3. We also conducted a MOTOR SKILL test in order to select people for the elicitation step, we measured motor ability levels in fingers. It is a developmental NEuroPSYchological assessment from NEPSY test batteries [22]. The test consisted in touching their thumb with the other fingers tip 8 times.
4. In relation to the gesture elicitation methodology, we calculated for each referent the central measure, the agreement rate (AR) as defined in [25]. This indication allows us to understand the extent to which the participants agreed on each gesture. Proposed gestures for specified tasks are analyzed and used in comparison in order to predict the most suitable gestures.
5. A questionnaire with subjective measures (7 questions) on a 7-point Likert scale [27] was developed to allow the participants to rate the quality of a gesture. They could indicate the ease of execution, the memorability, the good correspondence or the enjoyability in performing. [12]
6. Developed by John Brooke in 1980 [19], the System usability Scale (SUS) [1,19] is a questionnaire model on 5-point Likert scale [27] used to measure a system usability. 5 questions are framed in a positive way and the other questions in a negative way.
7. We use a second empirically-validated questionnaire model, the IBM Computer Satisfaction Usability Questionnaire (IBM CSUQ) [16,17], very closed to the SUS in order to standardize the results obtained and to check the results. It consists of 19 questions on a 7-point Likert scale, the questionnaire indicates the correlations with the usability of the system tested.

RESULTS

Participants viewing position

According to the analysis of the comfortable viewing positions (Table 2), we distinguished 17 different positions on 5 dimensions: the position of the upper body, the feet, the back, arms and items. Some were in a semi upright position, some used different items like cushions and blanket and some put their feet on the couch. This means that the participants were free to choose their own position but also that it seems difficult to categorize relaxed viewing positions.

Upper body	Feet	Back	Arm	Items
Sitting semi upright	On the couch, knees to right	On the back couch	Released on a cushion	On cushion and a blancket
Semi lying on the couch	On the couch, legs straight	On the back couch	Released on the couch	On cushion and a blancket
Sitting upright	On the ground	On the back couch	Released on the legs	Sitting on a cushion
Sitting upright	Crossed on the couch	On the back couch	Released on the legs	Cushion behind the back
Sitting semi upright	On the couch, knees to left	On the back couch	Released on the couch	Several cushions behind the back
Sitting semi upright	On the couch	On the back couch	Released on the legs	Cushion behind the back and blancket
Sitting semi upright	On the ground	On the back couch	Released on the couch	On cushion and a blancket
Semi lying on the couch	Crossed under body	On the back couch	Crossed on the belly	Multiples cushions and blanckets
Sitting semi upright	Crossed on ground	On the back couch	released on the belly	Cushion behind the back
Sitting semi upright	Crossed on ground	On the back couch	Released on the couch	Cushion behind the back
Sitting semi upright	On the ground	On the back couch	Released on the legs	none
Semi lying on the couch	On the ground	On the back couch	One on behind the head	Cushion behind the back
Sitting semi upright	One one the ground	on the corner or the couch	Released on the legs	Cushions
Sitting semi upright	On the ground	on the corner or the couch	One of the sofa corner	none
Sitting semi upright	Crossed on ground	On the back couch	One of the sofa corner	Cushion behing the back
Sitting upright	On the ground	Not touching the couch	on the legs	none
Sitting upright	On the ground	On the back couch	on the legs	Several cushions behind the back

Table 2. Overview of the participants viewing position.

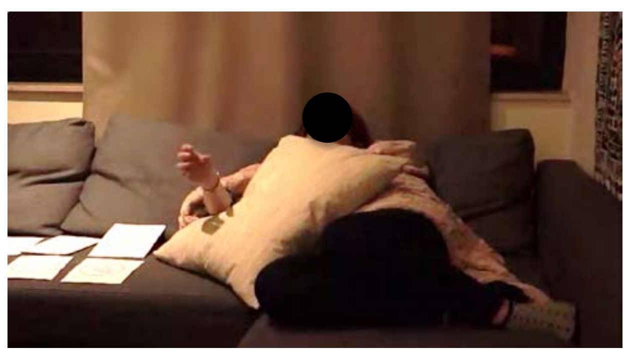

Figure 5. Upper body semi lying on the couch, back crossed under body, arm on the belly with multiple cushions and blankets.

[1] www.testmycreativity.com

Figure 6. Sitting upright position, feet on the ground, back not touching the couch, arm on the legs with no items.

Figure 7. Sitting semi-lying on the couch position, feet on the ground, back on the couch, one arm behind the head, cushions behind the back.

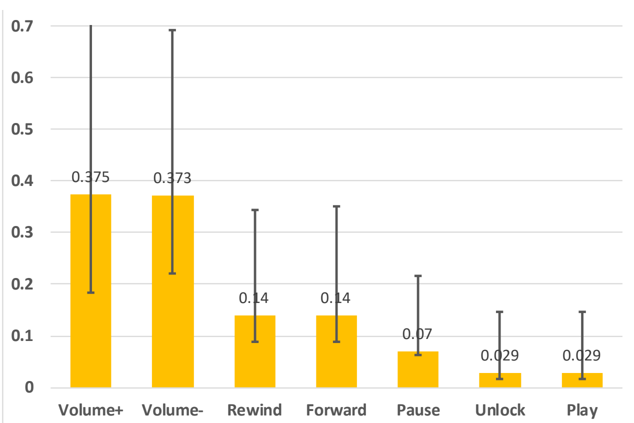

Figure 8. Agreement Rates for each referent calculated in phase 1 by Agate. Error bars show 95% confidence intervals computed with the AGATe tool [25].

Phase 1

We calculated the agreement rates using the AGATe (AGreement Analysis Toolkit) tool [25,26]. The resulting gesture set – taking the gesture with the highest agreement rate for all participants – is presented in Figure 8. There are three referents for which there are very low agreement scores; for these referents, we have provided two or three elicited gestures in the Table 3. Except for the volume referents, we do not have a consensus on gestures proposals for other referents. These results could be explained by our observations. The participants were very creative during the elicitation test. Overall, we listed 57 gestures for the 7 referents by the 19 people. Most of the elicited gestures are performed with the hand and the arm, only 4 out of 19 participants performed micro-gestures (with only fingers). Even if most of the gestures are spontaneous (less than 5 minutes for the conception) and natural, the variety of gestures is too wide and we cannot predict any agreement. Depending on their viewing position, the majority of the proposed gestures has been validated as easy to execute. The participants adapted their gestures to their position and had no other constraint to execute them.

Referent	Description of the elicited gesture
Volume up	Raise the arm vertically with the palm up
Volume down	Put the arm down vertically with the palm down
Rewind	1) Move the arm and the hand to the left 2) Wave left
Forward	1) Move the arm and the hand to the right 2) Wave right
Pause	1) Extend your arm and point your hand face to the TV 2) Get your hands down by tapping 3) Make a fist
Unlock	Snap
Play	Extend your arm and point your hand face to the TV

Table 3. Overview of the gestures with the highest agreement rate for each referent. Two or three gestures are given when agreement rates were very low for the referent.

Phase 2

The vast majority kept all the gestures elicited for this test, 7 people changed their gestures to control music on Spotify. 18 gestures changed on all proposed gestures (133). According to the results provided by the elicitation of the gestures, we can confirm our hypothesis that people keep the same gestures set for the same referents for the navigation in both the video or the music.

Phase 3

When using the Myo system, the participants followed all the instructions and progressively (practiced) all the tasks we gave them. Calibration requires in general 5 minutes to define the profile. Thanks to the gesture memory cards, the participants used the predefined gestures easily. In each interview we followed the whole experiment of the learning process. We completed the whole experiment with all the participants even if some signal recognition troubles appeared. In some cases, the gestures recognitions were not as precise as the natural users' interactions. Furthermore, we discovered a particular inversion in some gestures. For the Rewind and Forward referents as well as the Volume Up and Volume Down, it occurred that the actions were in the opposite direction to the one expected, even if the Myo gesture recognition interfaces (Figure 4) were correct. In other words, in some cases the correct Rewind referent (by waving the hand to the right) and the correlated gesture recognition interface activated the Forward action. That inversion effect for the Rewind and Forward referents occurred for 14 participants, and for the Volume Up and Down it occurred 19 times (for all the participants).

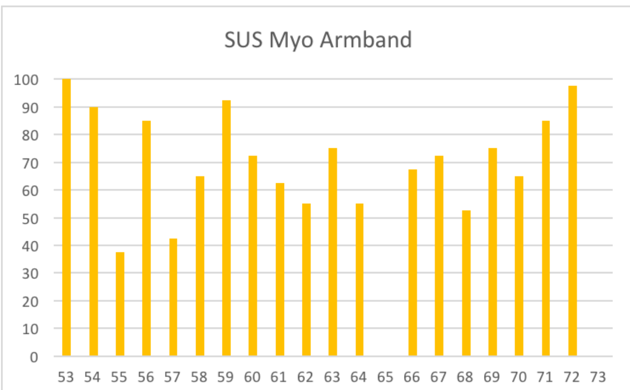

Figure 9. Bar charts of participants (p1 to p21) SUS overall score.

We calculated the SUS questionnaire data for all participants (19) of our study. The total score of each participant is illustrated in the Figure 9. The maximum score is 100 when the minimum score achieved is 37.5. The average score is 70.92. To analyse the SUS score, we considered score below 68 as not acceptable [1]. Figure 10 demonstrates that 9 out of 19 participants had a not acceptable score. Half of the users were not convinced by the Myo system in the proposed context. On the other hand, the other half of the participants sometimes very much appreciated the armband as presented.

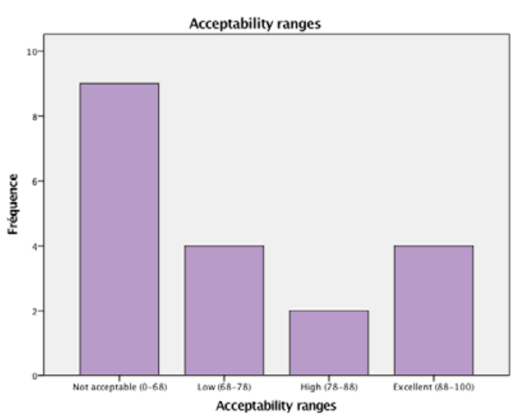

Figure 10. Bar charts of SUS analysis with adjective rating scales.

SUS results are to be associated with the series of subjective measures calculated on each gesture. In addition, the qualitative data and comments given during the experiment are also included in the analysis of the different criteria.

Referent: Gesture	**Good match**		**Ease of execution**		**Memorability**		**Enjoyable to execute**	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Unlock : Double tap	5,47	1,83	5,39	1,81	6,16	1,42	5,39	1,88
Play : Spread fingers	5,68	1,49	6,11	1,34	5,84	1,38	5,66	1,55
Pause : Spread fingers	5,47	1,38	6,21	1,02	6	1,45	6,079	1,06
Volume up : Make fist, turn left	3,42	2,48	5,05	1,91	4,11	2,26	5,05	1,51
Volume down : Make fist, turn right	5,47	1,39	6,21	1,02	6	1,45	6,08	1,05
Rewind : Wave right	3,74	2,33	4,92	2,16	5,26	1,93	4,08	1,76
Forward : Wave left	3,84	2,19	4,68	1,85	5,32	1,97	4,08	1,66

Table 4. Result of the subjective measures questionnaire.

Good match

Gestures proposals seem to have a good match for some referents like Unlock, Pause and Play. Figure shows that referents like volume down-up and rewind-forward in a video sequence are to be discussed. Indeed, the vast majority of people were disturbed by the direction that worked. "Oh but it's upside down" was a common expression recorded for 9 participants during the experiment.

Going back in the sequence linked to the wave left gesture was not interpreted in a natural way, as much as the gesture given to increase the sound by turning the fist to the left, which disrupted many participants. This comes from confusion about which gestures does what explained previously. Those gestures that actually worked during the experiment were in the opposite direction of what Myo had planned in the setup.

Ease of execution

The participants measured their effort in the performance of predefined gestures and repeatedly highlighted the difficulty of wrist twist to the left or to the right. This average is lower than the other gestures (Table 4). Moreover, the participants repeatedly tried not to go too far in waving their wrist. Therefore it became more difficult to detect the executed gesture and made the users even more frustrated. About the execution of other gestures, they were rather positive. No difficulties were significant in the performance even if the armband sometimes did not have a precise answer or when the system misinterpreted the executed gesture. The participants just tried again until the system figured it out.

Memorability

When asked whether or not the gestures used in Netflix were memorable, participants did not really need extra help to be able to recall them. However, according to the Table 4, we can see that to make a fist and turn left (to volume up) is less memorable for some participants. This can be explained because this gesture requires two particular movements with the hand to perform a single action.

Enjoyable to execute

All the gestures suggested by Thalmic Labs entertained most of the users. A similar result was found with the ease of execution. It was the rewind and forward navigation gestures that did not entertain the participants. Stress for the execution of these gestures was constantly present.

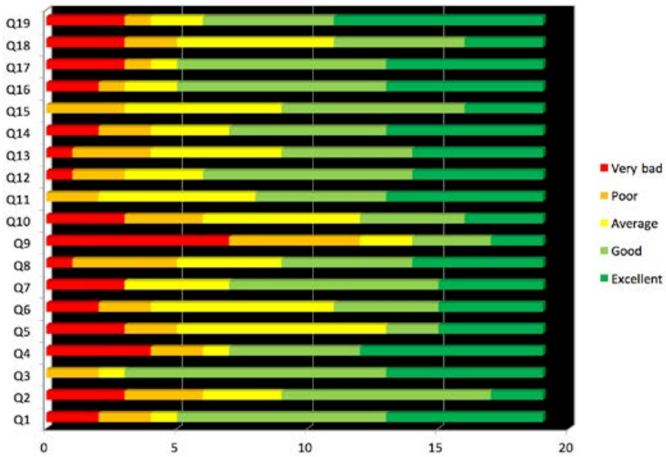

Figure 11. Distribution of participants' answers to the IBM CSUQ questionnaire.

Figure 11 depicts the distribution of the answers provided by the participants on the 19 IBM CSUQ questions. This test is interpreted like [4] as follows: a score between 6 and 7, represented with dark green, is considered as excellent; a score of 5, represented with light green, is considered as good; a score of 4, represented with yellow, is considered as average; a score of 3, represented in orange, is considered as poor; and a score between 1 and 2, represented in red, is considered as very bad. Figure 11 suggests that the global subjective satisfaction of the participants involved in the experiment follows a rather positive trend since Q19 is interpreted positively by 15 users out of 19 (Q19, $\mu = 5.68$, M = 6, s = 1.60). Question 9, the most negatively assessed question, raises a particular concern about the information quality to solve the users problems (Q9, $\mu = 4.11$, M = 4, s = 1.76): during the experiment, the visual information guidance to solve problems recognition and to readjust the armband are not clear and helpful for the users. However, regarding the Q10 to Q15 the information quality system is evaluated positively. About the interaction quality (Q16-Q18, $\mu = 5.49$, M = 6, s = 1.45): all questions have some negative answers, but indicate that there is a positive agreement among the respondents. Finally, "Q1. Overall, I am satisfied with how easy it is to use this system" ($\mu = 5.68$, M = 6, s = 1.45) proposes that participants are satisfied with the process supported by the Myo system. Moreover, the most positively assessed question (Q9, $\mu = 6.05$, M = 6, s=0.91) demonstrates that participants think that the system is very simple to use.

Phase 4

This last task allowed us to calculate specific agreement rates in which the participants could have included all the gestures learned with the Myo and those created in the elicitation process. Figure 12 demonstrated higher agreement rates for each referent than the first one we had calculated. As shown in Table 5, the gestures the Myo proposed were the most elicited for the Unlock, Play, Pause, Forward and Rewind referents. According to the results, we discovered that the gestures the Myo proposed can be an effective gestures set to satisfy most people using an interactive armband.

Figure 12. An overview of the Agreement Rates for each referent calculated in phase 4, computed by Agate [].

We saw that people have insisted on restoring the direction for the gestures: Rewind, Forward, Volume Up and Volume Down. They proposed a simplified gesture for the volume control in order to avoid making mistakes and created a new version of Wave gesture (left or right) to have less effort and more ease to execute the command. In our sample, 3 users out of 19 kept their own gestures that they had defined at the beginning, without taking any gesture of the Myo system. They liked so much their own gestures that they could not choose any other. The other users have integrated at least one gesture or several from the gestures the Myo proposed. To reduce effort or recognition problems, the participants often decided not to be as precise as the Myo system would recommend. They definitely chose gestures with greater ease of execution.

Referent	**Description of the elicited gesture**
Unlock	Double tap
Play	Spread fingers
Pause	Spread fingers
Forward	3) Wave right 4) Move the arm and the hand to the right
Volume up	1) Raise arm vertically with the palm up 2) Make a fist and rotate right
Volume down	1) Put arm down vertically with the palm down 2) Make a fist and rotate left
Rewind	1) Wave left 2) Move the arm and the hand to the left

Table 5. Overview of the gestures with the highest agreement rate for each referent calculated in phase 4.

DISCUSSION

During the experiment, we shaped a new hypothesis about the age and the ease of execution. The youngest participants (under 22 years) have a deeper understanding of the proposed gestures and we suppose they adapted themselves to improve the gesture recognitions of the armband while the older users (+58) did not seem to adapt their own gestures. It requires future research to test this new proposal.

We thought about the technical trouble we had while using the Myo and we tried to figure it out by testing the system on other computers. Even after the calibration (and having specified with which arm the Myo was used), the main problem with the inversion effect occurs with the wrong recognition of the used arm. Even if the correct gestures were recognised by the system.

We confirmed the usage of the same gestures for music and video control. Future studies should find out whether it is still the case when different rooms are used and whether the participant interacts the same way with the presence of other people near them, in a multitask context or in a smart environment.

CONCLUSION

In this paper, we presented the results of an elicitation study in an entertainment environment with an interactive armband. We also conducted a usability test with the device in the same context of use in order to align the results between the gesture elicitation study (which returns the most preferred gestures for the tasks envisioned) and the usability study (which returns the most usable gestures for the tasks performed). Again, this is the traditional debate of performance versus preference: participants may want to elicit some gestures that are not performant or not usable enough to be really incorporated in the final consensus set. Similarly, participants may return usable gestures (thus maximizing usability), but that are not also returning from the guessability study (thus minimizing the preference). Our findings revealed low agreement rates among our participants gesture proposals. Conversely, the participants went to a global higher agreement rate and satisfaction of use after learning and performing the predefined Myo gestures. The final consensus set resulting from this study expands the set of initial system-defined gestures. This new set could therefore become a candidate for an improved gesture recognition.

The next step of this gesture elicitation study consists in two actions: (1) feeding a new gesture recognizer with the consensus set resulting from this study. Kerber et al. [10] dit the same with their own recognizer; (2) extending the UsiXML UI description language [31] with the specification and mapping [32] of these gestures for supporting gesture user interface development and semi-automatic evaluation of guidelines based on them, similarly to [33,34] for web usability guidelines.

REFERENCES

1. Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4, 3, May 2009, 114–123. URL: http://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Bangor_May2009.pdf
2. Aaron Tabor, Scott Bateman, Erik Scheme, David R. Flatla, and Kathrin Gerling. 2017. Designing Game-Based Myoelectric Prosthesis Training. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '17). ACM Press, New York, NY, 1352–1363. DOI: https://doi.org/10.1145/3025453.3025676
3. Akilesh Rajavenkatanarayanan, Yoga Vignesh Surathi, Ashwin Ramesh Babu, and Michalis Papakostas. 2016. MyoDrive: A new way of interacting with mobile devices. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (PETRA '16). ACM Press, New York, NY, Article 18, DOI: https://doi.org/10.1145/2910674.2935863
4. Amna Khan and Zareena Kausar. 2016. Intention Realization of Intermittent Angular Positions of Elbow using Myo Armband Sensor. In *Proceedings of the 5th International Conference on Mechatronics and Control Engineering* (ICMCE '16). ACM Press, New York, NY, 108–112. DOI: https://doi.org/10.1145/3036932.3036950
5. Christine Kühnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Müller, and Sebastian Möller. 2011. I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies* 69, 11, 693–704. DOI: http://doi.org/10.1016/j.ijhcs.2011.04.005
6. Christopher Munroe, Yuanliang Meng, Holly Yanco, and Momotaz Begum. 2016. Augmented Reality Eyeglasses for Promoting Home-Based Rehabilitation for Children with Cerebral Palsy. In *Proceedings of ACM/IEEE International Conference on Human Robot Interaction* (HRI'16). IEEE Press, Piscataway, NJ, USA, 565–565. DOI: https://doi.org/10.1109/HRI.2016.7451858
7. David Dalmazzo and Rafael Ramirez. 2017. Air Violin: A Machine Learning Approach to Fingering Gesture Recognition. In *Proceedings of the 1st ACM International Workshop on Multimodal Interaction for Education* (MIE '2017). ACM Press, New York, NY, USA, 63–66. DOI: https://doi.org/10.1145/3139513.3139526
8. Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-hand Microgestures. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '16). ACM Press, New York, NY, 3403–3414. DOI: https://doi.org/10.1145/2858036.2858589
9. Finn Kensing and Jeanette Blomberg. 1998. Participatory

Design: Issues and Concerns. *Computer Supported Cooperative Work* 7, 3-4, January 1998, 167–185. DOI: http://dx.doi.org/10.1023/A:1008689307411

10. Frederic Kerber, Michael Puhl, and Antonio Krüger. 2017. User-independent real-time hand gesture recognition based on surface electroMyography. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '17). ACM Press, New York, 2017, Article 36. DOI: https://doi.org/10.1145/3098279.3098553

11. George Lucas. *Star Wars Original Trilogy*. Lucasfilm, 20th Century Fox. 1977-1983.

12. Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. 2015. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access* 3: 543–555. DOI: http://doi.org/10.1109/ACCESS.2015.2432679

13. Heli Koskimäki, Pekka Siirtola, and Juha Röning. 2017. MyoGym: Introducing an Open Gym Data Set for Activity Recognition Collected Using Myo Armband. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Symposium on Wearable Computers* (UbiComp '17). ACM Press, New York, NY, USA, 537–546. DOI: https://doi.org/10.1145/3123024.3124400

14. Huiyue Wu, Jianmin Wang, and Xiaolong (Luke) Zhang. 2016. User-centered gesture development in TV viewing environment. *Multimedia Tools and Applications*, Volume 75, Issue 2, pp 733–760. DOI: http://doi.org/10.1007/s11042-014-2323-5

15. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the Guessability of Symbolic Input. In *Proceedings of ACM Conference on Human Factors in Computing Systems, Extended Abstract* (CHI EA '05). ACM Press, New York, 1869–1872. DOI: http://dx.doi.org/10.1145/1056808.1057043

16. James R. Lewis. 1995. IBM computer usability satisfaction questionnaires : Psychometric evaluation and instructions for use. International Journal of Human–Computer Interaction 7, 1 (1995), 57–78. DOI: http://dx.doi.org/10.1080/10447319509526110

17. James R. Lewis. 2006. Sample Sizes for Usability Tests : Mostly Math, Not Magic. Interactions 13, 6 (Nov. 2006), 29–33. DOI: http://dx.doi.org/10.1145/1167948.1167973

18. Jarg Bergold and Stefan Thomas. 2012. Participatory Research Methods: A Methodological Approach in Motion. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 13, 1, 2012. DOI: http://dx.doi.org/10.17169/fqs-13.1.1801

19. John Brooke. SUS - A quick and dirty usability scale. In *Usability evaluation in industry* 189, 194, 1996. 4-7. URL: https://hell.meiert.org/core/pdf/sus.pdf

20. L. Lafevers. Why fantasy matters. 9th July, 2010. URL: https://www.wired.com/2010/09/why-fantasy-matters/

21. Lucas S. Figueiredo, Mariana G.M. Gonçalves Maciel Pinheiro, Edvar X.C. Vilar Neto, and Veronica Teichrieb. 2015. An Open Catalog of Hand Gestures from Sci-Fi Movies. In *Proceedings of the ACM Conference on Human Factors in Computing Systems, Extended Abstracts* (CHI EA '15). ACM Press, New York, NY, 1319–1324. DOI: https://doi.org/10.1145/2702613.2732888

22. M. Korkman, U. Kirk, and S. Kemp. 1998. NEPSY: A Developmental Neuropsychological Assessment. Psychological Corporation. San Antonio, Texas.

23. M. Montoya, O. Henao, and J. Muñoz. 2017. Muscle fatigue detection through wearable sensors: a comparative study using the Myo Armband. In *Proceedings of the XVIII International Conference on Human Computer Interaction* (Interacción '17). ACM, New York, NY, USA, Article 30. DOI: https://doi.org/10.1145/3123818.3123855

24. Mithileysh Sathiyanarayanan and Sharanya Rajan. 2016. Myo Armband for physiotherapy healthcare: A case study using gesture recognition application. In *Proceedings of the 8th International Conference on Communication Systems and Networks* (COMSNETS'2016). IEEE, Los Alamitos, 2016. DOI: https://doi.org/10.1109/COMSNETS.2016.7439933

25. Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM Press, New York, NY, USA, 1325–1334. DOI: http://dx.doi.org/10.1145/2702123.2702223

26. Radu-Daniel Vatavu. 2013. A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. *Journal of Ambient Intelligence and Smart Environments* 5, 2: 187–211. DOI: http://doi.org/10.3233/AIS-130200

27. Rensis Likert. 1932. A technique for the measurement of attitudes. Archives of Psychology 22, 140 (1932), 55–. http://psycnet.apa.org/record/1933-01885-001

28. Thalmic Labs. 2018. Myo Gesture Control Armband | Wearable Technology. Thalmic Labs. Waterloo, Canada. URL: https://www.Myo.com/

29. Tobias Mulling and Mithileysh Sathiyanarayanan. 2015. Characteristics of hand gesture navigation: a case study using a wearable device (Myo). In *Proceedings of the 2015 British HCI Conference* (British HCI '15). ACM Press, New York, NY, 283–284. DOI: http://dx.doi.org/10.1145/2783446.2783612

30. Wan-Lun Tsai, You-Lun Hsu, Chi-Po Lin, Chen-Yu Zhu, Yu-Cheng Chen, and Min-Chun Hu. 2016. Immersive

Virtual Reality with Multimodal Interaction and Streaming Technology. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (ICMI' 2016). ACM Press, New York, NY, USA, 416–416. https://doi.org/10.1145/2993148.2998526

31. Quentin Limbourg, Jean Vanderdonckt, Benjamin Michotte, Laurent Bouillon, Murielle Florins, USIXML: A User Interface Description Language Supporting Multiple Levels of Independence. Proc. of ICWE'2004 Workshops, 2004, pp. 325-338.

32. Francisco Montero, Víctor López-Jaquero, Jean Vanderdonckt Pascual González, María Lozano, Quentin Limbourb, Solving the Mapping Problem in User Interface Design by Seamless Integration in IdealXML. In: Gilroy S.W., Harrison M.D. (eds) Interactive Systems. Design, Specification, and Verification. DSV-IS 2005. Lecture Notes in Computer Science, vol 3941. Springer, Berlin, Heidelberg, DOI: https://doi.org/10.1007/11752707_14

33. Jean Vanderdonckt, Abdo Beirekdar, Automated Web Evaluation by Guideline Review. Journal of Web Engineering 4(2): 102-117 (2005)

34. Jean Vanderdonckt, Abdo Beirekdar, Monique Noirhomme-Fraiture (2004). Automated Evaluation of Web Usability and Accessibility by Guideline Review. In: Koch N., Fraternali P., Wirsing M. (eds) Web Engineering. ICWE 2004. Lecture Notes in Computer Science, vol 3140. Springer, Berlin, pp. 17-30. DOI: https://doi.org/10.1007/978-3-540-27834-4_4