Advances in Intelligent Systems and Computing 795

Tareq Z. Ahram *Editor*

# Advances in Human Factors in Wearable Technologies and Game Design

Proceedings of the AHFE 2018 International Conferences on Human Factors and Wearable Technologies, and Human Factors in Game Design and Virtual Environments, Held on July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA

Springer

# Advances in Intelligent Systems and Computing

Volume 795

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.



More information about this series at http://www.springer.com/series/11156

Tareq Z. Ahram
Editor

# Advances in Human Factors in Wearable Technologies and Game Design

Proceedings of the AHFE 2018 International Conferences on Human Factors and Wearable Technologies, and Human Factors in Game Design and Virtual Environments, Held on July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA

Springer

*Editor*
Tareq Z. Ahram
University of Central Florida
Orlando, FL, USA

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-94618-4 ISBN 978-3-319-94619-1 (eBook)
https://doi.org/10.1007/978-3-319-94619-1

Library of Congress Control Number: 2018947364



Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Advances in Human Factors and Ergonomics 2018

***AHFE 2018 Series Editors***

*Tareq Z. Ahram, Florida, USA*
*Waldemar Karwowski, Florida, USA*

***9th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences***

***Proceedings of the AHFE 2018 International Conferences on Human Factors and Wearable Technologies, and Human Factors in Game Design and Virtual Environments, held on July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA***

| | |
|---|---|
| *Advances in Affective and Pleasurable Design* | *Shuichi Fukuda* |
| *Advances in Neuroergonomics and Cognitive Engineering* | *Hasan Ayaz and Lukasz Mazur* |
| *Advances in Design for Inclusion* | *Giuseppe Di Bucchianico* |
| *Advances in Ergonomics in Design* | *Francisco Rebelo and Marcelo M. Soares* |
| *Advances in Human Error, Reliability, Resilience, and Performance* | *Ronald L. Boring* |
| *Advances in Human Factors and Ergonomics in Healthcare and Medical Devices* | *Nancy J. Lightner* |
| *Advances in Human Factors in Simulation and Modeling* | *Daniel N. Cassenti* |
| *Advances in Human Factors and Systems Interaction* | *Isabel L. Nunes* |
| *Advances in Human Factors in Cybersecurity* | *Tareq Z. Ahram and Denise Nicholson* |
| *Advances in Human Factors, Business Management and Society* | *Jussi Ilari Kantola, Salman Nazir and Tibor Barath* |
| *Advances in Human Factors in Robots and Unmanned Systems* | *Jessie Chen* |
| *Advances in Human Factors in Training, Education, and Learning Sciences* | *Salman Nazir, Anna-Maria Teperi and Aleksandra Polak-Sopińska* |
| *Advances in Human Aspects of Transportation* | *Neville Stanton* |

(continued)

(continued)

| | |
|---|---|
| *Advances in Artificial Intelligence, Software and Systems Engineering* | *Tareq Z. Ahram* |
| *Advances in Human Factors, Sustainable Urban Planning and Infrastructure* | *Jerzy Charytonowicz and Christianne Falcão* |
| *Advances in Physical Ergonomics & Human Factors* | *Ravindra S. Goonetilleke and Waldemar Karwowski* |
| *Advances in Interdisciplinary Practice in Industrial Design* | *WonJoon Chung and Cliff Sungsoo Shin* |
| *Advances in Safety Management and Human Factors* | *Pedro Miguel Ferreira Martins Arezes* |
| *Advances in Social and Occupational Ergonomics* | *Richard H. M. Goossens* |
| *Advances in Manufacturing, Production Management and Process Control* | *Waldemar Karwowski, Stefan Trzcielinski, Beata Mrugalska, Massimo Di Nicolantonio and Emilio Rossi* |
| *Advances in Usability, User Experience and Assistive Technology* | *Tareq Z. Ahram and Christianne Falcão* |
| *Advances in Human Factors in Wearable Technologies and Game Design* | *Tareq Z. Ahram* |
| *Advances in Human Factors in Communication of Design* | *Amic G. Ho* |

# Preface

Successful interaction with products, tools, and wearable technologies depends on usable designs and accommodating the needs of potential users without requiring costly training. In this context, this book is concerned with emerging technology of wearable devices with respect to concepts, theories, and applications of human factors knowledge focusing on the discovery, design, and understanding of human interaction and usability issues with products and systems for their improvement.

The game industry has been rapidly expanding in the past decades, and games became more appealing to a wider audience. The level of complexity in games' control interfaces and graphics has increased exponentially, in addition to the growing interest in integrating augmented reality in gaming experience. As a result, there is a growing demand for human factors and ergonomics practitioners to ensure the users' engagement in game design. The purpose of the AHFE International Conference on Human Factors in Game Design and Virtual Environments is to bring together researchers and practitioners from different fields who broadly share the study of game design applications.

This book focuses on the human aspects of Wearable Technologies and Game Design. It shows how user-centered practices can optimize wearable experience, thus improving user acceptance, satisfaction, and engagement toward novel wearable gadgets. It describes both research and best practices in the applications of human factors and ergonomics to sensors, wearable technologies, and game design innovations, as well as results obtained upon integration of the wearability principles identified by various researchers for aesthetics, affordance, comfort, contextual awareness, customization, ease of use, intuitiveness, privacy, reliability, responsiveness, satisfaction, subtlety, and user-friendliness. The book is organized into three sections that focus on the following subject matters:

*Section 1: Wearable Technologies and Sensors*
*Section 2: Game Design for Learning and Training*
*Section 3: Game Design Methodology, Usability, and Applications*

This book will be of special value to a large variety of professionals, researchers, and students in the broad field of game design, human modeling, human–computer interaction, and human systems integration, who are interested in feedback of devices' interfaces (visual and haptic), user-centered design, and design for special populations, particularly the elderly.

Each section contains research papers that have been reviewed by members of the International Editorial Board. Our sincere thanks and appreciation to the board members as listed below:

## Wearable Technologies



## Game Design

We hope this book is informative, but even more—that it is thought-provoking. We hope it inspires, leading the reader to contemplate other questions, applications, and potential solutions in creating good designs for all.

July 2018 Tareq Z. Ahram

# Contents

## Game Design Methodology, Usability and Applications

# Wearable Technologies and Sensors

# The Virtual Penetrating the Physical and the Implication for Augmented Reality Head-Up Displays

Matthias Walter(✉), Vanessa Seitz, and Klaus Bengler

Chair of Ergonomics, Technical University of Munich,
Boltzmannstr. 15, 85748 Garching b. München, Germany
{matthias.mw.walter,vanessa.seitz,bengler}@tum.de

**Abstract.** Augmented Reality is one of the upcoming topics in the development of human machine interfaces (HMI). The most promising technology using Augmented Reality to assist the user in his driving task is the Augmented Reality Head-Up Display (ARHUD). By reflecting virtual information at the windscreen into the driver's eyes, the impression of a hovering image is created. The virtual image can be overlaid on the road surface or penetrate into an object in which case an observer has to focus upon a distance further away than – or through – the object he is looking at. So far, the industry wide opinion on how to handle this has been either to change the HMI or to switch off the augmentation. Though it has not yet been investigated whether this influences spatial perception, usability or acceptance of the ARHUD [1]. In the current study we investigated whether penetration of the virtual image with a fixed image distance of 10 m into a leading vehicle was perceived as disturbing or influenced usability or cognitive workload. Navigation arrows were displayed in a fixed distance of 12 m using only monocular depth cues and superimposed by a vehicle in 6.6 m (low penetration) and 3.8 m (high penetration) distance. As a baseline, the leading vehicle was positioned at 13.8 m distance and thereby not superimposing the virtual image (no penetration). On both sides of the leading vehicle assistants presented visual cues which the subject was asked to count. Additionally subjects performed a visual Detection Response Task to evaluate the subjects' reaction times and cognitive workload [2]. Usability was evaluated using the System Usability Scale [3]. High penetration led to a significant change in acceptance when compared to no penetration. Usability, number of errors and reaction times were not significantly influenced. For low penetration no significant effects were recorded. The results suggest that adapting the virtual information displayed in an ARHUD in order to mediate penetrations between the virtual image and physical objects is not necessary. This has the potential to revolutionize the approach the automotive industry takes when implementing Augmented Reality in Head-Up Displays.

**Keywords:** Head-up Display · Augmented Reality · Contact Analog
Optical Penetration · Navigation


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 3–10, 2019.
https://doi.org/10.1007/978-3-319-94619-1_1

## 1 Introduction

The driving task has changed over time by an increase of assistance and information systems in the vehicle. While these systems are designed to enhance safety and comfort, the user is still tasked with supervising them [4]. One technical solution to control the complexity of information is the Head-up Display. Information can be displayed in the driver's primary field of view by projecting a virtual image over the bonnet of the car. This way the driver can monitor relevant information without head movement and with minimal focal accommodation or eye movement. Consequently distraction and reaction times to unexpected traffic events are reduced [5–8].

A possible next step in the development of Head-up Displays is to integrate Augmented Reality (AR). AR is the superimposition of the real world with virtual objects [9]. Similar to the conventional HUD, information in an ARHUD is displayed in the driver's primary field of vision, yet it is directly superimposed to or linked with real objects. As already shown AR efficiently guides the driver's attention and improves the detection of objects [10, 11]. Utilizing the spatial link between virtual information and physical world can help to reduce cognitive workload and reaction times in critical traffic events and increase the understanding of advanced driver assistance systems.

The ARHUD technology has the potential to improve how information is presented in the automotive environment and to improve road safety, yet no market ready solution has been presented so far. Numerous technical challenges must be overcome in order to provide an acceptable user experience. Conventional Head-up Displays sold in today's production cars manage with the constricted room between dashboard and steering column. ARHUDs on the other hand provide an increased field of view and virtual image distance and therefore need to incorporate more and larger mirrors or lenses. Hence, packaging poses a more serious problem for ARHUDs, which occupy 4 to 10 times the volume. Secondly, sensors, e.g. GPS, and data, e.g. maps, in cars are not able to provide the accuracy augmented reality applications depend on. The third aspect to consider is the human machine interface (HMI). Since the virtual image distance in an ARHUD is at least 7 m, at times the virtual image will penetrate through an object in the real world. One example is the waiting period at a traffic light. While waiting in line drivers have to focus through the cars standing in front of them in order to take in the information presented in the ARHUD. Since this effect does not occur naturally, it is unknown whether this influences spatial perception, usability or acceptance of the ARHUD.

In the presented study we investigated whether penetration of the virtual image with a fixed image distance of 10 m into a leading vehicle was perceived as disturbing or influenced usability or cognitive workload.

## 2 Method

### 2.1 Participants

Thirty-four participants (nine women, 25 men, $M_{age} = 36$, $SD_{age} = 8.6$, range = 25 – 61 years) partook in the study. All participants were employees of AUDI AG and we recruited them via e-mail. All participants drive more than 5000 km per year, 91.18%

drive more than 10,000 km per year. Out of 34, 32 reported to have experience in driving with conventional HUDs, 16 participants reported to use HUDs multiple times a month. Thirty-five percent had experience in AR applications. All participants had normal or corrected-to-normal vision at the time. They participated voluntarily, gave written consent and did not receive compensation.

### 2.2 Experimental Setup

The experiment setup comprised of a car placed in front of a junction (see Fig. 1). In 12 m distance in front of the car, a navigational prompt telling the subjects to turn left was presented. The car was equipped with an ARHUD, the visible resolution was 800 $\times$ 480 pixel and the distance to the virtual image was 10 m. The subjects witnessed the scene from the driver's seat. During the experiment the car did not move.

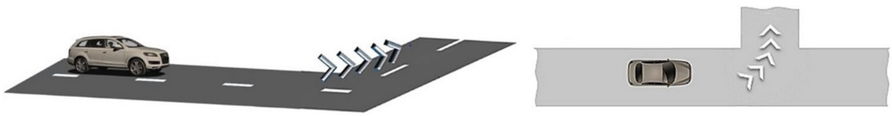


**Fig. 1.** Experimental setup with ARHUD vehicle (grey) and the displayed HMI; the arrows shown were standing upright as shown on the left, the depiction on the right is for illustration purpose only.

A second vehicle (white) was placed in front of the ego-car. We moved the white car to three different positions during the experiment presenting the subject with three stages of penetration (see Fig. 2):

No penetration, distance to vehicle 13.8 m
Minor penetration, distance to vehicle 6.6 m
Maximum penetration, distance to vehicle 3.8 m.

The HMI contained speed, distance to junction, navigational instructions to turn left and a symbol indicating the state of an adaptive cruise control system (ACC) (see Fig. 3). We used a warning symbol as a visual trigger for a detection response task (DRT). In order to provide a more realistic scenario by directing the driver's attention away from the ARHUD HMI and out of the car, two assistants presented black and green cards on both sides of the leading vehicle (see Fig. 4).

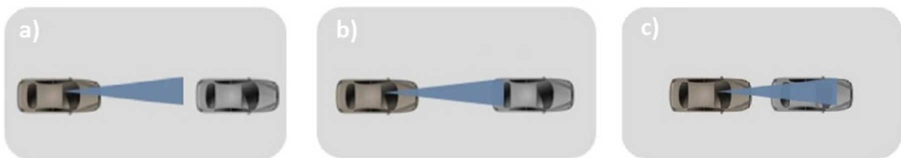



**Fig. 2.** Experimental setup. Three different distances between ego-vehicle and leading vehicle were presented. (a): no penetration, 13.8 m distance; (b): minor penetration, 6.6 m distance; (c): maximum penetration, distance: 3.8 m.

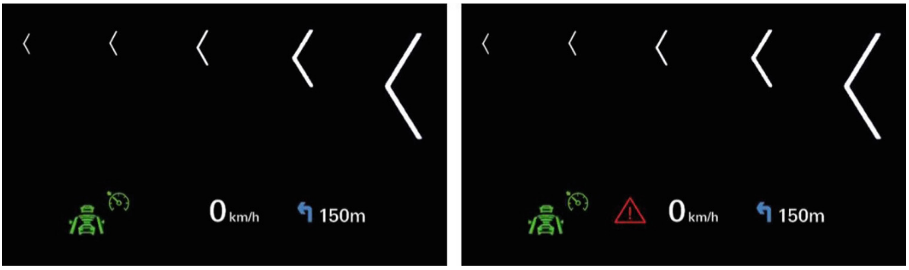



**Fig. 3.** Experimental HMI. On the right, the visual stimulus for the DRT is active.

**Fig. 4.** Main task diverting the subject's attention away from the ARHUD HMI.

## 2.3 Procedure

The participants executed two tasks during each variation. We instructed them to count and remember the cards shown by the assistants. The cards were shown in the same order for all subjects but in a different order for each variation. With this we led the subject's attention out of the car and to the presented scene thereby creating more realistic experimental conditions. Additionally the subjects performed a DRT. For a visual stimulus we used a red triangle around a red exclamation mark (see Fig. 3). The symbol was randomly presented every two to five seconds, the subjects pressed any key on a keyboard in response. The reaction times were recorded.

All participants witnessed all three variations of penetration in randomized order. Each variation was shown for 60 s. After having seen each variation the subjects filled in a questionnaire, i.e. three times in total.

The questionnaire contained the system usability scale (SUS) [3] and a section designed to measure acceptance. This second part subjects rated precision ("The symbols were placed at the right position in the environment"), ("The positioning in the surroundings was annoying"), intuitiveness ("The symbols were confusing"), ("The HMI was intuitively interpretable"), annoyance ("I perceived the HMI as distracting"), ("The HMI was helpful") on a 5-point Likert scale with the endpoints *fully disagree* and *fully agree*. These six questions were added to an ARHUD score from 0 to 24 pts. Additionally the questions were analyzed individually.

# 3 Results

Repeated measures ANOVAs were calculated to analyze all items separately. Greenhouse-Geisser corrections were applied in case of violations of sphericity. Post-hoc t-tests with Bonferroni-correction were calculated for all significant main effects of the factor variation.

## 3.1 System Usability Scale

Penetrations into the leading vehicle did not significantly influence the usability of the system (see Fig. 5). However a trend between the scores for no penetration (SUS score = 86.25) and the stages minor (SUS score = 83.82) and maximum penetration (SUS score = 83.46) can be identified (Fig. 6).

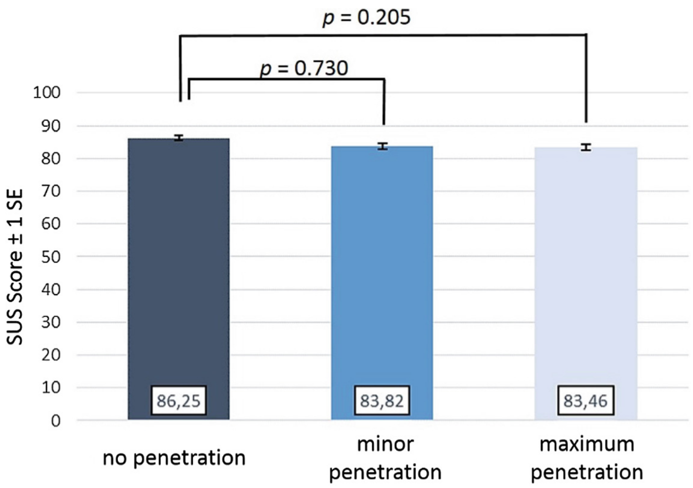



**Fig. 5.** SUS Score ± 1 SE. No statistically significant influences were found.

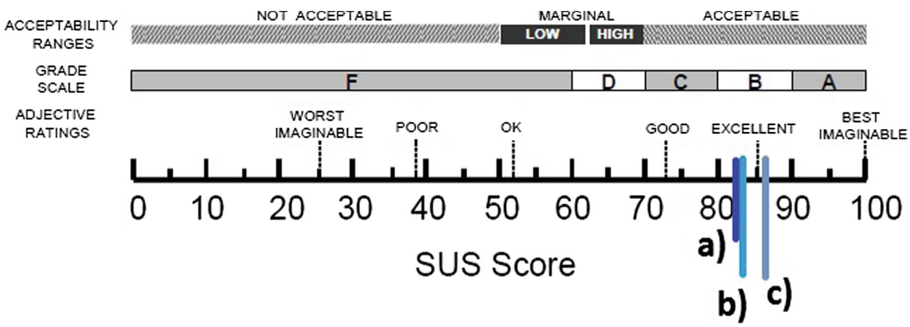



**Fig. 6.** SUS Score on the scale according to Brooke [3]. (a): no penetration; (b): minor penetration; (c): maximum penetration.

### 3.2 ARHUD Score

The six questions of the ARHUD questionnaire were added to result in a score from 0 to 24 pts (see Sect. 2.3). A trend is visible (see Fig. 7) and the effects did prove to be statistically significant when comparing “no penetration” to “maximum penetration” $F(2, 66) = 4.84$, $p = .018$. When analyzing the questions individually however, we found a significant effect of the degree of penetration on the spatial perception of the navigational cue. The effect occurred in both questions aimed at spatial perception between the stages “no penetration” and “maximum penetration” Q1: $F(2, 66) = 8.084$, $p < .001$ and Q2: $F(2, 66) = 4.466$, $p < .015$ as well as between “no penetration” and “minor penetration” within the question “The symbols were placed at the right position in the environment” (see Fig. 8).

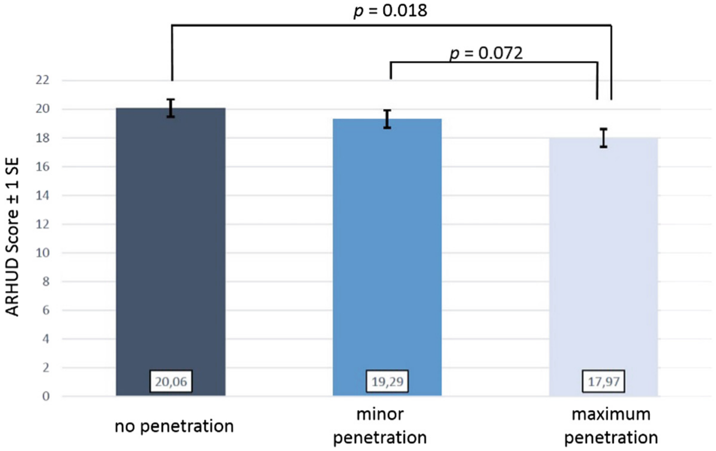



**Fig. 7.** ARHUD score on a scale from 0 to 24.

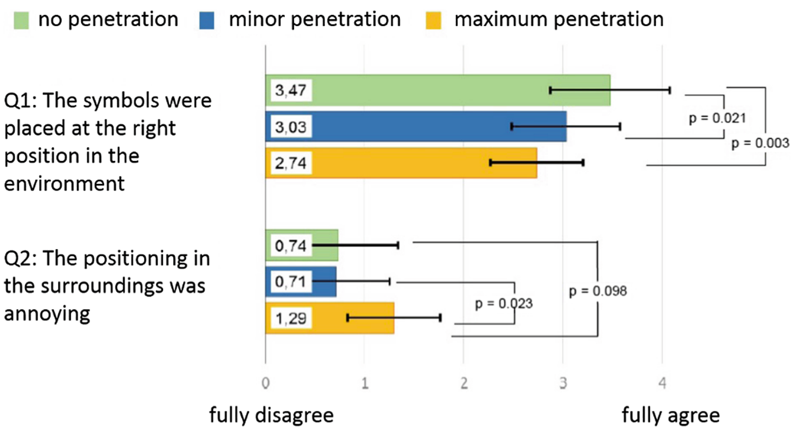



**Fig. 8.** Influence of the degree of penetration onto spatial perception of navigational cues.

### 3.3 Detection Response Task

The DRT was performed with red triangle around a red exclamation mark as a visual stimulus (see Fig. 3). The reaction times showed no significant influence by the grade of penetration (see Fig. 9).

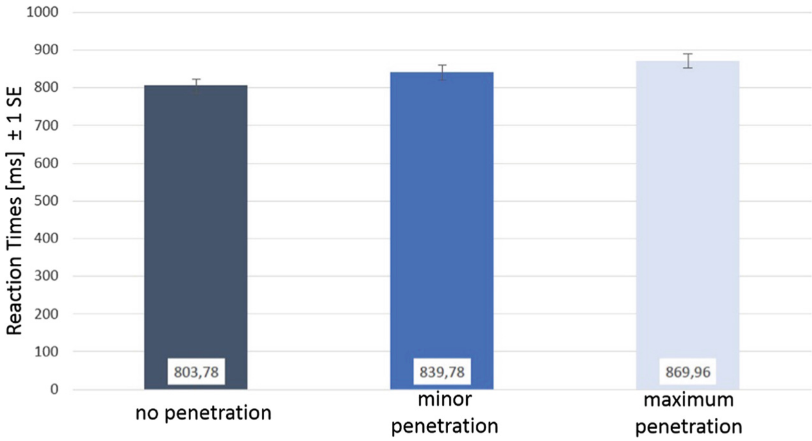



**Fig. 9.** Reaction times [ms] depending on the degree of penetration.

# 4 Discussion

In the current study, different stages of penetration of a virtual image, presented in an automotive ARHUD, into a physical object were compared in the aspects of usability, acceptance and reaction times of a DRT. Participants were not actively driving and therefore we were able to reproduce the distances between the cars with a tolerance <.1 m (see Fig. 2).

Pfannmüller et al. [1] reported an effect of obscured navigational cues on spatial perception in an automotive ARHUD. As an extension of this effect, we expected an increase of the penetration to significantly affect usability, acceptance and reaction times. No effects were found regarding the usability or reaction times. However, we found a significant effect of the degree of optical penetration on the acceptance of the automotive ARHUD system. Furthermore, in context of spatial perception a significant influence was revealed. This is in line with the findings of Pfannmüller et al. [1].

Since the experiment was conducted without the subjects driving, a misinterpretation of the distance to the AR content did not lead to a navigational error. Although the penetrations displayed did not affect the before mentioned criteria it still can effect navigational performance while driving. This must be verified in a future study, which takes into account objective measures of driving performance and distraction.

For this experiment we used a minimalistic HMI design but a comparison of different ARHUD HMI concepts for ACC has revealed surprisingly little differences between a minimalistic and a very salient, stimulating concept [12]. It is possible that a repetition of the experiment with a more salient display concept can reveal an effect of the degree of penetration onto the before mentioned variables. This needs to be addressed in future investigations.

## 5 Conclusion

In the presented study an optical effect which does not occur naturally was examined. Although an influence of the visual penetration of AR content into a leading vehicle on the acceptance was revealed, no effect regarding usability or reaction times were found. Therefore, the results suggest that adapting the virtual information displayed in an ARHUD in order to mediate penetrations between the virtual image and physical objects is not necessary i.e. an automotive ARHUD HMI does not need to change when closing on an object.

However, the current study was conducted without the participants actually driving the car. Furthermore, the HMI used was of a minimalistic design. Future investigations should focus on a naturalistic driving study to validate the results.

## References

1. Pfannmüller, L., Walter, M., Senner, B., Bengler, K.: Depth perception of augmented reality information in an automotive contact analog head-up display. J. Vis. **15**(12), 1078 (2015). https://doi.org/10.1167/15.12.1078
2. Bubb, H., Bengler, K., Grünen, R.E., Vollrath, M.: Automobilergonomie. Springer Fachmedien, Wiesbaden (2015)
3. Brooke, J.: SUS-a quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
4. Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., Winner, H.: Three decades of driver assistance systems: review and future perspectives. IEEE Intell. Transp. Syst. Mag. **6**, 6–22 (2014)
5. Gish, K.W., Staplin, L.: Human Factors Aspects of Using Head Up Displays in Automobiles: A Review of the Literature. U.S. Department of Transportation - National Highway Traffic Safety Administration, Washington D.C. (1995)
6. Kiefer, R.J.: Defining the "HUD benefit time window". In: Vision in Vehicles - VI, pp. 133–142, Derby, England, North Holland, Amsterdam, New York (1998)
7. Kiefer, R.J.: Older drivers' pedestrian detection times surrounding head-up versus head-down speedometer glances. In: Vision in Vehicles - VII, pp. 111–118. Elsevier, Marseille, France, Amsterdam, New York (1999)
8. Horrey, W.J., Wickens, C.D., Alexander, A.L.: The effects of head-up display clutter and in-vehicle display separation on concurrent driving performance. Proc. Hum. Factor Ergon. Soc. Annu. Meet. **47**, 1880–1884 (2003)
9. Azuma, R.T.: A survey of augmented reality. Presence-Teleoper. Virtual **6**, 355–385 (1997)
10. Yeh, M., Wickens, C.D.: Display signaling in augmented reality: effects of cue reliability and image realism on attention allocation and trust calibration. Hum. Factors **43**, 355–365 (2001)
11. Rusch, M.L., Schall, M.C., Gavin, P., Lee, J.D., Dawson, J.D., Vecera, S., Rizzo, M.: Directing driver attention with augmented reality cues. Transp. Res. Part F Traffic Psychol. Behav. **16**, 127–137 (2013)
12. Israel, B., Seitz, M., Bubb, H., Senner, B.: Contact analog information in the head-up display – how much information supports the driver? In: Khalid, H., Hedge, A., Ahram, T. (eds.), Advances in Ergonomics Modeling and Usability Evaluation, pp. 163–171. CRC Press (2010)

# Determining Comfortable Pressure Ranges for Wearable EEG Headsets

Stijn Verwulgen[1(✉)], Daniel Lacko[1], Hoppenbrouwers Justine[1], Siemon Kustermans[1], Stine Moons[1], Falk Thys[1], Sander Zelck[1], Kristof Vaes[1], Toon Huysmans[2,3], Jochen Vleugels[1], and Steven Truijen[4]

[1] Department of Product Development, Faculty of Design Sciences, University of Antwerp, Antwerp, Belgium
stijn.verwulgen@uantwerpen.be

[2] Section on Applied Ergonomics and Design, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

[3] Imec - Vision Lab, Department of Physics, Faculty of Science, University of Antwerp, Antwerp, Belgium

[4] MOVANT, Faculty of Health and Health Sciences, University of Antwerp, Antwerp, Belgium

**Abstract.** Measuring and interpretation of brain wave signals through electroencephalography (EEG) is an emerging technology. The technique is traditionally applied in a clinical setting with EEG caps and conductive gels to ensure proper contact through a subject's hair, and anticipate inter-subject anthropometric variations. Development of dry electrodes offers the potential to develop wearable EEG headsets. Such devices could induce medical and commercial applications. In this paper, we evaluate a prototype EEG headset that actively places electrodes at standardized positions on the subject's head, where each electrode is applied with equal pressure. The system is designed for use with dry electrodes. Our research delivers a better understanding on the link between general level of comfort and possible useful clear data signals, that can be used in brain computer interfaces (BCI). The present study is confined to the impact of adjustable electrodes pressure on level of user comfort only. Levels of discomfort are assessed in twelve participants, wearing an EEG headset with controllable electrode pressure exerted at 14 locations. Of-the-shelf dry electrodes are used. In a first session, evenly distributed pressure is increased and afterwards decreased in fixed time intervals, going from 10 kPa to 30 kPa and vice versa with steps of 2 kPa. In a second session, a subject specific acceptable pressure level is retrieved from the data of the first session and constantly applied for 30 min. During this intervention, level of discomfort is assessed in a VAS-scale. Additional observation and surveys yields insights on user experience in wearing a pressure exerting EEG headset.

**Keywords:** Brain computer interface · Electroencephalography
Wearable EEG headset · Pressure · Dry electrodes · Discomfort


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 11–19, 2019.
https://doi.org/10.1007/978-3-319-94619-1_2

## 1 Introduction

A brain Computer Interface reads out the subject's brain signals and tune them to control actions performed by a system that is external to the subject's body [1]. Non-invasive BCI is a relative new technology, by which electrical fluctuations are detected onto the subject's head. Fluctuations are typically in a range between 2 and 100 Hz with normal peak-to-peak magnitude between 0.5 and 100 μV [2]. The first EEG assessment dates from 1924, through needles that punctured the upper skin layer, and read-out through a Galvanometer [2]. Since then, signal quality kept being improved using vacuum tubes and transistors. Nowadays, one can get clear signals in a relatively user-friendly and comfortable way. Brain computer interfaces have gained much interest for their potential medical and consumer applications, but current research is mainly confined to lab setting. Main applications in the medical field are monitoring and diagnostics, conducted with classical EEG caps with conductive gel [2, 3]. In order to unlock the full potential that EEG based BCIs offer, wearable EEG headsets should be developed [4]. Preliminary results indicate that a commercial headset can control a robotic arm in four directions, thereby bypassing the neuromuscular system [5]. Such advanced applications of wearable headsets are limited to a proof-of-concept.

An extensive research focus on end user and application is required to accelerate real-live applications [6]. Miniaturization of EEG electrodes, enhanced sensitivity and conductivity, active amplification, electric shielding, wireless data and miniaturization of electronics and improved signal processing and classifiers are all promising technologies that could facilitate such breakthrough. Connection with the scalp and electrodes, adaptability and standardization of electrode positions, comfort and acceptance, ease of use in mounting and un-mounting the headset, are main design drivers in the development of wearable EEG headsets [7]. Gel based electrodes, although they provide proper electrical contact even in the presence of hair, have a low acceptance for the end user to this respect [8]. Sensitive dry electrodes are increasingly offered as an alternative for wearable headsets.

A challenging usability and functional factor in the deployment of dry electrodes is the pressure of the electrodes exerted on the subject's skull [8]. On one hand, this pressure should not be too high, not to induce discomfort or annoyance by the wearer. At the other hand, a sufficiently amount of pressure is required for making stable contact with subjects' head [7]. Also, the presence of hair could be an important factor to take account of in the design of wearable EEG headsets [9] and exerting pressure could be a solution to ensure proper contact of electrodes protruding the hair layer [10, 11]. Increasing electrode pressure evidently increases the chance of electrodes making proper contact with the subject's skin, thus increasing conductivity and thus increasing signal quality. So, the electrode pressure should be not too low, to ensure proper signal quality.

Pressure requirements should be integrated in other requirements for wearable EEG headsets [7]. A particular challenge in the design of EEG head caps is placement of electrodes at pre-defined anatomical locations. These locations are geometrically inferred from four anatomical points: nasion, inion, left pre-auriculair point and right pre-auriculair point (respectively A1 and A2 in Fig. 2, left), along the so called 10–20 system [12]. They should be incorporated in the design of EEG headcaps to ensure accurate, standardized and repeatable EEG recording locations. In clinical applications, these positions are provided through the configuration of the textile cap, and contact is ensured by conductive gel. In the design of wearable headsets, the challenge remains to integrate standardized and accurate positioning with accurate pressure range, bound from above by usability and comfort issues and bound from below by functional requirements.

This pressure range is not yet systematically investigated in scientific literature. The problem is that no research instruments exist that automatically position electrode locations at pre-defined 10–20 locations, independent of the individual user's head size and geometry, at the same time ensures that the same pressure is exerted, uniformly at each particular electrode location, and moreover, that pressure can be controlled to assess subject's discomfort and signal quality.

## 2 Materials and Methods

Dedicated equipment was developed to acquire correct electrode positioning and controllable pressure, to simulate the behavior of dry electrodes in a wearable headset (patent pending).

This test setting was applied on healthy volunteers in a three phases. Firstly, comfort levels were assessed under subsequently increasing the pressure on the electrodes. Secondly, comfort levels were assessed under subsequently decreasing the pressure on the electrodes. Thirdly, key comfort levels were retrieved from these sequences and the time was assessed that comfort level persisted under corresponding pressure levels.

### 2.1 Test Equipment

The research and development of the equipment that allows conducting experiments has a long history, with many iterations and the generation of new scientific knowledge.

The first step was a high school design assignment where a group of four students was challenged to design a wearable EEG headset with dry electrodes: easy to mount and clean, comfortable for operator and subject and to be used in a clinical setting. They designed a modular headset, adaptable to multiple purposes in a low cost concept [13], see Fig. 1.

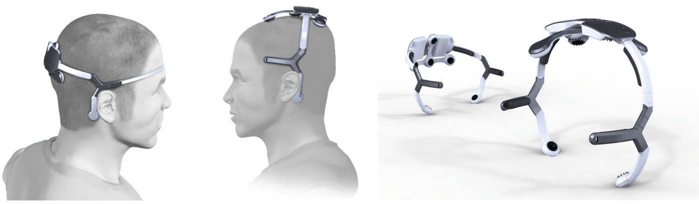


**Fig. 1.** Modular EEG headset: visually evoked event related potentials can be detected on the back of the subject's skull (left), and cognitive evoked potentials at the top (right). The concept allows for the braces to be configured, thereby using the same components for shells housing electrodes and electronics [13].

Geometry and sizing were based on the British adult population for male and female [14]. Breadth of the headset was adjusted to the breadth of the head at the maximum level above and behind the ears. The maximum circumference above the brow ridges was taken as guideline for head circumference. The bitragion coronal arc measured across the crown of the head (arcwidth) provided the curvature in the frontal plane. As sagittal arc, the length of the occipital-frontal curvature from the external occipital protuberance to the glabella in the sagittal plane was taken (arclength).

The modular concept and design was appreciated through a Red Dot Design award [15], but physical prototyping revealed sub-optimal fit and lack of functionality. Size and shape of the headset were based on available classical anthropometric data of the human head and naive underlying geometrical shapes representing the human head, but the design of wearable products require 3D anthropometric models and methods to link univariate measurements to non-trivial geometrical shapes that accurately represents the human body shape [16].

A shape model of the human skull was constructed from 100 medical images, CT and MRI scans [17], thereby omitting the presence of hair [9]. The model was shown saturated for adding new skull models. Anatomical landmarks were annotated manually, e.g. inion, nasion, glabella, on which anthropometrical measurements can be inferred geometrically, e.g. head length as distance between inion and glabella. Thus, the shape model of the head was enriched with univariate measurements that allows parameterization of the human head shape. Most influencing parameters on global head shapes were retrieved: head length, bitragion width, circumference and arcwidth. Such parameters allows linking product dimensions to head shape to ensure proper fit and function [16, 18].

The enriched shape model of the skull was used to design an EEG headsets at 14 electrode locations (Fig. 2, left) commonly used in of-the-shelve consumer headsets [19]. The electrode variations were mapped with the shape model (Fig. 2, middle) to achieve a one-size fits all prototype (Fig. 2, right). This results in 10–15% improvement in fit and accuracy compared to of-the-shelf available headsets [20].

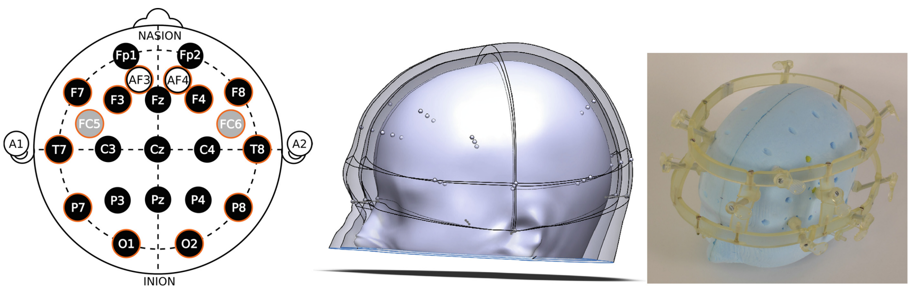



**Fig. 2.** Adjustable non functional EEG headset. Left: those electrodes at 10–20 locations that are assessed by the design (orange circles). Middle: variation of these locations along the main shape (first principal component), mean head shape and P5 and P95 displayed as ghost surfaces. Right: non functional prototype that anticipates these variations [4].

The test setting in this work is built upon the geometry of this prototype [20]. An additional technology is supplied to ensure automatic electrode positioning at these 14 standardized, pre-defined locations and in which pressure can be controlled in function of e.g. time, signal quality or discomfort. In this work, we present a first pilot study to assess discomfort depending on electrode pressure. The test setting was equipped with dry spikey electrodes from the OpenBCI platform (Florida Research Instruments).

### 2.2 Subject Assessments

A group of 12 healthy volunteers was enrolled in this study after informed consent. Study was approved by the ethical committee of UA/UZA (16/11/132). Before the experiments, subject completed a small survey on their age, gender, mood and experience with brain computer interfaces.

Pressure is required to ensure proper connection between the electrodes and the scalp. For thicker and more voluminous the hair, more pressure is expected to ensure a good connection between scalp and the electrodes. Therefore, subjects were balanced for amount of hair. A combination of hair thickness and hair volume was assessed to that end, by making a ponytail of maximal length and measuring the circumference of the string. Three categories were distinguished: low hair volume: 5 cm and below, medium hair volume: between 5 cm and 10 cm and high hair volume: 10 cm and up. Each category contained four subjects.

The level of discomfort was measured with a ten point Likert scale, ranging from 1: no discomfort, 5: discomfort, up to 10: very painful.

Tests were conducted in a space with smooth walls where the participant was sitting behind a desk in front of a laptop or handling a smartphone. In each assessment, the test device was initially placed on the subject's head without exerting any pressure. Subjects could control the pressure, ranging from 0 kPa up to 30 kPa. This corresponds to a weight of 600 g on each individual electrode.

### 2.3 Initiating Pilot

A first pilot was conducted to map bottlenecks and pinpoint a smooth study protocol. While increasing the pressure, a first participant could play on this laptop the puzzle game 'rush hour' online, to simulate a low involvement task. The test started at 0 kPa and every minute, the participant was asked to raise the pressure with 2 kPa. The participant was interested in the game until a pressure level rose above 24 kPa. After this value, she was more focused on the timer than on the game. The last minute, the electrodes were pushing on the test subject with 30 kPa. Then participant had no interest anymore in the game and was waiting for the test to end. In a second part, the participant was exposed to the maximum of 30 kPa and every minute the operator would lower the pressure with 2 kPa.

It was observed that the participant focused on the numbers related to the pressure on the computer and the time on the stopwatch. Being in charge of the test setting stressed the participant.

Both tests were repeated when an operator was in charge of the pressure levels and alerted the participant when altering the pressure. When decreasing the pressure, unpleasantness was instantly noticeable but more bearable than when the pressure was increased. Afterwards, the participant felt better than being in charge of pressure control. Notable, imprints from electrode spikes were more visual than in the first test.

Finally, to pinpoint an optimal duration, the participant was asked to wear the headset for as long as possible under a pressure of 15 kPa. This test was conducted to measure how long the maximum duration of tests could be. The participant could endure this pressure for 15 min. It was estimated that 12 min would be acceptable for participants to spread pressure between 10 kPa to 30 kPa.

## 2.4 Test Sessions

In the first session, the pressure was gradually increased in 12 min from 10 kPa to 30 kPa. In the second session, the pressure was decreased in the same time and over the same range. In both sessions, the test subject was inquired every minute on the experienced level of discomfort. At any moment, the subjects could press a panic button that elevated the pressure at once.

Between the two sessions, subjects were asked to fill in a short survey on overall satisfaction with the prototype BCI headset.

In the third session, a subject specific pressure was applied for 12 min. The pressure was optimized for each subject, by taking the mean values of the pressures of the first quartiles discomfort. Again, subjects were inquired every minute on the experienced level of discomfort. While the test proceeded, subjects played Tetris on a smartphone, to simulate a task with high involvement. This could correspond to a real world use of wearable EEG headsets.

Tests were conducted by four researchers: two observers, a moderator and a researcher responsible for welcoming the participants and handing out the surveys.

## 2.5 Research Questions

The setup was used to answer the following research questions.

- Is there a significant difference between the comfort levels of increasing pressure and the comfort levels of decreasing the pressure?
- Is there a significant difference between the comfort levels measured when the pressure is kept constant at subject specific optimized pressure and the comfort levels when the pressure is decreasing or increasing?
- What is the influence on hair volume on the level of discomfort at optimized pressure?
- Does the level of discomfort at optimized pressure remains constant?

Results were statistically analysed using SPSS IBM Statistics. Confidence intervals of 95% were used to define significant differences.

# 3 Results

## 3.1 Comfort Levels Under Increasing and Decreasing Pressure

The pressure range where the headset was not inducing discomfort was retrieved from Likert scale levels scoring discomfort between 1 and 5. For each participant, median discomfort score was calculated. A Wilcoxon Signed Ranks Test revealed that mean of median comfort levels in sessions with increasing pressure (A) was significantly lower than in sessions with decreasing pressure (B), respectively 3.5 and 4.9. Two participants had a median of B lower than median of A, eight participants had a median of B higher than median of A and two participants had the same median for A and B.

## 3.2 Comfort Levels Under Constant and Changing Pressure

Median discomfort level measured when the pressure is kept constant was calculated on the entire time range from 0 tot 12 min. A Wilcoxon Signed Rank Test showed that mean median values were significantly lower when the pressure is kept constant at subject-specific acceptable level than mean median discomfort when the pressure is increasing or decreasing, respectively 2.6, 3.5 and 4.9. Eight participants had lower medians under constant pressure than under increasing pressure, three participants had medians that scored the opposite and one participant had the same median. Two had medians at constant pressure that were higher than their medians under decreasing pressure.

Mean value of subject specific pressure was 12 kPa.

## 3.3 Influence of Hair Thickness

The hypothesis is that high hair volume results in lower level of discomfort. A Kruskal-Wallis Test was used on the mean level of discomfort on three different hair types. With $p = 0.4$, the null hypothesis was not rejected. So with pressure optimized at subject-specific levels, at first quartile levels of discomfort, no influence of hair could be detected. Further research with a dedicated study design and a sufficiently amount of participants is recommended to pinpoint the effect of hair.

## 3.4 Fluctuation of Discomfort at Optimized Pressure

An ordinal regression analysis on the acquired data was performed. The null hypothesis that the variation in comfort doesn't differ significantly over time when the same pressure is applied was maintained, with 18% of variance explained by the regression model and a 0.45 goodness of fit (Pearson). So variation in discomfort over time is not significant when constant optimized pressure is applied. However, other tests on longer the time range could yield different results.

## 4 Discussion and Conclusion

Study was limited to discomfort pain ranging from 1–10, with 1 being marked as 'comfortable', 5 'uncomfortable' and 10 as 'painful'. Most participants marked down '1' as their first reading. So they might see '1' as a baseline, to increase as the pressure went up. Others started at 5, indicating that even the first pressure level wasn't comfortable. Another limitation is that the order of test sessions was not randomized for increasing and decreasing pressure. Subjects were aware of the order in which the tests would be conducted. So there might be a co-founding factor between both test sessions.

Increasing pressure on the participants scalp is - measured by our pain scale – less discomfortable than decreasing pressure on the participants scalp. An evenly applied pressure of 12 kPa through the electrodes on the scalp falls within an acceptable comfort level for most subjects, during 12 min. This corresponds to a weight of 72 g on spike electrodes from openBCI (Florida Research Instruments), exerted at electrode locations used in commercial wearable headsets [19]. Pressure for minimized discomfort could be influenced by the hair thickness of the user but more research is required to pinpoint relation between pressure and signal quality in function of hair type. When the pressure stays within the acceptable range, the comfort level of the user won't change significantly over a time range of 12 min. To make statements about a longer time range, further research is recommended.

**Acknowledgments.** This work was financially supported by the University of Antwerp, Industrial Research Fund (Industrieel onderzoeksfonds-IOF).

## References

1. Wolpaw, J.R., Wolpaw, E.W.: Brain-computer interfaces: something new under the sun. Brain-Comput. Interfaces: Princ. Pract. 3–12 (2012)
2. Teplan, M.: Fundamentals of EEG measurement. Meas. Sci. Rev. **2**(2), 1–11 (2002)
3. Tautan, A.-M., Mihajlovic, V., Chen, Y.-H., Grundlehner, B., Penders, J., Serdijn, W.A. (eds.): Signal quality in dry electrode EEG and the relation to skin-electrode contact impedance magnitude. In: BIODEVICES (2014)
4. Lacko, D.: The application of 3D anthropometry for the development of headgear-a case study on the design of ergonomic brain-computer interfaces (2017)
5. Bousseta, R., El Ouakouak, I., Gharbi, M., Regragui, F.: EEG based brain computer interface for controlling a robot arm movement through thought. IRBM **39**, 129–135 (2018)
6. Mihajlović, V., Grundlehner, B., Vullers, R., Penders, J.: Wearable, wireless EEG solutions in daily life applications: what are we missing? IEEE J. Biomed. Health Inform. **19**(1), 6–21 (2015)
7. Hairston, W.D., Whitaker, K.W., Ries, A.J., Vettel, J.M., Bradford, J.C., Kerick, S.E., et al.: Usability of four commercially-oriented EEG systems. J. Neural Eng. **11**(4), 046018 (2014)
8. Nijboer, F., Van De Laar, B., Gerritsen, S., Nijholt, A., Poel, M.: Usability of three electroencephalogram headsets for brain–computer interfaces: a within subject comparison. Interact. Comput. **27**(5), 500–511 (2015)

9. Verwulgen, S., Vleugels, J., Lacko, D., Haring, E., De Bruyne, G., Huysmans, T. (eds.): Thickness of compressed hair layer: a pilot study in a manikin. In: Proceedings of the 7th International Conference on 3D Body Scanning Technologies, 3DBST 2016, Lugano, Switzerland, 30 November–1 December 2016 (2016)
10. Chi, Y.M., Wang, Y., Wang, Y.-T., Jung, T.-P., Kerth, T., Cao, Y. (eds.): A practical mobile dry EEG system for human computer interfaces. In: International Conference on Augmented Cognition. Springer (2013)
11. Estepp, J.R., Christensen, J.C., Monnin, J.W., Davis, I.M., Wilson, G.F. (eds.): Validation of a dry electrode system for EEG. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, Los Angeles (2009)
12. Jurcak, V., Tsuzuki, D., Dan, I.: 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage **34**(4), 1600–1611 (2007)
13. Faes, N., Vandemergel, M., Van Hove, F., Willems, D.: Solution. Internally available: Artesis Hogeschool Antwerpen 2009–2010
14. Pheasant, S., Haslegrave, C.M.: Bodyspace, Anthropometry, Ergonomics and the Design of Work. Taylor and Francis Group, London (2006)
15. GmbH RD: A Fireworks of Creativity - Singapore Celebrates red dot Winners Gelsenkirchener Straße 181, 45309 Essen, Germany (2010). https://en.red-dot.org/4253.html. Accessed Feb 2018
16. Verwulgen, S., Lacko, D., Vleugels, J., Vaes, K., Danckaers, F., De Bruyne, G., et al.: A new data structure and workflow for using 3D anthropometry in the design of wearable products. Int. J. Ind. Ergon. **64**, 108–117 (2018)
17. Lacko, D., Huysmans, T., Parizel, P.M., De Bruyne, G., Verwulgen, S., Van Hulle, M.M., et al.: Evaluation of an anthropometric shape model of the human scalp. Appl. Ergon. **48**, 70–85 (2015)
18. Lacko, D., Huysmans, T., Vleugels, J., De Bruyne, G., Van Hulle, M.M., Sijbers, J., et al.: Product sizing with 3D anthropometry and k-medoids clustering. Comput.-Aided Des. **91**, 60–74 (2017)
19. Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., Dutoit, T.: Performance of the Emotiv Epoc headset for P300-based applications. Biomed. Eng. Online **12**(1), 56 (2013)
20. Lacko, D., Vleugels, J., Fransen, E., Huysmans, T., De Bruyne, G., Van Hulle, M.M., et al.: Ergonomic design of an EEG headset using 3D anthropometry. Appl. Ergon. **58**, 128–136 (2017)

# Can Personal Activity Trackers Be Used to Provide Insight into Sit-to-Stand Workstation Usage and Benefits?

Trace Forkan(✉), Theresa Stack, and Daniel Autenreith

Safety, Health and Industrial Hygiene Department, MT Tech,
1300 West Park Street, Butte, MT 5970, USA
tforkan@mtech.edu

**Abstract.** We investigated whether activity trackers could be used to differentiate between sitting and standing at adjustable workstations and to determine if sit-to-stand workstation usage was associated with higher activity levels. A paired-t-test was used to assess the difference between the mean step counts measured using the activity trackers during sitting and standing periods among six office workers. Twelve office workers also wore activity trackers while self-reporting their standing as a percentage of total work time every week for six weeks. Spearman correlation was used to assess the relationship between standing percentage and step count. The difference in mean step count between sitting and standing was not statistically significant ($p = 0.113$) and the Spearman correlation between standing percentage and step count was weak ($\rho = 0.301$) and not statistically significant ($p = 0.342$). These findings suggest that basic activity trackers may not be useful in measuring sit-to-stand workstation usage.

**Keywords:** Wearable technology · Sit to stand workstations · Activity tracking

## 1 Background

Prolonged sitting at work is associated with health problems as is a sedentary level of physical activity. Sit-to-stand workstations have gained popularity as a means to decrease sitting and promote more physical activity. However, it is unclear whether sit-to-stand workstations are actually being used on a regular basis by those they are designed to help. There are also questions about whether the use of these devices is leading to a reduced risk of adverse health outcomes, which might be expected with an increase in physical activity.

Today's wearable activity trackers are no longer bulky plastic push button gadgets with a single function. Activity trackers are stylish and claim to monitor all your daily activities 24/7 to provide useful insights into your health. Activity trackers vary in design, but most are capable of tracking steps and calculating activity levels. Certain models have features that include stairs climbed, sleep monitoring, calorie consumption, and heartrate.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 20–28, 2019.
https://doi.org/10.1007/978-3-319-94619-1_3

According to the Consumer Technology Association, 33 million devices had been purchased in the US by the start of 2016 [3]. Across all emerging consumer tech categories, wearable activity trackers have the highest anticipated household purchase intent, at 11%, for 2016.

Activity trackers, largely, track your daily activities in terms of steps and calories burned through movement using GPS tracking and accelerometers. Accuracy depends on many factors to include the make and model of the activity tracker, where it is placed on the body, and the activity it is measuring. Most companies do not claim be highly accurate, instead they point to trending. The success of the products comes from empowering people to see their overall health and fitness trends over time—it is these trends that matter most in achieving their goals [1].

University of Nebraska researchers examined various fitness monitors (Jawbone, Fitbit, Nike Fuel Band, and Actigraph) and found they overestimated or underestimated energy expenditures by 10–15%, depending on the monitor [4]. A study published in the Journal of Physical Activity and Health compared the Fitbit to a lab-based method for estimation of energy expenditure for various activities during six-minute periods. The researchers found the Fitbit to either underestimate or overestimated energy expenditure by 1–25 calories. The smallest differences were seen in waking and treadmill running. The greatest underestimations were in cycling and stair stepping [4].

While 25 calories may not seem like a large amount, extend this out to an hour, and you are looking at 250-calorie underestimation for cycling or stair stepping. This obviously makes weight management and planning dietary intake for specific goals difficult and may deter individuals from continued efforts to improve their lifestyle. For example, when setting goals to move more by increasing step count.

How much of our day is spent seated? Staker [7] in an objective measure found that over 75% of the office workday is seated and much of that is accumulated in unbroken periods of 30 min or more [7]. The study linked physical inactivity to a premature mortality rate of 5.5% [2]. A 2008 Vanderbilt University study of 6,300 people published in the American Journal of Epidemiology estimated that the average American spends 55% of waking time (7.7 h per day) in sedentary behaviors such as sitting [6]. A solution is increased movement.

There is more muscular exertion in the lower extremities when standing vs sitting. In fact, the University of Minnesota Undergraduate Research Opportunities Program funded a study published in the *Journal of Physical Activity and Health* that analyzes the caloric expenditure when standing verses sitting. The results are quite convincing for advocating standing vs sitting. The University of Minnesota found that when assuming a standing posture, an additional 114 kcal are expended each day. This translates to approximately 20,461 kcal, or 5.85 lb, of additional caloric expenditure over a one-year period.

The *Take a Stand Study* published by in 2011 addresses the question regarding the benefits of short duration standing incorporated into sedentary work and found that with one hour of increased standing time: 87% of participants felt more energized; 71% felt more focused; and 50% felt reduction in pain [2].

These finds are supported in an unpublished study conducted by Stack in 2009 at a Navy worksite. Eighty workers that received sit to stand work stations between 2005 and 2009 were surveyed, 76 responded to the study, a 95% response rate. Seventy percent felt more productive, 40% reported a dramatic decrease in pain or discomfort and 55% reported a decrease in discomfort or pain. One hundred percent would recommend an adjustable workstation to their peers. A study published in 2015 by the Cochrance Database of Systematic reviews found a low quality evidence that sit-stand desks can reduce sitting time at work. They note a need for high quality cluster-randomized trials to assess the effects of different types of interventions on objectively measured sitting time [5].

Given the popularity and availability of both sit-to-stand workstations and personal activity trackers, we investigated whether or not activity trackers could be used to (1) differentiate between sitting and standing at adjustable workstations, and (2) determine if sit-to-stand workstation usage was associated with higher activity levels at work.

## 2 Methods

Two studies were conducted to determine the reliability of an activity tracker to measure changes in sedentary activity. Subjects documented personal standing time while using a sit to stand workstation and seated time during the regular workday. Additional data were collected directly from a Fitbit, on a minute to minute basis, with a synchronized software program called Fitabase®. The software provides similar outputs on activity as the regular Fitbit application but provides greater fidelity. The Fitbit application averages data over a 15 min period where as the Fitabase® software records minute by minute data points. The two studies were approved by the University of Montana Institutional Review Board.

A group of 41 potential candidates were identified as having a sit-to-stand workstation. Out of the 41 individuals, 16 responded and agreed to partake in the study, a resulting 39.0% (16/41) response rate. Participants were recruited using a University of Montana Institutional Review Board approved email. Recruitment was based on having largely sedentary occupations and similar occupational environments. Any interested candidates were able to partake in the study if they met the following inclusion criteria:

1. Full time employees (32 h per week, minimal).
2. 50.0% of their workday consisted of sedentary work (sitting or standing).
3. Had a sit-to-stand workstation.
4. Agreed to complete a weekly sit-to-stand workstation survey.
5. Agreed to wear a Fitbit r every day at work during the study period.

**Equipment.** The studies focused on step count data collected via Fitbits. Sit-to-stand workstation designs varied amongst participants. Some participants utilized sit-to-stand workstations previously provided by the employer, that allowed for the workstation monitor and keyboard to be raised or lowered via an articulating arm. Other participants utilized makeshift stations that allowed them the capability of working unhindered in either position. An example of a makeshift station was the utilization of a metal desk organizer to elevate the workstation to allow a standing work position. For study 2, weekly self-reporting surveys were sent out via Qualtrics, an online research software. Participants without an Fitbit were provided with a Fitbit Flex for their participation in the study. Participants that successfully completed the study were allowed to keep the accelerometer for personal use upon the conclusion of the study. If participants already possessed a Fitbit activity tracker, they were allowed to enroll their device in the study and received a gift card of equivalent value for their participation.

**Study 1.** A self-reporting study was conducted on six office workers. Participants conducted sixteen 30 min sessions of sitting and standing, noting to the minute the time spent in either position. Participants were instructed to record data only while in their workstation setting. If excursions out of the work area occurred, participants would cease logging data for the time gone and resume upon arrival back to their workstation. Standing was again performed at a sit-to-stand workstation and Fitbit data was synced using the Fitabase® software. A total of 4 h of standing and 4 h of sitting data were collected for each participant to determine if an activity tracer\ is capable of detecting differences between sitting and standing sedentary positions. These data sets were used to calculate a sitting step total and standing step total for each participant.

**Study 2.** An observational study was conducted on 16 Montana Tech faculty and staff members. The objective was to determine if the duration of standing time using a sit-to-stand workstation while at work is associated with daily activity level. Participants were requested to comply by wearing a Fitbit daily, continue normal work behavior, and to report the amount of time spent sitting and standing per day on a weekly basis. Data were collected via an online survey created and tracked by Qualtrics. Fitabase® collected Fitbit step and activity data, on a minute by minute basis.

Study 2 was conducted over a 6-week period from April 2016 to mid-May 2016. The study examined activity during a 5-day work week, Monday through Friday, from 8:00 AM to 5:00 PM. Participants were requested to sign a consent form and were assigned randomly generated identification numbers to protect their identity and keep personal information confidential. Participants were asked not to alter their normal work behavior and to report the amount of time spent sitting and standing each day on a weekly basis.

Fitbit step data were used to calculate mean daily step count per individual, by dividing the total number of steps per person by the number of days in the study. Data from the Qualtrics surveys were used to determine the overall average percent standing and percent sitting times per week per participant. Percent standing times were first averaged per week, then averaged over the total number of days for which a questionnaire was successfully completed per individual, providing an overall average standing time percentage.

**Analysis Study 1.** Statistical analyses of the data were processed utilizing Minitab 17 Statistical Software, State Collage PA Minitab, Inc. A final sitting step total and standing step total for each participant was calculated and paired for each subject. The mean number of steps registered while sitting was compared with the mean number of steps registered while standing utilizing a paired t-test. A paired t-test was chosen to test that more activity would occur with a standing position as opposed to a sitting position (Ha: $\mu_d > 0$). The difference between the number of registered steps while sitting vs. while standing was calculated and included in the analysis. The assumptions of the t-test were met by having a continuous dependent variable and the independent variable were matched pairs. The difference between the two variables contained no significant outliers and was normally distributed.

**Analysis Study 2.** Statistical analyses of the data were preformed utilizing Minitab 17 Statistical Software, State Collage PA Minitab, Inc. Self-reported standing percentages and mean daily step counts were first calculated in Microsoft Excel. Standing percentages and mean daily step counts for each participant were analyzed using Spearman correlation. The Spearman correlation method was chosen given that the data did not meet the assumption of being linearly distributed. The Spearman correlation was performed under the null hypothesis that no relationship would exist between standing percentages and mean daily step counts (Ho: r = 0). The Ho hypothesis would be rejected if a relationship was found to exist between the two variables (Ha: $r \neq 0$). The p-value will determine the statistical significance of the relationship between variables if found to be less than $\alpha$ (0.05).

## 3 Results

**Study 1.** Seven individuals were recruited to participate. Of the seven, six successfully completed all required components of the study, 85.7% (6/7). One participant failed to comply with the study parameters, thus resulting in exclusion from data analysis. The total number of registered steps while sitting and the total number of registered steps while standing were calculated and paired per participant. The total number of steps registered while standing varied substantially between participants, ranging from 0 steps to 814 steps. The total number of steps registered while sitting varied less, ranging from 0 steps to maximum of 314 steps. The difference in the total number of registered steps was calculated for each participant, ranging from −23 steps to 500 steps. The negative value here resulted from a participant having more while sitting than while standing *(0 (standing) − 23 (sitting) = −23 steps).*

Analysis of the total number of steps registered while standing found a mean of 309, standard error mean of 139, and a standard deviation of 340. The total number of steps registered while sitting for participants revealed a mean of 67, a standard error mean of 49, and a standard deviation of 122. Based off of the paired t-test analysis, no statistically significant difference was detected between the total number of steps registered while sitting and the total number of steps registered while standing (p-value = 0.113). The distribution for the number of steps registered while standing and number of steps registered while sitting can be seen in Fig. 1.

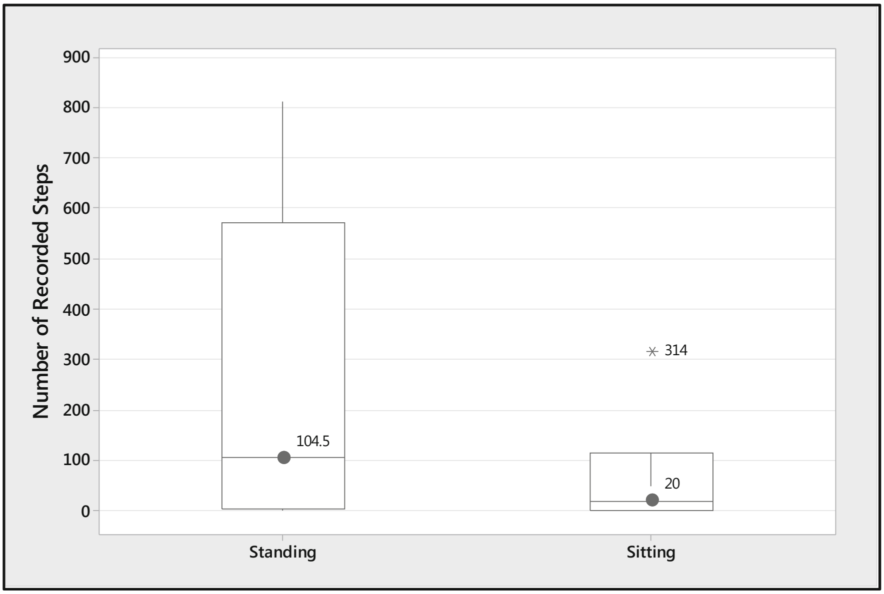



**Fig. 1.** Box and whisker plot of total number of steps registered while standing compared to total number of steps registered while sitting.

**Study B.** Of the 16 individuals that agreed to participate, 12 successfully completed all required components of the study, 75.0% (12/16). Four participants failed to either register their activity tracker or synchronize their daily sit-to-stand workstation usage with Fitabase, thus resulting in exclusion from data analysis. The 12 remaining participants were monitored for the 30-day (5 days × 6 weeks) study period. The mean standing percentage, mean daily step count, and number of days with questionnaire entries were paired by participant for the 12 participants that successfully completed the study parameters. Mean standing percentages varied noticeably between participants, ranging from 27–86%. Mean daily step counts ranged from as low as 3,033 steps to as high as 7,094 steps. The number of recorded self-reporting surveys also varied, with 58.3% (7/12) participants having 30 recorded questionnaire entries, 25.0% (3/12) having 25 recorded entries, and 16.7% (2/12) having only 20 recorded entries; see Table 1. The comparison between mean standing percentage and mean daily step count per participant can be seen in Fig. 2.

Analysis of the self-reported mean standing percentages revealed a mean of 49, a standard error mean of 6, and a standard deviation of 20. Mean daily step count data for the participants revealed a mean of 5,052, and a standard error mean of 444, a standard deviation of 1539. The Spearman correlation analysis revealed a weak relationship ($\rho = 0.301$) between participants' self-reported standing time percentages and their respective average daily step value. Furthermore, the correlation was not statistically significant between a participant's mean standing percentage and mean daily step count

**Table 1.** Self-reported standing percentage, daily step average, and number of recorded questionnaire days observed per participant.

| Participant | Mean standing percentage (%) | Mean daily step count | # Days with questionnaire entries |
|---|---|---|---|
| 1 | 47 | 5149 | 30 |
| 2 | 86 | 4399 | 30 |
| 3 | 31 | 5183 | 30 |
| 4 | 54 | 3515 | 20 |
| 5 | 83 | 5354 | 30 |
| 6 | 69 | 6452 | 30 |
| 7 | 29 | 7094 | 30 |
| 8 | 28 | 3730 | 25 |
| 9 | 53 | 5291 | 20 |
| 10 | 35 | 3448 | 25 |
| 11 | 52 | 7979 | 25 |
| 12 | 27 | 3033 | 30 |

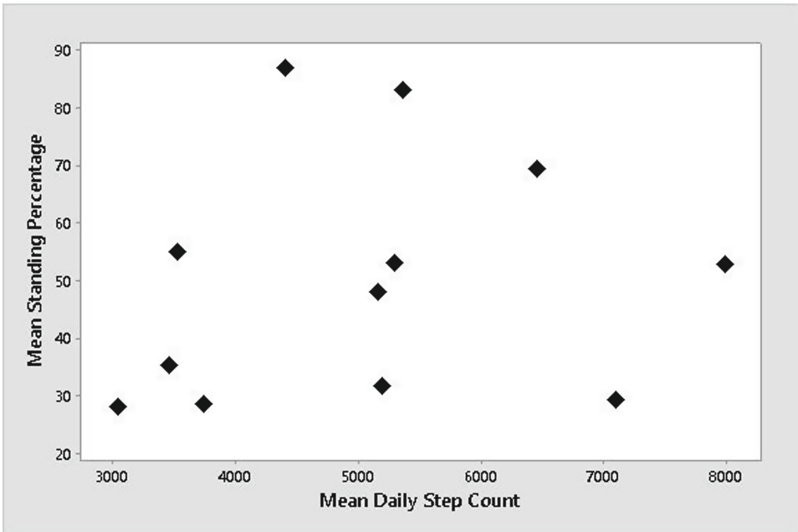



**Fig. 2.** Scatterplot graph of mean daily step count compared to participant mean standing percentages for each subject.

(p-value = 0.342). Only a 3.5% ($R^2$) of the variability in step count was accounted for due to percent standing. As shown in the scatterplot in Fig. 2, no clear relationship between the participant's mean standing percentage and mean daily step count is evident.

## 4 Discussion

Activity trackers such as Fitbit have been determined capable of distinguishing between varying levels of activity, categorized as low, medium and high. The differences in activity level classification is based on step count and a cut point, 99 steps in 5 min is low, whereas 100 steps in 5 min may be medium. Standing still does not generally produce a high enough steps to meet the cut point between low and medium, therefore a user will not see an increase in their percentage of activity in the medium category with increased standing time. They may also not see an increase in steps overall. On the other hand, activity trackers erroneously record steps where none have been taken, this was apparent in our studies and is in line with other findings. This finding should be conveyed to workers who are using activity trackers to monitor their activity level. Standing may have physiological benefits but using step count and activity level on a Fitbit does not appear to measure them. Heartrate may be a better metric.

The main strength of the studies was investigating whether an activity tracker could distinguish between sedentary sitting vs sedentary standing positions, which has not been previously attempted. Despite our results showing that no significant difference in the standing steps and sitting steps could be seen, the amount of steps registered while standing was generally higher than the number of steps while sitting. This finding is important as it suggests that activity trackers, in this case Fitbits, may be useful in ergonomic studies monitoring sit-to-stand workstation usage, but further research is needed to determine if a statistically significant difference can be detected with larger data sets. Additional research will also help establish how researchers and employers can utilize the data to measure the degree of sit-to-stand workstation usage and to determine if sit-to-stand workstations result in changes in activity of the users.

Suggestions for future research include repeating Study 2, but account for time spent sitting, standing and walking in the self-reporting questionnaire, so that activities can be correctly correlated with the mean daily step counts; expanding Study 2 to include a larger sample size to verify the significance of the Fitbit's ability to distinguish between sitting and standing positions. Activity trackers capable of physiological measure should also be explored.

Our findings suggest that basic activity trackers may not be useful in measuring sit-to-stand workstation usage. Furthermore, there was no evidence to indicate a meaningful relationship between increased sit-to-stand workstation utilization, as a percentage of time at work, and increased physical activity while at work. Activity trackers capable of measuring heart rate may hold promise as a means to differentiate between sitting and standing during office work. Further research is needed to determine if greater sit-to-stand workstation usage increases physical activity in general, not just at work.

**Acknowledgments.** Two entities provided support for this research project. The National Institute for Occupational Safety & Health's Montana Tech Training Program Grant (2016) and the Montana Tech of the University of Montana's Faculty Development Grant Initiative (2016).

## References

1. Chemi, E.: Here's what happened when I wore 10 fitness trackers at once. Retrieved from CNBC, 7 June 2016. https://www.cnbc.com/2016/05/26/heres-what-happened-when-i-wore-10-fitness-trackers-at-once.html
2. Pronk, N.: Reducing occupational sitting time and improving worker health: the take-a-stand project. Prev. Chronic Dis. (2012)
3. Sarasohn-Kahn, J.: Heathcare IT news. Retrieved from wearable activity tracking device purchasing expected to grow 11 percent in 2016, 05 May 2016. http://www.healthcareitnews.com/blog/wearable-activity-tracking-device-purchasing-expected-grow-11-percent-2016
4. Sasaki, J., Hickey, A., Mavilia, M.: Validation of the fitbit wireless activty tracker for prediction of energy expendature. J. Phys. Act. Health **12**(2), 149–154 (2014)
5. Shrestha, N., Kukkonen-Harjula, K., Verbeek, J., Ijaz, S.: Workplace interventions for reducing sitting at work. Cochrane Database Syst. Rev. (2016)
6. Simons, C., Hughes, L., Engeland, M., Goldbohm, R.: Physical activity, occupational sitting time, and colorectal cancer risk in the Netherlands cohort study. Am. J. Epidemiol. **177**, 514–530 (2013)
7. Straker, L., Atherton, G., Dunstan, D.: Excessive occupational sitting is not a 'safe system of work': time for doctors to get chatting with patients. Med. J. Aust. **3**, 138–140 (2014)

# Application of Wearable Technology for the Acquisition of Learning Motivation in an Adaptive E-Learning Platform

Mathias Bauer[1(✉)], Cassandra Bräuer[1], Jacqueline Schuldt[1], Moritz Niemann[2], and Heidi Krömker[1]

[1] Ilmenau University of Technology, Ilmenau, Germany
{Mathias.Bauer,Cassandra.Braeuer,Jacqueline.Schuldt, Heidi.Kroemker}@tu-ilmenau.de
[2] Medical School Hamburg, Hamburg, Germany
moritz.Niemann@medicalschool-hamburg.de

**Abstract.** Motivated learning is the prerequisite for a deep processing of learning content and a long retention performance, as well as the basis for joy of learning and persistent interest. The SensoMot-project ("Sensor Measures of Motivation for Adaptive Learning") aims at identifying critical motivational incidents during adaptive e-learning sessions in the context of university courses of micro- and nano-technology through sensory acquisition with current consumer wearables. These critical motivational incidents will be used to adapt learning content at runtime and thus enhance motivation.

**Keywords:** Adaptive e-learning · Motivation · Physiological data Sensory acquisition · Wearable technology

## 1 Introduction

Wearable technology in the form of fitness-trackers or smartwatches is used within the scope of the SensoMot-project because of the increasing user-awareness for self-reflective data analysis and for the potentials of this technology itself. The quantified self-trend emerged from a many-faceted group of life-hackers, data analysts and early adopters [1] and shall now be made accessible to a wider group of potential users – within the project the focus is especially on learning situations. The main objective is to use the unobtrusive sensors for measuring physiological data provided by wearables instead of overly expensive medical equipment as well as intrusive time and cognitive resources consuming self-reporting methods. Available wearable sensors and corresponding physiological indicators cover electroencephalography (EEG), heart rate and electrodermal activity [2]. Within the scope of this paper studies of two universities involved in the SensoMot-project are presented which examine the usability and sensory data quality of current wearable devices. Paragraph two includes fundamentals of wearable technology focusing on wrist-worn devices and their means of sensory data acquisition. Furthermore, the usage of the wearable devices in the scope of the


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 29–40, 2019.
https://doi.org/10.1007/978-3-319-94619-1_4

SensoMot research project will be described. Paragraph three depicts the research design and implementation of the two studies at the universities in Ilmenau and Hamburg. Subsequently, chapter four outlines the results of the described studies.

## 2 Wearable Technology in E-Learning Contexts

Learners in e-learning are not always able to express their feelings and thoughts in words. Often, they may not even be aware that they are excited, motivated or bored, and sometimes they may not accurately report during or after an e-learning session which content they have paid less attention to. Psychophysiological methods are independent of speech or memory, and they measure accurately in time [3]. They are used to circumvent the disadvantages of subjectively verbal data collection and to supplement it [4]. In order not to block the learning flow, indicators for motivational adaptation processes shall be obtained from data patterns obtained with non-reactive sensors.

### 2.1 State of the Art in Wearable Technology

Within the scope of the conducted studies the focus of the examination was on current consumer wearables, especially wrist-worn fitness wearables like activity trackers and smartwatches. These devices typically are equipped with a varying number of sensors and other technologies. In comparison to other mobile devices the scale, size and energy consumption of the different components of a wristband-wearable have to be reduced [5]. An ideal wearable should come with the following physical and functional attributes:

- Physical attributes: lightweight, aesthetically pleasing, nearly invisible, shape conformable.
- Functional attributes: multifunctional, configurable, responsive, bandwidth [6].

In comparison to other mobile devices wearables tend to be more convenient for users to use and can further be distinguished from then according to the following five characteristics:

- The device can be used while the wearer is in motion.
- The device can be used while the user's hands are free or occupied with other tasks.
- The device is located within the corporal envelope of the user.
- The user should be able to maintain control over the device.
- The device should be constantly within the user's reach. [7]

Wearables are able to execute several core use cases like perceiving, processing and analyzing, storing, transmitting or displaying of information which require a personalized information processing that transforms sensory data into perceivable data [6]. The system components of common fitness wearables are shown in Fig. 1 and consist of: system-on-chip with low-power microprocessor and wireless interface, actuators (e.g. visual, auditory or haptic feedback), external flash data storage, Li-ion battery, system power management and different classes of sensors like motion-sensors (e.g. accelerometer, gyroscope, magnetometer), environmental sensors (e.g. temperature, pressure, humidity) or physiological sensors (e.g. heart rate, respiration rate) [5].

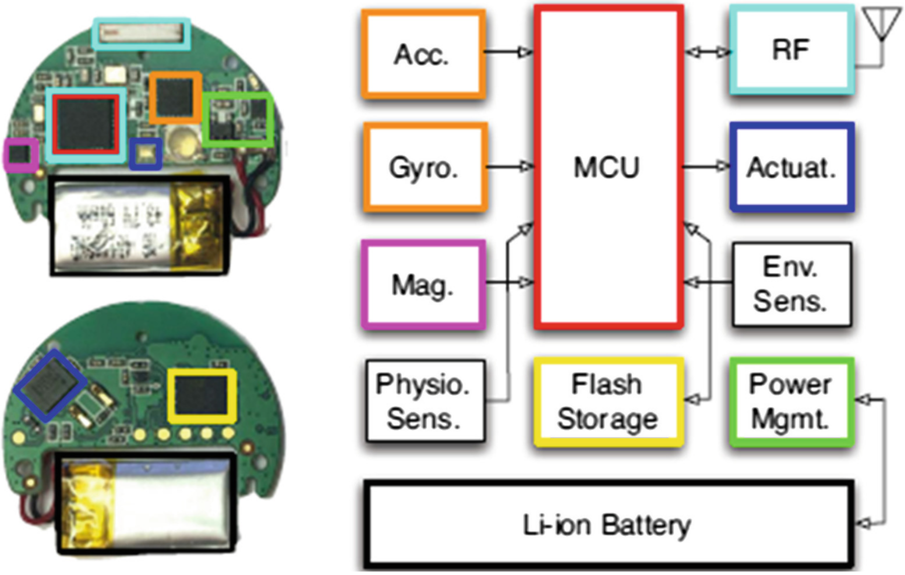



**Fig. 1.** System components of common fitness wearables [5]

The data processing and stream of system activities to achieve the goal of measuring vital parameters of the user in everyday life situations can be described as follows: first, the input is acquired via active or passive sensors of different classes. The focus of the current studies was on physiological data like heart rate, skin conductance or brain signals. Afterwards, the sensory data is stored, interpreted and processed before being displayed to the user via a user interface. This usually also includes a companion device, typically a smartphone, and a dedicated companion app for elaborated data analysis and logging [8].

Figure 2 shows a current taxonomy of wearables, beginning with the scope of functionality and followed by type, deployment mode, communication mode and field of use [6]. The majority of current consumer wearables try to adopt the concept of “quantified self”, that means the use of personal data to improve a person’s health and well-being. Quantified self especially addresses the growing trend of self-monitoring or self-tracking which is the process of recording and monitoring one’s behavior, thoughts or vital parameters [1]. Since this concept relies on appropriate sensors, wearable devices are well suited because of their already named attributes and functionalities.

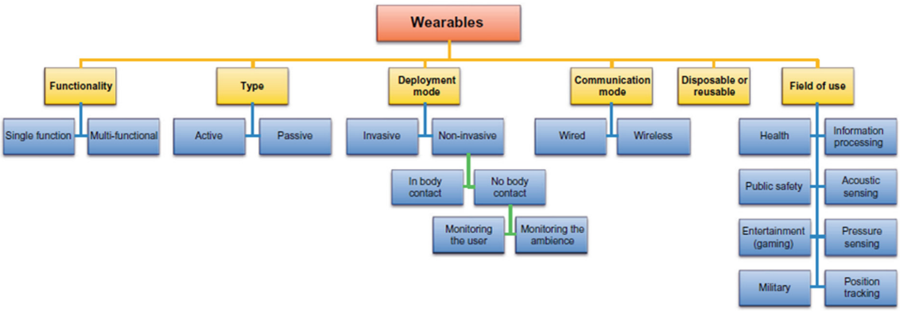



**Fig. 2.** Taxonomy of wearables [6]

## 2.2 Wearable Devices as Sensors for Psychophysiological Data Acquisition

Psychophysiology examines the interaction between the user's mind and body, and relies on the assumptions that changes in the human body are based on changes in human behavior, affective or motivational states and that there is an optimal cognitive and physical user state for a given task [2]. Advantages of psychophysiological measures are that they provide objective quantitative results unbiased by subjective ratings. Also, sensory devices like wearables are unobtrusive and noninvasive, which means, they do not influence the user's task performance. Data convergence is another benefit and states that multiple psychophysiological measures can increase the accuracy of the mental construct being examined. A key advantage especially for adaptive systems is that these measures provide immediate and continuous results that can be interpreted at runtime and being used for adjusting a system the current user needs. However, psychophysiological measures also have some limitations like the high requirements regarding an accurate and high-quality raw-data that is free from artefacts. The processing and analysis of the amount of acquired sensory data requires computational power especially concerning the real-time demand.

In contrast to self-reporting methods psychophysiological measures do not provide reasons for users' reactions towards stimuli of an interactive system [2, 3]. Commonly used psychophysiological measures in the scope of adaptive systems are electroencephalography, cardiovascular activity (heart rate, heart rate variability), electrodermal activity, skin temperature and electromyography.

Current wrist-worn wearables measure cardiovascular activity typically by means of optical procedures – most commonly photoplethysmography (PPG). PPG consists of the measurement of changes in blood volume and optical means. The measurement principle is that living tissue is illuminated with a light beam that captures and analyzes a proportion of the propagated light in the tissue. The analysis depicts functional und structural information of the tissue. Blood is an inhomogeneous liquid medium so the absorption and dispersion of the light beam varies during the cardiac cycle. Changes in blood volume contribute to observable PPG signal variations that are separated in two components: AC and DC. The choice of the light's wavelength has to match the application context and the skin color of potential users. Studies showed that green light is preferred over infrared light in combination with light skin pigmentation and normal ambient temperature [9]. Furthermore, comparisons of electrocardiogram (ECG) and various PPG-signals showed that green PPG-signals provided the most valid results [10]. PPG Measurement requires the usage of at least one light-emitting and one light-receiving source. Wrist-related measurement of PPG signals like in current wearable devices is advantageous because wrist and forearm are less prone to movement artefacts and it is well accepted by users because of its similarity to wearing a classical watch [9]. A cardiovascular indicator of particular interest is heart rate variability which underlies variations based on environmental stimuli that can lead to emotional or physiological reactions. Psychological stress or arousal for example can lead to an increased heart rate and a decreased heart rate variability [11].

Some recent wrist-worn wearables like the Empatica Embrace or the Microsoft Band 2 are equipped with sensors for the measurement of electrodermal activity and its attached parameters like galvanic skin resistance (GSR), skin conductance or skin potential. Skin conductance as a central parameter of electrodermal activity is mainly influenced through the function of the eccrine sweat glands that are responsible for the regulation of the temperature of the human body [12]. Electrodermal activity is influenced by various external and internal parameters like ambient temperature as well as psychological and emotional processes and is very well suited for the measurement of stress as a state of increased arousal accompanied by negative emotions. Considering these psychophysiological signals studies were conducted regarding the assessment of user emotions during the usage of computer systems [13]. The measurement of the electrodermal activity with wrist-worn wearables relies on exosomatic means using direct current in combination with accelerometers for addressing body movements and posture. Exosomatic measures require two electrodes on active sites, usually using palms or fingers because of their sensitivity regarding distinguished electrodermal activity [12].

As an alternative to wrist-worn wearables, wearable headbands like Emotiv Insight or MUSE were released for the measurement of brain signals in accordance to clinical electroencephalography (EEG). These devices aim at measuring the electrical activity of the brain as means for research on relations regarding human behavior. EEG allows for precise time measurement of mental processes and divides brain activity according to different frequency bands indicating different levels of mental activity or arousal. EEG can especially be used regarding the study of awareness, emotions or mental workload [14, 15]. By dividing an EEG band that corresponds to brain activation by one that corresponds to brain idling, an EEG index can be obtained that represent a general measure of brain activation and correspond to user workload and engagement [16].

### 2.3 Wearable Devices and Their Use in the SensoMot Project-Framework

Wearable devices shall be used as means for sensory acquisition of physiological data within the scope of the project-framework. The result of the intend will be an adaptive e-learning platform that adjusts its instruction, presentation and navigation according to changes of the learner´s current motivation. The fit between learner and learning

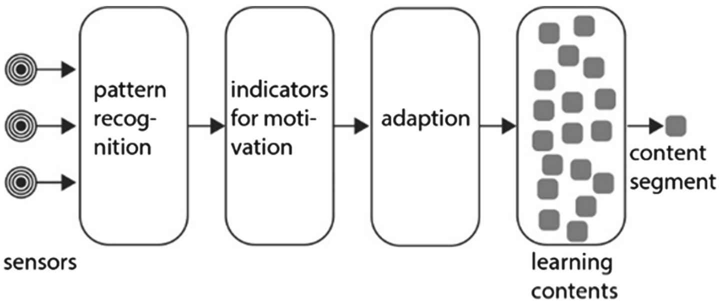



**Fig. 3.** SensoMot process chain: sensor data is condensed by means of pattern recognition to indicate motivational states, which are used for the adaptation of learning content [17].

situation should be improved through reacting on motivationally critical incidents during learning via adaptation [17]. The process chain is depicted in Fig. 3. The wearable devices take the role of the sensors within this chain.

The sensory data is recorded via wrist-worn or head-worn wearable devices and then merged, processed, interpreted and analyzed by means of pattern recognition via machine learning. Based on recorded changes in indicators for motivation the adaptation of specific learning content is executed [17].

## 3 Method

To exclude usability of wearables as an influencing variable on the learning process, a first formative usability evaluation of current consumer devices was conducted in 2016 at the Ilmenau University of Technology. Besides, an experiment using wearable devices in a learning session for acquiring sensory data while learning was conducted at the Medical School Hamburg in 2017.

The focus of the formative evaluation was on task performance, ease of use, usability metrics like efficiency and effectiveness as well as on qualitative statements concerning the potential of wearable technology in learning contexts.

In a second step, different wrist-worn wearables and EEG-headsets were employed in an experiment where subjects studied while wearing the devices. Subjects made self-reports about their motivational state during and after the experiment. Sensor data collected from the devices was successfully used to draw inference about the subjects' motivational states. A usability questionnaire was presented afterwards.

From 2016 until 2017 a multi-stage iterative user centered evaluation of current wrist-worn wearables was conducted at the University of Technology Ilmenau by the help of student projects. The focus was on laboratory usability tests [18] with users, online surveys towards wearable usage and acceptance and requirements elicitation. Within the first stage of the evaluation in summer semester of 2016 core use cases of current wrist-worn wearables were identified with several student projects that resulted in typical benchmark tasks that should be executed in subsequent usability tests. The specific wrist-worn wearables being evaluated were one smartwatch – the Samsung

Gear S2 and three fitness trackers – the Samsung Gear Fit, the Sony Smartband 2 and the Garmin Vivosmart HR. The devices were tested along with a mutual android smartphone in a usability lab to minimize external influencing variables.

The usability tests were part of the second stage of the evaluation and took place in the winter semester of 2016. Nine participants tested the four devices, looking for critical incidents during usage and using a post-test questionnaire as well as the thinking aloud method during the test. According to [19] this number of participants should be sufficient in a formative evaluation to identify the majority of critical incidents or usability issues.

Within the Medical School Hamburg-study 64 subjects participated in a learning experiment. All subjects were majoring in psychology and 14 were male. Participants studied material from a higher semester of educational psychology for 60 min. While studying, all subjects wore the Microsoft Band 2 (MSB2) on their offhand (left hand for right-handed individuals and right hand for left-handed individuals). MSB2 data was available for 50 participants. The MSB2 continuously measured GSR, heart rate and RR-intervals derived from a PPG sensor, as well as movement data from a gyroscope and accelerometer. In addition, 31 subjects wore the Emotiv Insight EEG headset, and 32 subjects donned an alternative EEG headset, the InterAxon Muse. The headsets continually measured scalp EEG throughout the whole learning session. Data from the Muse headset had to be discarded due to technical problems with data collection. While learning, subjects were interrupted every 4.5 min by a parsimonious questionnaire regarding their motivational state, for a total of 13 experience samples. The questionnaire was presented on a smartphone and response time averaged 16.3 s. At the end of the 60 min, learning outcome was assessed with a multiple choice questionnaire. Subjects reported on their experiences during the learning session by filling out the “Mood and Affect” and “Motivation” portion of the Dundee Stress State Questionnaire (DSSQ) [20]. The “Mood and Affect” portion of the DSSQ is equivalent to the UWIST Mood Adjective Checklist, and is comprised of 29 affective adjectives. Subjects are asked to state to which degree they felt the given affective state over the course of the learning session. These items are then summed up to the four scales “Energetic Arousal”, “Tense Arousal”, “Hedonic Tone”, and “Anger/Frustration”. The “Motivation” part of the DSSQ consists of the “Intrinsic Motivation” and “Workload“subscales, the latter of which is the NASA-TLX questionnaire in modified form [21].

## 4 Results and Discussion

The core use cases identified in the first evaluation stage serving as benchmark tasks and being evaluated during the usability tests were:

- Putting on the wrist-worn wearable.
- Synchronizing the wearable with the companion device and app.
- Starting the wearables’ heart frequency measurement and displaying the result on the companion device and app.

- Starting the wearable's step counting measurement while walking around the lab and displaying the results on the companion device and app (walked steps, distance, burnt calories).
- Free navigation tasks across the devices´ different functionalities [8].

These functions are available in all four tested wearable devices and serve as starting point for a comparative analysis. The questionnaire regarding the user experience and perceived usability of the devices that had to be filled in after the usability tests was a combination of different common usability and user experience questionnaires like the SUS or the USE questionnaire [22] expanded by wearable-specific items. The items were rated on a five-point scale. Means of the ratings as a combined overview of all four devices can be seen in Fig. 4.

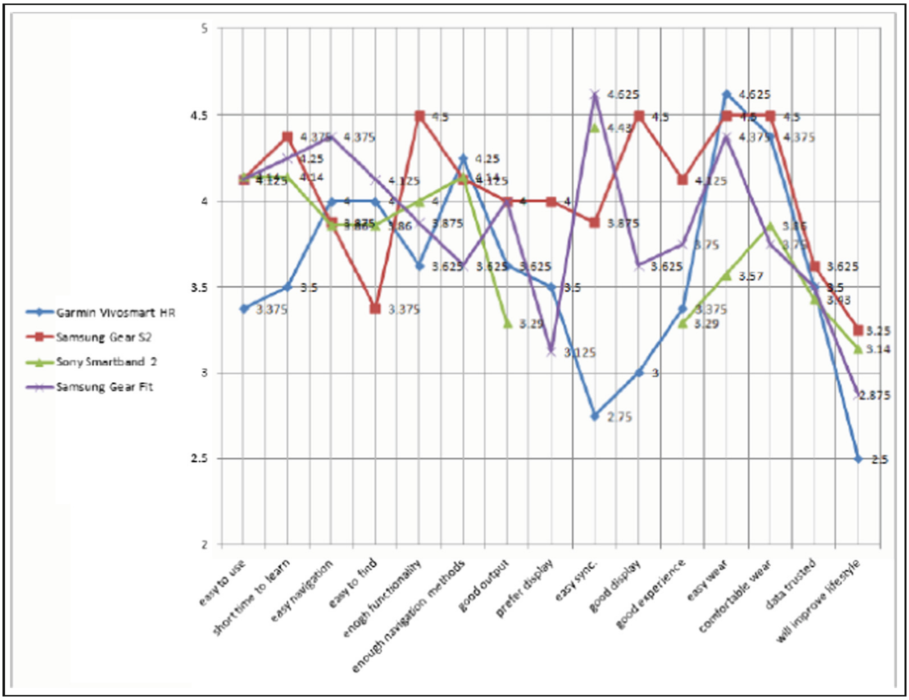



**Fig. 4.** Combined results of the perceived user experience of all four tested wearable devices [8]

Some general findings were that all of the devices provided enough functionality in the opinion of the users and were easy and comfortable to wear. Regarding the Garmin Vivosmart HR many users had trouble synchronizing the wearable with the companion device and the lack of a colored display affected the users' ratings. The Gear S2 smartwatch was generally rated well because of the appealing touch display. It was said though, that a device like this should provide more health sensing features and some of the gathered data was hard to find in the companion app. This feature was rated well in case of the Sony Smartband 2 but the device is lacking a display which affected user's ratings and opinions noticeable. The Samsung Gear Fit was said to be a device without significant flaws whose navigation concept is easy to learn. A drawback however was the bulky rectangular display that affected the aesthetic aspect of wearability. A conspicuity that was observed during the tests of the different devices were the different number of steps counted while walking around the lab. Although the participants walked the same route with every device the number of counted steps differed partly drastically [8].

Altogether, the devices achieved an acceptable to good perceived user experience, especially for being easy to use, wear and learn. Additionally, the participants were asked to rate the importance of incorporated functionalities on a five-point scale which should result in stating functional requirements based on the MoSCoW requirements analysis [23].

The identified must-requirements were the existence of a display (ideally a colored touch display), long battery life, easy wearability and the ability to save long-term statistics to meet the requirements of the quantified-self trend. The different sensors, especially health related ones like heart rate or step counter are should-requirements that were desired by most participants but the preference for particular sensors depends on the user's goals.

In 2017 an online survey regarding the acceptance and prevalence of wearables amongst students was conducted in combination with a market and environmental analysis [24]. The survey asked whether participants own a wearable device and if so which type and brand as well as since when and how frequently. There were also items concerning the perceived usefulness according to the technology acceptance model [25]. The online survey had 115 participants (N = 115) – 71.3% of them did not own a wearable device. Amongst the participants that owned a wearable the fitness tracker was the most frequently named device. More than half of the participants that own a wearable (51.5%) owned their device less than one year. 69.7% of these participants wear their device daily. Most of the participants who do not own a wearable seem to be reluctant concerning the purchase of such a device.

Regarding the functionality of wearables, the actual usefulness is mainly perceived for the already mentioned core use cases like health, fitness, data acquisition and analysis, route planning and time indication. 63.7% of the participants that own a wearable are satisfied or very satisfied with their devices. Amongst the participants without wearable devices the relatively high acquisition costs and the lack of a noticeable added value were named reasons for staying away from this technology. Major perceived advantages provided by wearables were according to the participants′ opinions fostering "being online", information being immediately available, health care and replacing the need to constantly checking one's smartphone. Major perceived disadvantages were concerns about data security and privacy issues as well as the feeling of being more and more dependent on technology [24].

For analysis of the Hamburg study, the median GSR and an index of brain activity (Power in the Beta Band divided by Power in the Alpha Band across all electrode sites) were computed over the whole learning session. The 13 intervals where subjects were busy filling out the parsimonious questionnaire were excluded.

Real-world psychophysiological data frequently violates assumptions requiring normally distributed data, as was also the case here. Therefore, the Kendall tau-b correlation coefficient ($\tau_b$) as a non-parametric measure of bivariate association was used. A tau-b of 1 represents a perfect relationship between two variables, while 0 represents no relationship at all.

As can be seen in Table 1, there was no statistically significant relationship between the EEG index Beta/Alpha and the DSSQ scales "Intrinsic Motivation" ($\tau_b$ = .20, p = .137), "Workload" ($\tau_b$ = − .12, p = .365), "Tense Arousal" ($\tau_b$ = − .05, p = .690), "Anger/Frustration" ($\tau_b$ = − .02, p = .847), "Hedonic Tone" ($\tau_b$ = .07, p = .600), and "Energetic Arousal" ($\tau_b$ = .22, p = .106). These results are based on n = 27 people for which EEG data was available.

There were significant positive relationships between GSR and the DSSQ subscale "Intrinsic Motivation" ($\tau_b$ = .29, $p$ = .002), as well as "Energetic Arousal" ($\tau_b$ = .27, $p$ = .006). There were no significant relationships between GSR and the subscales "Workload" ($\tau_b$ = .10, $p$ = .294), "Tense Arousal" ($\tau_b$ = − .01, $p$ = .906), "Anger/Frustration" ($\tau_b$ = − .16, $p$ = .113), and "Hedonic Tone" ($\tau_b$ = .15, $p$ = .118). These results are based on n = 50 people for which GSR data was available.

**Table 1.** Bivariate relationships between sensor measures and DSSQ scales. Significant results (α = .01) are marked with asterisk.

| Sensor measure | DSSQ scale | Kendall $\tau_b$ |
|---|---|---|
| EEG beta/alpha | Intrinsic motivation | .20 |
| | Workload | −.12 |
| | Tense arousal | −.05 |
| | Anger/frustration | −.02 |
| | Hedonic tone | .07 |
| | Energetic arousal | .22 |
| GSR | Intrinsic motivation | .29* |
| | Workload | .10 |
| | Tense arousal | −.01 |
| | Anger/frustration | −.16 |
| | Hedonic tone | .15 |
| | Energetic arousal | .27* |

## 5 Conclusion

The usability and user experience of the wrist-worn devices was rated very positive while EEG-headsets received mixed approval both in terms of user acceptance for learning support and regarding feel of comfort while wearing. Especially the smartwatch exceeds the fitness-trackers as to functionality and the more appealing user interface. However, besides being a convenient extension to a user's smartphone a true added value regarding the application of wearables has yet to be shown to increase user acceptance.

The Hamburg study showed that wearables can be used to measure psychological states of users during a learning session. GSR measured from the wrist-worn Microsoft Band 2, but not an index of electroencephalographic activity, was correlated with users' self-reports about motivational and affective states.

No relationship was found between subjective workload and this EEG index in the current research. While experimental tasks are usually made up of a very narrowly defined set of demands, a broad learning session such as the one employed in this study, consists of many different tasks such as reading, writing, remembering, memory retrieval, reflection, among others. Further analysis will have to explore more complicated EEG indices, as well as their connection to moment-to-moment motivational states.

These results highlight that wearable technology is in the process of opening up whole new avenues of research, enabling studies that were thought impossible just a few years ago. We urge decision makers to make raw data accessible via application programming interfaces and software development kits to facilitate widespread adoption of this new technology by researchers.

**Acknowledgments.** Part of the authors' work has been supported by the German Federal Ministry for Education and Research (BMBF) within the joint project SensoMot under grant no. 16SV7516, within the program "Tangible Learning".

## References

1. Choe, E.K., Lee, N.B., Lee, B., Pratt, W., Kientz, J.A.: Understanding quantified-selfers' practices in collecting and exploring personal data. In: Jones, M., Palanque, P., Schmidt, A., Grossman, T. (eds.) CHI 2014, One of a CHInd. Conference proceedings: Toronto, Canada, 26 April – 1 May 2014, the 32nd Annual ACM Conference on Human Factors in Computing Systems, the 32nd annual ACM conference, Toronto, Ontario, Canada, pp. 1143–1152. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2556288.2557372
2. Hou, M., Banbury, S., Burns, C.: Intelligent Adaptive Systems: An Interaction-Centered Design Perspective. CRC Press, Boca Raton (2015)
3. Trepte, S., Reinecke, L.: Medienpsychologie. Kohlhammer-Urban-Taschenbücher, vol. 726, 1st edn. Kohlhammer, Stuttgart (2013)
4. Ravaja, N.: Contributions of psychophysiology to media research: review and recommendations. Media Psychol. **6**, 193–235 (2004). https://doi.org/10.1207/s1532785xmep0602_4
5. Williamson, J., Liu, Q., Lu, F., Mohrman, W., Li, K., Dick, R., Shang, L.: Data sensing and analysis: challenges for wearables. In: The 20th Asia and South Pacific Design Automation Conference, 2015 20th Asia and South Pacific Design Automation Conference (ASP-DAC), Chiba, Japan, 19 January 2015 – 22 January 2015, pp. 136–141. IEEE (2015). https://doi.org/10.1109/aspdac.2015.7058994
6. Park, S., Chung, K., Jayaraman, S.: Wearables. In: Sazonov, E. (ed.) Wearable Sensors, pp. 1–23. Academic Press, Cambridge (2014)
7. Jiang, H., Chen, X., Zhang, S., Zhang, X., Kong, W., Zhang, T.: Software for wearable devices: challenges and opportunities. In: 2015 IEEE 39th Annual Computer Software and Applications Conference, 2015 IEEE 39th Annual Computer Software and Applications Conference (COMPSAC), Taichung, Taiwan, 01 July 2015 – 05 July 2015, pp. 592–597. IEEE (2015). https://doi.org/10.1109/compsac.2015.269

8. Tarakji, A.M.: Analysis of fitness wearables. Unpublished Master Thesis, Technische Universität Ilmenau (2016)
9. Lemay, M., Bertschi, M., Sola, J., Renevey, P., Parak, J., Korhonen, I.: Application of optical heart rate monitoring. In: Sazonov, E., Neuman, M.R. (eds.) Wearable Sensors. Fundamentals, Implementations and Applications, pp. 105–129. Academic Press, Cambridge (2014)
10. Lee, J., Matsumura, K., Yamakoshi, K.-i., Rolfe, P., Tanaka, S., Yamakoshi, T.: Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion. In: Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society (2013). https://doi.org/10.1109/embc.2013.6609852
11. Appelhans, B.M., Luecken, L.J.: Heart rate variability as an index of regulated emotional responding. Rev. Gen. Psychol. **10**, 229 (2006). https://doi.org/10.1037/1089-2680.10.3.229
12. Boucsein, W.: Electrodermal Activity. Springer, Boston (2012)
13. Westerink, J.H.D.M., van den Broek, E.L., Schut, M.H., van Herk, J., Tuinenbreijer, K.: Computing emotion awareness through galvanic skin response and facial electromyography. In: Toolenaar, F., Westerink, J.H.D.M., Ouwerkerk, M., Overbeek, T.J.M., Pasveer, W.F., Ruyter, B. de (eds.) Probing Experience, vol. 8, pp. 149–162. Springer, Dordrecht (2008). Philips Research
14. Radüntz, T.: Kontinuierliche Bewertung psychischer Beanspruchung an informationsintensiven Arbeitsplätzen auf Basis des Elektroenzephalogramms. Humboldt-Universität zu Berlin (2016)
15. Birbaumer, N., Schmidt, R.F.: Biologische Psychologie, 7th edn. Springer, Heidelberg (2010)
16. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biol. Psychol. **40**(1–2), 187–195 (1995)
17. Schneider, O., Martens, T., Bauer, M., Ott-Kroner, A., Dick, U., Dorochevsky, M.: SensoMot – sensorische erfassung von motivationsindikatoren zur steuerung adaptiver lerninhalte. In: Igel, C., Ullrich, C., Wessner, M. (eds.) Bildungsräume 2017, pp. 267–272. Gesellschaft für Informatik, Bonn (2017)
18. Rubin, J., Chisnell, D.: Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd edn. Wiley, Indianapolis (2008)
19. Nielsen, J.: Usability Engineering. Kaufmann, Amsterdam (2010)
20. Matthews, G., Joyner, L., Gilliland, K., Campbell, S.E., Falconer, S., Huggins, J.: Validation of a comprehensive stress state questionnaire: Towards a state big three. Pers. psychol. Eur. **7**, 335–350 (1999)
21. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Human Mental Workload, Advances in Psychology, vol. 52, pp. 139–183. Elsevier (1988)
22. Hartson, R., Pyla, P.S.: The UX Book: Process and Guidelines for Ensuring a Quality User Experience, 1st edn. Morgan Kaufmann, Amsterdam (2012)
23. Cline, A.: Agile Development in the Real World. Apress, Berkeley (2015)
24. Göring, K.: Analyse der Umwelt von Wearables mittels Techniken des strategischen Managements. Unpublished Master Thesis, Technische Universität Ilmenau (2017)
25. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology. a comparison of two theoretical models. Manage. Sci. **35**, 982–1003 (1989). https://doi.org/10.1287/mnsc.35.8.982

# What Are You Waiting for? – Perceived Barriers to the Adoption of Fitness-Applications and Wearables

Chantal Lidynia(✉), Eva-Maria Schomakers, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{lidynia, schomakers, ziefle}@comm.rwth-aachen.de

**Abstract.** To get a better insight into perceived barriers and motives to use fitness apps and wearables, a mixed-method study-design, consisting of both a qualitative and a quantitative part, has been used. In an online questionnaire answered by N = 166 participants, the perceived usefulness of fitness-apps as well as perceived benefits and barriers were evaluated. Additional factors such as experience with such apps and wearables, technical self-efficacy, and privacy concerns were also taken into account. Results show that fitness apps and wearables are met with approval. They are deemed useful and provide necessary information to start or keep a healthy lifestyle. Demographic variables also had an impact on the intention to use such devices. One of the biggest barriers seems to be the concern for one's privacy, the collected data seen as rather sensitive. Also, the additional use of a wearable changes the perception and intention to use a fitness app.

**Keywords:** Human factors · Wearables · Life-logging · Motives
Barriers · Activity tracker

## 1 Introduction

Nowadays, computing is ubiquitous. There is basically no venue of life that cannot be tracked, enhanced, or supported by technology. One such example are so-called fitness-apps that can be used on their own and sometimes offer the addition of a wearable, a small computer that includes different sensing mechanisms to add more data to the app input. As the saying goes, "knowledge is power," and these small fitness-trackers offer a lot of knowledge about daily activities and biometrics. For example, it is possible to get around-the-clock measurements of one's heart rate, caloric expenditure, but also steps taken, distance and route travelled, as well as stairs climbed, to name but a few.

These technologies – meaning apps and wearables that enable keeping track of and putting numbers to one's daily life and thereby discovering habits and behaviors, thereby facilitating changes to more desirable behaviors – are termed life-logging technologies cf. [1]. The intention behind wearables and fitness-apps is to inform their user about the finished activities or exercises but also to encourage more activity.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 41–52, 2019.
https://doi.org/10.1007/978-3-319-94619-1_5

According to recommendations of the World Health Organization (WHO), adults should take a minimum of 10,000 steps a day and exercise at least 150 min a week at medium intensity to lay the groundwork for a healthy lifestyle and keep the body, and by extension the mind, in good working order [2]. Because the digitalization has also led to a lot of sedentary behavior, sitting in front of a computer or hunched over a smartphone or tablet, the idea is to get people moving again, e.g., [3, 4]. To prevent common maladies such as bad posture, obesity, cardiac problems, hypertension, and diabetes, to name but a few, the information given by fitness-apps and wearables are a possible solution to not gravely miscalculate one's own physical activity and caloric intake. Also, the apps can send reminders to encourage more movement, for example, to take at least 250 steps every hour while at work.

However, actual usage is still low, at least in Germany [5]. To understand what drives actual users but also prevents others from adopting life-logging technologies a mixed-method study was conducted to quantify possible motives and barriers.

## 2 Related Work

The study of human factors and perceived barriers to technology use is nothing new. Several models such as the Technology Acceptance Model (TAM, [6]) or the Unified Theory of Acceptance and Use of Technology (UTAUT, UTAUT2; [7, 8]) have been developed to explain the acceptance and subsequent use of new technologies. However, those models were designed with the use of information and communication technologies in the workplace in mind. Wearables and fitness-apps, on the other hand, are still fairly uncommon and predominantly used for leisure time. Therefore, additional motives and barriers might influence acceptance and usage of these personal devices.

There are a number of studies that have addressed the technical side of fitness-trackers. A main focus of those studies has been the accuracy of the devices, e.g., [9–11]. Oftentimes, the scientists used a direct comparison of commercially available trackers and medical grade equipment. The results show that the measurements can differ immensely, depending on brand and model [12, 13], but also activity [10, 12, 14]. Nevertheless, some devices are fairly accurate, cf. [15], and wearables such as wristbands or even rings are smaller and less obtrusive than smartphones.

Another prominent feature of research was that of abandonment. While subjects had interest in the technology and its potential, studies have shown that the duration of use was limited, oftentimes to less than a year [16, 17]. Arguments for the cessation of use range from a mismatch in expectation to a change in health status that no longer necessitates the use, as well as technology complexity or failure [18].

Only recently have (potential) users been surveyed as to their predominant motives but also barriers to using activity trackers. Lidynia et al. [19] have found that while there is a general interest in the data that can be tracked by small wearables, one of the main barriers is that of privacy and the violation thereof.

Privacy in the online world has been a well-researched subject so far [20]. The general consensus being that stated attitude and actual behavior diverge immensely. This means that, whilst people state to want to protect their online data and feel reluctant to share, their actual behavior on social network sites, for example, shows a completely different picture. For a more detailed explanation and overview of this phenomenon, the privacy paradox, see, for example, [21–23]. While the concern for one's data is still present, actual users of life-logging technologies have obviously weighed the pros of receiving the data against the cons of granting the manufacturers access to them and decided their benefits outweighed the possible risks. This privacy calculus, described, for example, by [20, 24, 25], offers a good starting point for understanding the rejection of life-logging technologies.

What are the perceived benefits of these apps and wearables, and perhaps more importantly, what are the perceived barriers? While these technologies offer an easy way of keeping track of one's health, or at least the adherence to a healthy lifestyle, what prevents a wider spread use? To answer these questions, the following study was conducted.

## 3 Research Design

To identify perceived barriers but also motives for the use of fitness applications and wearables, first, a focus group study was conducted. Based on the qualitative input garnered there and from a wide literature research, the quantitative study was designed and an online questionnaire was drafted.

The questionnaire first surveyed demographic data. This included age, gender, and education. Additionally, constructs such as self-efficacy when interacting with technology (*SET*, $\alpha = .858$, based on [26]), online privacy concerns (*OPC*, $\alpha = .668$), and healthy lifestyle (*HL*, $\alpha = .635$) were polled.

After a brief introduction of the most commonly available functions actual fitness applications offer, participants were asked if they already used such an app. With this, expertise groups were distinguished as experience has been shown to influence perception in other contexts, for example, [27, 28]. Next followed different sections in which participants were asked to rate statements concerning the motives, perceived value, perceived usability, but also barriers to the use of a fitness application. They did so on 6-point Likert scales ranging from 0 ("do not agree at all") to 5 ("fully agree").

The data was then analyzed using both parametric as well as non-parametric methods. The level of significance was set to .05 and calculated via two-tailed tests. For better readability, mean values are presented ± the standard deviation.

## 4 Sample

The online questionnaire was answered by 173 participants in Germany during summer in 2017. Excluding incomplete answers, a total of 166 datasets were included in the presented analysis, with 73 (44%) female and 93 (56%) male. The age within the sample ranged from 18 to 65 years of age ($M = 29.3 \pm 9.8$ years). Concerning experience with the technology, 38 (23%) participants reported the current use of a fitness application (15 men and 20 women), including 16 (9.6%) participants who use a wearable fitness tracker (8 men and women, respectively). In general, the sample felt rather self-efficient when interacting with technology (M = 3.3 ± .96, min = 0, max = 5), but a slight gender bias could be detected as men felt more competent than women (see Table 1).

Concern for online privacy was also present (M = 3.16 ± .93), with women being slightly more concerned than men. According to the reports, the sample was following a somewhat healthy lifestyle (M = 3.0 ± .93), with women more conscious of their activity level and balanced diet than men.

**Table 1.** Characteristics of the sample ($^{**}p < .01$, $^{*}p < .05$, $^{+}p < .1$). SET: Self-Efficacy when interacting with Technology, OPC: Online Privacy Concerns, HL: Healthy Lifestyle, AU: Interest in App Use, WU: Interest in Wearable Use

| | Age | Gender | SET | OPC | HL | AU | WU |
|---|---|---|---|---|---|---|---|
| Age | — | | | $.299^{**}$ | | $-.296^{**}$ | $-.239^{**}$ |
| Gender | | — | $.488^{**}$ | $.184^{*}$ | $-.347^{**}$ | | $-.158^{*}$ |
| SET | | | — | $.287^{**}$ | $-.329^{**}$ | | $-.67^{+}$ |
| OPC | | | | — | $-.264^{**}$ | $-.417^{**}$ | $-.422^{**}$ |
| HL | | | | | — | $.405^{**}$ | $.452^{**}$ |
| AU | | | | | | — | $.753^{**}$ |
| WU | | | | | | | — |

# 5 Results

The following section will present the results of our study. First, the evaluation of motives for the use of fitness apps will be presented. Next, the perceived barriers to the use will be introduced in more detail. Both evaluations will be differentiated by expertise group. Finally, barriers and motives for the addition of a wearable will be presented, also with a closer view on usage experience.

## 5.1 Motives for the Use of Fitness Applications

Participants were asked to evaluate seven possible motives that have been derived from previous focus group studies. Experience or expertise as it may be, does distinguish the evaluation of the items significantly in all but one case, see Fig. 1. That is, users of fitness-apps agree to all motives but that of prevention of illnesses more than the non-users do. For all t-tests, the absolute value of Cohen's d exceeds .8. Still, current non-users evaluated the majority of motives for fitness app use rather neutrally, rating on average close to the center of the scale. Even the usefulness of a fitness application for their specific needs, which has the largest discrepancy in evaluation between the groups, achieved a mean agreement of M = 2.11 ± 1.25 by current non-users, $t_{usefull}(164) = 7.736$, $p < .001$, d = 1.472. Active users, on the other hand, perceive all motives but the prevention of illnesses to be very prevalent, especially the improvement or maintenance of well-being and fitness ($M_{user} = 4.26 \pm .82$). This is the most important motive also for non-users ($M_{non\text{-}users} = 3.1 \pm 1.33$), ($t_{well\text{-}being}(88) = 6.418$, $p < .001$, d = .933). The prevention or treatment of an illness is rejected as valid motivator by both groups ($M_{users} = 2 \pm 1.11$; $M_{non\text{-}users} = 1.73 \pm 1.2$), ($t_{illness}(164) = 1.187$, n.s.).

Unsurprisingly, also the general interest in the use of a fitness app is much more pronounced in active users (M = 4.0 ± 1.14) than it is in non-users (M = 2.1 ± 1.25), $t_{interest}(58) = 8.603$, $p < .001$, d = 1.39. However, this non-interest is not overly marked.

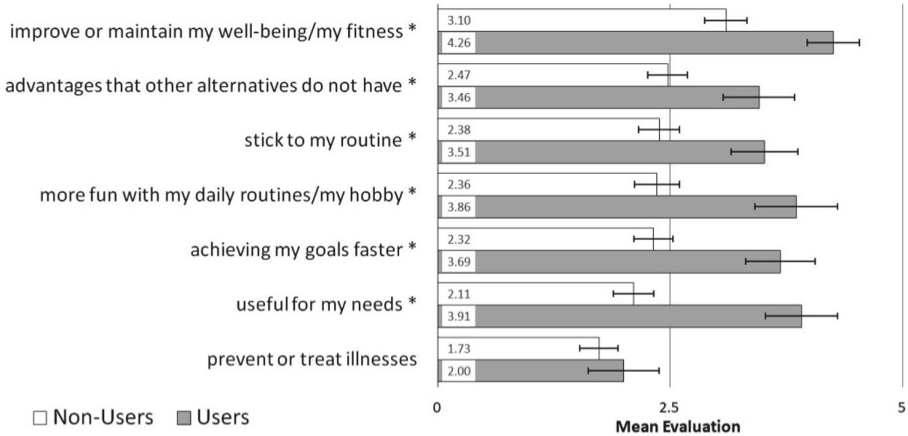



**Fig. 1.** Mean agreement to motives for the use of fitness applications, distinguished by experience ($n_{user} = 35$; $n_{non\text{-}user} = 131$) (0 = "do not agree at all," 5 = "fully agree"), with 95% confidence intervals. An * indicates significant differences between the groups, p < .05.

## 5.2 Barriers to the Use of Fitness Applications

The participants were also asked to indicate their agreement to a total of 12 possible concerns that could arise when using fitness applications. The results show significant differences between users and non-users of fitness apps in 9 out of 12 statements (for $p < .01$, the absolute value of Cohen's d > .7 for the t-tests, while Cohen's d > .5 in case of $p < .05$). That is, most barriers are of more concern to non-users than to users. A complete comparison of the user groups is presented in Table 2. The data is presented in ascending order from least relevant barrier to most relevant barrier to the use of fitness applications from the perspective of current non-users.

The most important barrier for non-users is that they lose control over their data ($t_{control\text{-}loss}(164) = -4.615$; $p < .001$; d = .878) whereas the greatest barrier for active users of fitness apps is the unspecificity of the app's content ($t_{too\text{-}generalized}(164) = .247$, n.s.). This is also the only barrier that is more important to users than non-users. All other possible barriers are of more concern to non-users.

Although both groups rank a possible negative impact on their health due to erroneous data or device malfunctions as least valid concern, they do differ significantly in the markedness of this concern ($t_{heath\_impact}(84) = -2.969$, $p < .01$, d = .565).

Within the most important barriers, many relate to privacy and data security concerns. Still, this sample indicated that the sensitivity of the collected data was not the biggest concern, with non-users only nearing the arithmetical neutral level of

**Table 2.** Average agreement of users and non-users on evaluation statements (0 = "do not agree at all" and 5 = "fully agree") and related standard deviations according to experience group

| I am concerned… | Non-users M ± SD | Users M ± SD |
|---|---|---|
| …errors within the app or the device operating the app can have a negative effect on my health* | 1.2 ± 1.17 | .71 ± .75 |
| …I will become dependent on the technology | 1.21 ± 1.17 | .94 ± 1.11 |
| …the collected data is too sensitive* | 2.12 ± 1.4 | 1.31 ± 1.28 |
| …the information and data gathered within the app will be erroneous | 2.5 ± 1.12 | 2.2 ± 1.16 |
| …that the app will keep reminding me (negatively) to do more for my health/fitness* | 2.63 ± 1.55 | 1.46 ± 1.29 |
| …the content of the app is too generalized to apply to me | 2.73 ± 1.02 | 2.77 ± .84 |
| …using the app will be time-consuming** | 2.87 ± 1.26 | 1.63 ± .97 |
| …the app will forward my personal data to third parties** | 3.22 ± 1.21 | 2.34 ± 1.26 |
| …that my data will be misused* | 3.24 ± 1.42 | 2.2 ± 1.23 |
| … that the app will become annoying after a while* | 3.35 ± 1.27 | 2.17 ± 1.1 |
| …the app will interfere too much with my privacy (e.g., via unauthorized sharing or frequent feedback.)** | 3.38 ± 1.15 | 2.49 ± 1.12 |
| …that I have no control over what will happen to my data* | 3.6 ± 1.23 | 2.51 ± 1.27 |

An * indicates significant differences between the user groups with p < .05; ** indicate significant differences with p < .01.

agreement ($t_{sensitivity}(164) = 33.09$, $p < .01$, d = .588). Users are more concerned than non-users about the possible violation of one's privacy, be it by unauthorized sharing and frequent feedback ($t_{feedback}(164) = -4.106$), $p < .001$, d = .781), possible data misuse ($t_{misuse}(164) = -3.929$, $p < .001$, d = .748), or third party acquisition ($t_{third\text{-}party}(164) = -3.829$, $p < .001$, d = .729).

No differences between experience groups could be found for the evaluation of possible dependency on technology ($t_{dependency}(164) = -2.231$, n.s.) and the concern of erroneous data ($t_{erroneous}(164) = -1.488$, n.s.). However, the groups differ significantly in their evaluation of time expenditure for the use of a fitness application ($t_{time}(164) = -5.406$, p < .001, d = −1.029) as well as becoming annoyed after prolonged usage ($t_{annoyance}(164) = -5.013$, p < .001, d = .954). Again, the group with higher concerns here are the non-users. People who are active users do not feel as bothered.

### 5.3 Evaluation of Supplementing a Fitness-App with a Wearable

For the analysis of the evaluation of supplementing a fitness app with a wearable, user groups were newly divided into non-users (n = 123) – who do not use a wearable nor a fitness app – app users (n = 22) – who use a fitness app without a wearable – and wearable users (n = 16). To examine differences in the evaluation between these expertise groups, analysis of variance (ANOVA) procedures were calculated using

Gabriel's and Games-Howell post-hoc tests as group sizes were unequal. Games-Howell is reported when homogeneity of variance could not be assumed. To assess the effect size, eta square $\eta^2$ was calculated.

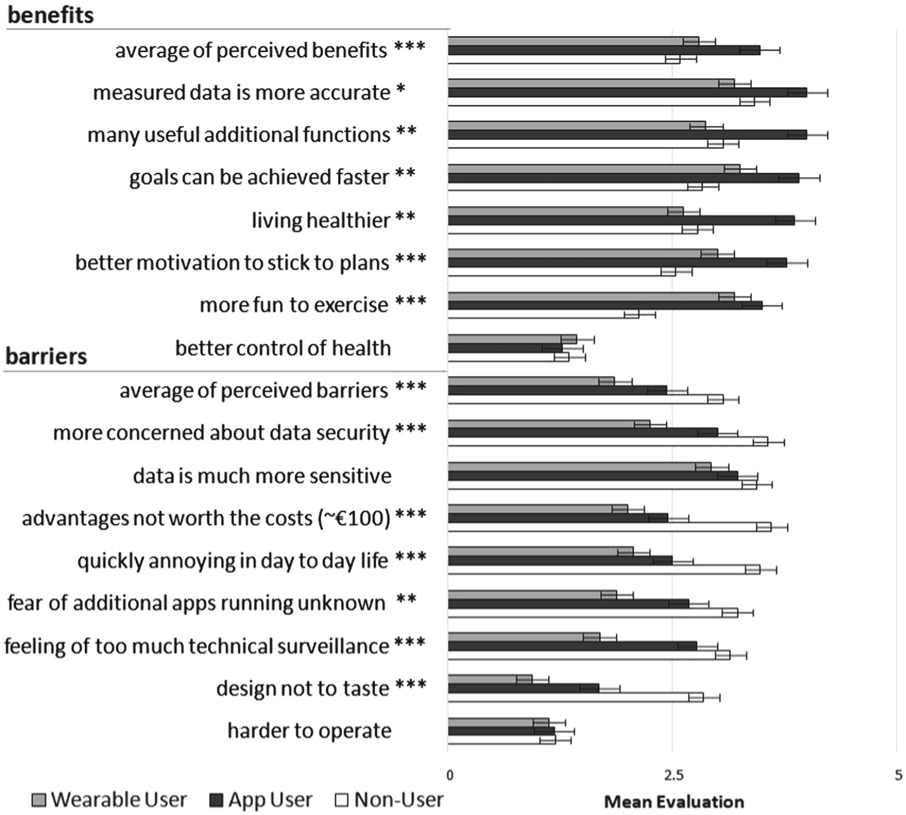



**Fig. 2.** Mean agreement to motives and barriers to the use of a wearable in addition to a fitness application, distinguished by experience (do not agree at all = 0, agree completely = 5), including SEM. The asterisks indicate significant differences between the groups, * = p < .05, ** = p < .01, *** = p < .001 (non-users n = 123, app users n = 22, wearable user n = 16).

The participants evaluated seven benefits of adding a wearable compared to the use of a fitness app without wearable. The benefits included aspects of more accurate data, additional functions, and increased support of health and fitness goals. On average, the benefits were slightly agreed on by all groups (M = 2.73 ± .99, cf. Fig. 2). The expertise groups rated the benefits significantly differently ($F(2, 158) = 8.174, p < .001, \eta^2 = .094$). App users showed the highest agreement to the benefits on average (M = 3.47 ± 0.67) and post-hoc tests revealed this evaluation to differ significantly from non-users ($p < .001$). The evaluation of non-users (M = 22.59 ± .91) and wearable users (M = 2.79 ± 1.48) did not differ. Most important benefits concerned more accurate data (M = 3.47 ± 1.11) and additional functions (M = 3.17 ± 1.3). Better control of health and less dependency on medical assistance is rejected on average (M = 1.35 ± 1.22).

Moreover, nine disadvantages of wearables were assessed that regarded data security and privacy, annoyance with the wearable, design, costs, and usability. These were also averagely agreed on (M = 2.86 ± .95) with large differences between expertise groups ($F(2, 158) = 16.642$, $p < .001$, $\eta^2 = .174$, cf. Fig. 2). As could be expected, wearable users rejected the barriers (M = 1.86 ± .77), and in contrast, non-users agreed to most barriers (M = 3.06 ± .89) with a highly significant difference between these perceptions ($p < .001$). App users are quite neutral in their evaluation of the barriers (M = 2.44 ± .77) and differ significantly from non-users ($p < .01$).

Barriers regarding security of data (M = 3.36 ± 1.37), data sensitivity (M = 3.36 ± 1.12), and costs (M = 3.29 ± 1.45) are the most important ones. Least important and rejected are barriers concerning design (M = 3.36 ± 1.12) and usability (M = 1.19 ± 1.23).

Overall, the evaluation patterns of the different expertise groups regarding barriers and benefits, respectively, remain the same: Non-users have higher concerns than app-users and wearable users reject the concerns the most. The benefits of an additional wearable find the most agreement by app users, followed by wearable users. Of the three groups, non-users are the least convinced of the positive effects of a wearable. With the exception of control of health, data sensitivity, and usability, the expertise groups differ significantly in their evaluations.

## 5.4 Influencing Factors on the Interest to Use Fitness App and Wearable

To understand which user factors influence the interest in using a fitness app or a wearable, stepwise multiple linear regressions were performed. Independent variables were age, gender, self-efficacy in interacting with technology (SET), online privacy concerns (OPC), and healthy lifestyle (HL). In both cases, self-efficacy when dealing with technology showed no influence, see Fig. 3. The regression model for the interest in using a fitness app contained 4 factors and explained 36% of the variance. The addition of a wearable included only HL and OPC as relevant factors to explain 30% of variance in the willingness to use a wearable. In both cases, a healthy lifestyle is the strongest predictor for interest.

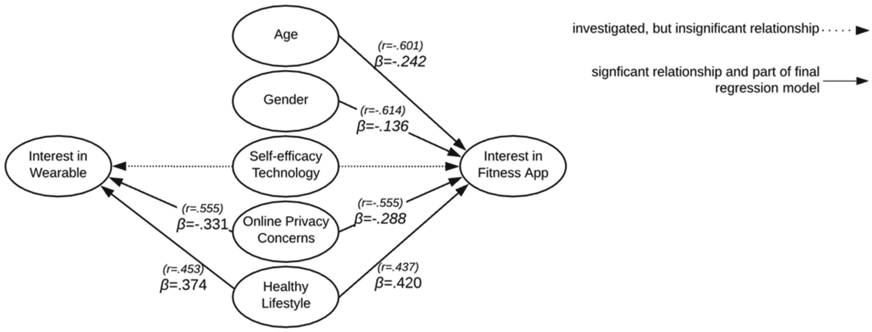



**Fig. 3.** Determinants for the interest to use a wearable ($r^2 = .309$) or a fitness app ($r^2 = .362$).

## 6 Discussion

The digitalization marches forward, offering more and more opportunities and possibilities of new technical devices, be it for work or leisure time. But the use of these devices oftentimes necessitates sedentary behavior and results in hunched over postures, thereby limiting physical activity tremendously, e.g., [3]. At the same time, new devices and applications for mobile devices are introduced that are meant to facilitate more activity, for example by tracking the status quo and also sending reminders to walk a few more step each hour. The benefits of physical activity have been proven numerous times, see, for example, [29, 30]. Nevertheless, the actual use of life-logging technologies and activity trackers is still low. The aim of this study was to identify relevant motives and barriers to the use of these technologies.

Whereas actual users, unsurprisingly, agree with most motives for the use, be it the improvement of health, more fun with daily routines, or advantages other alternatives do not offer, non-users are not convinced that those motives can be delivered by using a so-called fitness app. The only motive clearly agreed upon by current non-users is maintenance or improvement of the current fitness-level or well-being. On other possible motives, current non-users are rather undecided, oftentimes rating near the neutral center of the scale. This raises the question of what came first? The perception that the benefits predominate and therefore one decides to use a fitness app, or does the use of the fitness app result in the benefits being perceived as more prevalent? As this is only a snapshot of the status quo, this question cannot be answered, but a longitudinal study to reveal insights into these temporal developments could uncover relevant insights into the study of life-logging acceptance.

Adding another device, namely a wearable, into the equation, it is interesting to note that possible benefits are not rated best by those who have experience with similar devices. It is the app users that see the most benefits and believe the measured data to be more accurate and therefore using a wearable would be easier and more effective to reach goals or live healthier. In contrast, the barriers of a wearable in addition to the fitness app are rejected the most by actual users. Perhaps, current non-users would reject barriers also more after having tried such a device for some time. This again is the question of causality that should be addressed in future studies.

Based on previous findings, the sensitivity of the collected data was one of the most important barriers to the use of fitness applications and wearables [19, 31]. In the present study, the sensitivity of the data was not as important as the perceived loss of control over the data or the unauthorized forwarding to third parties. This again illustrates that one of the main factors preventing the use of life-logging applications is that of privacy concerns. The uncertainty of what happens to the data and who will gain access to it is a large impediment to using the benefits and opportunities to their fullest.

Furthermore, both users and non-users reject the prevention or treatment of illnesses as a valid motive for the use of fitness applications and wearables. This shows that people are not convinced that the use of apps or wearables can help prevent illnesses or deceases. Here, better communication strategies of proven benefits of more physical activity, but also the potential of using apps or wearables as reminders need to be further developed.

Yet another valuable insight of the present study is that another obstacle to using life-logging technologies seems the fear of constant reminders to be more active and with it the misgiving of the device becoming an annoyance instead of a helpful tool. The use of life-logging technologies and fitness trackers as an active motivator instead of a passive recording device therefore should be given more attention in future studies.

## 7 Limitations and Outlook

This study provided valuable insights into the motives and barriers perceived by users and non-users to the adaptation of fitness apps and wearables and particularly into the differences between these expertise groups. Unfortunately, the sample size of actual users of life-logging devices including wearables was rather small. Additionally, the mean age was roughly 30 years of age, therefore surveying only a young sample. In future studies, a larger and more diverse sample should be addressed to better understand actual users and also a wider variety of potential users.

Furthermore, the present survey used a fictitious example of an app and wearable. While based on commercially available examples, the answers of the participants were mostly based on guesses as, with the large group of non-users, they lacked any real insight into the topic. Therefore, it would be helpful to get the insight of more people who have at least tried out different fitness apps or are still using them. Another possibility would be to include a short mock-up of a fitness app that participants could interact with to better understand concerns they have or introduce benefits they might have previously not been privy to.

While a questionnaire can gather a lot of information, it might also be helpful to use different methods, such as conjoint analysis, to find out which benefits and barriers are indeed the most beneficial or detrimental to the adoption of life-logging technologies.

**Acknowledgements.** We thank all participants of the focus group and the survey for their willingness to share their thoughts and feelings about persuasive technologies, fitness trackers, and their privacy concerns. We also thank Niklas Kunstleben for his research support. Parts of this work have been funded by the German Ministry of Education and Research (BMBF) under project “myneData” (KIS1DSD045).

## References

1. Lupton, D.: Self-tracking cultures: towards a sociology of personal informatics (2014)
2. World Health Organization: Global Recommendations on Physical Activity for Health. WHO Press, Geneva (2010)
3. Hadgraft, N.T., Lynch, B.M., Clark, B.K., Healy, G.N., Owen, N., Dunstan, D.W.: Excessive sitting at work and at home: correlates of occupational sitting and TV viewing time in working adults. BMC Public Health **15**, 899 (2015)
4. de Rezende, L.F.M., Lopes, M.R., Rey-López, J.P., Matsudo, V.K.R., Do Carmo Luiz, O.: Sedentary behavior and health outcomes: an overview of systematic reviews. PLoS ONE **9**, e105620 (2014)

5. Ernsting, C., Dombrowski, S.U., Oedekoven, M., O'Sullivan, J.L., Kanzler, M., Kuhlmey, A., Gellert, P.: Using smartphones and health apps to change and manage health behaviors: a population-based survey. J. Med. Internet Res. **19**, e101 (2017)
6. Davis, F.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**, 319–340 (1989)
7. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **27**, 425–478 (2003)
8. Venkatesh, V., Walton, S.M., Thong, J.Y.L., Xu, X.: Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q. **36**, 157–178 (2012)
9. Kaewkannate, K., Kim, S.: A comparison of wearable fitness devices. BMC Public Health **16**, 433 (2016)
10. Sasaki, J.E., Hickey, A., Mavilia, M., Tedesco, J., John, D., Keadle, S.K., Freedson, P.S.: Validation of the Fitbit wireless activity Tracker(R) for prediction of energy expenditure. J. Phys. Act. Health **12**, 149–154 (2015)
11. An, H.-S., Jones, G.C., Kang, S.-K., Welk, G.J., Lee, J.-M.: How valid are wearable physical activity trackers for measuring steps? Eur. J. Sport Sci. **17**, 360–368 (2017)
12. Storm, F.A., Heller, B.W., Mazzà, C.: Step detection and activity recognition accuracy of seven physical activity monitors. PLoS ONE **10**, e0118723 (2015)
13. O'Connell, S., Ólaighin, G., Kelly, L., Murphy, E., Beirne, S., Burke, N., Kilgannon, O., Quinlan, L.R.: These shoes are made for walking: sensitivity performance evaluation of commercial activity monitors under the expected conditions and circumstances required to achieve the international daily step goal of 10,000 steps. PLoS ONE **11**, e0154956 (2016)
14. Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A.M., Houghtaling, P., Javadikasgari, H., Desai, M.Y.: Variable accuracy of wearable heart rate monitors during aerobic exercise. Med. Sci. Sports Exerc. **49**, 1697–1703 (2017)
15. Case, M.A., Burwick, H.A., Volpp, K.G., Patel, M.S.: Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA **313**, 625–626 (2015)
16. Meyer, J., Schnauber, J., Heuten, W., Wienbergen, H., Hambrecht, R., Appelrath, H.-J., Boll, S.: Exploring longitudinal use of activity trackers. In: Proceedings of IEEE ICHI - International Conference Healthcare Informatics, pp. 198–206 (2016)
17. Ledger, D., McCaffrey, D.: Inside wearables: how the science of human behavior change offers the secret to long-term engagement. Endeavor Partn. **200**, 1–17 (2014)
18. Clawson, J., Pater, J.A., Miller, A.D., Mynatt, E.D., Mamykina, L.: No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In: UbiComp 2015, pp. 647–658 (2015)
19. Lidynia, C., Brauner, P., Ziefle, M.: A step in the right direction – understanding privacy concerns and perceived sensitivity of fitness trackers. In: Ahram, T., Falcão, C. (eds.) Advances in Human Factors in Wearable Technologies and Game Design. Advances in Intelligent Systems and Computing, vol. 608, pp. 42–53. Springer, Cham (2018)
20. Bélanger, F., Crossler, R.E.: Privacy in the digital age: a review of information privacy research in information systems. MIS Q. **35**, 1017–1041 (2011)
21. Norberg, P.A., Horne, D.R., Horne, D.A.: The privacy paradox: personal information disclosure intentions versus behaviors. J. Consum. Aff. **41**, 100–126 (2007)
22. Hallam, C., Zanella, G.: Online self-disclosure: the privacy paradox explained as a temporally discounted balance between concerns and rewards. Comput. Hum. Behav. **68**, 217–227 (2017)
23. Nurse, J.R.C., Williams, M., Nurse, J.R.C., Creese, S.: The perfect storm: the privacy paradox and the Internet-of-Things. In: International Conference on Availability, Reliability and Security (2016)

24. Dinev, T., Hart, P.: An extended privacy calculus model for e-commerce transactions. Inf. Syst. Res. **17**, 61–80 (2006)
25. Trepte, S., Reinecke, L., Ellison, N.B., Quiring, O., Yao, M.Z., Ziegele, M.: A cross-cultural perspective on the privacy calculus. Soc. Med. + Soc. **3**, 1–13 (2017)
26. Beier, G.: Kontrollüberzeugungen im Umgang mit Technik [Locus of control when interacting with technology]. Rep. Psychol. **24**, 684–693 (1999)
27. Schomakers, E.-M., Lidynia, C., Vervier, L., Ziefle, M.: Of guardians, cynics, and pragmatists – a typology of privacy concerns and behavior. In: 3rd International Conference on Internet of Things, Big Data and Security. SCITEPRESS - Science and Technology Publications (2018)
28. Lidynia, C., Philipsen, R., Ziefle, M.: Droning on about drones - acceptance of and perceived barriers to drones in civil usage contexts. In: Savage-Knepshield, P., Chen, J. (eds.) Advances in Human Factors in Robots and Unmanned Systems, pp. 317–329. Springer, Heidelberg (2017)
29. Warburton, D.E.R., Nicol, C.W., Bredin, S.S.D.: Health benefits of physical activity: the evidence. CMAJ **174**, 801–809 (2006)
30. Lee, I.-M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., Katzmarzyk, P.T.: Impact of physical inactivity on the world's major non-communicable diseases. Lancet **380**, 219–229 (2012)
31. Li, H., Wu, J., Gao, Y., Shi, Y.: Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. Int. J. Med. Inform. **88**, 8–17 (2016)

# Exploring the Acceptance of mHealth Applications - Do Acceptance Patterns Vary Depending on Context?

Eva-Maria Schomakers(✉), Chantal Lidynia, and Martina Ziefle

Human-Computer Interaction Center, RWTH Aachen University,
Campus Boulevard 57, 52074 Aachen, Germany
{schomakers,lidynia,ziefle}@comm.rwth-aachen.de

**Abstract.** In the present study, we investigate influencing factors on the acceptance of mHealth smartphone apps, using an extended UTAUT model. N = 165 participants evaluated use intention, performance expectancy (PE), effort expectancy (EE), social influence (SI), facilitating conditions (FC), as well as privacy concerns for a fitness app (lifestyle context) and a diabetes app (medical context). Structural equation modeling is used to assess the relevance of influences on adoption intention in these contexts. Results show that acceptance factors indeed differ strongly between lifestyle and medical contexts. For the latter, only PE and SI determine intention to use, although privacy concerns are higher than in the lifestyle context. In contrast, intention to use the fitness app is predicted by PE, SI, FA, and privacy concerns. The extended UTAUT model showed very good predictive relevance for use intention in both contexts. These findings reveal that technology acceptance needs to be examined depending on context.

**Keywords:** Technology acceptance · mHealth · Life-logging · UTAUT
Privacy concerns · Human factors

## 1 Introduction

The healthcare systems of Western countries are under pressure by the demographic developments that result in costs deficits and health personnel shortages, among other things. Additionally, due to changing lifestyles, reduced physical activity, and increased obesity chronic diseases occur more frequently, e.g., diabetes mellitus [1]. At the same time, the digitization is marching on. Mobile technologies become more prevalent. For example, in Germany, 95% of the population own a smartphone [2]. It seems inevitable that a device that is almost always and everywhere present will also be used to track, maintain, or improve one's health. App stores offer a wide variety of applications (apps) to keep an eye on one's caloric intake, energy expenditure, steps and distances walked, to name but a few [3]. Additionally, apps have the potential to support chronically ill patients in their therapy and self-management (e.g., [4, 5]). These apps are part of the area of mobile Health Technologies (mHealth).

One prominent example for disease specific apps are diabetes apps that support patients in self-management tasks [6]. In 2010, 12% of the German population suffered


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 53–64, 2019.
https://doi.org/10.1007/978-3-319-94619-1_6

from diabetes, which is projected to rise to 13.5% in 2030 [7]. These patients need to self-monitor their dietary intake, medication, and glucose level. Studies have shown that diabetes apps have a positive impact on diabetes self-management and therapy adherence [4, 8].

Furthermore, with the ever-increasing amount of sedentary behavior, be it at work or at home, the need for more activity to enhance quality of life, especially in later age, is not really in question [9]. The World Health Organization (WHO) recommends that a healthy adult should at least walk 10,000 steps a day and participate in 150 min of moderately intense workouts per week [10]. Keeping track of that and thus supporting an active lifestyle is made infinitely easier by smartphone apps and additional tools, such as pedometers, heart-rate monitors, and activity trackers.

However, current surveys have revealed that these tools are not as frequently used as they could be. In fact, in Germany, for instance, only a third of the population takes advantage of fitness or mHealth apps and devices [11]. The question then arises, why? What are factors that prevent people from making use of the opportunities of mHealth applications? The present study aims at a closer look to find the relevant drivers behind acceptance or rejection of mHealth applications and to compare acceptance patterns in two areas of mHealth, a fitness app compared to a diabetes app. To do so, we draw on the validated *Unified Theory of Acceptance and Use of Technology* (UTAUT) and include privacy concerns into the model, as these represent an important barrier to technology use in other areas (e.g., [12, 13]). In the next section, we will introduce the theoretical background of our study in detail. After that, we will present our research model and methodological approach.

## 2 Theoretical Background

First, the basis for the current study will be laid by giving a short overview of technology acceptance and privacy research in general. Next, the framework of previous acceptance studies on health and mHealth technologies will be briefly outlined.

### 2.1 Technology Acceptance

The research field of technology acceptance seeks to understand the factors that determine users' intentions to use technology and the actual use behavior. Multiple technology acceptance models have been proposed in the last decades. Most influential and widely used is the *Technology Acceptance Model* (TAM, [14, 15]. It proposes that behavioral intentions to use ICTs are determined by the *perceived usefulness* and the *perceived ease of use*. Despite the fact that the TAM has been successfully applied in many areas including medical contexts (e.g., [16, 17]), the theory has been revised and extended. Venkatesh et al. integrated eight competing acceptance models into the *Unified Theory of Acceptance and Use of Technology* (UTAUT) [18]. The UTAUT model assumes behavioral intentions to be determined by *performance expectancy, effort expectancy,* and *social influence*, and behavioral intentions and *facilitating conditions* to predict actual behavior.

In many studies applying the UTAUT model, performance expectancy – the usefulness of a technology – is the strongest predictor of use intentions (e.g., [18, 19]). Effort expectancy describes the perceived effort of using the technology and is derived from the TAM variable ease of use [18]. Social influence is defined as "the degree to which an individual perceives that important others believe he or she would use the new system" ([18] p. 451). Facilitating conditions describe conditions in the environment that facilitate using the technology, e.g., assistance with problems [18]. The UTAUT model and its extended versions have been successfully applied to different contexts (e.g., [20, 21]) and also in the context of medical technologies and mHealth (e.g., [19, 22, 23]).

### 2.2 Privacy Concerns

Privacy concerns are debated strongly in regard to ICT and emerging technologies as users' concerns are generally high and depict a barrier against the adoption of technologies and services [13]. Privacy can be defined as the right to control information about oneself [24]. Traditionally, informational privacy could be distinguished from physical, psychological, and interactional privacy [25]. With the advent of connected technologies, the Internet of Things, and data collection in more and more areas of life, information privacy has become another aspect within other types of privacy [26]. That is, as data is collected about conversations, location, and medical characteristics and activities, information privacy is key to protect all of these areas of life. Correspondingly, it is not surprising that privacy concerns are increasing [27].

The type of information collected and its perceived sensitivity affects privacy concerns (e.g., [12, 28]). Particularly, medical information is perceived as highly sensitive (e.g., [29, 30]). Previous research showed that the formation of privacy concerns differs between contexts and that users perceive the highest level of risk for a healthcare context [31]. Moreover, privacy concerns show increased influence on behavioral intentions in a medical context [32]. At the same time, most health apps show poor information privacy practices [33]. This suggests that users could be even more concerned about their privacy when using a mHealth app that collects sensitive medical data.

### 2.3 Related Works in mHealth Adoption and Acceptance

The adoption of mHealth technologies is still a fairly young subject in the research arena. Regarding other health technologies, e.g., health-related Internet sites, perceived usefulness and perceived ease of use have been found to have a positive relationship to the intention to use and actual use behavior [34]. Lidynia et al. could show that a general interest in the monitoring of one's own body and caloric intake exists [35]. They also worked out that a very important factor of not using these technologies was the concern for sensitive data and the unwillingness to share data.

Studies concerning mHealth acceptance are fewer. Hoque and Sorwar analyzed mHealth adoption in Bangladesh and found performance expectancy, effort expectancy, and social influences to be important antecedents of behavioral intentions to use mHealth [22]. In contrast, Boontarig et al. found that elderly's intention to use electronic health (eHealth) services is significantly influenced by facilitating conditions and effort expectancy besides perceived value, but not by performance expectancy and social influence [36]. Sun et al. found all constructs of the UTAUT model as well as threat appraisals to be significant, with performance expectancy showing the strongest influence [19]. In a qualitative study, Parker et al. identified privacy concerns as a barrier to mHealth use by older adults [37], but Guo et al. found privacy concerns not to influence adoption intentions [38]. These deviating results regarding mHealth acceptance factors could be explained by the different examples for mHealth that the studies applied – besides cultural and sample differences. As the field of mHealth is broad, acceptance patterns may differ between applications.

## 3 Method

In this study, we compare the acceptance patterns of potential mHealth users between two contexts, a lifestyle context and a medical context. Using and extending the validated UTAUT model, we analyze which variables are relevant for the intention to use a fitness and a diabetes app, and how the perception of these variables by potential users differ. A questionnaire was distributed online, asking all participants to evaluate the fitness and the diabetes app in a repeated-measures design. To control serial position and carry-over-effects, the order of the apps and their evaluation was varied.

### 3.1 The Research Model

Our research model is illustrated in Fig. 1. We hypothesize that the central constructs of the UTAUT model as well as privacy concerns predict the intention to use mHealth in both contexts, but that they differ in their importance between the contexts.

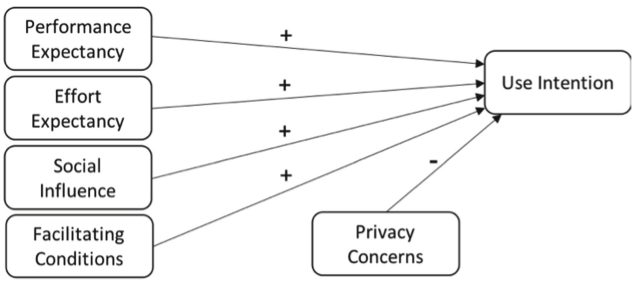



**Fig. 1.** Research model for both contexts.

### 3.2 The Questionnaire

The questionnaire consisted of three parts. The first part assessed demographic data: age, gender, education level as well as occupation. In the second part, the participants answered questions to their subjective health status and knowledge about mHealth applications. In the third part, the participants were introduced to a hypothetical fitness application as well as a diabetes app and then evaluated these. The description of the applications included their functions, pictures, and, in the case of diabetes app, a short explanation about diabetes. The participants were questioned, whether they use a similar application. In the following evaluation of the applications, the formulation of the items were adapted to either user or non-user. Non-diabetes participants were asked to put themselves in the scenario of suffering from diabetes to evaluate the diabetes app.

For the evaluation of the applications, the constructs *performance expectancy, effort expectancy, social influence,* and *facilitating conditions* based on the UTAUT model [18] were assessed. The original items were translated to German and were adapted to the context of fitness and diabetes app, were necessary. Additionally, privacy concerns regarding the applications were assessed with five items developed specifically (e.g., *"I am concerned, that the application violates my privacy.",* and *"I am concerned that I cannot control what happens to my data."*). All constructs were measured on a 6-point Likert scale ranging from "I do not agree at all" (1) to "I fully agree" (6).

### 3.3 The Statistical Analysis

To test context differences in the evaluation of the tested constructs, ANOVA procedures were used. For reliability analysis of the scales, Cronbach's Alpha was calculated. All scales exceeded $\alpha > .7$ (cf. Tables 1 and 2). Significance level was set at 5% for all analyses.

Partial Least Squares (PLS), a component-based structural equation modeling was used to test the research model using the software smartPLS [39]. It is a second-generation multivariate method that can be used to test and validate models and the relationships between constructs. In comparison to covariance-based structural equation modeling, PLS is more suitable with relatively small sample sizes and has no strict requirements regarding the data distribution [40]. The reliability and validity of the scales, the so-called measurement model, is evaluated during the process. The evaluation results are described in more detail in the next section.

**Table 1.** Reliability and validity measures and correlations of the constructs of the fitness app model.

| | CR | CA | AVE | PE | EE | SI | FA | PC | UI |
|---|---|---|---|---|---|---|---|---|---|
| PE | .931 | .900 | .772 | **.879** | | | | | |
| EE | .877 | .786 | .707 | .230 | **.841** | | | | |
| SI | .867 | .769 | .685 | .593 | .179 | **.828** | | | |
| FA | .849 | .732 | .654 | .145 | .713 | .195 | **.809** | | |
| PC | .920 | .891 | .699 | −.317 | .025 | −.301 | .070 | **.836** | |
| UI | .926 | .880 | .806 | .751 | .248 | .598 | .277 | −.363 | **.898** |

Bolded diagonal elements denote the square root of AVE.
CR = Composite Reliability, CA = Cronbach's Alpha, AVE = Average Variance Extracted, PE = Performance Expectancy, EE = Effort Expectancy, SI = Social Influence, FA = Facilitating Conditions, PC = Privacy Concerns, UI = Use Intention.

### 3.4 Evaluation of the Measurement Model

The measurement models for the two contexts were evaluated individually. All constructs were measured reflectively. To assess adequate quality of the measurement model, reliability, convergent validity, and discriminant validity were confirmed following [40]. Composite Reliability (CR) were used to assess the internal consistency reliability and Average Variance Extracted (AVE) to assess the convergent validity of the constructs. All constructs reached adequate values of CR < .7 and AVE < .5 in both contexts (cf. Tables 1 and 2). To confirm adequate indicator reliability for each scale, items loadings were checked to be higher than > .708. The conservative Fornell-Larcker criterium was used to assess discriminant validity of the constructs.

**Table 2.** Reliability and validity measures and correlations of the constructs of the diabetes app model.

| | CR | CA | AVE | PE | EE | SI | FA | PC | UI |
|---|---|---|---|---|---|---|---|---|---|
| PE | .917 | .880 | .735 | **.857** | | | | | |
| EE | .918 | .865 | .790 | .471 | **.889** | | | | |
| SI | .914 | .858 | .779 | .561 | .191 | **.883** | | | |
| FA | .820 | .717 | .605 | .431 | .686 | .269 | **.778** | | |
| PC | .849 | .902 | .597 | .210 | .181 | .135 | .266 | **.773** | |
| UI | .855 | .747 | .664 | .769 | .481 | .644 | .470 | .161 | **.815** |

Bolded diagonal elements denote the square root of AVE. CR = Composite Reliability, CA = Cronbach's Alpha, AVE = Average Variance Extracted, PE = Performance Expectancy, EE = Effort Expectancy, SI = Social Influence, FA = Facilitating Conditions, PC = Privacy Concerns, UI = Use Intention.

### 3.5 The Sample

In total, 165 German internet users between the age of 18 and 65 years ($M$ = 29.31, $SD$ = 9.83) completed the online questionnaire. Gender distribution was quite balanced (44.2% women). The educational level of the sample was high on average, with 61% having completed a Bachelor or Master degree. 90.9% of the participants perceived their health status to be rather good, good or very good, with no participant reporting a very bad health status ($M$ = 4.58, $SD$ = .88, $min$ = 1, $max$ = 6).

One participant indicated to be already using a diabetes app. About one fifth of the participants (21.2%) used fitness applications comparable to the one presented. Questioned about their knowledge about mHealth and fitness applications, 51.5% reported to know (rather) little until that survey ($M$ = 3.27, $SD$ = 1.21, $min$ = 1, $max$ = 6).

# 4 Results

In the following part, we will first examine the absolute evaluation of the constructs and their differences between the two contexts. Thereafter, we compare the results of the structural equation models, studying how the influences on use intention differ between the contexts.

## 4.1 Differences in the Evaluation of Fitness App vs. Diabetes App

In Fig. 2, the differences in absolute evaluation of the constructs in comparison between the contexts are illustrated. The use intention in the case of the diabetes app differs significantly from the context of the fitness app ($F(1, 164) = 38.93, p < .001$). Potential users are more willing to use a diabetes app in case they suffer from diabetes ($M = 4.13$, $SD = 0.93$) than to use a fitness app ($M = 3.60$, $SD = 1.11$). In the evaluation of effort expectancy, the two contexts do not differ. But the diabetes app is perceived as being more useful (cf. performance expectancy, $F(1, 164) = 218.16, p < .001$). As well, social influences to use the diabetes app are considered to be moderately high in case of the diabetes app ($M = 4.46$, $SD = 1.13$), whereas it is slightly rejected in the case of the fitness app ($M = 3.14$, $SD = 1.12$) ($F(1, 164) = 218.69$, $p < .001$). In contrast, facilitating conditions to use the app are perceived to be quite high in the case of the fitness app ($M = 5.05$, $SD = 0.94$) and lower in the case of the diabetes app ($M = 4.51$, $SD = 0.87$), ($F(1, 164) = 100.32$, $p < .001$). Moderate privacy concerns exit in both contexts, but are more developed in the case of fitness app ($M_{FA} = 4.31$, $SD_{FA} = 1.07$ vs. $M_{DA}= 3.87$, $SD_{DA} = 1.1$), ($F(1, 164) = 56.14$, $p < .001$).

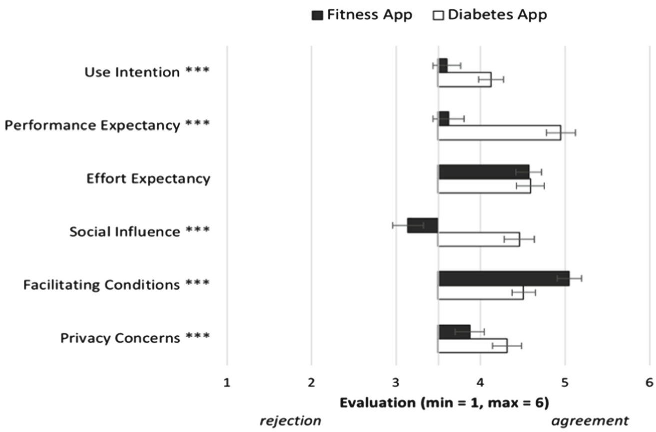



**Fig. 2.** Mean evaluation of the constructs in comparison between the contexts (*** = p < .001, with 95% confidence intervals, N = 165).

## 4.2 Comparing the Extended UTAUT Models

The structural models were assessed by checking the significance of the path coefficients, calculating the effect size $f^2$ to assess the contribution of the constructs to

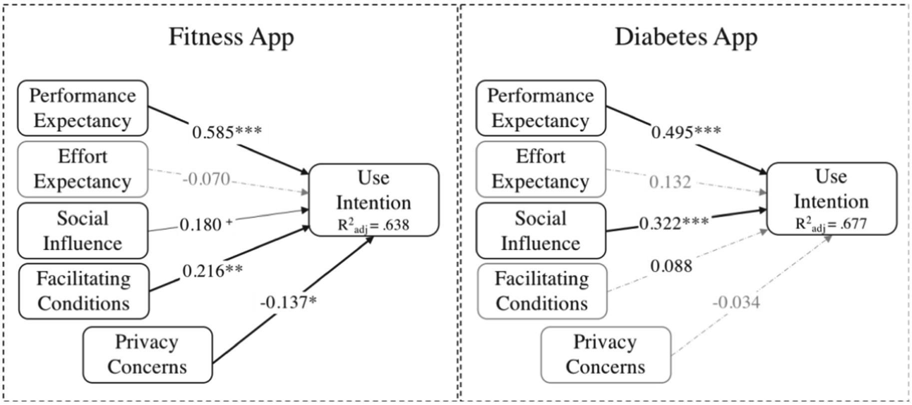



**Fig. 3.** Results of the smartPLS SEM with path coefficients and significance ($^{+}$ = p < .1, * = p < .05, ** = p < .01, *** = p < .001, N = 165).

explaining the variance in use intention, and by calculating the predictive relevance $Q^2$ of the constructs. As illustrated in Fig. 3, not all path coefficients were significant, especially in the case of the diabetes app. Still, the remaining constructs can explain a good amount of variance in the intention to use the respective application ($R^2_{adj,FA} = .638$ and $R^2_{adj,DA} = .677$) and have large predictive relevance ($Q^2_{FA} = .476$ and $Q^2_{DA} = .405$).

The intention to use a fitness app is strongly predicted by performance expectancy ($\beta = .585$, $t = 6.66$, $p < .001$, $f^2 = .574$). Additionally, facilitating conditions ($\beta = .216$, $t = 2.77$, $p < .01$, $f^2 = .063$) have a small positive influence on use intention. Privacy concerns do in contrast lessen the intention to use a fitness app ($\beta = -0.137$, $t = 2.41$, $p < .05$, $f^2 = .045$). Social influence shows a tendency to be positively related to intention to use a fitness app that is only significant at 10% ($\beta = .18$, $t = 1.87$, $p < .1$, $f^2 = .055$). Effort expectancy has no influence.

Performance expectancy also shows a strong predictive power for the intention to use a diabetes app ($\beta = .495$, $t = 5.22$, $p < .001$, $f^2 = .424$). Social influence is the only other construct having an influence on use intention in this context ($\beta = .322$, $t = 3.97$, $p < .001$, $f^2 = .221$). All other constructs cannot predict the intention to use a diabetes app.

## 5 Discussion

This study applied an extended UTAUT model to contrast acceptance patterns between two mHealth contexts, a medical context using a diabetes smartphone app as example vs. a lifestyle context with a typical fitness smartphone app. We could show, that the contexts indeed differ in absolute perception of the measured constructs as well as in the factors influencing the intention to use the exemplary apps.

Confirming previous research in medical and other contexts (e.g., [18, 19, 22]), performance expectancy was the strongest predictor of use intentions for fitness and diabetes app in our study. Performance expectancy is perceived to be higher in the case of the diabetes app and use intentions are also higher. That the participants perceive the diabetes app as more useful could be explained by the higher necessity to use the app, as the costs for non-compliance, e.g., for inaccurate insulin dosages, are high. The use of fitness apps can motivate a more active lifestyle which contributes to a better quality of life [9]. But in contrast to the diabetes app, the consequences for, e.g., not walking 10 000 steps a day, are not observed as timely and therefore may seem less costly.

Effort expectancy did not predict the intentions to use diabetes nor fitness app in this study, contradicting older research (e.g., [18]). Generally, the ease of use of both applications was perceived as very good. One explanation to these results could be that smartphone applications are very wide-spread and thus users are experienced by now. At the same time, our sample was quite young on average. In most areas, older people show less self-efficacy in interacting with technologies [41]. Also, the study by Hoque and Sorwar [22] focusing on older adults found effort expectancy as well as technology anxiety to be quite strong predictors of the intention to use mHealth applications. It would be revealing to compare our results to an older sample to examine age effect. Maybe older participants would perceive effort expectancy still as more relevant.

Social Influence shows a stronger influence on use intention in the medical context, confirming previous results [19, 22]. Additionally, the participants perceive social influence to be higher in the medical context. That social influence is slightly rejected and no significant predictor of use intention in case of the fitness app, is in line with the findings of Lidynia et al. [35] that users do not like to share their fitness app information with peers. Thus, data sharing with peers is no relevant driver for fitness app use. But recommendations by healthcare personnel can increase intention to use diabetes apps.

Facilitating conditions were perceived higher in the case of the fitness app. We can only speculate that the participants thought that they would find more helpful assistance when using a fitness app because these are more prevalent than diabetes apps. At the same time, facilitating conditions were a moderately strong predictor of use intention for fitness apps, but showed no influence on the intentions to use the diabetes app. As the participants regarded the medical app as very much more useful than the fitness app, missing facilitating conditions could be overridden in their relevance for use intention by the usefulness of the apps. Further research should break down the concept of usefulness or performance expectancy and analyze it in more detail, e.g., by incorporating the perceived necessity of using a system and available alternatives. Moreover, the trade-offs between barriers and benefits are an important research focus.

Privacy concerns have been analyzed excessively for their influence on use behavior (e.g., [13, 32]), but only showed a weak influence on use intention in case of the fitness app and no influence in case of the diabetes app. Still, the participants agreed to be concerned about their privacy, even more in the medical context. Thus, the more sensitive medical information that is collected by a diabetes tracker results in stronger concerns, confirming previous empirical results (e.g., [31]). That the higher concerns did not negatively influence use intention in the medical context could be explained by the phenomenon of the privacy paradox and the theory of the privacy calculus. The often observed discrepancy between high privacy concerns and contradicting low privacy preserving behavior, is labelled the privacy paradox (e.g., [42]). To explain this phenomenon, the privacy calculus theory postulates that users perform a calculus between perceived privacy risks and the benefits they gain by their behavior (e.g., [31, 43]). In case of the diabetes app, performance expectancy is very high and may outbalance the perceived privacy concerns. This phenomenon should be studied further in order to understand user decisions in privacy-sensitive areas like health technologies. With the use of conjoint studies, the trade-offs between privacy concerns on the one hand and benefits on the other hand can be analyzed.

Our empirical research provided valuable insights into the context-dependent acceptance of mHealth technologies. However, limitations regarding the applied method and the sample need to be taken into account. On one hand, this study examined the intentions to use the presented applications. In further research, the actual use behavior needs to be taken into account. Additionally the participants evaluated the applications after being presented with a short description of their functions and a picture instead of having the possibility to test them. Moreover, in the case of the diabetes app, all questions were hypothetical in nature depending on the participants to pretend to be suffering from diabetes. We cannot assume that the participants could empathize with the situation and needs of a diabetes patient. Thus, it would be very interesting to study the perceptions of diabetes patients and to examine the reasons why some of them are using mHealth apps to support their self-management and others do not.

In general, our sample was quite young and well educated. Fitness apps are mostly used by younger adults [44], but the occurrence of diabetes increases with age [45]. Thus, especially older adults are an interesting user group who could benefit largely from mHealth apps but who are also known to adopt new technologies slower [46]. The needs and wishes of this special user group should be taken into account in the development of medical technologies and the research into technology acceptance.

Finally, this study was conducted with a solely German sample. Technology acceptance and privacy concerns have been shown to vary depending on culture. For future studies, acceptance patterns of mHealth should be systematically compared between cultures.

**Acknowledgements.** The authors thank all participants for sharing their thoughts and opinions and Niklas Kunstleben for research support. This research was funded by the German Ministry of Education and Research (BMBF) under the project MyneData (KIS1DSD045).

## References

1. Whiting, D.R., Guariguata, L., Weil, C., Shaw, J.: IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diab. Res. Clin. Pract. **94**, 311–321 (2011)
2. Statistisches Bundesamt: Statistisches Jahrbuch 2017. Statistisches Bundesamt (Destatis), Wiesbaden (2017)
3. Cowan, L.T., Van Wagenen, S.A., Brown, B.A., Hedin, R.J., Seino-Stephan, Y., Hall, P.C., West, J.H.: Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ. Behav. **40**, 133–139 (2013)
4. Sieverdes, J.C., Treiber, F., Jenkins, C.: Improving diabetes management with mobile health technology. Am. J. Med. Sci. **345**, 289–295 (2013)

5. Wang, Y., Xue, H., Huang, Y., Huang, L., Zhang, D.: A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. Adv. Nutr. Int. Rev. J. **8**, 449–462 (2017)
6. IMS Health: Distribution of disease specific apps available worldwide in 2013 and 2015, by category https://www.statista.com/statistics/623981/healthcare-apps-worldwide-by-disease-category/
7. Shaw, J.E., Sicree, R.A., Zimmet, P.Z.: Global estimates of the prevalence of diabetes for 2010 and 2030. Diab. Res. Clin. Pract. **87**, 4–14 (2010)
8. Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B.B., Ginsburg, A.S.: Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J. Med. Internet Res. **17**, e52 (2015)
9. Krug, S., Jordan, S., Mensink, G.B.M., Müters, S., Finger, J., Lampert, T.: Physical activity results of the german health interview and examination survey for adults (DEGS1). Bundesgesundheitsblatt - Gesundheitsforsch. - Gesundheitsschutz **56**, 765–771 (2013)
10. World Health Organization: Global Recommendations on Physical Activity for Health. WHO Press, Geneva, Switzerland (2010)
11. Rasche, P., Wille, M., Bröhl, C., Theis, S., Schäfer, K., Knobe, M., Mertens, A., Medic, R.: Prevalence of health app use among older adults in Germany: national survey. JMIR mHealth uHealth **6**, e26 (2018)
12. Li, H., Wu, J., Gao, Y., Shi, Y.: Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. Int. J. Med. Inform. **88**, 8–17 (2016)
13. Baruh, L., Secinti, E., Cemalcilar, Z.: Online privacy concerns and privacy management: a meta-analytical review. J. Commun. **67**, 26–53 (2017)
14. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**, 319–340 (1989)
15. Davis, F., Bagozzi, R., Warshaw, P.: User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. **35**, 982 (1989)
16. King, W.R., He, J.: A meta-analysis of the technology acceptance model. Inf. Manag. **43**, 740–755 (2006)
17. Klein, R.: Internet-based patient-physician electronic communication applications: patient acceptance and trust. e-Serv. J. **5**, 27–52 (2007)
18. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **27**, 425–478 (2003)
19. Sun, Y., Wang, N., Guo, X., Peng, Z.: Understanding the acceptance of mobile health services: a comparison and integration of alternative models. J. Electron. Commer. Res. **14**, 183–200 (2013)
20. Williams, M.D., Rana, N.P., Dwivedi, Y.K.: The unified theory of acceptance and use of technology (UTAUT): a literature review (2015)
21. Attuquayefio, S., Addo, H.: Review of studies with UTAUT as conceptual framework. Eur. Sci. J. **10**, 1857–7881 (2014)
22. Hoque, R., Sorwar, G.: Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int. J. Med. Inform. **101**, 75–84 (2017)
23. Yuan, S., Ma, W., Kanthawala, S., Peng, W.: Keep using my health apps: discover users' perception of health and fitness apps with the UTAUT2 model. Telemed. e-Health **21**, 735–741 (2015)
24. Westin, A.F.: Privacy and Freedom. Am. Sociol. Rev. **33**, 173 (1967)
25. Burgoon, J.: Privacy and communication. Ann. Int. Commun. Assoc. **6**, 206–249 (1982)
26. Koops, B., Newell, B.C., Timan, T., Skorvanek, I., Chokrevski, T., Galic, M.: A typology of privacy. Univ. Pennsylvanica J. Int. Law. **38**, 1–93 (2017)

27. European Commission: Special Eurobarometer 431 - Data Protection, Cologne (2015)
28. Mothersbaugh, D.L., Foxx, W.K., Beatty, S.E., Wang, S.: Disclosure antecedents in an online service context: the role of sensitivity of information. J. Serv. Res. **15**, 76–98 (2012)
29. Anderson, C.L., Agarwal, R.: The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. Inf. Syst. Res. **22**, 469–490 (2011)
30. Rohm, A.J., Milne, G.R.: Just what the doctor ordered: the role of information sensitivity and trust in reducing medical information privacy concern. J. Bus. Res. **57**, 1000–1011 (2004)
31. Xu, H., Dinev, T., Smith, J., Hart, P.: Information privacy concerns: linking individual perceptions with institutional privacy assurances. J. Assoc. Inf. Syst. **12**, 798–824 (2011)
32. Bansal, G., Zahedi, F.M., Gefen, D.: Do context and personality matter? Trust and privacy concerns in disclosing private information online. Inf. Manag. **53**, 1–12 (2016)
33. Huckvale, K., Prieto, J.T., Tilney, M., Benghozi, P.-J., Car, J.: Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med. **13**, 214 (2015)
34. Or, C.K.L., Karsh, B.T.: A systematic review of patient acceptance of consumer health information technology. J. Am. Med. Informat. Assoc. **16**, 550–560 (2009)
35. Lidynia, C., Brauner, P., Ziefle, M.: A step in the right direction – understanding privacy concerns and perceived sensitivity of fitness trackers. In: AHFE 2017: Advances in Human Factors in Wearable Technologies and Game Design, pp. 42–53 (2018)
36. Boontarig, W., Chutimaskul, W., Chongsuphajaisiddhi, V., Papasratorn, B.: Factors influencing the Thai elderly intention to use smartphone for e-Health services. In: 2012 IEEE Symposium on Humanities, Science and Engineering Research, SHUSER 2012, pp. 479–483 (2012)
37. Parker, S.J., Jessel, S., Richardson, J.E., Reid, M.C.: Older adults are mobile too! Identifying the barriers and facilitators to older adults' use of mHealth for pain management. BMC Geriatr. **13**, 43 (2013)
38. Guo, X., Sun, Y., Yan, Z., Wang, N.: Privacy-personalization paradox in adoption of mobile health service: the mediating role of trust. In: Proceedings PACIS 2012 Paper 27 (2012)
39. Ringle, C., Wende, S., Becker, J.-M.: SmartPLS 3. Bönningstedt, SmartPLS (2015)
40. Hair Jr., J.F., Hult, G., Ringle, C., Sarstedt, M.: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Sage Publications (2011)
41. Ziefle, M., Wilkowska, W.: Technology acceptability for medical assistance. In: 4th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) (2010)
42. Kokolakis, S.: Privacy attitudes and privacy behaviour: a review of current research on the privacy paradox phenomenon. Comput. Secur. **2011**, 1–29 (2015)
43. Dinev, T., Hart, P.: An extended privacy calculus model for e-commerce transactions. Inf. Syst. Res. **17**, 61–80 (2006)
44. Fittkau & Maaß Consulting: Share of Smartphone Users that Used Fitness APS in Germany in May 2015, by Age Group https://www.statista.com/statistics/452454/fitness-app-usage-among-smartphone-users-in-germany-by-age/
45. Guariguata, L., Whiting, D.R., Hambleton, I., Beagley, J., Linnenkamp, U., Shaw, J.E.: Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res. Clin. Pract. **103**, 137–149 (2014)
46. Czaja, S.J., Charness, N., Fisk, A.D., Hertzog, C., Nair, S.N., Rogers, W.A., Sharit, J.: Factors predicting the use of technology: findings from the center for research and education on aging and technology enhancement (CREATE). Psyhol Aging. **21**, 333–352 (2006)

# Using Non-invasive Wearable Sensors to Estimate Perceived Fatigue Level in Manual Material Handling Tasks

Liuxing Tsao[1], Liang Ma[1(✉)], and Christian-Thomas Papp[1,2]

[1] Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
liangma@tsinghua.edu.cn

[2] Laboratory for Machine Tools and Production Engineering, RWTH Aachen University, 52056 Aachen, Germany

**Abstract.** Physical fatigue in manual material handling (MMH) may cause musculoskeletal disorders (MSDs), which threatens the well-being of workers. However, conventional techniques for measuring fatigue have their limitations and are difficult to implement in realistic working conditions without sufficient resources. In this study, we proposed a method utilizing non-invasive wearable sensors to collect bio-signals (respiration, photo-plethysmography, and electrodermal activity) and estimate perceived physical fatigue. Six participants volunteered in two MMH tasks at two paces. Subsets of five bio-signal measures were selected to estimate perceived fatigue levels using a universal regression model and six individualized regression models. We classified perceived fatigue into three levels and examined the correct classification rate of the estimated fatigue levels. Correct classification rates for the general model and the individualized models were 67% and 80%, respectively. Results confirm the feasibility to predict fatigue level using wearable sensors, but the regression models should be used with caution.

**Keywords:** Wearable sensors · Physical fatigue · Fatigue measurement Individualized regression models

## 1 Introduction

Physical fatigue in manual material handling (MMH) may cause musculoskeletal disorders (MSDs) [1], which threatens the well-being of workers. In 2015, work-related MSDs accounted for 31% of the total cases of nonfatal injuries and illness which caused days off work in the United States [2]. Surveys in major cities of China showed that lots of frontier workers are suffering from injuries and discomforts of MSDs [3, 4]. It is no doubt that physical fatigue influences the well-being of the workers and also reduces the effectiveness and efficiency of production. Ergonomics interventions can play a cost-efficient way in assessing and managing the fatigue and disorders of workers, further ensure the health of workers and bring benefits to the industry [5].

Previous researchers have developed various methods to assess exposures in industrial sectors for fatigue management. These methods fall into three categories:


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 65–74, 2019.
https://doi.org/10.1007/978-3-319-94619-1_7

self-report, observational methods, and direct measurements [6]. Self-reports are the methods that using interviews, questionnaires or other techniques to collect workplace risk factors from the workers or the managers. Observational techniques require the input of ergonomics experts to watch the working process and evaluate the exposure levels. Direct methods rely on sensors to collect physiological measures of workers and assess the status of workers. Although these techniques have improved the situation of fatigue management, there are still limitations in the existing methods. As reviewed by Li and Buckle [7], the interaction effects of the leading factors of fatigue and MSDs are still not well understood. Some techniques require a great effort of expertise in data collection and data analysis and are not easy to implement in realistic working scenarios. Moreover, monitoring the workers or interviewing them frequently can be disturbed to the workers. There is an emerging need to find out a non-invasive and cost-efficient way for easier fatigue management in MMH tasks.

During the last years, the development of wearable technologies has gained lots of attention in the scientific community and the industry [8]. Wearable systems can comprise various types of small physiological sensors, transmission modules, and processing capabilities, thus facilitate low-cost unobtrusive solutions for activity status monitoring. For example, inertial sensor systems have been developed to analyze the activities of wearers in fitness and sports training [9, 10]; Integrated health-related wearable systems can be used to monitor the health condition of patients in hospital or elderlies at home [11]; inertial sensors and heart rate monitor can provide useful features for physical fatigue evaluation and prediction [12]. Besides the above-mentioned applications, wearable sensors that can track bio-signals will also provide valuable information for fatigue management. Wearable sensors show great potentials to be utilized in fatigue management, for their ease to be attached to the human body and the capabilities of collecting data continuously and non-invasively.

In this study, we proposed a method that using non-invasive wearable sensors to collect bio-signals for estimating fatigue of workers in MMH tasks. Two typical tasks, lowering and lifting loads and turning loads were performed by six participants at various paces (quick/slow) to simulate the MMH tasks. Three representative bio-signals that can reflect the activity level were selected in the preliminary study, including respiration (RSP), photo-plethysmography (PPG), and electrodermal activity (EDA). Borg RPE scale was used to collect the perceived exertion level [13]. A universal regression model and six individualized regression models were trained to estimate fatigue level for the participants. Fatigue scores were further classified into three levels (low, medium, and high) and the correct classification rate of the estimated fatigue level was examined for model validation.

## 2 Methods

### 2.1 Task Design

In the experiment, we selected two typical tasks in MMH, including lifting and lowering loads and turning loads. To trigger different levels of fatigue, the tasks were executed at two level of paces, quick (five seconds per movement in the lifting and lowering task, two seconds per movement in the turning task) and slow (ten seconds per movement in the lifting and lowering task, five seconds per movement in the turning task). A total of four sessions (two tasks * two paces) were included in the experiment, each session for five minutes. Before the four sessions, a two-minute rest period was set for baseline data collection. After each session, a five-minute rest was provided to allow recovery of the participants. The experiment setup is demonstrated in Figs. 1 and 2.

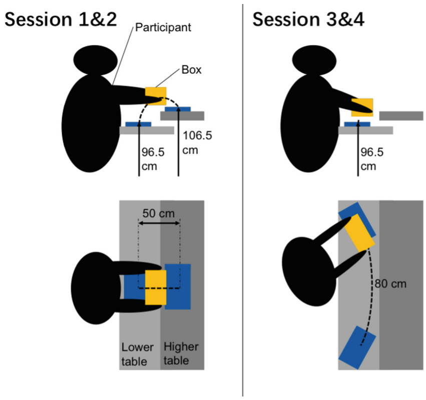



**Fig. 1.** The experiment setup (session 1&2, lifting and lowering at five seconds per movement and ten seconds per movement; session 3&4, turning the box at five seconds per movement and two seconds per movement).

**Fig. 2.** The overview of the lab.

In the first and second sessions, the task was to lift the box from a lower table to a higher table, then lower it back to the lower table. The lower table had a height of 90 cm and the higher table was at a height of 100 cm (as we used pedestals of 6.5 cm, the final heights were 96.5 cm and 106.5 cm, respectively). The target locations were marked in a straight line and the participants did not need any rotational movements. In the third and fourth session, the participants were asked to rotate and turn the box back and forth for a distance of around 80 cm in a standing position. The height of the table was 90 cm with an additional 6.5 cm of the pedestals. Depending on the stature of the participants, a stretch forward might be needed. The box had a dimension of 36 cm (length) * 13 cm (width) * 26 cm (depth) and a weight of 6 kg. In the first and third session, the participants moved the box every five seconds (quick pace for lifting and lowering, slow pace for turning). In the second session, the participants moved the box every ten seconds (slow pace). In the fourth session, the participants moved the box every two seconds (quick pace).

## 2.2 Equipment

A wireless measurement system BioNomadix (MP150) and selected modules were used to collect the bio-signals during the simulated MMH tasks: RSPEC module was selected to track abdominal or thoracic expansion and contraction while breathing, providing information of respiration (RSP). PPG channel measured blood volume pulse via optical plethysmographic methods and provided heart rate, interbeat interval, and vasodilation/constriction data. The electrodermal response (EDA) was also included in the test to track the activation level of participants. The disposable electrodes for EDA were attached to the participants' neck, left and right, and on the left shoulder. The placement of the sensors is shown in Fig. 3.


**Fig. 3.** The placement of the sensors (RSP belt around the breast, PPG attached on the ear, and EDA electrodes on the neck, left and right, and left shoulder).

In addition, the Borg RPE scale was used to collect the perceived exertion of participants every one minute. A smartphone was used to take pictures during the experiment and it was also used as a digital metronome to indicate the pace.

### 2.3 Data Analysis

Five features were extracted from the bio-signals using the software AcqKnowledge 4.0. These features were: the peak voltage of RSP, the peak voltage of PPG, the rate of RSP, the rate of PPG, and the peak microsiemens of EDA. They were further re-sampled at 10 Hz using the software so that these reduced samples can be easier to handle in the regression models. The start and the end of the four sessions were marked and signals were further segmented into one-minute interval to match the Borg score collection interval (600 samples for each feature per minute after re-sampling).

After segmenting the signals, there would be 21 data points of each participant (1 baseline + 4 sessions * 5 min). Each data point contained the perceived fatigue measure, and five predictive measures (averaged of the 600 samples in every minute for each measure). Considering the units and scales of the bio-signal measures were different, we normalized the data into 0–1, using Eq. (1), where $i$ is the participant ID, $j$ is the feature ID, and $t$ indicates the time.

$$Norm.\,data_{j,t}^{i} = \frac{Abs.\,data_{j,\,t}^{i} - \min\left(Abs.\,data_{j,\,all\,t}^{i}\right)}{\max\left(Abs.\,data_{j,\,all\,t}^{i}\right) - \min\left(Abs.\,data_{j,\,all\,t}^{i}\right)}. \quad (1)$$

As a preliminary test to estimate fatigue level based on the selected features, we chose linear regression modeling to describe the relationship between fatigue and the collected bio-signals. Stepwise regressions (backward and both directions) were used to select the best subsets of features to be included in the regression, using Akaike Information Criterion (AIC) as a measure of fit.

Considering the influences of individual factors, we developed two sets of models, one was a universal model combining all the participant's data and the other set contained six individualized models fitted by each participant's data. The forth minute data of each session (the 4$^{th}$, 9$^{th}$, 14$^{th}$, 19$^{th}$ min for each participant) were selected as testing data set and the remaining 17 data points for each participant were used to train the models. A detailed description of the training set and testing set is summarized in Table 1.

**Table 1.** The assignment of training and testing set for the regression models

| | Training set | Sample size | Test set | Sample size |
|---|---|---|---|---|
| Universal model | Rest + 1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20 min<br>All participants | 102 | 4, 9, 14, 19 min<br>All participants | 24 |
| Individualized model<br>Participant $i$ | Rest + 1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20 min<br>Participant $i$ | 17 | 4, 9, 14, 19 min<br>Participant $i$ | 4 |

# 3 Results

## 3.1 Participants

A total of six participants volunteered in the study. As a preliminary test, we covered a wide range of participants for better discovering the possibilities and challenges of the proposed method. The demographic information of the six participants is summarized in Table 2.

**Table 2.** Demographic information of the six participants

| Participant | Gender | Age | Height (cm) | Weight (kg) | BMI |
|---|---|---|---|---|---|
| P01 | Male | 24 | 186 | 105 | 30.4 |
| P02 | Female | 27 | 173 | 58 | 19.4 |
| P03 | Male | 22 | 183 | 78 | 23.3 |
| P04 | Female | 28 | 165 | 57 | 20.9 |
| P05 | Male | 22 | 178 | 62 | 19.6 |
| P06 | Female | 23 | 160 | 48 | 18.8 |
| Average | 3 males and 3 females | 24.3 | 174.2 | 68.0 | 22.0 |

## 3.2 Perceived Fatigue

We used Borg RPE to describe the perceived level of participants. The trends of Borg scores for the six participants are demonstrated in Fig. 4.

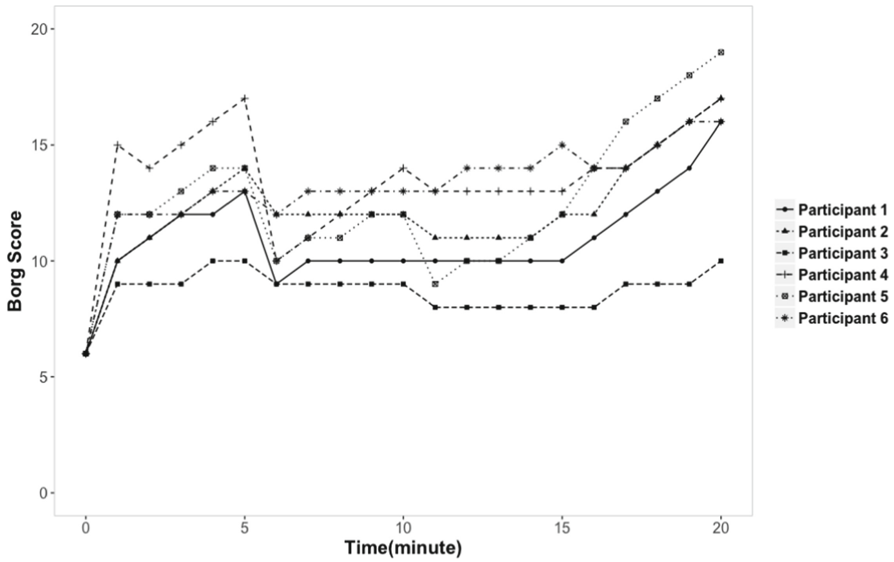



**Fig. 4.** Borg scores for the participants

The average Borg scores for each participant during each session are summarized in Table 3. Due to the small sample size, we did not perform statistical analysis, but we can still find a general pattern that quick pace leads to higher fatigue level (except for Participant 6 in the lifting and lowering task).

**Table 3.** Borg score for each participant during each session

| Participant ID | 1 | 2 | 3 | 4 | 5 | 6 | Average |
|---|---|---|---|---|---|---|---|
| Quick, lifting and lowering | 11.6 | 12.0 | 9.4 | 15.4 | 13.0 | 12.4 | 12.3 |
| Slow, lifting and lowering | 9.8 | 12.0 | 9.0 | 12.0 | 11.2 | 12.8 | 11.1 |
| Slow, turning | 10.0 | 11.2 | 8.0 | 13.0 | 10.4 | 14.0 | 11.1 |
| Quick, turning | 13.2 | 14.8 | 9.0 | 15.2 | 16.8 | 15.0 | 14.0 |

### 3.3 The Bio-signals

In Fig. 5, we demonstrate the changes of bio-signals along with time. Although it is difficult to figure out a universal pattern to describe the changes of the signals during different tasks, in some measures for some specific participants, the values are higher during the quick pace task (e.g., RSP voltage for participant 1, PPG voltage for participant 4, PPG rate for participant 3). The detailed relationship between these measures and fatigue level will be explained using regression models in the next section.

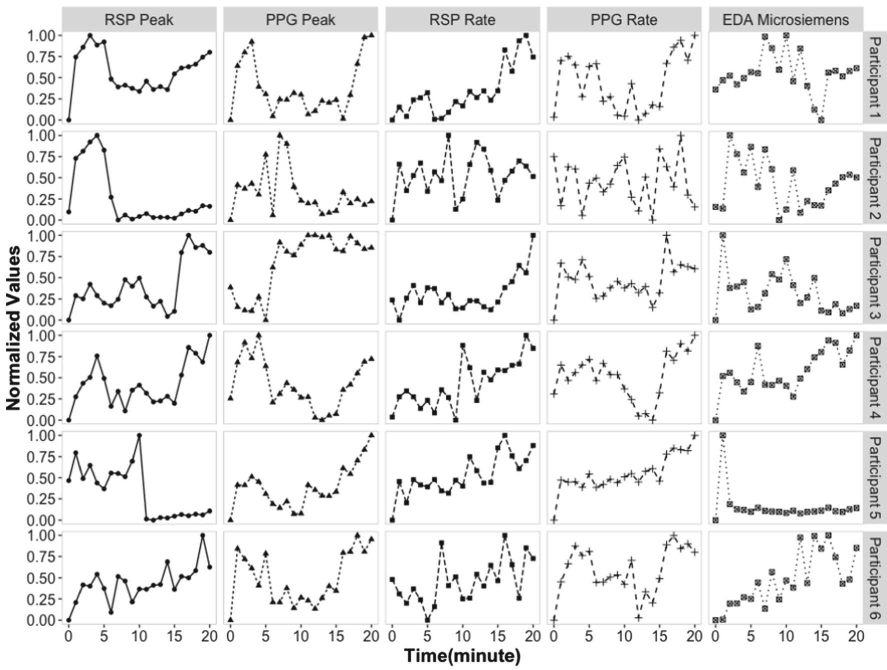



**Fig. 5.** The bio-signals of the participants

### 3.4 Regression Models and Validation

To estimate perceived fatigue levels using the bio-signals, we trained two sets of regression models, one is a universal model that utilizing data points of all participants, another set contains six individualized models (refer to Table 1 for the assignment of training and testing data points). The obtained models with the testing statistics are listed in Table 4. The inclusion and exclusion of measures vary a lot across different participants, and no single measure is included in all the models. The respiration rate is included in most of the models, indicating the importance of this measure in estimating fatigue.

**Table 4.** Regression model coefficients and statistics (training set)

| | Intercept | RSP peak | PPG peak | RSP rate | PPG rate | EDA peak | Adjusted $R^2$ |
|---|---|---|---|---|---|---|---|
| Universal model | 7.13 | / | / | 3.81 | 3.59 | 2.50 | 0.39 |
| Participant 1 | 5.56 | 3.77 | 1.70 | 3.67 | / | 2.26 | 0.84 |
| Participant 2 | No model | | | | | | |
| Participant 3 | 6.94 | / | / | 3.06 | / | 2.30 | 0.41 |
| Participant 4 | 8.79 | / | 5.18 | 5.47 | / | / | 0.49 |
| Participant 5 | 4.17 | 2.82 | 4.86 | / | 9.70 | / | 0.85 |
| Participant 6 | 9.01 | 9.30 | / | −2.39 | / | 3.02 | 0.77 |

To test the models, we used the data points at the $4^{th}$, $9^{th}$, $14^{th}$, and $19^{th}$ min for each participant as the testing set (see Table 1 for the assignment of data points). The absolute differences between the estimated fatigue score and the perceived score were listed in Table 5 for both models. Furthermore, we classified the fatigue level by the Borg scores into low (score $\leq$ 9.5), medium (9.5 < score $\leq$ 13.5), and high (score > 13.5) and calculated the correct classification rates for the testing data points. Results of the classification performance are also listed in Table 5. It is noticed that the individualized models perform better than the universal model with smaller estimation differences and higher correct classification rates.

**Table 5.** Performance of the models using testing set

| | Universal model | | Individualized model | |
|---|---|---|---|---|
| | Absolute difference | Correct rate | Absolute difference | Correct rate |
| Participant 1 | 1.0 | 75.0% | 0.8 | 75.0% |
| Participant 2 | 2.3 | 75.0% | No model | No model |
| Participant 3 | 1.8 | 50.0% | 0.6 | 75.0% |
| Participant 4 | 2.6 | 75.0% | 1.7 | 100.0% |
| Participant 5 | 2.5 | 50.0% | 1.5 | 75.0% |
| Participant 6 | 1.3 | 75.0% | 2.1 | 75.0% |
| Average | 1.9 | 66.7% | 1.3 | 80.0% |

## 4 Discussion

In this study, we used non-invasive wearable sensors to collect bio-signals of participants during simulated MMH tasks and extracted information from the bio-signals to estimate the perceived fatigue using regression models. Two MMH tasks at two paces were included so as to trigger different levels of fatigue.

Regarding the selection of bio-signals, we included the respiration, the photoplethysmography, and the electrodermal activity to reflect the physiological status of the participants. From the descriptive results (see Fig. 5), we find that these measures do not follow any general trends, though some measures increase when the work pace is high. From the regression models, although no single measure is significant in every model, the RSP rate is significant in the universal model and three of the individualized models, while RSP peak voltages, PPG peak voltages, and EDA peak microsiemens are included in three individualized models. These measures seem to be closely correlated with perceived fatigue, though whether to include them in fatigue estimation needs more careful analysis.

Regarding the selection of universal model or the individualized models, we are not surprised to see that individualized models perform better than the universal model almost for every participant. Further examining the models, we notice that whether to include a specific measure in the models depends on the participant, and the coefficients of the same measure vary across models a lot. The results address the importance of considering individual differences when constructing regression models for fatigue evaluation. When the target population is not identical, great caution should be put to balance the cost of generating numerous individualized models and the need to describe the variation among different persons.

## 5 Conclusion

In this study, we proposed a method to estimate physical fatigue during MMH tasks using data collected from non-invasive wearable sensors. A preliminary feasibility test of the proposed method was performed using a sample size of six participants. Two sets of linear regression models were calculated, one was a universal model utilizing all participant's data while another set contained six individualized models for each participant. Results show that both the universal model and the individualized models can be used to estimate fatigue, however, the individualized models performed better in our case due to the variation of individual factors across the participants. We suggest a more structured experiment that better considering the individual differences and covering more bio-signals with a larger sample size.

## References

1. Armstrong, T.J., Buckle, P., Fine, L.J., Hagberg, M., Jonsson, B., Kilbom, A., Kuorinka, I. A., Silverstein, B.A., Sjogaard, G., Viikari-Juntura, E.R.: A conceptual model for work-related neck and upper-limb musculoskeletal disorders. Scand. J. Work Environ. Health. **19**, 73–84 (1993)
2. Bureau of Labor Statistics.: Nonfatal Occupational Injuries and Illness Requiring Days Away From Work (2015). https://www.bls.gov/news.release/pdf/osh2.pdf
3. Yu, W., Yu, I.T.S., Li, Z., Wang, X., Sun, T., Lin, H., Wan, S., Qiu, H., Xie, S.: Work-related injuries and musculoskeletal disorders among factory workers in a major city of China. Accid. Anal. Prev. **48**, 457–463 (2012)
4. Yu, S., Lu, M.-L., Gu, G., Zhou, W., He, L., Wang, S.: Musculoskeletal symptoms and associated risk factors in a large sample of Chinese workers in Henan province of China. Am. J. Ind. Med. **55**, 281–293 (2012)
5. Helander, M.G., Burri, G.J.: Cost effectiveness of ergonomics and quality improvements in electronics manufacturing. Int. J. Ind. Ergon. **15**, 137–151 (1995)
6. David, G.C.: Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. Occup. Med. **55**, 190–199 (2005)
7. Li, G., Buckle, P.: Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods. Ergonomics **42**, 674–695 (1999)
8. Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **40**, 1–12 (2010)
9. O'Reilly, M.A., Whelan, D.F., Ward, T.E., Delahunt, E., Caulfield, B.M.: Classification of deadlift biomechanics with wearable inertial measurement units. J. Biomech. **58**, 155–161 (2017)
10. Reenalda, J., Maartens, E., Homan, L., Buurke, J.H.(Jaap): Continuous three dimensional analysis of running mechanics during a marathon by means of inertial magnetic measurement units to objectify changes in running mechanics. J. Biomech. **49**, 3362–3367 (2016)
11. Kakria, P., Tripathi, N.K., Kitipawang, P.: A real-time health monitoring system for remote cardiac patients using smartphone and wearable sensors. Int. J. Telemed. Appl. **2015**, 1–11 (2015). Article no. 8
12. Maman, Z.S., Yazdi, M.A.A., Cavuoto, L.A., Megahed, F.M.: A data-driven approach to modeling physical using wearable sensors. Appl. Ergon. **65**, 515–529 (2017)
13. Borg, G.: Borg's Perceived Exertion and Pain Scales. Human Kinetics, Champaign (1998)

# Crafting Usable Quantified Self-wearable Technologies for Older Adult

Jayden Khakurel[1(✉)], Antti Knutas[2], Helinä Melkas[1], Birgit Penzenstadler[3], and Jari Porras[1]

[1] Lapppeenranta University of Technology, Skinnarilankatu 34, 53850 Lappeenranta, Finland
{jayden.khakurel,helina.melkas,jari.porras}@lut.fi
[2] Lero, The Irish Software Research Centre, Glasnevin, Dublin, Ireland
antti.knutas@lut.fi
[3] California State University, Long Beach, CA, USA

**Abstract.** Commercially off-the-shelf (COTS) quantified self-wearable technologies (QSWT) have enabled younger individuals to adopt a measurable living style [49] through the collection of "quantifiable data". However, the adoption of wearables remains lowest among the older adult, and the question of what is holding adoption back remains. The purpose of this study is to: (i) explore and present the device characteristics of smartwatches and pedometers that affect the adoption of wearables across different cultures; (ii) study country-specific older adult's importance weights on identified issues; and (iii) provide informal usability guidelines for manufacturers, researchers, and application developers. The results revealed that the usability issues such as screen size, tapping detection, device size, interaction techniques, navigation, and typography were some of the reasons for the low adoption of wearables among the older adult. Further, device and screen size were significantly more essential for the Finnish compared to US older adult participants, demonstrating that culture might influence the perception of some device characteristics.

**Keywords:** Wearables · Usability · Culturability · Older adult Framework · User interface · Elderly · Quantified self-technologies Smartwatches · Pedometers

## 1 Introduction

Quantified self-wearable technologies (QSWTs), such as commercial off-the-shelf (COTS) products including smartwatches and pedometers, and their associated applications have enabled individuals to adopt a measurable living style [49] through the collection of "quantifiable data," such as sleep patterns, calorie intake, and steps taken. However, smartwatches and pedometers have the lowest adoption of wearables among the older adult[1] to enhance quantifiable data practices [38] due to: (i) various design dimensions, complex interfaces, and extensive functionalities [14] or (ii) perceptions of

[1] http://www.emarketer.com/Chart/US-Wearable-User-Penetration-by-Age-2017-of-population-each-group/202360.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 75–87, 2019.
https://doi.org/10.1007/978-3-319-94619-1_8

being unable to learn new things because of insufficient cognitive capability, vision, or motor function [11]. Almost none of the QSWTs available on the market in their current form are suited for the older adult [39]. For example, Angelini et al. [3] reported that most interfaces proposed to date for smartwatches offer limited accessibility to older adult: the screens are small, the information is often shown with small characters, and small buttons are used to navigate the interface [3].

Even though the low adoption of wearables has become increasingly visible, there is still a lack of (i) research on older adult experiences with the adoption of wearing smartwatches [6]; (ii) knowledge about what features the older adult desire when using COTS smartwatches and pedometers, which is critical for wearable device and service design [19]; and (iii) knowledge of usability issue variances related to cultural dimensions, leading to non-adoption among the older adult. Angelini et al. [3] show that one reason could be that designing for the older adult implies several additional challenges concerning products with a generic target; in addition, in our society, "older adult" are classified generally as a single separate group. Furthermore, country-specific older adult importance weights given to the specific cause of non-adoption from the perspectives of device characteristics may vary because of cultural origins and different traditions, custom, ethics, and values [9, 44]. This is a setback for the older adult in using and taking advantage of COTS QSWTs. Therefore, the research question (RQ) is posed to obtain a more comprehensive overview of the gap:

*RQ: What should be considered by technology designers and the research community to enhance the device characteristics related to QSWTs and to improve older adult adoption traits? Rationale: This provides information that can serve as a basis for improving adoption by enhancing device characteristics.*

To answer the research question, we identify which types of usability issues related to COTS QSWTs persist across different cultures. The study begins with previously identified reasons associated with the adoption and withdrawal of wearables by the older adult, and it continues with identifying the current usability issues of various stages evaluated using contextual action theory (CAT), as presented by Stanton [43], and a usability evaluation method from Ivory and Hearst [18]. Second, we apply the Mann–Whitney U test [48] (also known as the Mann–Whitney Wilcoxon Test) to analyze country-specific differences in older adults' identified usability issues. Third, we provide informal usability guidelines for technology designers, researchers, and application developers to broaden the scope of their designs and interfaces for upcoming devices and applications to provide a richer end-user experience so that wearables can also be adopted by the older adult.

## 2 Related Work

Related work on crafting usable quantified wearables falls within two areas: a focus on reasons associated with the adoption and withdrawal of wearables by the older adult and a focus on "Culturability" to understand the importance of the relationship been culture and usability [5].

**Wearable Technology Adoption and Withdrawal Among the older adult.** Rasche et al. [35] conducted a usability evaluation with an activity tracker of the older adult (60 + years) to understand whether activity trackers are emphasized to stigmatize the older adult, their intention to use the devices, and their positive and negative experiences. Results from their study show that the older adult were motivated and felt comfortable adopting activity trackers as a motivational support tool in their lives because of the motivational aspects and objective control. However, usability remained challenging concerning the device's wearing position on the body. Similarly, Fausset et al. [13] evaluated activity-monitoring devices among the older adult over two weeks to understand the cause of adoption and withdrawals. Interestingly, for the older adult, the initial interest was "positive;" however, some participants continued for only a short time due to "lack of usefulness of wearable devices," "data inaccuracy," and "wearability." Their results indicate that (i) despite being initially receptive to using the technology, participants do not always accept and use the technologies unconditionally [13]; and (ii) there is an "interplay of usability issues, such as inaccuracy of data, wearability, and adoption which kept them from not using activity tracker for long-term." Another study that conducted usability experiments using fitness trackers among older adult was described by Schlomann et al. [39]. In this study, "consequences of use" and "device functionality" were the two main concerns for older adults to adopt wearable devices.

**Culturability.** Wallace et al. [45] distributed a survey to 144 subjects from four countries to understand whether usability attributes vary among them and whether these variances were related to cultural dimensions. According to their results, usability attributes across countries vary in terms of efficiency and satisfaction, whereas no differences were noticed concerning effectiveness. The influence of culture on usability and design, even within western nations, is also emphasized by Khaslavsky [23], who asserted that users between two western countries might display different culturally motivated problems when interacting with the same application localized only through translation. To offset such differences, the authors presented a series of guidelines for integrating culture into design. The guidelines are as follows: (i) Consider more in-depth conceptual problems with your design when localizing; (ii) Culture-specific localization is necessary for every country, not only Asia; and (iii) Use the package of variables, such as speed of message, context, personal space, time, power distance, collectivism, diffuse vs specific, and particularism vs universalism, to drive your search for more information from users.

## 3 Methodology and Procedure

### 3.1 Experimental Setup

Data for this study were derived from the four-week-long usability experiments on COTS wearable devices, which were carried out in Finland and the US among individuals aged 60 years or over. The first evaluation was carried out in Finland (2017) with 13 elderly participants (age M = 62.23, SD = 1.921), and the second study was carried out in the US (2018) with 20 elderly participants (age M = 61.92, SD = 1.6062), which is considered sufficient based on Macefield's [26] recommendation that i.e. a group size of 3–20 participants is typically valid, with 5–10 participants being a sensible baseline range. Participants from both countries were living independently and were keen to use new technology to improve their well-being. Both countries followed the same methodology for usability experiments, and participants were recruited through direct contact, advertisement, and networking. The two countries were selected because of the overall similarity of the cultures, except in aspects most relevant to this study, such as in how welfare and healthcare are arranged, including in elderly care. The Finnish system is mostly based on public funding and healthcare system is centrally funded, whereas US system is mostly based on private funding and private medical insurances.

As a first step, a general presentation about the particular research was provided to each participant, followed by a recruitment form that collected preliminary information from participants, such as technological knowledge, current use of external devices including smartphones, age, and consent to participate. All participants in the study were presented with an ethical review statement and informed consent, and in return, a signed consent form was obtained. The entire questionnaire was reviewed by two reviewers before submission for ethical committee approval. The Lappeenranta University of Technology and California State University, Long Beach, institutional review board approved the study.

**Procedure.** Contextual Action Theory (CAT), as explained by Stanton [43], and the usability evaluation method [18] were used as the foundational methodologies to enhance our understanding of the cause of the non-adoption of COTS QSWTs among fit older individuals across different culture.

Stanton [43] point out that contextual action theory explains human actions in terms of coping with technology within a context, and five phases are associated with contextual actions: (i) the user is presented with the actual demands and resources of the device design, the tasks to be performed on the device, environmental constraints (e.g. time), and so on; (ii) those demands and resources are appraised by the actor; (iii) perceived demand and resources are compared; (iv) the possible degradation of pathways might occur; and (v) these responses' effects on device interactions. During the first phase of CAT, we present the actual demands and resources to the participants, consisting of the following: (i) **Devices:** Functioning wearable COTS devices, i.e. smartwatches and pedometers, to help us to explore the significance of various types of data for future design, as identified by Kanis [21]. No requirements were provided for device selection. (ii) **Timeline:** Participants were asked to participate in two one-hour controlled environments (i.e. first meet-up session and final meet-up session), with two weeks of each category of device (i.e. every day) use between the sessions in a semi-controlled environment. (iii) **Experimental tasks**: During the first and final meet-up session, we assigned experimental tasks[2] to be performed to test usability. For the semi-controlled environment, participants were asked to use the devices in real conditions and to complete the daily log in the provided diary. No specific pre-defined activities, such as sleep, walk, and exercise, were presented to the participants. This

[2] https://doi.org/10.5281/zenodo.832159.

semi-controlled environment aimed to make participants comfortable with using the device and to gather influential data from their dairies. The diary method has been applied because it forces participants to record all activities for the period covered, and data reported in the diary are arguably more reliable [37].

**Measurements.** The diary included several kinds of data, such as (i) whether the devices were worn (*if not, why?*); (ii) activities undertaken (*i.e. walking, hiking, running, cycling, etc*.); (iii) motivation in doing physical activities because of device use (*i.e. if yes, reason for motivation; if not, why?*); (iv) used applications (*which ones; if not used, why?*); (v) usability issues (i.e. screen size; icons; interaction techniques; interaction with screen; font size; button location; data accuracy; screen resolution; device weight, shape, and size; lack of screen; and battery life, with options to add any missing usability issues); and (vi) additional comments that asked participants for "other comments that should be specified." A list of usability issues from the diaries was derived based on issues previously identified issues [3, 13, 16, 32, 33, 41]. For analysis, the usability issues of COTS QSWTs have been clustered into two categories: hardware and user interface (UI). Especially, hardware concerns issues related to the external look and feel and internal components, such as sensors, processor, memory, power supply, and transceiver [2, 25]. Meanwhile, the UI concerns issues with the various ways in which users interact with the device [2].

## 4 Results

In the interest of the study, we focus our discussion on two results: identifying usability issues from usability evaluations to focus on what types of usability issues participants reported and comparing significant relations across cultures reveals that the importance weights given to specific usability issues concerning device characteristics significantly vary across cultures among the older adult.

**Identifying Usability Issues from a Usability Evaluation.** We identified 14 usability issues that were reported during the studies (see Fig. 1) between Finland and the US, of which six were related to UI and eight to hardware.

Participants also reported usability issues with interaction techniques, such as that feedback on smartwatches was irritating: "*Having all the notifications on my watch with vibration feels so irritating and like getting an electric shock*," or "*I pressed the screen on the device but it doesn't respond; is this device broken?*" Similarly, other participants reported that it was too annoying to receive notifications: "*I received the notification while I was sleeping at night, it was annoying*." Another participant who used the smartwatch said the "*touchscreen reacts so fast when I press on it*."

Connectivity issues appeared when participants tried to connect with external devices using Bluetooth: "*Trouble pairing with computer. After much research on computer, figured out I needed dongle. Once dongle connected, able to connect to laptop, never to table top*". Regarding iconography on pedometers, "*I don't remember all the icon. Preference should be given what we used the most*." Inaccuracy regarding sleep and walk data was also reported by the older adult, which was in line with the previous study of [13]. Further, it was mentioned that the font size was too small to read: "*I cannot read the text with my reading glasses, can I make this font larger?*"

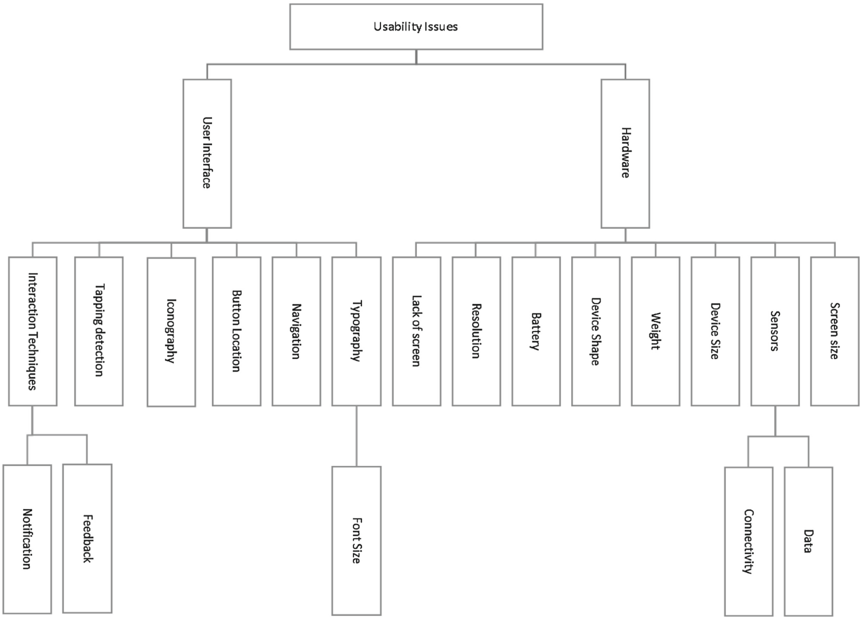



**Fig. 1.** Reported usability issues for smartwatches and pedometers

After the final data coding, the data were analyzed further based on the number of times the participants reported each usability issue. UI and hardware sub-categories, such as interaction techniques; tapping detection; iconography; button location; navigation; lack of screen; typography; screen resolution; battery; device shape, device size, weight, and size; sensors; and screen size, were considered. Statistics show (Figs. 2 and 3) the mean and standard deviations of the scores for the usability issues for both Finland and US.

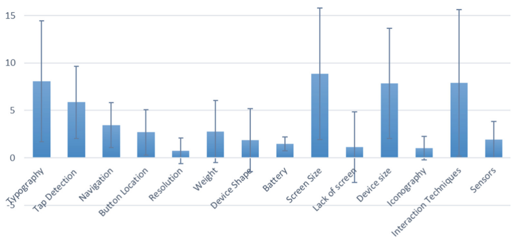



**Fig. 2.** Mean and standard deviation of usability issues for Finland

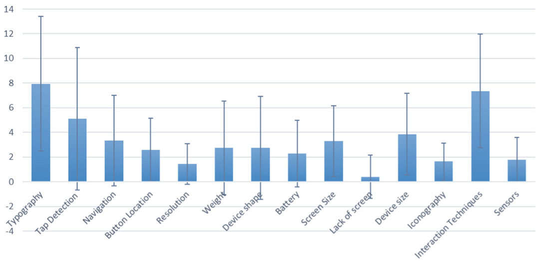



**Fig. 3.** Mean and standard deviation of usability issues for US

This outcome can point to the fact that for QSWTs, screen size, device size, interaction techniques, tapping detection, typography, and navigation were the most influential in both countries.

**Comparing Significant Relations Across Cultures.** A survey scale was constructed based on the identified most dominant usability to understand further whether the importance weights exist in the culture. Participants were asked to rate how much the identified usability issues correspond to the questionnaire on a 7-point Likert scale (0 = strongly agree to 7 = strongly disagree). The hypothesis was based on screen size, device size, interaction techniques, tapping detection, typography, and navigation. To gain insight, we surveyed with same participants from the usability studies. Data from the survey responses were downloaded from the Webropol online platform into an MS Excel spreadsheet, and they were analyzed with the R statistical language and its statistics ("stats") library [34]. Descriptive statistics were generated by the psych R library [36]. The Mann–Whitney U test for difference in means was used to test the differences between datasets. When analyzing the interval data with the Mann–Whitney U statistical test, a continuity correction was enabled to compensate for non-continuous data [7]. The Bonferroni correction was used to adjust the p-value to compensate for the family-wise error rate in multiple comparisons [1].

The Mann–Whitney U test results are summarized in Table 1. A total of six variables were tested and the difference of means between the two groups was significant between (i) the device size and (ii) the screen size. Other tested variables related to typography, navigation and interaction were not significant. For example, the Mann–Whitney U test indicated that issues related to device sizes has a significant difference ($U = 228$, $p < 0.001$) for Finnish ($Mdn = 4$), compared to US older participants ($Mdn = 6$).

The significant relationship result obtained from the two survey datasets (Finland and US) shows that some of the usability issues concerning wearable device characteristics are significantly essential and vary across cultures among older adult participants from Finland more than those from the US, which is in line with Khaslavsky's [23] statement. Khaslavsky [23] stated that two western countries may display different culturally motivated problems. For example, Mallenius et al. [27] found that Finnish individuals are interested in ease of use and value services provided by the devices that can make their everyday lives and tasks easier and safer.

**Table 1.** Results from Mann–Whitney U tests.

| Variable | U-value | Mdn (US) | Mdn (Finnish) | p | Corrected p (Bonferroni method)/significance |
|---|---|---|---|---|---|
| Device size | 228 | 6 | 4 | 0.0001727 | 0.001036233/Yes |
| Screen size | 284.5 | 5 | 4 | 0.003384 | 0.020305537/Yes |
| Typography | 539 | 2 | 3 | 0.5336 | 1.000000000/No |
| Interaction technique | 587.5 | 4 | 4.5 | 0.1909 | 1.000000000/No |
| Tapping detection | 654.5 | 5 | 6 | 0.02295 | 0.137708336/No |
| Navigation | 364 | 4 | 3 | 0.07015 | 0.420873408/No |

Therefore, one reason the results varied is because Finnish older individuals prefer more perceived comfort and convenience attributes [20] that could also include value services. For example, Finnish participants perceived comfort from wearability, convenience attributes from size and weight, and the value services from wearable devices as tools to facilitate behaviour changes to increase and maintain physical activity levels more than their US counterparts. Another reason could be the participants' cultural backgrounds; particularly, users' perceptions of effectiveness, efficiency, and satisfaction [46] may differ with regard to device and screen size.

In contrast, usability issues with wearables characteristics (i.e. UI), such as font size and feedback, showed no significant differences between both countries. One possible reason is that the importance weight tied to typography and feedback among older participants could usually be countervailed by (i) individual characteristics, such as age and health [10] or (ii) the local font characters, such as Chinese font characters, which are composed of strokes affecting readability [17]. However, when it came to the tapping detection threshold and navigation, the results were close to significant, and the closeness in the results can be attributed to either the differences in older adult or cultural characteristics. Therefore, this warrants further investigation with larger datasets.

## 5 Discussion

In his inspiring work, Carmien and Manzanares [8] state, "Identifying the cause of the non-adoption is the first step towards ameliorating this situation; having identified the problem the next step would be to design around the obstacles that were designed into the systems" (p. 28). This research has dealt with identifying the cause of adopting COTS QSWTs among the older adult (60+). Our results raise concerns from many angles of device characteristics, and they were in line with previously identified usability issues [3, 13]. To offset such concerns, the authors have recommended informal guidelines based on identified usability issues, qualitative feedback from the older adult, and existing literature reviews to help technology designers, developers, and researchers to design upcoming QSWTs targeted to the older adult. The most important set of informal guidelines is followed for both hardware and UI.

### 5.1 Enhancing Usability for Hardware

**Consider Configure-to-Order (CTO) Products.** During usability studies, the older adult indicated that the device and screen sizes presented the most dominant usability issues. This may be because screen size decreased the efficiency and processing of conveyed information [24] by limiting input modalities, navigation behavior, and readability, while the device shape limited wearability because of individual characteristics, such as variation in wrist shape and size due to the age, gender, or body structure. To reduce the impact on usability caused by device and screen size, technology designers should ensure both are sized appropriately for the older adult. We propose it is highly important for technology designers to consider applying CTO products with wide device and screen size measurement and shape variations, such as large/round, small/round, large/square, and small/square [24], depending on preferences.

**Consider Maximum Magnitude of Effect for Minimal Means.** Likewise, a visual pattern is pleasing to the eye when relatively simple design features reveal a wealth of information [15], so the hardware design of smartwatches and pedometers should be pleasing to the body through wearability and comfort that improve the aesthetic experience of older adult users. When it comes to the older adult, we found that wearable devices were used mainly for sleep analysis and counting heartbeats and steps, and those data were used to facilitate behaviour change. Therefore, technology designers should give careful consideration while designing the device characteristics (hardware and User Interface) that have a maximum magnitude of effect for minimal means. For example, removing unwanted hardware, such as the near field communication (NFC) and radiofrequency (RF) chips handling calls and texts, might reduce the shape, size, and weight of smartwatches and pedometers to be light and comfortable and to not affect older adults' daily behaviour.

**Consider Improving Sensor Precision.** Quantitative feedback showed that the older adult had difficulties trusting the reliability of notified sleep analysis data, such as wake after sleep onset (WASO), sleep efficiency (SE), and total sleep time (TST) with current smartwatches and pedometers. This may be because (i) current wearable devices measure the binary presence of sleep or waking states by measuring wrist movements [28] using wrist-worn accelerometer sensors [50] or because (ii) consumer health wearables are based on simple descriptive statistics [31]. Therefore, technology designers and researchers should consider alternative techniques to could improve notifications of sleep analysis data. This could be done by, for example: (i) identifying sleep stages: awake, light sleep, deep sleep, and rapid eye movement (REM) through RF [51]; (ii) measuring skin temperature, light, and activity across days to detect internal circadian rhythms [42, 47]; (iii) capturing entire body movements, rather than focusing on one specific body part; and (iv) detecting the complete set of motion-related parameters [30].

**Consider Culture While Designing the Devices.** From the Mann–Whitney U tests, we found that culture may have an influence when it comes to some user characteristics. For example, Finnish users place more importance on device and screen size than their US counterparts. Therefore, we recommend technology designers look into cross-culture design requirements and get feedback from local older adult using local usability evaluators before designing devices. Shi [40] states, “When cultural differences exist between the evaluator and test user, some usability problems might be masked, instead of being uncovered.” This may help to understand local users’ attitudes and intentions to use, as well as to build an effective relationship with them and their devices.

### 5.2 Enhancing Usability for UI

**Consider Alternative UIs.** The older adult indicated that the UI characteristics of devices, such as typography, button location, and interaction techniques, affected their daily usage. This may be because individual characteristics, such as age, disabilities, and environmental context, might have influenced the UI. A previous study [39] and qualitative feedback from the older adult also indicated, “There are so many things, which I do not need.” As [12] pointed out, simple interface manipulation can contribute to positive preference outcomes, and one basic approach to improving the UI’s usability is for technology designers and application developers to consider an alternative user interface (AUI) approach. We believe such a consideration could support User personalization, which the older adult desire. For example, adding age and any impairments during first time device or application start up could allow devices to personalize the typographical variables automatically, such as font size, font color, and background color, which could reduce the demands placed on accommodation (the eye’s ability to change its optical power for a better focus) and vergence (eye movement for focusing on near and far objects) [4]. As Morrison et al. [29] state, “Notifications appear to be most acceptable when users are provided with control over if, when, and how they are received, and when notifications are delivered at convenient times that do not disrupt daily routine.” Therefore, personalized information, daily behaviour data, and the environmental context of the UI should alert on device screen with any of the following interaction techniques: text, audio, graphic, tactile, and haptic [49]. The consideration of an alternative UI has implications for the “how information” for older adult at first; therefore, specific instructions should be presented during device or application startup.

## 6 Conclusion

Previous studies [13, 35, 39] have identified issues such as data inaccuracy, functionality of devices, consequences of use, wearability because of device characteristics as reasons associated with the withdrawal from the use of wearables by older adult. In this study, we first identified usability issues related to smartwatches and pedometers among the older adult between two countries. Second, we looked at whether culture weighs on usability issues. Finally, based on those issues, we provided informal usability guidelines that aim to help technology designers and application developers craft usable future COTS QSWTs for the older adult. This study involved a small number of participants in both geographic locations, meaning the results may not be generalizable; thus, all stakeholders, including device manufacturers and application developers, should take the findings as suggestions rather than conclusive evidence [22].

This study suggests several potential areas of improvement of QSWT for technology designers, researchers, and application developers. First, technological designers must be sensitive to individual and device characteristics and cultures that might impact device adoption by the older adult. Second, both technology designers and researchers must be sensitive to improving sensors and algorithms to avoid the potential consequences of inaccurate data that are currently occurring through wearable devices. Third, technology designers, researchers, and application developers must consider an AUI on both embedded operating systems and applications, so the older adult feel comfortable and develop a high degree of satisfaction, motivation, and enjoyment regarding these devices' usefulness [22].

While more research is needed to offset the usability issues caused by the smartwatches and pedometers among the older adult, future research could investigate the impact of: (i) low and high context cultures; (ii) local font characters, such as Chinese font characteristics, which are composed of strokes and which affect the readability [17] and usability of wearable devices; (iii) evaluate the significance of an AUI through a task-based experiment; (iv) analyze empirically the weights of different interaction techniques between smartwatches and pedometers; and (v) study how long it takes the older adult to learn device functionalities. The research can also be extended to analyze empirically older adults' perceptions of adopting the CTO approach.

**Acknowledgments.** First author would like to thank Miina Sillanpää Foundation, LUT Research Platform on Smart Services for Digitalisation (DIGI-USER) and Second author would like to thank Ulla Tuominen Foundation for their generous support of research.

## References

1. Abdi, H.: The Bonferonni and Šidák corrections for multiple comparisons. In: Salkind, N. (ed.) Encyclopedia of Measurement and Statistics. SAGE Publications Inc., Thousand Oaks (2007)
2. Ally, M., Gardiner, M.: Application and device characteristics as drivers for smart mobile device adoption and productivity. Int. J. Organ. Behav. **17**(4), 35–47 (2012)
3. Angelini, L., et al.: Designing a desirable smart bracelet for older adults. In: Proceedings of the 2013 ACM Conference Pervasive and Ubiquitous Computing Adjunct Publication, pp. 425–434 (2013)
4. Bababekova, Y., et al.: Font size and viewing distance of handheld smart phones. Optom. Vis. Sci. **88**(7), 795–797 (2011)
5. Barber, W., Badre, A.: Culturability: the merging of culture and usability. In: Proceedings of the 4th Conference on Human Factors and the Web, pp. 1–14 (1998)
6. Batsis, J.A., et al.: Use of a wearable activity device in rural older obese adults. Gerontol. Geriatr. Med. **2**, 233372141667807 (2016)
7. Bergmann, R., Ludbrook, J.: Different outcomes of the Wilcoxon—Mann—Whitney test from different statistics packages. Am. Stat. **54**(1), 72–77 (2000)
8. Carmien, S., Manzanares, A.G.: Elders using smartphones – a set of research based heuristic guidelines for designers. In: Proceedings of the 8th International Conference, UAHCI 2014 Held as Part of HCI International 2014 Heraklion, Crete, Greece, 22–27 June 2014, Part II, pp. 26–37 (2014)
9. Chen, K., et al.: Kansei design with cross cultural perspectives. In: Usability and Internationalization. HCI and Culture, pp. 47–56 Springer, Heidelberg (2007)

10. Chen, S.Y., Fu, Y.C.: Leisure participation and enjoyment among the elderly: individual characteristics and sociability. Educ. Gerontol. **34**(10), 871–889 (2008)
11. Chiu, C.-J., Liu, C.-W.: Understanding older adult's technology adoption and withdrawal for elderly care and education: mixed method analysis from national survey. J. Med. Internet Res. **19**(11), e374 (2017)
12. Cockburn, A., et al.: The effects of interaction sequencing on user experience and preference. Int. J. Hum. Comput Stud. **108**, 89–104 (2017)
13. Fausset, C.B., et al.: Older adults' use of and attitudes toward activity monitoring technologies. In: Proceedings of the Human Factors Ergonomics Society Annual Meeting, vol. 57, no. 1, pp. 1683–1687 (2013)
14. Gudur, R.R., et al.: Ageing, technology anxiety and intuitive use of complex interfaces. In: Lecture Notes in Computer Science (including Subseries, Lecture Notes in Artificial Intelligence Notes Bioinformatics), vol. 8119, Part 3, pp. 564–581 (2013)
15. Hekkert, P.: Design aesthetics: principles of pleasure in design. Psychol. Sci. **48**(2), 157–172 (2006)
16. Holzinger, A., et al.: Perceived usefulness among elderly people: experiences and lessons learned during the evaluation of a wrist device. In: Proceedings of the 4th International ICST Conference Pervasive Computing Technologies for Healthcare, pp. 1–5 (2010)
17. Huang, D.L., et al.: Effects of font size, display resolution and task type on reading Chinese fonts from mobile devices. Int. J. Ind. Ergon. **39**(1), 81–89 (2009)
18. Ivory, M.Y., Hearst, M.A.: The state of the art in automating usability evaluation of user interfaces. ACM Comput. Surv. **33**(4), 470–516 (2001)
19. Jeong, H., et al.: Smartwatch wearing behavior analysis. In: Proceedings of the ACM Interactive, Mobile, Wearable Ubiquitous Technologies, vol. 1, no. 3, pp. 1–31 (2017)
20. Kalantari, M.: Consumers' adoption of wearable technologies: literature review, synthesis, and future research agenda. Int. J. Technol. Mark. **12**(3), 274–307 (2017)
21. Kanis, H.: Usage centred research for everyday product design. Appl. Ergon. **29**(1), 75–82 (1998)
22. Khakurel, J. et al.: Living with smartwatches and pedometers: the intergenerational gap in internal and external contexts. In: GOODTECHS Conference Proceedings, Lecture Notes of ICST (LNICST), Pisa, Italy, pp. 31–41. Springer, Heidelberg (2018)
23. Khaslavsky, J.: Integrating culture into interface design. In: CHI 98 Conference Summary on Human Factors in Computing Systems, CHI 1998, pp. 365–366 (1998)
24. Kim, K.J.: Shape and size matter for smartwatches: effects of screen shape, screen size, and presentation mode in wearable communication. J. Comput. Commun. **22**(3), 124–140 (2017)
25. Liu, C.C., Wu, D.W., Jou, M., Tsai, S.J.: Development of a sensor network system for industrial technology education. In: Lytras, M.D., Ordonez De Pablos, P., Ziderman, A., Roulstone, A., Maurer, H., Imber, J.B. (eds.) WSKS 2010. CCIS, vol. 111, pp. 369–374. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16318-0_43
26. Macefield, R.: How to specify the participant group size for usability studies: a practitioner's guide. J. Usability Stud. **5**(1), 34–45 (2009)
27. Mallenius, S., et al.: Factors affecting the adoption and use of mobile devices and services by elderly people–results from a pilot study. In: 6th Annual Global Mobility Roundtable, vol. 31, p. 12 (2007)
28. Marino, M., et al.: Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep **36**(11), 1747–1755 (2013)
29. Morrison, L.G., et al.: The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial. PLoS ONE **12**, 1 (2017)

30. Ni, H., et al.: Non-intrusive sleep pattern recognition with ubiquitous sensing in elderly assistive environment. Front. Comput. Sci. **9**(6), 966–979 (2015)
31. Piwek, L., et al.: The rise of consumer health wearables: promises and barriers. PLoS Med. **13**, 2 (2016)
32. Preusse, K.C., et al.: Activity monitoring technologies and older adult users. In: Proceedings of the International Symposium Human Factors Ergonomics Health Care, vol. 3, no. 1, 23–27 (2014)
33. Pulli, P., et al.: User interaction in smart ambient environment targeted for senior citizen. Med. Biol. Eng. Comput. **50**(11), 1119–1126 (2012)
34. R Core Team. R: A Language and Environment for Statistical Computing (2017). https://www.r-project.org/
35. Rasche, P., et al.: Activity tracker and elderly. In: 2015 IEEE International Conference on IS - SN - VO - VL, Computing Information Technology Ubiquitous Computing Communications Dependable, Autonomic Secure Computing Pervasive Intelligence Computing, pp. 1411–1416 (2015)
36. Revelle, W.: psych: Procedures for Psychological, Psychometric, and Personality Research (2017). https://cran.r-project.org/package=psych
37. Rieman, J.: The diary study: a workplace-oriented research tool to guide laboratory efforts. In: Proceedings of the SIGCHI Conference Human Factors Computing System, CHI 1993, pp. 321–326 (1993)
38. Rosales, A., et al.: Older people and smartwatches, initial experiences. El Prof. la Inf. **26**(3), 457 (2017)
39. Schlomann, A., et al.: Means of motivation or of stress? The use of fitness trackers for self-monitoring by older adults. HeilberufeScience **7**(3), 111–116 (2016)
40. Shi, Q.: Cultural usability: the effects of culture on usability testing. In: Human-Computer Interaction, INTERACT 2007, Pt 2, Proceedings, pp. 611–616 (2007)
41. Sin, A.K., et al.: A wearable device for the elderly: a case study in Malaysia. In: Proceedings of the 6th International Conference on Information Technology and Multimedia, pp. 318–323. IEEE (2014)
42. Smarr, B.L., et al.: A wearable sensor system with circadian rhythm stability estimation for prototyping biomedical studies. IEEE Trans. Affect. Comput. **7**(3), 220–230 (2016)
43. Stanton, N.: Ecological Ergonomics: Understanding Human Action in Context. Taylor & Francis, Abingdon-on-Thames (1994)
44. Tractinsky, N.: Aesthetics and apparent usability: empirically assessing cultural and methodological issues. In: Proceedings of the Conference Human Factors in Computing System, pp. 115–122 (1997)
45. Wallace, S., et al.: Culture and the importance of usability attributes. Inf. Technol. People. **26**(1), 77–93 (2013)
46. Wallace, S., Yu, H.-C.: The effect of culture on usability: comparing the perceptions and performance of Taiwanese and North American MP3 player users. J. Usability Stud. **4**(3), 136–146 (2009)
47. Wirz-Justice, A.: How to measure circadian rhythms in humans. Medicographia **29**, 84–90 (2007)
48. Wohlin, C., et al.: Experimentation in Software Engineering. Springer, Heidelberg (2012)
49. Wu, Q., et al.: How fitness trackers facilitate health behavior change. In: Proceedings of the Human Factors Ergonomics Society Annual Meeting, vol. 60, no. 1, 1068–1072 (2016)
50. Yang, C.-C., Hsu, Y.-L.: A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors **10**(8), 7772–7788 (2010)
51. Zhao, M., et al.: Learning sleep stages from radio signals: a conditional adversarial architecture. ICML **70**, 4100–4109 (2017)

# An Immersive Environment for Experiential Training and Remote Control in Hazardous Industrial Tasks

Nicholas Caporusso[1,2(✉)], Luigi Biasi[1], Giovanni Cinquepalmi[1], and Vitoantonio Bevilacqua[1]

[1] Polytechnic University of Bari, Via Edoardo Orabona, 4, Bari, Italy
luigi.biasi91@gmail.com, cinquepalmigianni@gmail.com, vitoantonio.bevilacqua@poliba.it
[2] Fort Hays State University, 600 Park Street, Hays, USA
n_caporusso@fhsu.edu

**Abstract.** Inadequate training and lack of risk awareness are among the main causes of work-related accidents in manufacturing facilities and during hazardous tasks. In the last decade, Virtual Reality was introduced as an effective tool for improving individuals' skills in their tasks in several different contexts by simulating operating scenarios. Nowadays, advances in interaction paradigms and novel devices create opportunities to further enhance training and work safety. In this paper, we introduce an immersive system based on affordable wearable devices for providing on-the-job training. In addition to its advantages for instructional purposes, we discuss the results of an experimental study about the performance of system as an assessment tool for evaluating the presence of incorrect movements that lead to work-related conditions.

**Keywords:** Wearable · Haptics · Training · Safety · Industry 4.0

## 1 Introduction

In 2015, over 2.8 million workplace injuries were reported in the US alone, rating work-related nonfatal adverse events at approximately 3% of workers' population [1]. Similar figures have been reported in other countries, with an average rate of 1.12 fatal accidents per 100.000 full-time equivalent workers [2]. Recent statistics show a continuously descending trend since the last decade. However, specific sectors in the goods-producing industry, such as, construction and manufacturing, maintain a high-risk profile, with an average rate of 24.6 nonfatal accidents per 100 workers [1].

Indeed, hazardous tasks, which are inherently associated with specific risk factors, have a direct impact on the occurrence of work-related injuries. Furthermore, the literature clearly shows that employees are more likely to have an accident at work in their initial period of employment in a job; also, turnover is associated with higher rate of adverse events [3]. Among work-related illnesses and complaints, the most common self-reported accidents are musculoskeletal disorders, with conditions involving the upper limbs (and especially the hands) ranking second [Muggleton]. Additionally,


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 88–97, 2019.
https://doi.org/10.1007/978-3-319-94619-1_9

young employees and newcomers are a high-risk population especially because inadequate learning leads to habits and behaviors which potentially cause injuries and clinical conditions (e.g., carpal tunnel syndrome) [4].

As lack of awareness on potential risks is a major, yet preventable, cause of accidents and fatalities, proper instruction is crucial for reducing adverse events [5]. In this regard, Virtual Reality (VR) has been introduced as a viable tool for completing traditional training. The advantages of simulated environments include hands-on experience, fast design and scalable usage, and long-term cost reduction. In addition, VR can increase opportunities for delivering accurate training, and it has been demonstrated to improve learning performance.

In this paper, we introduce an immersive system based on wearable technology and on a novel haptic interface. The system is especially designed for delivering task-specific, experiential training in a simulated on-the-job situation. In addition to providing users with a more realistic interaction with the work environment, the proposed system can be utilized to acquire and analyze individuals' movements and to assess potentially hazardous behaviors which can lead to work-related conditions of the hand. Finally, we detail the results of a preliminary study about the performance of our system.

## 2 Related Work

In the last decade, several projects introduced the use of Virtual Reality (VR) and Augmented Reality (AR) for training purposes in different mission- and life-critical contexts. Applications, such as, fight and flight simulators, driving simulators, and virtual intensive-care units, have been successfully adopted in the military, automotive, and healthcare industries, respectively [6, 7]. Moreover, in the recent years, VR and AR have increasingly been utilized for job training, in other sectors, such as, public safety. For instance, the authors of [6] propose a simulated environment for improving spatial navigation in firefighters. The use of VR-based simulation in safety training programs has been explored in the construction and mining industry [8], and it has been subsequently adopted by the manufacturing sector, where simulations are utilized to deliver just-in-time training for operations and maintenance purposes in addition to safety. Several studies, such as, [9], focused on healthcare, demonstrated the effectiveness of practicing on virtual equipment and in training the workforce on complex procedures [6]. More sophisticated applications include cooperation with and remote control of equipment, industrial machinery, and robots, or for learning additive manufacturing processes [10]. Nevertheless, most of the actual implementations of VR and AR in the industry regarded environments and models in which trainees can watch instructional demonstrations and interact with elements in the 3d scene. Unfortunately, although they increase the outcome of training, their performance might suffer from limited engagement, which, in turn, might undermine their applicability to actual scenarios. For instance, the authors of [11] proposed the use of VR as a tool for personnel selection in the early 2000, though other systems (e.g., Social Networks) received more attention as a screening tool for human resources. Moreover, the advantages of implementing collaborative multi-team VR environments was described in [12], though further research and development are required to enable physical participation of members from multiple teams in the same space.

Recent development in interactive technology and the introduction of physically immersive systems, such as, Cave Automatic Virtual Environments (CAVE) enhanced simulations, as they provide trainees with a completely different experience with respect to conventional screen-based VR software and improved users' engagement level with tasks, individually and in teams [13]. CAVE systems enable navigating a virtual scenario by actually walking in it, which is especially beneficial in training for hazardous operations, and harsh and extreme environments. However, they suffer from several drawbacks, such as, issues in depth perception [14], high maintenance costs, and complex logistics. As a result, they are mainly employed in experimental studies. Conversely, head-mounted-displays (HMD) are the preferred technology for work-related contexts. In [15], the authors present an immersive virtual environment for training and collaboration based on low-cost technology. The system utilizes an HMD and a gesture detection device to create a portable, affordable interactive solution that might be utilized for on-the-job and just-in-time tasks.

Especially in workforce training, feedback is among the crucial aspects in the use of VR and AR [16, 17]: the presence of realistic response from the environment results in improved engagement and performance. Several research projects introduced devices especially designed for custom tasks, which have issues in industrialization and affordability. Conversely, we propose an immersive training system that integrates VR and low-cost wearable devices to improve workers' training and risk awareness. Our system is especially focused on evaluating wrong behavior in tasks to improve safety, reduce the risk of injuries, and prevent long-term consequences of incorrect movement.

## 3 System Design

The proposed system aims at enhancing standard protocols for VR-based on-the-job training by using affordable immersive technology (i.e., head-mounted-displays) to increase engagement in simulated work-related tasks. Moreover, our system integrates wearable devices to improve the outcome of current technology based on learning-by-doing approaches with a two-fold objective: (1) increase the realism of the simulation, and (2) accurately measure factors in the execution of tasks (e.g., the movement of the hand) which are relevant for preventing work-related injuries and long-term conditions. To this end, the system consists in a hardware/software platform that integrates the design of scenario-based VR simulations, devices for acquiring movement and for delivering feedback over multiple sensory channels (e.g., tactile), and a risk assessment module for the analysis of hazardous and repetitive tasks. As a result, trainees can enter the simulation using an HMD, they can actually realize the task using wearable devices that capture their movement and give them tactile feedback from the environment, and they can receive information about their proficiency. Simultaneously, the system analyzes their movement, evaluates the potential short- and long-term risks associated with their performance, and suggests prevention measures.

The architecture of the system is structured in a stimulation and an acquisition component. The former consists in a VR engine (i.e., Unity) that enable creating and navigating scenarios that represent the task to be accomplished by the trainee. Simulations can be displayed using standard screens or HMDs. Moreover, the VR engine includes easy-to-use modeling software that can be utilized for generating realistic simulations, and it has already been utilized in different scenarios, including hazardous tasks.

The acquisition component consists of motion tracking systems that can capture the movement of subjects realizing the task. For instance, images and videos of the torso or the entire body can be acquired using infrastructure-based motion tracking systems in combination with wearable reflective markers that enable collecting data points from specific regions of interest (e.g., the upper limbs), and they can be utilized for quantitative movement analysis. In addition to motion tracking, the proposed system supports multiple input/output devices, including ad hoc technology and, specifically, wearable peripherals that capture orientation and movement of the hand and of the fingers. For instance, in its first implementation, the system supports dbGLOVE [18], a wearable device that consist in a pad equipped with inertial sensors to capture acceleration and orientation of the hand over 9 degrees of freedom, and bending sensors that detect grasping and flexion of fingers (see Fig. 1). The device can accurately acquire the movement of multiple joints of the hand and of the fingers, and convert them into control signals. Users can interact with the environment by directly grasping, manipulating, and holding simulated objects with their hands.

**Fig. 1.** Acquisition device. Orientation and movement are converted into control signals that enable reproducing the hand in the simulation, in real-time.

Simultaneously, data acquired using the wearable device enable extracting motion and inertial patterns that represent the typical conditions that increase occurrence and damage of musculoskeletal conditions. The system supports in-presence and remote analysis: in addition to data points, a 3d model of the hand can replicate the movement of the subject, so that experts and physicians have qualitative and quantitative information to support their diagnosis and to suggest prevention strategies and correct movements. Also, the proposed system can integrate Myo [19], which acquires inertial components of motion and myoelectric signals from the forearm to calculate the movement of the hand and the fingers. Furthermore, the wearable interface acts as a stimulation device, because it provides detailed haptic feedback on sixteen points over the palm which can be utilized to represent objects, pressure, touch cues, and tactile icons. As a result, users can interact and operate in a more natural and realistic fashion compared to clicking on standard controllers. Simultaneously, they can test and learn the actual movements which will be required by their activity. Furthermore, they can get feedback on their actions, and learn how to realize them better, to avoid situations associated with higher risk of adverse events, injuries, and long-term conditions.

The system utilizes low-cost technology that enables its adoption directly at the workplace, where the availability of training tools has several advantages in addition to the possibility of using them for instructional purposes. Also, the system can be utilized continuously for task rehabilitation, after the training phase, to identify and mitigate the potential presence side effects of incorrect behavior caused by work experience. Furthermore, the system might be utilized for rehabilitation therapy of work-related conditions.

## 4 Experimental Study

In this section, we detail the results of a preliminary study about the efficacy of the proposed system. The objective of the study was two-fold: (1) evaluate the accuracy of the control device in acquiring and discriminating correct movement patterns from potentially hazardous actions; and (2) validate the performance of the proposed system as a training tool in terms of risk awareness and task proficiency over time. To this end, we designed two experimental tasks that were realized in a virtual environment simulating a simple scenario involving operational constraints and risk factors (Fig. 2).

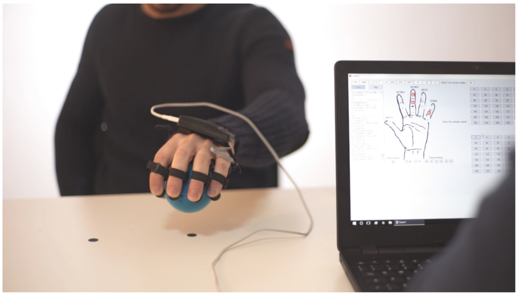


**Fig. 2.** Application of the system to in-presence risk assessment or rehabilitation contexts.

In Task 1, subjects were asked to operate a control room consisting of a knob which they had to close by rotating it clockwise in the least amount of time, to prevent spilling of hazardous material. In a preliminary experiment, we recruited 30 male participants who matched the characteristics of a novice worker. The purpose of the task was to evaluate the tendency of individuals of realizing movements that could lead to potential musculoskeletal conditions. We utilized the wearable component of the system to measure how subjects moved their hand during the tasks and we added some physical constraints to the virtual environment: users were required to grasp the virtual knob to rotate it (Fig. 3).

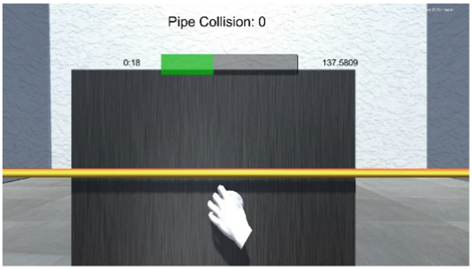



**Fig. 3.** The simulated environment of Task 2. Users are required to interact with a control panel in which they must accomplish the rotation of a knob placed relatively close to a hot metal bar.

In Task 2, participants were asked to accomplish the same operation as in the previous experiment. In addition, we included a hazardous component: a metal pipe was placed in close contact with the knob in the simulated environment, to simulate the presence of an element that would lead to an immediate accident at work. Therefore, subjects were required to realize more accurate movements and rotations of the knob to avoid incurring in an accident. 15 male participants aged 24–30 were recruited for this task.

Each subject repeated the task 5 times to evaluate the effectiveness of the system in delivering training. Acceleration and orientation acquired from the wearable device were converted into quaternions representing the rotation matrix, which was then utilized for real-time visualization of the hand and for calculating the following parameters:

- number, degree, and time of positive (clockwise) rotations of the hand, which counted towards accomplishing the task (both Task 1 and Task 2);
- number, degree, and time of negative (counter-clockwise) rotations of the hand, which involved additional effort and time to compensate the lost progress (both Task 1 and Task 2);
- number of collisions with the pipe (only Task 1), which represented incidents.

## 5 Results and Discussion

All participants successfully understood and completed the experiment. The device showed high accuracy in recognizing the rotation of the hand and the movement of fingers, though some modifications to its parameters were needed to increase its responsiveness with respect to the requirements of the task and to the individual configuration of the hand (i.e., changes in values acquired using bending sensors due to different length of participants' fingers).

In addition, by analyzing in detail subjects' behavior during the tasks, we were able to cluster individuals with respect to their probability of having an adverse event in an actual work environment. By doing this, we could suggest additional training to subjects depending on their identified risk class. As shown in Fig. 4, we identified 4 different types of behavior, which can be associated with specific intervention measures (Table 1).

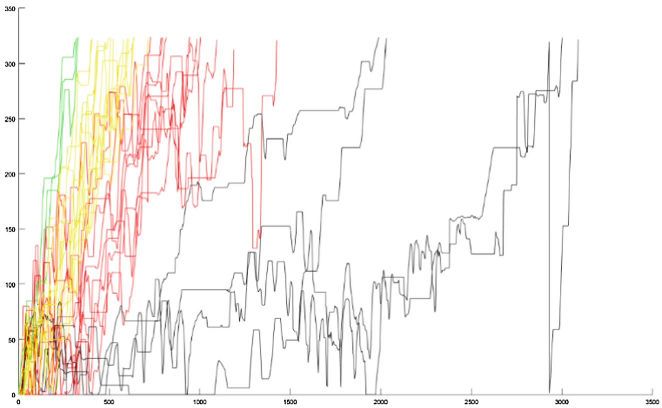



**Fig. 4.** Task advancement plotted as users' rotation patterns. Rotation angle is represented on the y axis, whereas the x axis shows time. Steady lines indicate intervals between rotations. The four groups are identified by different colors.

**Table 1.** Users' clusters and their centroids, with the distribution of the population, in relation to the angular coefficients of regression lines.

| User | Value | Users |
|---|---|---|
| C1 | 0.08 | 4 |
| C2 | 0.29 | 10 |
| C3 | 0.54 | 13 |
| C4 | 1.006 | 3 |

Also, the effectiveness of the system in training subjects is demonstrated by the number of collisions with the hot pipe, which is reduced to almost 0 after 5 trials, as shown in Fig. 5, which represents the performance improvement of participants.

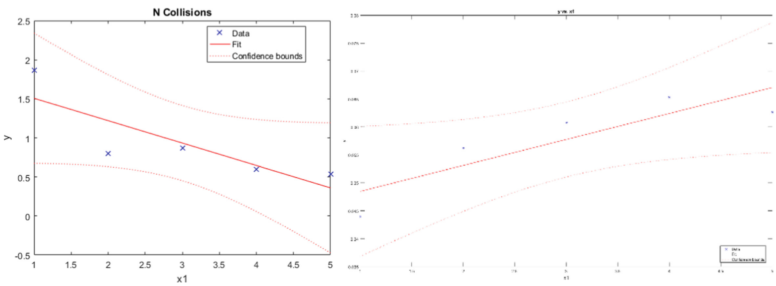



**Fig. 5.** Number of collisions with the metal bar over the 5 trials of Task 2 (left), and average overall performance of participants, which includes the time for accomplishing the task.

Table 2 summarizes the results for task 2: 54% of rotations involved smaller angles, associated with a lower risk, 7% regarded angles from 90° to 120°, associated with medium risk, and 39% of rotations were larger than 120°, which stress the wrist the most and require a rotation of the elbow, also. Interestingly, participants did not realize angles larger than 180°, which would involve a rotation of the torso, with additional risk for the back.

**Table 2.** Results of Task 2. Movements are represented as a percentage of rotation in regard to the angle.

| User | <90° | 90°–120° | 120°–180° | Time |
|---|---|---|---|---|
| s1 | 56% | 13% | 31% | 11 |
| s2 | 17% | 17% | 67% | 14.8 |
| s3 | 54% | 0% | 46% | 11.6 |
| s4 | 86% | 14% | 0% | 30 |
| s5 | 60% | 13% | 27% | 17 |
| s6 | 50% | 0% | 50% | 12.2 |
| s7 | 59% | 6% | 35% | 13.2 |
| s8 | 47% | 0% | 53% | 17.4 |
| s9 | 45% | 0% | 55% | 11.6 |
| s10 | 81% | 11% | 8% | 18.83 |
| s11 | 57% | 7% | 36% | 21 |
| s12 | 50% | 0% | 50% | 16.2 |
| s13 | 53% | 18% | 29% | 19.2 |
| s14 | 42% | 8% | 50% | 10.2 |
| s15 | 59% | 0% | 41% | 17 |

On average, participants were able to complete the experiment in 16 s. As shown in the table, most of subjects realized smaller rotations (i.e., smaller than 90°), which have low risk of musculoskeletal conditions, on the short term. Nevertheless, more than 30% of participants (e.g., s2, s6, and s8) are at risk, because they realize larger movements, which involve excessive rotations of the wrist and, thus, stress on ligaments that can potentially cause permanent conditions, on the long term. Moreover, by analyzing the time and the frequency of the movement, we could identify potential problems due to the repeated quick rotations, which also might be dangerous. Overall, we registered 66% improvement in task speed, and 83% increase in achieving the objective more safely.

## 6 Conclusion

According to the literature and to recent reports about workplace safety, of work-related injuries could be avoided by improving task training and by increasing workers' risk awareness [4]. Nevertheless, current training systems based on Virtual Reality and on the analysis of movement are not suitable for being employed in work settings due to their cost and lack of context adaptability.

In this paper, we described an immersive system consisting of wearable devices and, specifically, a novel haptic interface, for providing individuals with VR-based hands-on training on hazardous tasks. Moreover, we discussed the experimental study about the efficacy of the proposed system: preliminary data from a group of users in a simulated environment support its viability as a tool for improving individuals' safety using a learning-by doing approach.

Future work will include extensive trials in actual industrial scenarios, to evaluate the consistency of the training effect and the effectiveness of its portability to real-life contexts. Although in this work Task 2 simulated a highly dangerous maneuver, the system might have better applications in contexts where interaction involves less hazardous operations with risks are subtler and related to long-term conditions rather than immediate accidents. Moreover, several aspects of the proposed system will be investigated, such as, integration of the haptic component in infrastructure-based immersive environments, e.g., CAVE. Also, interesting directions for further research include human factors (e.g., evaluation of motion sickness) and physiological aspects, such as, comparison with movement patterns from individuals suffering from hand-related conditions.

## References

1. U.S. Bureau of Labor Statistics, Employer-reported workplace injuries and illnesses (2015)
2. UK Government Health and Safety Executive: Statistics on fatal injuries in the workplace in Great Britain (2016)
3. Burt, C.D.B.: New employee accident rates, pp. 9–22 (2015)
4. Breslin, F.C., Smith, P.: Trial by fire: a multivariate examination of the relation between job tenure and work injuries. Occup. Environ. Med. **63**, 27–32 (2006)
5. Muggleton, J.M., Allen, R., Chappell, P.H.: Hand and arm injuries associated with repetitive manual work in industry: a review of disorders, risk factors and preventive measures. Ergonomics **42**(5), 714–739 (1999)
6. Bliss, J.P., Tidwell, P.D., Guest, M.A.: The effectiveness of virtual reality for administering spatial navigation training to firefighters. Presence Teleoper. Virtual Environ. **6**(1), 73–86 (1997)
7. Ragazzoni, L., Ingrassia, P.L., Echeverri, L., Maccapani, F., Berryman, L., Burkle, F.M., Della Corte, F.: Virtual reality simulation training for Ebola deployment. Disaster Med. Public Health Prep. **9**(5), 543–546 (2015)
8. Van Wyk, E., De Villiers, R.: Virtual reality training applications for the mining industry. In: Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, pp. 53–63. ACM, February 2009
9. Patel, R., Dennick, R.: Simulation based teaching in interventional radiology training: is it effective? Clin. Radiol. **72**(3), 266-e7 (2017)
10. Renner, A., Holub, J., Sridhar, S., Evans, G., Winer, E.: A virtual reality application for additive manufacturing process training. In: ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. V01AT02A033–V01AT02A033. American Society of Mechanical Engineers, August 2015
11. Aguinas, H., Henle, C.A., Beaty Jr., J.C.: Virtual reality technology: a new tool for personnel selection. Int. J. Sel. Assess. **9**(1–2), 70–83 (2001)

12. Davis, M.C., Can, D.D., Pindrik, J., Rocque, B.G., Johnston, J.M.: Virtual interactive presence in global surgical education: international collaboration through augmented reality. World Neurosurg. **86**, 103–111 (2016)
13. Muhanna, M.A.: Virtual reality and the CAVE: taxonomy, interaction challenges and research directions. J. King Saud Univ. Comput. Inf. Sci. **27**(3), 344–361 (2015)
14. Ng, A.K., Chan, L.K., Lau, H.Y.: Corrective feedback for depth perception in CAVE-like systems. In: 2017 IEEE Virtual Reality (VR), pp. 293–294. IEEE, March 2017
15. Coburn, J.Q., Salmon, J.L., Freeman, I.: Effectiveness of an immersive virtual environment for collaboration with gesture support using low-cost hardware. J. Mech. Des. **140**(4), 042001 (2018)
16. Chang, T.P., Gerard, J., Pusic, M.V.: Screen-based simulation, virtual reality, and haptic simulators. In: Comprehensive Healthcare Simulation: Pediatrics, pp. 105–114. Springer, Cham (2016)
17. Zhao, D., Lucas, J.: Virtual reality simulation for construction safety promotion. Int. J. Inj. Control Saf. Promot. **22**(1), 57–67 (2015)
18. Caporusso, N., Biasi, L., Cinquepalmi, G., Trotta, G.F., Brunetti, A., Bevilacqua, V.: A wearable device supporting multiple touch-and gesture-based languages for the deaf-blind. In: International Conference on Applied Human Factors and Ergonomics, pp. 32–41. Springer, Cham (2017)
19. Taylor, K., Engsberg, J., Foreman, M.: Utility and usability of the MYO Gesture armband as a fine motor virtual reality gaming intervention. Arch. Phys. Med. Rehabil. **97**(10), e125 (2016)

# User Discomfort Evaluation Research on the Weight and Wearing Mode of Head-Wearable Device

Jun Zhuang(✉), Yue Liu, Yanyang Jia, and Yisong Huang

School of Design, Hunan University, Yuelu Mountain, Changsha, Hunan, China
zhangjun@hnu.edu.cn

**Abstract.** Augmented reality glasses are unlike regular glasses, its built-in camera, sensors and modules make it a small computing device capable of independent data processing. Because of the different product structures and wearing modes of AR glasses, its center of gravity and weight have become an important factor of the wearing experience. This research investigates mainly from the perspective of ergonomics and human-computer interaction, through the psychology and subjective experimental research under different structures and wearing modes, figuring out the relationship between users' subjective discomfort for AR glasses and weight, the threshold value for each wearing mode under different weight, also figuring out the relationship between discomfort rate and the load in each support points, finally put forward a design constraints guidance to enhance the AR glasses wearing experience. The experimental results show that users have different weight thresholds for each AR glasses with different structures and wearing modes. Different wearing modes under the same weight can affect users' feeling of discomfort.

**Keywords:** AR glasses · Discomfort rate · Human-computer interaction Wearing mode · Weight

## 1 Introduction

As of the development of augmented reality technology, Augmented Reality Glasses (AR glasses) have become the representative head-wearables in the consumer market. The experience of traditional head wearable devices is affected by the multidimensional effects of content, scene and human-computer interactive. However, during the design and evaluation process of product ergonomics, the weight of the product has become an important indicator for the experience and comfort of the head-wearable device due to the limitations of the existing technical architecture. AR glasses' batteries, display module, sensor and circuit boards and other hardware are usually installed on the front of the AR glasses [1]. The weight of module is large, user's nose, ears and the other support point of head will feel uncomfortable or compressed while wearing AR glasses, which affect the user's comfort experience. Similarly, in the use of head-wearable


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 98–110, 2019.
https://doi.org/10.1007/978-3-319-94619-1_10

device, the weight of the device can make user's head tilted down. User needs to use neck reverse power to keep one's normal head profile [2], and the excessive use of muscle force or pressure could cause discomfort or even damage cervical [3]. Bracklry et al. found that the weight and weight distribution of the schoolbag can change the standing and walking body posture of children, thus affecting the comfort of children when they are wearing school bags [4]. Therefore, weight is not only an important physical attribute of the product, but also one of the ergonomics elements that determine the degree of comfort in the process of using the product [5]. Changing et al. uses 3D glasses as an object to studies the user's subjective discomfort and nose affected by pressure, which found that user's subjective discomfort was reduced while the weigh-concentrated location moved from the front to the rear of temples [6]. However, the weight distribution and support point of the head of AR glasses is decided by wearing mode [7]. The existing research does not determine the wearing mode of AR glasses that with good user experience and the relationship between the weight change of AR glasses and subjective discomfort of users under different wearing mode. Therefore, if the wearing mode and weight of AR glasses can be determined in the early stage of new product development (NPD) to consider their effects on comfort, it will beneficial to speed up the efficiency of the development process and enhance the user experience.

The weight and weight distribution between support point of AR glasses will affect the user subjective comfort experience when using AR glasses [8], subjective evaluation is the only way to determine the user's comfort/discomfort and pain during use product [9]. In the related research, Kningt et al. used Borg-cr10 scale [10] to enable the testers to assess the neck muscle activity and discomfort and pain caused by head-mounted display (HMD) [11]. Takanori et al. obtained subjective discomfort during the test process by visual analogue scale (VAS) [12], and determined the influence of weight distribution of supporting point on user comfort when using 3D glasses [6].

Through two phases of experiment research. The preliminary experiment examined the effect of aggravating intervals on user subjective discomfort to determine a proper time for formal experiment. The formal experiment phase, in which continual weight adding experiments were conducted on experiment glasses models in four different wearing modes, yielded user endurance thresholds for actual glasses weight-adding in the different wearing modes. The experiment analyzed the correlation between user's subjective discomfort and glasses weight in the different wearing modes, meanwhile yielded pressure values of the corresponding supporting points, proving the correlation between pressure and users' subjective discomfort.

## 2 Preliminary Experimental Study – Confirm the Time of the Formal Experiment

### 2.1 Participants

10 subjects (female: 4, male: 6) were recruited. Their average age was 24.7. They didn't have any discomfort and medical problems on the experiment day. Eight of the subjects were spectacles wearers and two were not wearing spectacles. Before the experiment,

All the subjects fully understood the evaluation criteria of subjective discomfort, participants were provided with relevant documentation which including the purpose, process and precautions of the experiment, and signed informed consent form.

## 2.2 Experiment Materials

Through categorizing the AR glasses in the market and analyzing the feature of AR glasses that need to be accustomed to different scenes, 4 types of glasses with different structures and various wearing modes were selected to make models for the experiment (with different supporting points on the head): centers of gravity of the 4 types of glasses were calculated and weight increase components were made at the centers of gravity, meaning that several weight increase blocks with a weight of 9.25 g (which is the weight of a standard nut) were prepared when the basic weight of the experiment model was guaranteed at 120 g.

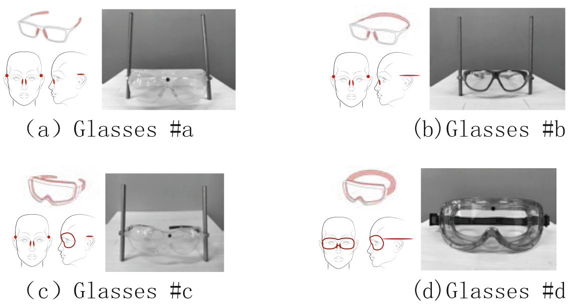



**Fig. 1.** Four different wearing mode glasses with different supporting points

As shown in Fig. 1, The support point of glasses #a (nose bridge, ear root points and head), the support point of glasses #b (nose bridge, ear root points and head), the support point of glasses #c (nose bridge, ear root points and head), the support point of glasses #d (nose bridge, ear root points, face and head).

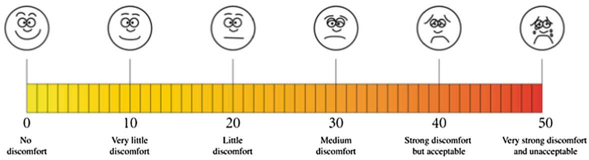



**Fig. 2.** Visual analogue scale/score (VAS)

As shown in Fig. 2, the user through the CP50 (50 = 0 = no discomfort, very strong discomfort, totally unacceptable) and the Visual Analogue Scale/Score (VAS) [11] on the subjective rating of discomfort.

### 2.3 Experiment Process

The evaluation of subjective discomfort of experimental subjects in the experiment may be related to the time of users in wearing glasses. In order to confirm that discomfort under different weights could be clearly differentiated during the experiment, glasses #a was taken as the pre-experiment model and the weight on 10 experiment subjects was increased continuously when they wore the experiment model under the three different time interval groups: t = 1 min, t = 3 min and t = 5 min. Next, evaluation scores of discomforts under each weight were obtained to analyze the trend of subjective discomfort in the three experimental groups, and therefore, to determine the most suitable time interval for the experiment.

### 2.4 Experimental Process

Experiment was carried out in a quiet room, with 10 experiment subjects numbered as #1, #2, #3, #4, #5… #10; subjects wore glasses #a respectively and were required to look straight ahead. Weight was increased at 1-min interval randomly and the overall subjective commenting scores of discomfort combined with cp50 and visual analogue scale (VAS) of pain under each weight were recorded respectively, and when the subjective discomfort reached 50, this group of experiment was over; these three experiments with different time intervals were carried out randomly so that the subjects would not be influenced by specific time length when making comments and it was also guaranteed that feelings of each group would not influence each other. There was a 30-min break between the groups to get the subjective evaluation scores of three groups of experiment subjects for the overall discomfort of glasses with different weight.

### 2.5 Results: Determine the Increased Time Interval

10 experiment subjects were assigned to do 3 groups of experiment respectively, and eventually, 218 pieces of subjective discomfort data were collected. As is shown in Fig. 3, values of discomfort corresponding to three weight increase time intervals show

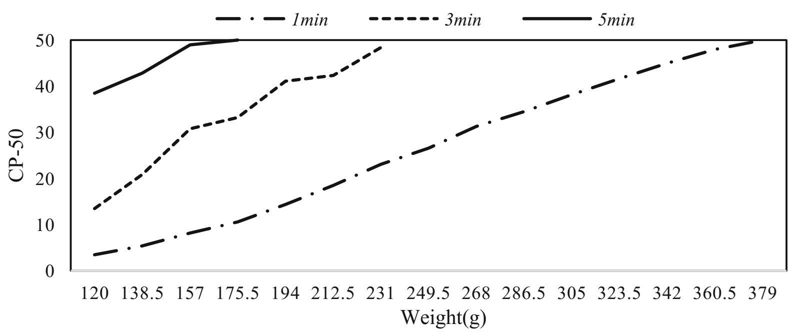



**Fig. 3.** The relationship between discomfort rating and weight under different time interval

different trends, in which the subjects failed to feel the two neighboring weights clearly when t = 1 min, and the score of discomfort reached 50 points under relatively low weight when t = 5 min. Since the experiment data samples were not enough and the discomfort of the subjects were greatly influenced by how long the glasses were worn, setting the weight increase time interval at 3 min was relatively suitable, which could make sure that the subjects can fully feel the discomfort under different weights while the wearing time is prevented from being the major factor of affecting users' discomfort.

## 3 Weight Evaluation Experiment - Determining the Relationship Between Subjective Discomfort and Weight

### 3.1 Participants

35 subjects (female: 16, male: 19) were recruited. Their average age was 25. They didn't have any discomfort and medical problems on the experiment day. 21 of the subjects were spectacles wearers and 14 were not wearing spectacles. Before the experiment, All the subjects fully understood the evaluation criteria of subjective discomfort, participants were provided with related instructions which including the purpose, process and precautions of weight evaluation experiment, and signed informed consent form.

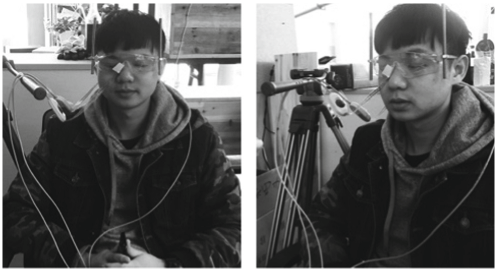


**Fig. 4.** FlexiForce standard model A201

### 3.2 Experiment Material

As shown in Fig. 2 of the 4 experimental glasses model, we use the standard nuts mass 9.25 g as experimental weighting, as shown in Fig. 4 of the three-channel pressure sensor of FlexiForce Standard Model A201, the thin film pressure sensor is fixed on the head supporting points, and the three channels corresponding to the pressure of glasses #a, #b, #c on nose bridge, ears and head, and glass #d on nose bridge, head and face, shown in Fig. 2 visual analogue scale.

### 3.3 Experimental Conditions

As shown in Fig. 1 of 4 experimental glasses #a, #b, #c, #d in 4 experimental group, which basis weight is 120 g, experimental subjects wear 4 glasses and continued to aggravate to t = 3 min time interval, then we gain the comfort evaluation scores and supporting points pressure value in 4 different weight groups.

### 3.4 Experiment Process

Experiment subjects were numbered as #a_1 (#a: the first type of glasses, 1: the first subject). Each experiment was carried out by 4 subjects synchronically, with each of them wearing one of the four glasses, glasses #a, glasses #b, glasses #c and glasses #d to conduct the experiment at the same time. Subjects were asked to remain seated on the chairs and keep their sight on a horizontal level to watch the video. The weight was increased or decreased randomly by n*18.5 g at an interval of t = 3 min. Experimenters made records of the overall subjective feeling combined with CP50 visual analog scale of discomfort as well as local discomfort on the supporting points, and measured the pressure on the supporting points at that time through pressure sensor. The added weight was not greater than the weight corresponding to 50 points of the maximum discomfort given by the subjects. After the subjects finished the test of a type of glasses and after they believed that they had already returned to a steady state both physically and mentally, they would test another wearing mode of glasses alternately until all the ten subjects completed the tests of all the four wearing modes of glasses.

## 4 Experimental Data and Analysis

### 4.1 Load-Bearing Thresholds Under Four Wearing Modes

Since scores of subjective discomfort were affected by individual differences, different subjects increased the weight for different times in the experiment, yet each subject would give the maximum score of discomfort D = 50 at certain weight. Therefore, mode is introduced to reflect the general weight-bearing level of the group, that is, when D = 50, the weight of the glasses at this point is chosen as the weight threshold Mmax of this type of glasses. As is shown in Fig. 5, subjective score D = 50 occurs for 10 times when M = 194 g for glasses #a. Therefore, weight threshold Mamax of glasses #a under the experiment conditions is (120, 194). Likewise, weight threshold Mbmax for glasses #b is (120, 212.5); weight threshold Mcmax for glasses #c is (120, 194) and weight threshold Mdmax for glasses #d is (120, 231).

### 4.2 Subjective Discomfort in Four Wearing Modes

In the experiment, overall discomfort score and that on each supporting point under the weight (M) of the glasses for each wearing mode of glasses were collected. D (overall discomfort), Dn (discomfort on the nose bridge), De (discomfort on the ear), Dh (discomfort on the head) and Df (discomfort on the face). In each experiment group, there were 35 groups of scores relating to discomfort. The relations of the weight (M) of glasses, subjective discomfort in general and on supporting points are analyzed according to mixed linear regression method:

$$\begin{cases} Da = -5.769 + 0.235 * Ma \\ Dan = 3.54 + 0.185 * Ma \\ Dae = -11.791 + 0.179 * Ma \\ Dah = -16.332 + 0.155 * Ma \end{cases} \tag{1}$$

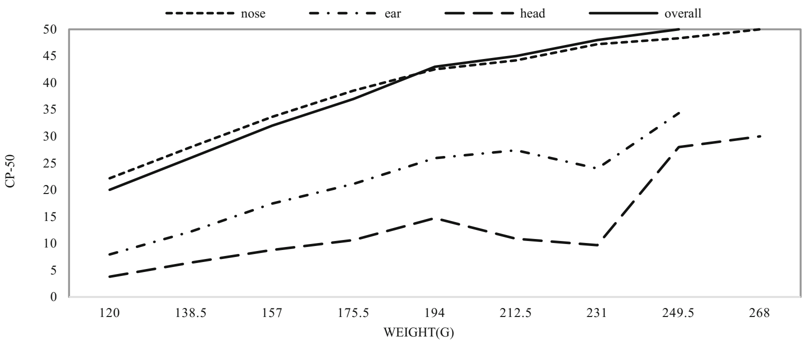
$$\begin{cases} Db = 1.112 + 0.181 * Mb \\ Dbn = -2.519 + 0.191 * Mb \\ Dbe = -15.387 + 0.184 * Mb \\ Dbh = -22.11 + 0.225 * Mb \end{cases} \tag{2}$$

$$\begin{cases} Dc = 5.777 + 0.173 * M \\ Dcn = 5.745 + 0.170 * Mc \\ Dce = -11.212 + 0.164 * Mc \\ Dch = -17.717 + 0.197 * Mc \end{cases} \tag{3}$$

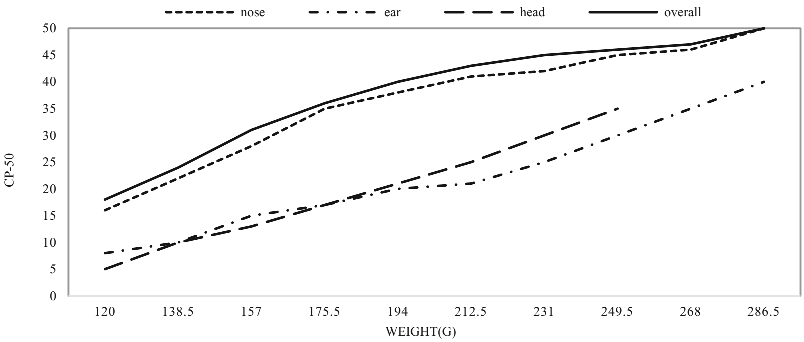
$$\begin{cases} Da = -5.769 + 0.235 * Ma \\ Dan = 3.54 + 0.185 * Ma \\ Dae = -11.791 + 0.179 * Ma \\ Dah = -16.332 + 0.155 * Ma \end{cases} \tag{4}$$

As shown in Fig. 5, the relationship between the subjective uncomfortableness and the weight of the various types of glasses is drawn into a function diagram. It is shown in the function relation diagram of linear regression that, influenced by the weight of the glasses model, overall subjective discomfort of experiment subjects increases gradually with the increase of the weight of glasses. In the meantime, discomfort on corresponding supporting points under different wearing modes also shows different trend of increase with the increase of model weight. We can thus infer that, users' discomfort when wearing AR glasses is influenced by the weight of glasses and also influenced by the weight distribution of supporting points.

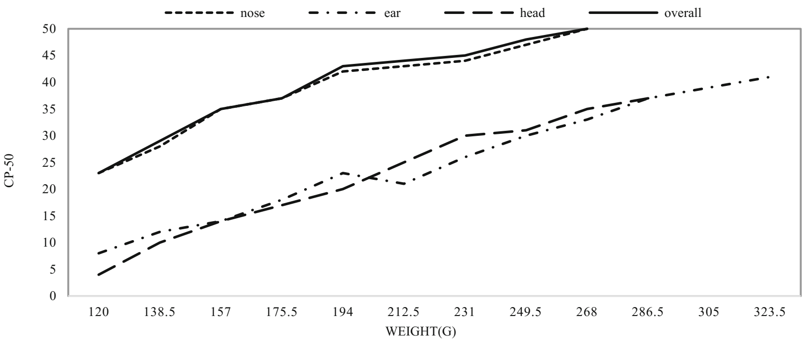
Besides, according to the function equation of regression and taking D = 50 as the limit value, corresponding theoretical limit weight (M) of glasses are calculated reversely. Ma = 237.3, Mb = 270.1, Mc = 255.6 and Md = 289.5, that is, Ma (120, 237.3), Mb (120, 270.1), Mc (120, 255.6) and Md (120, 289.5). The result is compared with the mode analysis result mentioned in the former part to get the intersection, which is users' weight bearing interval for experimental glasses: weight threshold Mamax for #a is (120, 194), weight threshold Mbmax for #b is (120, 212.5), weight threshold Mcmax for #c is (120, 194) and weight threshold Mdmax for #d is (120, 231).




(#a) Discomfort

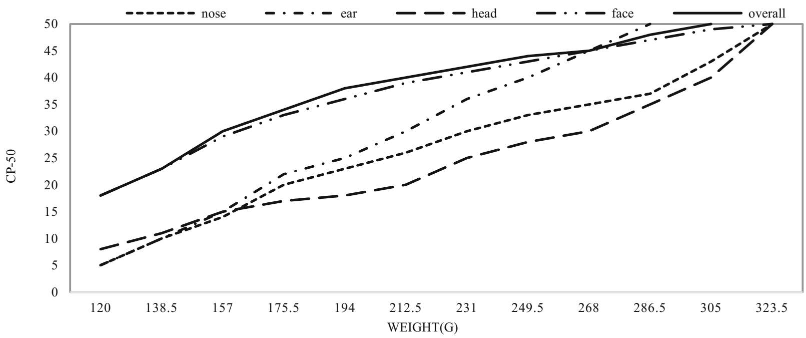



(#b) Discomfort




(#c) Discomfort

**Fig. 5.** Subjective discomfort and weight function of four glasses




(#d) Discomfort

**Fig. 5.** (*continued*)

### 4.3 The Correlation Between Subjective Discomfort and Corresponding Supporting Points Discomfort

Based on the aforementioned data, the overall subjective discomfort of the users' glasses wearing and the corresponding supporting point discomfort showed similar tendency. Therefore, it can be seen that as weight changes, overall subjective discomfort is affected by the corresponding supporting points discomfort - this is primarily shown by comparing average user rating of subjective discomfort on each supporting point, and using Pearson Correlation Analysis to compare and determine the statistical relations between subjective discomfort and glasses weight (Tables 1, 2, 3 and 4).

**Table 1.** Glasses #a discomfort rating and Pearson

| Object | Discomfort | Pearson |
|---|---|---|
| Overall | 37.63 | 0.982 |
| Nose | 39.401279 | 0.969 |
| Ear | 21.2975 | 0.949 |
| Head | 13.654444 | 0.84 |

**Table 2.** Glasses #b discomfort rating and Pearson

| Object | Discomfort | Pearson |
|---|---|---|
| Overall | 38.00 | 0.957 |
| Nose | 36.30 | 0.967 |
| Ear | 22.10 | 0.988 |
| Head | 19.50 | 0.998 |

**Table 3.** Glasses #c discomfort rating and Pearson

| Object | Discomfort | Pearson |
|---|---|---|
| Overall | 39.33 | 0.969 |
| Nose | 38.78 | 0.971 |
| Ear | 25.17 | 0.994 |
| Head | 22.30 | 0.994 |

**Table 4.** Glasses #d discomfort rating and Pearson

| Object | Discomfort | Pearson |
|---|---|---|
| Overall | 37.45 | 0.968 |
| Nose | 27.17 | 0.993 |
| Ear | 27.80 | 0.999 |
| Head | 24.75 | 0.975 |
| Face | 37.75 | 0.974 |

The comparison between the average subjective comfort score and overall comfort score of each supporting point can reflect the proportion of supporting point discomfort in overall discomfort experience, while the Pearson Coefficient shows the correlation between subjective discomfort rating and weight accordingly. By comparison, it is demonstrated that in the average subjective discomfort rating of Experiment Glasses Model #a, Dan = 39.4, Da = 37.63 - the two values are the largest and most similar. Also, r = 0.982 is the most similar to rn = 0.967 of nose bridge, which means there is a statistical relationship between the correlation of overall discomfort rating and weight, and the correlation between nose bridge discomfort and weight. This also means that the subjective discomfort of Glasses #a is mainly influenced by nose bridge discomfort. In the average subjective discomfort rating of Experiment Glasses Model #b, Dbn = 36.3, Db = 38 - the two values are the largest and most similar, and r = 0.957 is the closest to rn = 0.967 of nose bridge. Hence the subjective discomfort is mostly affected by nose bridge discomfort. In that of Glasses #c, it has been yielded that Dcn = 38.78, Dc = 39.33, and r = 0.969 while rn = 0.971 of nose bridge - thus the subjective discomfort is also mainly influenced by nose bridge discomfort; in that of Glasses #d, Ddf = 37.75, Dd = 37.45, in which r = 0.968, and rn = 0.974, so the subjective discomfort of Glasses #d is influenced by nose bridge discomfort as well.

### 4.4 The Relation Between Each Supporting Point Pressure of the Four Wearing Modes and the Subjective Discomfort

At the same time, in the experiment, the pressure values (P) of each supporting point of participants in different weight conditions were recorded - including Pn (pressure value on the nose bridge), Pe (pressure value on the ear root points), Ph (pressure value on the head), Pf (pressure value on the face). According to the Linear Mixed Regression Analysis, the relation between pressure value P of supporting points and glasses weight

M were analyzed. Integrated with Pearson Correlation Analysis, the influence of pressure on users' subjective discomfort was also identified.

$$\begin{cases} Pan = -0.345 + 0.015 * Ma \\ Pae = 4.647 + 0.005 * Ma \\ Pah = 0.661 + 0.006 * Ma \end{cases} \tag{5}$$

$$\begin{cases} Pbn = -1.213 + 0.017 * Mb \\ Pbe = 2.325 + 0.003 * Mb \\ Pbh = 1.625 + 0.010 * Mb \end{cases} \tag{6}$$

$$\begin{cases} Pcn = -0.329 + 0.016 * Mc \\ Pce = 1.732 + 0.009 * Mc \\ Pch = 1.046 + 0.003 * Mc \end{cases} \tag{7}$$

$$\begin{cases} Pdn = 0.195 + 0.005 * Md \\ Pdh = 1.695 + 0.004 * Md \\ Pdf = 2.859 + 0.014 * Md \end{cases} \tag{8}$$

By comparing the coefficients of P and M (i.e. comparing the correlations between the tendency of pressure values on different supporting points and weight changes), this research found that in Glasses #a: 0.015 > 0.006 > 0.005, so the influence of weight on the pressure change on each part for Glasses #a is ranked from the biggest to the smallest as: nose bridge, head, and ear root points; for Glasses #b: 0.017 > 0.010 > 0.003, therefore the weight influence on pressure changes on different parts, from big to small, is ranked as: nose bridge, head, and ear root points; that for Glasses #c is nose bridge, ear root points, and head as the coefficients yield 0.016 > 0.009 > 0.003; for Glasses #d, 0.014 > 0.005 > 0.004, so the weight influence rankings are: face, nose bridge and head.

According to the relationship of discomfort between the experimental subjects above and the supporting point, it verified that the overall subjective discomfort of glass #a, #b, #c is decided by the pressure of the main supporting point which is nose bridge, and the overall subjective discomfort of glass d is decided by the pressure of the main supporting point which is face.

## 5 Discussion and Discovery

This study verifies the relationship between subjective discomfort and the weight of the AR glasses which user wearing by the two phases experiment, and verify the relationship between overall discomfort and the pressure of main supporting points by the analysis of key supporting point pressure data. Thus, in the experimental conditions (basis weight 120 g), we gain the design standard of weight from the four kinds of wearing mode by the function relationship between subjective discomfort and the weight (Fig. 6).

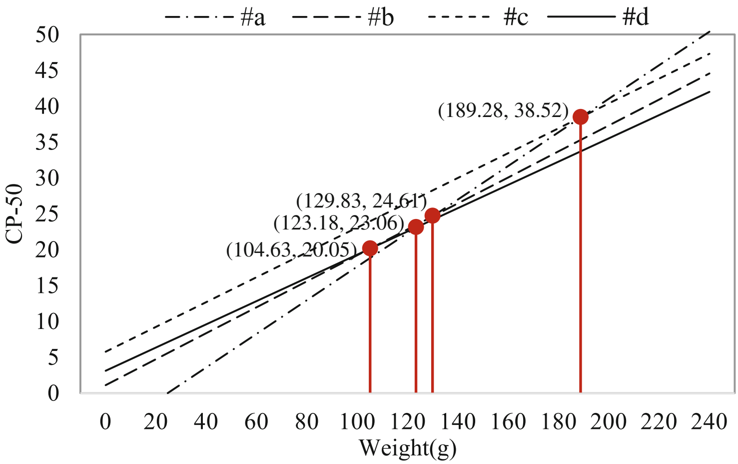



**Fig. 6.** The relationship between subjective discomfort and weight in four ways of wearing.

As the result shows, when M > 123.18 g in Glasses #d, the overall discomfort is the lowest among the four wearing modes, so the wearing mode in Glasses #d is deemed more reasonable when glasses weight is considerably heavy. META AR Glasses 2015 weigh 1.1 lb (=498 g), which is considered heavy, can be worn in the same mode as in Glasses #d - not only does this relieve some pressure of glasses weight on sensitive parts on the human face, but it also effectively distributes the weight, stabilizing the center of gravity thus ensuring the wearing stability. However, as hardware technology develops, products show a trend of light weight. For instance, Intel AR Glasses demonstrate the concept of light weight - hence designed according to the wearing mode in Glasses #a. As the experiment shows, when glasses main body M < 123.18 g, the wearing mode of Glasses #a has the lowest discomfort. In conclusion, when in different weight ranges, the priority in choosing a wearing mode (i.e. in choosing the lowest discomfort) also varies - this research provides design guidance and constraint conditions of using Dependent Linear Equation for improving wearing comfort of AR glasses and users' experience.

Besides, discomfort of wearing the glasses is also related to the pressure on each weight-bearing point, therefore, which can be relieved by optimizing and strengthening the product framework of AR glasses, adjusting the weight distribution of glasses and optimizing the material of the contacting surface between glasses and supporting points, and then making the glasses more comfortable to wear.

## 6 Conclusion

The present study identified the statistical relationships between the subjective disco fort and the weight of AR glasses under different wearing mode, and provided a design guidance for the appropriate wearing mode of AR glasses under different weights.

Besides, the present study investigated the pressure of supporting points in terms of subjective discomfort, determine the rational planning of the location of the various components to distract the pressure of the AR glasses which is the critical role to improving the comfort of AR glasses.

## References


1. Andrabi, S.J., Reiter, M.K., Sturton, C.: Usability of augmented reality for revealing secret messages to users but not their devices. In: SOUPS (2015)
2. Melzer, J., et al.: Helmet-Mounted Displays: Sensation, Perception, and Cognition Issues, pp. 805–848. U.S. Army Aeromedical Research Laboratory, Fort Rucker (2009)
3. Merkle, A.C., Kleinberger, M., Uy, O.M.: The effects of head-supported mass on the risk of neck injury in army personnel. Johns. Hopkins APL Tech. Dig. **26**(1), 75–83 (2005)
4. Brackley, H.M., Stevenson, J.M., Selinger, J.C.: Effect of backpack load placement on posture and spinal curvature in prepubescent children. Work **32**(3), 351–360 (2009)
5. Motti, V.G. Caine, K.: Human factors considerations in the design of wearable devices. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, Los Angeles (2014)
6. Chang, J., et al.: Effects of weight balance on a 3D TV shutter type glasses: subjective discomfort and physical contact load on the nose. Int. J. Ind. Ergon. **44**(6), 801–809 (2014)
7. Zhang, Z.-S., Cheng, X.-F.: Design of augmented reality glasses based on ergonomics. Packag. Eng. **38**(20), 61–66 (2017)
8. Spitzer, M.B., et al.: Eyeglass-based systems for wearable computing. In: First International Symposium on Wearable Computers. Digest of Papers. IEEE (1997)
9. Vergara, M., Page, A.: Relationship between comfort and back posture and mobility in sitting-posture. Appl. Ergon. **33**(1), 1–8 (2002)
10. Borg, G.: Borg's Perceived Exertion and Pain Scales. Human kinetics, Stockholm University (1998)
11. Chihara, T., Seo, A.: Evaluation of physical workload affected by mass and center of mass of head-mounted display. Appl. Ergon. **68**, 204–212 (2018)
12. Carlsson, A.M.: Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain **16**(1), 87–101 (1983)

# Preliminary Therapeutic Knee Band Development for Knee Osteoarthritis (KOA) Relief in the Elderly: A Pilot Study

Zidan Gong[1(✉)], Rong Liu[1], Winnie Yu[1], Thomas Wong[2], and Yuanqi Guo[3]

[1] Institute of Textiles and Clothing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
amygongzidan@163.com
[2] GINGER Knowledge Transfer and Consultancy Limited, Kowloon, Hong Kong
[3] Pok Qi Hospital-the Chinese University of Hong Kong Chinese Medicine Centre for Training and Research (Shatin), Shatin, Hong Kong

**Abstract.** Knee osteoarthritis (KOA) is the most common type of arthritis with a high prevalence especially among the elderly. A new concept of acu-magnetic therapy has been developed and designed into a therapeutic knee band for KOA relief. Therefore, aim of this study was to evaluate the effectiveness and feasibility of the acu-magnetic knee band for KOA relief in the elderly. A randomized controlled trial (RCT) has been conducted among 30 eligible subjects allocated to either a treatment group with six-week acu-magnetic intervention or a control group without intervention. The clinical effects was evaluated using the Western Ontario and McMaster Universities Osteoarthritis (WOMAC) Index, Berg Balance Scale (BBS) and knee range of motion (ROM). Results presented a superiority established in the treatment group on all measurements when compared with the control group. In conclusion, the acu-magnetic knee band was feasible and effective for KOA relief in the elderly.

**Keywords:** Acu-magnetic knee band · Knee osteoarthritis (KOA) The elderly

## 1 Introduction

Knee osteoarthritis (KOA) is a common type of arthritis with a high prevalence especially among the elderly and imposes a considerable medical and socioeconomic burden [1, 2]. More than one-third of the elderly aged 65 years old and above are suffering from the KOA with the main presenting symptoms of stiffness, pain, and mobility limitation [3, 4]. KOA could be the major cause of disability in the elderly and significantly influence their life quality [5, 6]. Moreover, due to the less social participation and the sedentary lifestyle resulted by KOA, those elderly are more likely to get mental or cardiovascular diseases [7]. Considering the particularity of the elderly including their poor health condition, rich medical history together with the limitations of conventional treatments such as side effect, drug dependence or drug interaction,


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 111–122, 2019.
https://doi.org/10.1007/978-3-319-94619-1_11

medical risk, etc., it is necessary to develop a suitable way for this group of people according to their need.

Acu-therapy is increasingly popular in treating many diseases including some musculoskeletal and connective tissue disorders and were reported to have a good therapeutic effect on KOA treating [8]. It belongs to the TCM (traditional Chinese medicine) system based on the theory of acupoint stimulation in order to achieve an internal balance of human body system, therefore, improve self-curative ability [9, 10]. It is anticipated that acu-magnetic garment therapy may help reliving the KOA suffering in the elderly. Comparing to the medical treatments, garment therapy is non-invasive, side-effect free, convenient and no need to apply external manual force. It is the first attempt to use acu-magnetic knee band to help the elderly to relieve KOA at home. This study aims to evaluate the effectiveness and feasibility of the acu-magnetic knee band for KOA suffering relieving in the elderly by using a newly designed therapeutic knee band.

## 2 Materials and Methods

### 2.1 Materials

In order to select the most suitable elastic fabric for the development of acu-magnetic knee band, the physical and mechanical properties of six different knitted fabrics (a)–(f) were tested. Table 1 indicates that Sample (e) is the lightest (251 g/m$^2$), the thinnest (0.90 mm) having the best thermal conductivity (Q-max: 0.09 W/cm$^2$) while Sample (c) is most air-permeable (25.08 cc/sec/cm$^2$). Both of these two knee band materials have good stretching performance. Moreover, a wearing test of knee bands in these two materials was conducted to evaluate the pressure (top, middle and bottom of the knee band), convenience, slippage, movement influence and general comfort sensation. 8 volunteers were required to wear the knee band for at least 2 h and do some daily life activities (sit down; stand up; sitting; walking; jogging; up and down stairs) before making assessment. The 11-box numerical rating scale was used for each item with the score ranging from "0" to "10" matching the scale of uncomfortable to comfortable. Considering the results from both the objective and subjective evaluation of the samples, Sample (e) was selected for the development of therapeutic knee band.

Based on both academic reviews and clinical advice, six acupoints including Ex-LE5 (Xiyan), SP 10 (Xuehai), LR 7 (Xiguan), GB34 (Yanglingquan), ST 35 (Dubi) and ST36 (Zusanli) on the human body were selected to apply acu-magnetic therapy [1, 11–21]. To apply the magnet therapy on selected acupoints in this study, medical nonwoven adhesive tape (Fig. 1a) was used to combine knee bands and oblate magnets (diameter: 18 mm; height: 2 mm; Fig. 1b) with the treatment points on the inner surface of the knee bands matching the acupoints on human body stated above. All of the magnets applied in this trial have been tested for their magnetic strength using the GM08 Gaussmeter (Hirst Magnetic Instruments Ltd. Fig. 1c) to ensure the strength ranging from 600 to 2000 Gauss. [22–26]

A preliminary acu-magnetic knee band was presented in Fig. 2 below. When subject wearing this therapeutic knee band Fig. 2a, locations marked in red in the knee

**Table 1.** Physical and mechanical properties of the fabrics tested

| Items | | | Sample | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | a | b | c | d | e | f |
| Weight (g/m$^2$) | | | 540 | 445 | 566 | 534 | 251 | 510 |
| Thickness (mm) | | | 1.94 | 1.77 | 2.05 | 2.07 | 0.90 | 1.52 |
| Air permeability (cc/sec/cm$^2$) | | | 18.43 | 18.82 | 25.08 | 18.82 | 17.60 | 21.69 |
| Thermal conductivity Q-max (W/cm$^2$) | | | 0.05 | 0.06 | 0.07 | 0.06 | 0.09 | 0.07 |
| Tensile | Lengthwise | LT (-) | 1.29 | 1.16 | 1.27 | 1.33 | 0.85 | 1.12 |
| | | WT (N/m) | 8.92 | 16.39 | 17.25 | 8.66 | 19.54 | 6.8 |
| | | RT (%) | 65.27 | 63.64 | 65.8 | 61.76 | 65.9 | 61.38 |
| | Widthwise | LT (-) | 1.26 | 1.08 | 1.26 | 1.25 | 0.74 | 1.24 |
| | | WT (N/m) | 10.23 | 5.39 | 13.25 | 11.31 | 17.03 | 9.29 |
| | | RT (%) | 71.26 | 64.73 | 74.70 | 68.63 | 69.28 | 70.04 |
| | Bias direction | LT (-) | 1.18 | 1.09 | 1.26 | 1.30 | 0.91 | 1.16 |
| | | WT (N/m) | 11.39 | 10.45 | 12.33 | 7.53 | 21.17 | 8.39 |
| | | RT (%) | 60.07 | 60.60 | 71.54 | 62.24 | 69.72 | 62.62 |

LT-Linearity; WT-Tensile energy per unit area; RT-Resilience
Testing instruments: Electronic balance, BX300, Shimadzu Corporation, Japan; Thickness gauge, Hans Baer AG, CH, Zurich; Air permeability tester, KDG instruments, Sussex, England; THERMOLABO II, KES-F7, KATO TECH CO., LTD; Tensile& shear tester, KESFB1-AUTO-A, KATO TECH CO., LTD.

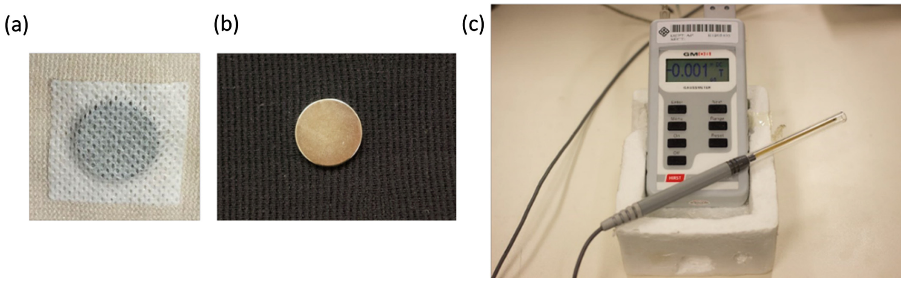



**Fig. 1.** Acu-magnetic treatment material: (1) Medical nonwoven adhesive tape to attach magnets to the knee band, (2) Oblate magnets, (3) GM08 Gaussmeter

band are treatment points which could apply magnetic and pressure stimulations on the selected acupoints on human body located on four meridians (Gall bladder, Stomach, Spleen and Liver) and an extra point of Ex-LE5 Xiyan (extra-meridian point). The distribution of treatment points on the acu-magnetic knee band was presented in Fig. 2b.

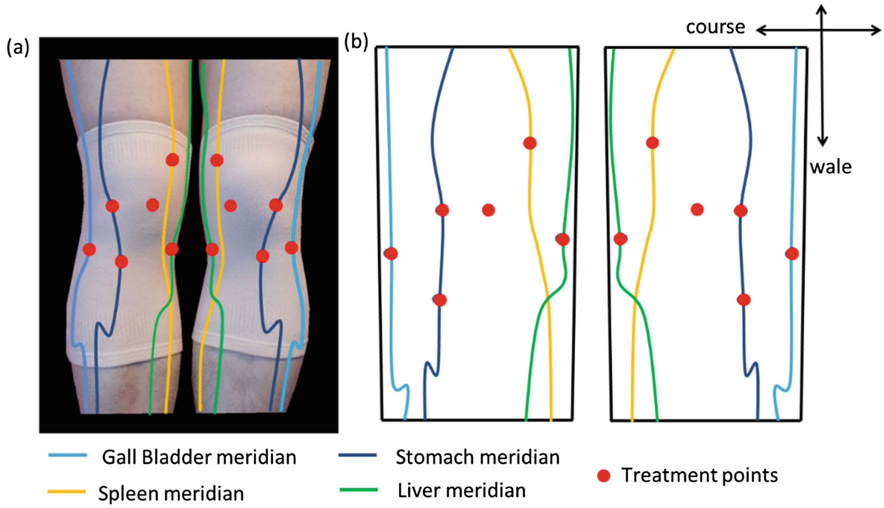



**Fig. 2.** The acu-magnetic knee band marked with acupoints along the four meridians of human leg

However, different KOA patients may have various bone sizes, leg circumferences, and meridian positions. For legs with larger or less dimensions, the magnets may not be always placed at the predetermined treatment points locations on the knee band. Therefore, different sizes of the acu-magnetic knee bands should be made with consideration of the dimensions of elderly legs, the elongation of fabric and the positions of acu-point to fit a wide range of KOA patients. After measuring subjects who were in small, medium and large size respectively to determine the locations of selected acu-points in different sizes, three sizes of therapeutic knee band have been developed. Ex-LE5 and ST 35 are located in the depression medial and lateral to the patella ligament when the knee flexed which are almost on the same horizontal line. Moreover, these two points are easy to locate accurately on human body and the distance between these two points have little difference (Mean + SD: 3.5 + 0.07 cm) among people with different leg sizes. Therefore, Ex-LE5 and ST 35 were set as the basic points with the same distance (3.5 cm) and height (12 cm from the top of the knee band) in all the three sizes (S, M and L) to help to identify other treatment points and ensure that all of other points are within the knee band base. For other four acupoints except Ex-LE5 and ST 35, location adjustment for various sizes was about 0.5 cm towards up or down and left or right moving from original selected points.

### 2.2 Methods

To evaluate the effectiveness and feasibility of the acu-magnetic knee band for KOA suffering relieving, a randomized controlled trial (RCT) was conducted involving a treatment group and a control group. Ethical approval has been obtained from the Hong Kong Polytechnic University. Participants were recruited strictly according to both inclusion and exclusion criteria then randomly (computer generated sequence) allocated to two groups with the ratio of 1:1. The KOA conditions in both groups including knee stiffness, knee pain, physical function, balance ability and knee range of motion were assessed at different time intervals (week 0, 2, 4 and 6) using a series of measurements.

**Participants**

The subjects in this study were recruited from an elderly home in Hong Kong, based on a set of criteria. Inclusion criteria were: (1) aged 65 years old and above; (2) have episodic pain caused by KOA in the previous six months; (3) have KOA-resulted lower limb mobility limitations. The exclusion criteria for participant recruitment were: (1) have other musculoskeletal issues that associated with the knee pain; (2) have inflammatory, metabolic or neuropathic arthropathies; (3) have severe concomitant illnesses that might interfere with the clinical evaluation of the patient; (4) have received acupuncture/acupressure therapy in the previous two weeks; (5) have cognitive impairment; (6) have skin allergy; (7) have wound or pressure sore on knee; (8) taking steroid drugs. The information sheet and consent form were signed by each subject and they were informed to have the right to withdraw at any stage of this study.

**Intervention**

Participants in the treatment group were instructed to wear the knee bands with inner acu-magnets at the six acupoints on their lower limbs. The researcher checked the fitting according to a guideline approved by the professor in Chinese medicine. Each subject was required to wear the therapeutic knee band for at least two hours every day and the therapy lasted for 6 weeks. In the control group, however, no acu-magnetic intervention was applied, patients just kept their usual care or previous treatment for KOA management.

**Measures and outcomes**

Before the intervention, each participant has completed a demographic questionnaire that contained three parts - demographic information, history of KOA and general health conditions. Any possible confounding factors such as past injury, medication or other therapies were identified. Primary outcomes were the self-reported knee conditions evaluated through the Western Ontario and McMaster Universities Osteoarthritis (WOMAC) Index. Considering the language factor of the samples in this research, a Chinese version WOMAC which has been validated by Xie et al. [27] was adopted. This validated questionnaire has 24 items focused on three domains-pain, stiffness and physical function. The higher score in each domain means the more pain and stiffness, and the poorer physical function. The balance ability of patients assessed by the Berg Balance Scale (BBS) was one of the secondary outcomes. This is a valid instrument to evaluate the effectiveness of any intervention and to make quantitative descriptions of function in clinical practice and research. Totally, there are 14 items in this measurement and a higher adding up score reflects a better balance ability. Another secondary outcome was the knee flexion range of motion (ROM) recorded using a goniometer with 30-cm movable arms. During the measuring process, the knee position was recorded where patients said ‘stop’ because of discomfort or threshold of pain, then read the degrees. Bigger value of recorded degree presents better knee range of motion.

**Data analysis**
Statistical analysis was performed using Statistical Product and Service Solutions (IBM SPSS Statistics 19). The normality of variables were previously evaluated and confirmed using the Shapiro–Wilk test. Paired T-test was conducted to test the significance of the score changes in WOMAC, BBS and ROM at different time intervals within or between groups. A significance level of .05 ($p < 0.05$) was set up for statistical threshold.

# 3 Results

## 3.1 Baseline Characteristics

There were 30 eligible patients completed the whole treatment from the baseline to the endpoint. Baseline characteristics of KOA patients in both treatment and control groups were presented in Table 2. In general, basic information of patients including age, BMI (Body Mass Index), KOA suffering duration was similar in both groups. Moreover, week 0 data in Table 3 and Table 5 showed that there was no significant difference between treatment and control group in stiffness ($P = 0.872$), pain ($P = 0.746$), physical function ($P = 0.831$), BBS ($P = 0.685$) and ROM (left, $P = 0.868$; right, $P = 0.597$) suggesting a similar baseline on all measurements.

**Table 2.** Baseline characteristics

| Variable | Treatment group (n = 15) | Control group (n = 15) |
|---|---|---|
| Age | 79.07 (12.39) | 76.47 (11.37) |
| BMI | 22.68 (2.70) | 23.66 (2.95) |
| Female sex, No. (%) | 9 (60) | 5 (33) |
| Symptom duration, year | 12.33 (9.11) | 12.67 (11.91) |
| 1 to $\leq 5$ y, No. (%) | 4 (27) | 6 (40) |
| 6 to $\leq 10$ y, No. (%) | 6 (40) | 4 (27) |
| >10 y, No. (%) | 5 (33) | 5 (33) |
| WOMAC-stiffness | 3.60 (1.35) | 3.67 (1.29) |
| WOMAC-pain | 8.13 (3.44) | 8.53 (2.53) |
| WOMAC-physical function | 29.47 (8.78) | 30.20 (8.16) |
| BBS | 29.07 (10.25) | 27.20 (12.22) |
| ROM-left knee, degree | 104.53 (11.38) | 103.80 (11.87) |
| ROM-right knee, degree | 105.53 (11.48) | 102.87 (13.63) |

Data were presented as mean (standard deviation, SD);
Data in Female sex and Symptom duration year range were presented as number (percentage, %)

### 3.2 Primary Outcomes

Table 3 presented the WOMAC score in each domain of both treatment and control groups at various time intervals and the difference between groups. After two weeks intervention of acu-magnetic therapy in treatment group, the results showed a significant difference between groups in stiffness (P = 0.043), pain (P = 0.006) but not in physical function (P = 0.276). Data recorded at week 4 indicated a more obvious difference between groups in stiffness (P = 0.000), pain (P = 0.000) and physical function (P = 0.023) and this significance persisted to the end sixth week with the result of P = 0.000, P = 0.000, P = 0.015 respectively in stiffness, pain and physical function. Moreover, score changes in WOMAC items during the whole treatment in both groups has been summarized in the Table 4. In the treatment group, there was a significant score change in stiffness (−2.07 ± 1.22, 95% CI [−2.74; −1.39], P = 0.000), pain (−4.40 ± 2.77, 95% CI [−5.94; −2.86], P = 0.000) and physical function (−7.33 ± 4.69, 95% CI [−9.93; −4.74], P = 0.000) implying a positive clinical effects on KOA relief. In the control group, however, no obvious improvement existed in stiffness (0.00 ± 1.51, 95% CI [−0.84; 0.84], P = 1.000), pain (−0.13 ± 2.33, 95% CI [−1.42; 1.15], P = 0.827) and physical function (−0.87 ± 4.10, 95% CI [−3.14; 1.41], P = 0.427).

**Table 3.** WOMAC score and difference between groups at different time intervals

| Week | WOMAC items | Treatment group Mean (SD) | Control group Mean (SD) | P value |
|---|---|---|---|---|
| Week 0 (baseline) | Stiffness | 3.60 (1.35) | 3.67 (1.29) | 0.872 |
| | Pain | 8.13 (3.44) | 8.53 (2.53) | 0.746 |
| | Physical function | 29.47 (8.78) | 30.20 (8.16) | 0.831 |
| Week 2 | Stiffness | 2.47 (1.13) | 3.40 (1.18) | 0.043 |
| | Pain | 5.40 (2.56) | 8.40 (2.26) | 0.006 |
| | Physical function | 25.13 (8.31) | 28.60 (7.35) | 0.276 |
| Week 4 | Stiffness | 1.80 (1.01) | 4.13 (0.99) | 0.000 |
| | Pain | 4.20 (1.82) | 9.47 (2.20) | 0.000 |
| | Physical function | 23.07 (8.20) | 29.93 (5.64) | 0.023 |
| Week 6 (endpoint) | Stiffness | 1.53 (0.74) | 3.67 (1.23) | 0.000 |
| | Pain | 3.73 (1.49) | 8.40 (1.84) | 0.000 |
| | Physical function | 22.13 (7.93) | 29.33 (5.51) | 0.015 |

**Table 4.** Changes in WOMAC items within groups from baseline to endpoint

| WOMAC items | Treatment group (n = 15) | | | Control group (n = 15) | | |
|---|---|---|---|---|---|---|
| | Mean (SD) | 95% CI | P | Mean (SD) | 95% CI | P |
| Stiffness | −2.07 (1.22) | (−2.74; −1.39) | 0.000 | 0.00 (1.51) | (−0.84; 0.84) | 1.000 |
| Pain | −4.40 (2.77) | (−5.94; −2.86) | 0.000 | −0.13 (2.33) | (−1.42; 1.15) | 0.827 |
| Function | −7.33 (4.69) | (−9.93; −4.74) | 0.000 | −0.87 (4.10) | (−3.14; 1.41) | 0.427 |

### 3.3 Secondary Outcomes

BBS score and ROM degree of treatment group and control group at different time intervals were showed in Table 5 as well as the difference between groups. The results indicated that there was no significant difference in BBS (P = 0.143) and ROM (left, P = 0.173; right, P = 0.147) between these two groups during the first two weeks of intervention. On the fourth week record, acu-magnetic intervention in the treatment group exerted significant impact on balance (P = 0.018) and knee range of motion (left, P = 0.003; right, P = 0.001) as compared with the control group. On the endpoint of the intervention, this significance still existed between groups in either BBS (P = 0.049) or ROM (left, P = 0.003; right, P = 0.003). After receiving the acu-magnetic therapy for six weeks, patients in the treatment group experienced an obvious improvement in balance ability (8.60 ± 6.06, 95% CI [5.25; 11.95], P = 0.000), left ROM (8.47 ± 6.60, 95% CI [4.81; 12.12], P = 0.000) and right ROM (7.47 ± 9.83, 95% CI [2.02; 12.91], P = 0.011) as presented in Table 6. However, in the control group, results showed that ROM decreased by 3.47° (P = 0.018) and 1.53° (P = 0.368) respectively on left and right knee. Surprisingly, the BBS score in control group was observed to increase significantly (2.13 ± 2.39, 95% CI [0.81; 3.45], P = 0.004) from baseline to endpoint suggesting a positive impact on balance ability.

**Table 5.** BBS and ROM difference between groups at different time intervals

| Week | Items | Treatment group Mean (SD) | Control group Mean (SD) | P |
|---|---|---|---|---|
| Week 0 (baseline) | BBS | 29.07 (10.25) | 27.20 (12.22) | 0.685 |
| | ROM-Left | 104.53 (11.38) | 103.80 (11.87) | 0.868 |
| | ROM-Right | 105.53 (11.48) | 102.87 (13.63) | 0.597 |
| Week 2 | BBS | 35.13 (8.58) | 28.40 (12.68) | 0.143 |
| | ROM-Left | 112.07 (10.98) | 105.67 (10.15) | 0.173 |
| | ROM-Right | 114.00 (9.82) | 108.00 (9.41) | 0.147 |
| Week 4 | BBS | 37.93 (7.28) | 27.47 (11.98) | 0.018 |
| | ROM-Left | 115.00 (8.86) | 102.00 (10.49) | 0.003 |
| | ROM-Right | 115.53 (7.41) | 101.20 (10.12) | 0.001 |
| Week 6 (endpoint) | BBS | 37.67(6.37) | 29.33 (12.51) | 0.049 |
| | ROM-Left | 113.00 (7.02) | 100.33 (10.77) | 0.003 |
| | ROM-Right | 113.00 (4.55) | 101.33 (11.41) | 0.003 |

**Table 6.** Changes in BBS and ROM within groups from baseline to endpoint

| Items | Treatment group (n = 15) | | | Control group (n = 15) | | |
|---|---|---|---|---|---|---|
| | Mean (SD) | 95% CI | P | Mean (SD) | 95% CI | P |
| BBS | 8.60 (6.06) | (5.25; 11.95) | 0.000 | 2.13 (2.39) | (0.81; 3.45) | 0.004 |
| ROM-Left | 8.47 (6.60) | (4.81; 12.12) | 0.000 | −3.47 (5.03) | (−6.25; −0.68) | 0.018 |
| ROM-Right | 7.47 (9.83) | (2.02; 12.91) | 0.011 | −1.53 (6.38) | (−5.07; 2.00) | 0.368 |

## 4 Discussion

Current outcomes extracted from the RCT suggested that acu-magnetic knee band was effective for KOA relief in the elderly. The result collected by the WOMAC, BBS and ROM test indicated that the 6-week acu-magnetic intervention has significant improvement on knee stiffness, knee pain, physical function, balance ability and knee range of motion. However, the effects on different items varied at different stages of the treatment. Compared with the control group, a more obvious improvement was obtained by the treatment group in each item on the last two records (week 4 and 6) than that recorded on week 2. Moreover, significant difference was first observed in stiffness (P = 0.043) and pain (P = 0.006) on week 2 while the significance in physical function (P = 0.023), BBS (P = 0.018) and ROM (left, P = 0.003; right, P = 0.001) began to appear in week 4. This may because the acu-magnetic therapy had quick effects on promoting vital energy (Qi) and blood circulation in meridians, relaxing muscles, remodeling microvascular, reducing oedema and inhibiting pain sensory. [9, 10, 28–31] Therefore, the knee stiffness and pain was significantly improved at the early stage of the intervention. With less stiffness and pain, the participants were more willing to move and do more daily activities, a better physical function ability was thus achieved as well as the balance ability and knee range of motion.

The stiffness has been measured subjectively in the WOMAC-stiffness domain, and measured objectively by the knee ROM measurements. [32] It showed a significant improvement in knee stiffness from a subjective perception after 6-week intervention and confirmed by the objective measurements from ROM test. However, the improvement on WOMAC stiffness was discovered to be more obvious than that on overall ROM. The subjective perception might bring some personal bias on the effect evaluation of the acu-magnetic therapy for KOA suffering relieving. Subjects who wore the knee band believing that they were receiving treatment and may expect to have clinical effects. Such psychological expectation may bring bias to the result especially to those self-reported outcomes. In addition, other factors such weather, number of clothes, health condition change, etc. could also influence the knee conditions and physical function of the patients, therefore, bring bias to objective measurements. As a result, future studies should set up more control conditions trying to rule out different kinds of research bias.

Although there was a superiority established in treatment group over the control group in all measurements, a significant improvement has been achieved by the control group in BBS score (P = 0.004) at the endpoint. This finding may partly due to the BBS questionnaire setting and test methods. BBS is widely used for the balance measurement of the elderly as a function performance task. This valid instrument was consisted of 14 fixed functional tasks such as standing to sitting, transfers, turning to look behind, placing alternate foot on stool, standing on one foot, etc. Participants in both groups were required to perform all these 14 tasks every two weeks and score for each task was given according to their performance finally adding up to a BBS score. Totally, four times need to be tested during the whole RCT conduction thus participants would be experienced in performing these tasks which may contribute the positive result found in BBS.

The current study has several potential limitations. Firstly, the RCT was not blinded to either the patients or the researchers. It is impossible to mask the patients because the treatment group wearing a knee band but no intervention adopted by the control group. In addition, researchers were not blinded since they need to give instruction to treatment group on how to wear the acu-magnetic knee band accurately and collect data from both groups. Secondly, psychological factors may play a role in the positive results of the treatment group. A sham group need to be set up in future RCTs in order to explore the placebo effect. Moreover, the therapeutic knee band used in this study was a preliminary design that needs to be further refined. In the future, the knee bands should be embedded with oblate magnets to increase the convenience and long-term usability. Although there is a size setting (S, M and L) for different leg sizes of KOA patients, further adjustment on the hem edges will be needed to fit more patients.

## 5 Conclusion

This study has demonstrated the effectiveness and feasibility of the acu-magnetic knee band for KOA relief in the elderly, in terms of knee stiffness, knee pain, physical function, knee range of motion and balance ability improvement. We believed that after refinement, this new application could have tremendous potential to be widely accepted in the market and significantly support the home treatment which is highly urged by the elderly. Moreover, this therapeutic garment may not only provide promising methods to relieve KOA suffering and improve mobility but also give future directions for symptom management and functional maintenance in KOA in order to improve the life quality of the elderly KOA sufferers.

**Acknowledgments.** The authors would like to thank Dr. Francis Wong Kam Hung for giving professional advice on RCT design and the use of measurement devices. We also acknowledged support from the Chang Qing Elderly Home and all the participants in this study. This work was supported by the postgraduate program 88011 of the Hong Kong Polytechnic University.

## References

1. Ashraf, A., et al.: Comparison the effect of lateral wedge insole and acupuncture in medial compartment knee osteoarthritis: a randomized controlled trial. Knee **21**, 439–444 (2014)
2. Cho, H., GN, K., Kang, J., Suh, K., Kim, T.: Epidemiological characteristics of patellofemoral osteoarthritis in elderly Koreans and its symptomatic contribution in knee osteoarthritis. Knee **23**, 29–34 (2016)
3. Tevald, M., Murray, A., Luc, B., Lai, K., Sohn, D.: The contribution of leg press and knee extension strength and power to physical function in people with knee osteoarthritis: a cross-sectional study. Knee **23**, 942–949 (2016)
4. Zhang, W., et al.: OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthr. Cartil. **15**, 981–1000 (2007)
5. Ringdahl, E., Pandit, S.: Treatment of knee osteoarthritis. Am. Fam. Physician **83**, 1287–1292 (2011)

6. Dessery, Y., Belzile, É., Turmel, S., Corbeil, P.: Comparison of three knee braces in the treatment of medial knee osteoarthritis. Knee **21**, 1107–1114 (2014)
7. Tse, M., Wan, V., Ho, S.: Physical exercise: does it help in relieving pain and increasing mobility among older adults with chronic pain? J. Clin. Nurs. **20**, 635–644 (2011)
8. Ezzo, J., Hadhazy, V., Birch, S., Lao, L.: Acupuncture for osteoarthritis of the knee: a systematic review. Arthritis Rheum. **44**, 819–825 (2001)
9. Mehta, P., Dhapte, V., Kadam, S., Dhapte, V.: Contemporary acupressure therapy: adroit cure for painless recovery of therapeutic ailments. J. Tradit. **7**, 251–263 (2017)
10. Kwan, R., Leung, M., Lai, C.: Acupressure for agitation in nursing home residents with dementia: study protocol for a randomized controlled trial. Trials **15**, 410 (2014)
11. Chen, X., Spaeth, R., Retzepi, K., Ott, D., Kong, J.: Acupuncture modulates cortical thickness and functional connectivity in knee osteoarthritis patients. Sci. Rep. **4**, 6482 (2014)
12. Appleyard, I., Crichton, N., Robinson, N.: Warm needle acupuncture vs. needle acupuncture for osteoarthritis of the knee: a pilot study protocol. Eur. J. Integr. **8**, 407–413 (2016)
13. Sorour, A., Ayoub, A., El Aziz, E.: Effectiveness of acupressure versus isometric exercise on pain, stiffness, and physical function in knee osteoarthritis female patients. J. Adv. Res. **5**, 193–200 (2014)
14. Chen, L., Mao, J., Fernandes, S.: Integrating acupuncture with exercise-based physical therapy for knee osteoarthritis: a randomized controlled trial. J. Clin. **19**, 308 (2013)
15. Spaeth, R., Camhi, S., Hashmi, J., Vangel, M.: A longitudinal study of the reliability of acupuncture deqi sensations in knee osteoarthritis. Evidence-Based 2013 (2013)
16. Mavrommatis, C., Argyra, E., Vadalouka, A., Vasilakos, D.: Acupuncture as an adjunctive therapy to pharmacological treatment in patients with chronic pain due to osteoarthritis of the knee: a 3-armed, randomized, placebo-controlled trial. PAIN® **153**, 1720–1726 (2012)
17. Jubb, R., Tukmachi, E., Jones, P.: A blinded randomised trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupunc. Med. **26**, 69–78 (2008)
18. Scharf, H., Mansmann, U., Streitberger, K.: Acupuncture and knee osteoarthritis: a three-armed randomized trial acupuncture and knee osteoarthritis. Ann. Intern. Med. **1**, 41–42 (2006)
19. Foster, N., Thomas, E., Barlas, P., Hill, J., Young, J.: Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. BMJ **355**, 436 (2007)
20. Berman, B., Lao, L., Langenberg, P.: Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized controlled trial. Ann. Intern. Med. **141**, 901–910 (2004)
21. Berman, B., Singh, B., Lao, L.: A randomized trial of acupuncture as an adjunctive therapy in osteoarthritis of the knee. Rheumatology **38**, 346–354 (1999)
22. Alfano, A., Taylor, A., Foresman, P.: Static magnetic fields for treatment of fibromyalgia: a randomized controlled trial. J. Altern. Complement. Med. **7**, 53–64 (2001)
23. Hinman, M., Ford, J., Heyl, H.: Effects of static magnets on chronic knee pain and physical function: a double-blind study. Altern. Ther. Health Med. **8**, 50 (2002)
24. Kanai, S., Taniguchi, N., Kawamoto, M., Endo, H.: Effect of static magnetic field on pain associated with frozen shoulder. Pain **16**, 173–179 (2004)
25. Wolsko, P., Eisenberg, D., Simon, L.: Double-blind placebo-controlled trial of static magnets for the treatment of osteoarthritis of the knee: results of a pilot study. Altern. Ther. Health Med. **10**, 36–43 (2004)
26. Harlow, T., et al.: Randomised controlled trial of magnetic bracelets for relieving pain in osteoarthritis of the hip and knee. BMJ **329**, 1450–1454 (2004)

27. Xie, F., Li, S., Goeree, R., Tarride, J., O'Reilly, D.: Validation of Chinese Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual. Life **17**, 595–601 (2008)
28. Melzack, R., Wall, P.: Pain mechanisms: a new theory. Surv. Anesthesiol. **11**, 89–90 (1967)
29. Moyer, C., Seefeldt, L., Mann, E., Jackley, L.: Does massage therapy reduce cortisol? A comprehensive quantitative review. J. Bodyw. Mov. Ther. **15**, 3–14 (2011)
30. Berman, B., Birch, S., Cassidy, C., Cho, Z., Ezzo, J.: Clinical Acupuncture: Scientific Basis. Springer, Heidelberg (2011)
31. Wong, M., Shen, H.: Science-based mechanisms to explain the action of acupuncture. J. Assoc. Tradit. Chinese Med. **17**, 5–10 (2010)
32. So, B., Kong, I., Lee, R., Man, R.: The effect of Ai Chi aquatic therapy on individuals with knee osteoarthritis: a pilot study. J. Phys. Ther. Sci. **29**, 884–890 (2017)

# Step by Step – Users and Non-Users of Life-Logging Technologies

Chantal Lidynia(✉), Philipp Brauner, Laura Burbach, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus Boulevard 57, 52074 Aachen, Germany
{lidynia, brauner, burbach, ziefle}@comm.rwth-aachen.de

**Abstract.** A pronounced deficit of physical activity is one of the challenges in today's societies. Lacking the minimum of activity recommended for a healthy lifestyle can be avoided by so-called life-logging technologies. However, usage is still low. To understand what factors contribute to an acceptance and use of these technologies, we conducted a quantitative online study with users and non-users. In total, 412 people have participated, 225 of them active users of life-logging technologies and 187 non-users. It was found that individual user characteristics shape its acceptance. For instance, the goals for possible behavior change, which the use of life-logging devices can support, differ significantly between users and non-users. Furthermore, the study reveals that factors such as age, motives for physical activity, and privacy concerns are key determinants for projected acceptance of life-logging technologies.

**Keywords:** Persuasive technology · Privacy · User modelling Quantified-self · Consumer Health Information Technology

## 1 Introduction

The spread of online services, be it the internet, mobile apps, or smart devices, has enabled a lot of positive and empowering opportunities. By now, almost everyone can consult the internet, book vacations, or reach friends anytime and anywhere.

However, the ongoing digitalization has also led to a lot of sedentary behavior, be it in the workplace or in leisure time. People spend a good amount of time sitting at a desk or hunched over a small mobile device like a laptop, a smartphone, or a tablet PC, cf. [1, 2]. Nevertheless, while readily available online services might be a facilitating factor to creating health risks such as obesity, hypertension, bad posture, and the like, they also offer a possible solution to avoid or at least minimize these risks. So-called Consumer Health Information Technology (CHIT) also allows for the opportunity to keep a close eye on one's own life. It has become easy and almost effortless to keep track of step count, calorie intake, and even the path one has walked during the day.

Although the benefits of physical activity to reduce health risks have been proven, see, for example, [3, 4], the use of apps or wearables to keep track of one's activities and caloric intake is still not widely spread, even though market estimates put the sale of wearables in Germany soaring in the past few years, e.g., [5, 6]. The question then


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 123–134, 2019.
https://doi.org/10.1007/978-3-319-94619-1_12

arises, why? What are reasons such devices and apps are not more widely used? What factors – be it demographic, psychological, or technological – prevent the adoption of so-called life-logging, i.e., regularly recording aspects of one's life to understand current behavior and improve habits [7, 8], even though its benefits have already been attested, e.g., [9]?

The present study aims at determining the answers to those questions. Our main contribution is threefold: First, we show which goals people want to pursue and for which goals they desire technological assistance. Second, we identify factors that differ between current users and non-users of life-logging technologies. Third, we show which user factors and which perceptions about life-logging technologies contribute to a likely adoption of life-logging technology by current non-users.

## 2 Related Work

A cornerstone of the present study is that of technology acceptance. Previous studies in different contexts have found that the adoption of a technology, that is, the later use of a product or service, is largely depended on its acceptance [10–12]. Acceptance, in turn, depends on several other factors: on the one hand, there are user factors such as age, gender, or technology affinity. On the other hand, there are factors pertaining to the technology itself such as ease of use or perceived usefulness. Davis et al. found a strong relationship between *intention* and *action*, meaning the intention to use a device or service can predict the actual use later [13]. This is the basis for the Technology Acceptance Model (TAM).

Earlier, the Theory of Reasoned Action (TRA) by Fishbein and Ajzen has shown a strong relationship between the *intention to perform* a specific behavior (i.e., adopting a specific product or service) and the actual behavior [14]. According to the TRA, the intention towards the behavior is governed by the individual's attitudes and subjective norms towards the behavior.

Another theoretical framework that links intention with action is the Theory of Planned Behavior (TPB) [15]. Although other factors directly influence the intention, once the intent to display a behavior, be it the use of a technology or physical exercise, is strong enough, the actual performance is more likely to follow.

While the previously cited works lay the theoretical groundwork for technology acceptance and behavior change, there have also been several studies within the specific context of wearables and fitness applications. Recurring observations have been privacy concerns [16, 17] or lack thereof [18], and abandonment or discontinuance of use [19–21]. Lidynia et al., for example, have found that a general interest in fitness trackers and wearables, especially in the data collected, is present [22]. However, they also found that the concern for privacy was one of the biggest barriers to actually using life-logging devices and apps. As mentioned before, perceived usefulness plays a large role in technology acceptance and adoption. Consequently, several studies have focused on the accuracy of commercially available devices, e.g., [23–26]. It was found that the accuracy largely depended on model, e.g., [25, 27]; anatomical placement, e.g., [25]; activity that is recorded, e.g., [23, 25, 26]; or a combination of these factors, e.g., [28].

## 3 Research Design

To understand the goals potential users of life-logging technologies might want to pursue and which perceived barriers they might encounter, we conducted an empirical study using an online survey. The survey consists of three parts: Firstly, we queried the participants' demographic data, explanatory user factors, as well as their current use of life-logging technologies. Secondly, we asked for the goals they might want to pursue and for which of these goals they seek or would accept digital assistance. Thirdly, we queried the participants about potential barriers to using life-logging technologies in the pursuit of their goals. The list of potential barriers, as well as the list of goals was collected in antecedent focus groups (see below).

**Explanatory Variables:** The first section included demographic data such as *age*, *gender*, and *education*. This was supplemented with the following factors:

*Subjective Vitality (VIT):* 4 items taken from [29], translated into German.

*Motives for Physical Activity (revised scale) (MPAM-R):* This scale by Ryan and Frederick captures individuals' motives for performing sportive physical activities [30]. The scale is subdivided into five dimensions: perception of one's own competence, appearance, fitness, enjoyment, and social motives.

*Self-Efficacy in Interacting with Technology:* The context-specific self-efficacy determines an individual's selection of tasks, their performance in these tasks, and their perseverance in case of difficulties. In the technology domain, self-efficacy in interacting with technology has been found to profoundly shape how people interact with interactive systems. To understand its influence on the use of wearable electronic devices, we surveyed this construct with 8 items based on a scale by Beier [31].

Another section was devoted to *Need for Privacy (NfP)* and *Privacy Concerns (PC)*, each determined by 3 items to be rated on 6-point Likert scales, respectively. While the *NfP* is used to evaluate the participants' general stance on information disclosure, the second concept *(PC)* deals specifically with concerns regarding information privacy in an online context.

*Use of Life-Logging:* Concluding the list of explanatory user factors, we asked for the participants' current life-logging behavior, i.e., whether they use life-logging technologies or paper-based diaries for logging (no, apps, wearable, web portals, diary).

**Dependent Variables:**
*Pursued Goals:* As the goals individuals want to pursue might differ, we collected a list of possible objectives through literature review and two focus groups with users and non-users (4 participants in each). In total, we included 12 different goals in our study and asked the participants to rate the importance of each goal on a 6-point Likert scale ranging from "*not important at all*" to "*very important*":

- quit smoking
- quit drinking
- achieve an exercise goal
- reduce resting heart rate
- increase discipline
- drink more water
- get better sleep
- lose weight
- more balanced nutrition
- do more sports/exercise more
- be more active in daily life
- increase overall well-being

*Desire for Support:* For each of the goals we also asked the participants if they desire (technological) support for achieving this goal, also on a 6-point Likert scale.

*Perceived Barriers:* To understand what prevents people from using life-logging technologies to support the pursuit of their goals, we collected a list of potential barriers by using the same methodology as above. Again, the 13 perceived barriers were surveyed on 6-point Likert scales ranging from "*do not agree at all*" to "*fully agree*":

- already active enough
- no benefit from life-logging
- goals beyond reach
- battery's service life too short
- device does not look nice
- not knowing what happens to measured data
- constant reminders to get moving
- forced sharing of results in social networks
- arduous usage
- device does not work correctly
- getting negative feedback
- device is expensive
- constantly under surveillance, even on a "lazy" day

**Statistical Procedures:** All answers to the items were captured on 6-point Likert scales ranging from 'not important at all' to 'very important.' For the statistical analyses, items were rescaled to 0 to 100% and scales were aggregated into arithmetic means. The results are analyzed with parametrical and non-parametrical methods, using bivariate correlations (Pearson's $r$ or Spearman's $\rho$) as well as single and repeated multi- and univariate analyses of variance (M/ANOVA). The type I error rate (level of significance) is set to $\alpha = .05$ (findings $.05 < p < .1$ are reported as marginally significant). Pillai's value is used for the multivariate tests. Arithmetic means are reported with standard deviations (denoted ±). The whiskers in the diagrams show the 95% confidence interval. All scales used were tested for their internal reliability by calculating Cronbach's α (see Table 1).

**Table 1.** Reliability of scales and sources.

| Scale | Items | n | Cronbach's α |
|---|---|---|---|
| SET self-efficacy technology interaction [31] | 4 | 412 | $\alpha = .876$ |
| VIT subjective vitality [29] | 4 | 412 | $\alpha = .859$ |
| MPAM motives for physical activity [30] | 15 | 412 | $\alpha = .889$ |
| NfP need for privacy | 3 | 412 | $\alpha = .701$ |
| PC privacy concern | 3 | 412 | $\alpha = .636$ |
| GLC goals for life-changes | 12 | 412 | $\alpha = .779$ |
| BLL barriers to life-logging | 13 | 412 | $\alpha = .864$ |

### 3.1 Description of the Sample

A total of 412 data sets was collected via technology mediated social networks or email. 214 participants of the study are women (51.9%) and 198 are men (48.1%). The age in the sample ranges from 17 to 78 years (mean 36.1 ± 12.2 years). Age and gender are not correlated ($\rho = .061$, $p = .213 > .05$), indicating a heterogeneous sample.

In our sample, higher *age* is associated with a slightly higher *subjective vitality* ($r = .116$, $p = .016 < .05$), slightly lower *motives for physical activities measure* ($r = -.175$, $p < .001$), and lower *self-efficacy in interacting with technology* ($r = -.099$, $p = .044 < .05$).

*Gender* is linked to *technical self-efficacy* ($r = .312$, $p < .001$) and men report a significantly higher *self-efficacy in interacting with technology* (82.3 ± 17.6%) than women (69.3 ± 21.6%). No influence of *gender* was found on *subjective vitality* ($r = .078$, $p = .114 > .05$), whereas men reported higher *motives for physical activities* than women ($r = .163$, $p < .001$). Neither the *privacy concerns* nor the *need for privacy* were affected by *gender* ($p_{PC} = .345 > .05$, $p_{NfP} = .589 > .05$). In addition, we found a strong relationship between *subjective vitality* and the *motives for physical activities* ($r = .381$, $p < .001$), as well as a very strong relationship between the reported *privacy concerns* and *need for privacy* ($r = .652$, $p < .001$).

In our sample, 225 (54.6%) participants reported to be *users* of life-logging technologies and 187 (45.4%) participants reported to be *non-users*. Specifically, 169 (41.0%) stated to use smartphone apps for life-logging, 118 (28.6%) conveyed to use an extra device, such as a wristband, for life-logging, 32 (7.7%) are using a fitness portal, and 18 reported to keep a diary (4.4%). In the following, we focus on electronically mediated life-logging systems and only consider users of apps, portals, or wearables as life-loggers in our analyses.

## 4 Results

This section presents the findings of our survey. First, we show how users and non-users of life-logging technologies differ in regard to the investigated user factors. Second, we present a ranked list of goals that users and non-users of life-logging technologies want to pursue. Third, we illustrate the perceived barriers of adopting life-logging technologies. Finally, we investigate how personality factors and perception of life-logging technologies relate to the projected adoption.

### 4.1 Contrasting Users and Non-Users of Life-Logging Technologies

This section presents the similarities and differences between users and non-users of life-logging technologies. First, we show the similarities and differences regarding the participants' personality profiles. Next, we illustrate how users and non-users differ in regard to their goals and then in regard to the perceived barriers to using life-logging technologies.

First, *gender* is not related to the use of life-logging technologies ($\chi^2 = .385$, $p = .535 > .05$). Of the 214 women, 120 reported to be users with the remaining 94 being non-users. Of the 198 men, 104 indicated to be users and 93 stated to be non-users.

To understand if the surveyed personality factors influence the use of life-logging technologies, we calculated a MANOVA with usage as independent and *VIT*, *MPAM*, *SET*, *PC*, *NfP*, and *age* as dependent variables. The MANOVA revealed an overall significant effect ($V = .131$, $F_{6,405} = 10.159$, $p < .001$). In this sample, the use of logging technologies was neither related to *age* ($F_{1,410} = .030$, $p = .863 > .05$), *PC* ($F_{1,410} = .409$, $p = .523 > .05$), nor *VIT* ($F_{1,410} = 2.152$, $p = .143 > .05$). Differences emerged for the *NfP* ($F_{1,410} = 7.327$, $p = .005 < .005$), *MPAM* ($F_{1,410} = 22.327$, $p < .001$), and *SET* ($F_{1,410} = 26.421$, $p < .001$). Specifically, users of life-logging technologies reported a higher *motivation for physical activities* and a higher *self-efficacy in interacting with technology*, whereas non-users reported a higher *need for privacy*.

Figure 1 and Table 2 illustrate the influence of life-logging usage on these measures. In regard to the self-reported steps per day ($p = .023 < .05$) users and non-users of life-logging technologies also differ, although the self-reported hours of sportive activities per week is not significantly different ($p = .079 > .05$).

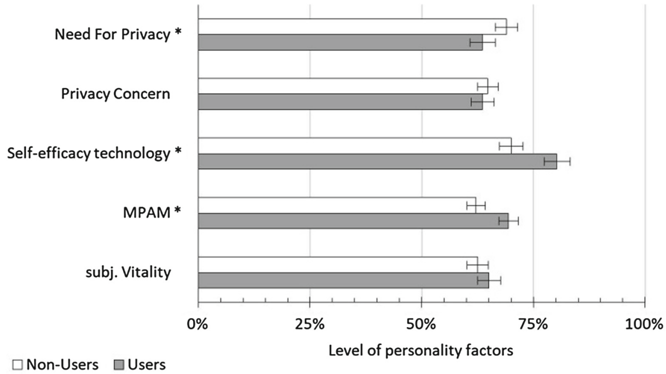



**Fig. 1.** User factors for users and non-users of life-logging technologies (whiskers show the 95%-CI, * indicates significant differences of $p < .05$).

**Table 2.** Differences between life-logging users and non-users.

| | Life-loggers (n = 225) | Non life-loggers (n = 187) | p |
|---|---|---|---|
| Age | 36.2 ± 11.0 | 36.0 ± 13.6 | n.s. |
| Gender | 120 f; 105 m | 94 f; 93 m | n.s. |
| SET | 80 ± 17% | 70 ± 22% | p < .001 |
| VIT | 65 ± 18% | 62 ± 18% | n.s. |
| MPAM | 69 ± 14% | 62 ± 17% | p < .001 |
| PC | 64 ± 17% | 65 ± 18% | n.s. |
| NfP | 64 ± 20% | 69 ± 19% | p = .005 < .05 |
| Est. steps per day | 9794 ± 21252 | 6048 ± 8162 | p = .023 < .05 |
| Est. h sport per week | 4.8 ± 3.6 | 4.0 ± 5.2 | p = .079 > .05 |
| Goals for life-changes | 58 ± 14% | 51 ± 16% | p < .001 |
| Desire for support | 51 ± 17% | 38 ± 22% | p < .001 |
| Barriers to life-logging | 55 ± 18% | 56 ± 20% | n.s. |

### 4.2 Goals for Life-Changes

As Fig. 2 illustrates, the items 'quit smoking' and 'drink less alcohol' are given the lowest priority by our subjects, whereas a higher level of sporting activity, more activity in everyday life, and increased overall well-being are given the highest priority.

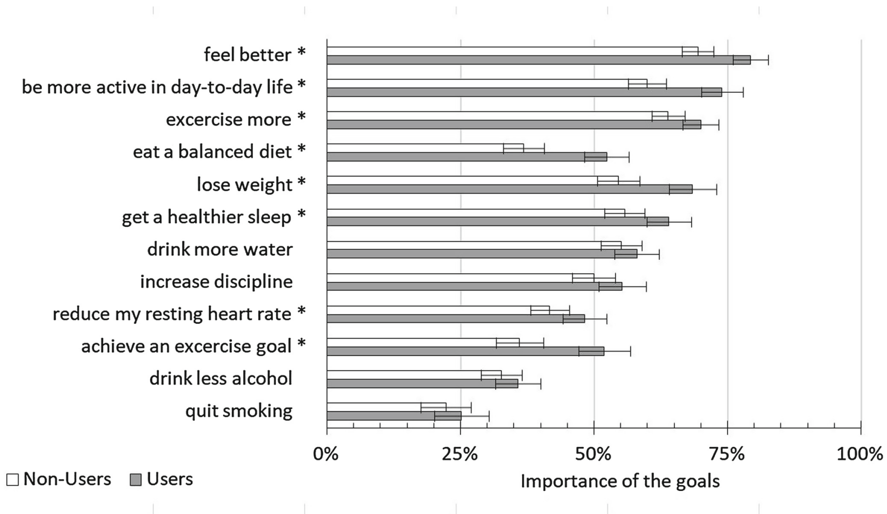



**Fig. 2.** Importance of goals by users and non-users of life-logging technologies (whiskers show the 95%-CI, * indicates significant differences of $p < .05$).

Next, we analyzed if the pursued goals differ depending on experience with life-logging technologies. A MANOVA with usage as independent variable and the goals as dependent variable affirmed a significant overall effect of usage on the goals ($V = .163$, $F_{12,399} = 6.495$, $p < .001$). However, differences in goal importance emerge only for some of the 13 surveyed goals. Specifically, these differences concern achieving a defined athletic goal ($p < .001$), resting heart rate ($p = .019 < .05$), sleep quality, ($p < .001$), weight loss ($p < .001$), nutrition ($p < .001$), increase of overall sportiness ($p = .009 < .005$), more activity in everyday life ($p < .001$), and overall well-being ($p < .001$). No differences were found for smoking cessation ($p = .409$), less alcohol consumption ($p = .284$), being more disciplined ($p = .083$), and drinking more water ($p = .305$).

### 4.3 Perceived Barriers to Life-Logging

Figure 3 shows that some aspects of life-logging technologies are perceived as stronger barriers than others. Also, the perception differs only for some of the queried barriers between users and non-users of life-logging technologies.

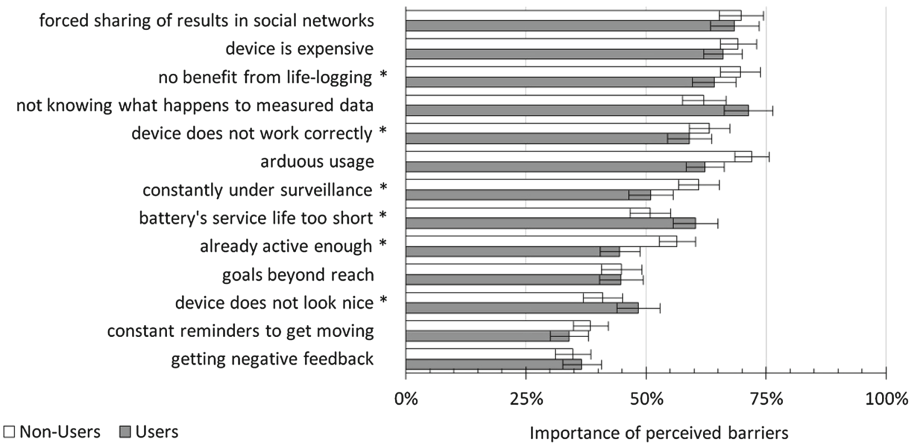



**Fig. 3.** Comparing barriers for users and non-users of life-logging technologies (whiskers show the 95%-CI, * indicates significant differences of $p < .05$).

The participants evaluated potential (negative) feedback and feedforward to get moving as the least important barriers, whereas the price of the device, the limited effectiveness, or the perceived necessity to share one's activities in social networks were rated as the most important barriers.

Next, we wanted to understand how users and non-users differ in regard to the perceived barriers to the use of life-logging technologies. A MANOVA with the usage of life-logging technologies as independent variable and the 13 potential barriers as dependent variables attests a significant overall effect ($F_{13,398} = 5.234$, $V = .148$, $p < .001$).

## 4.4 A Closer Look at Non-Users of Life-Logging

As shown above, actual usage of life-logging technologies significantly relates to most of the investigated independent and dependent variables. But what factors relate to the *projected adoption* of life-logging technologies by non-users?

To answer this question, we now focus only on the group of non-users of life-logging technologies ($n = 178$). Here, the average intention to use life-logging technologies is just 44 ± 27% and thus below the center of the scale. The large standard deviation indicates a rather strong disagreement among the members of this group.

To understand what causes this disagreement, we calculated two stepwise multiple linear regressions with the intention to use as dependent variable. First, we focus on the perceived benefits and barriers of this technology. Thus, we used the goals, desire for support, and barriers as independent variables for the stepwise multiple linear regression, see Fig. 4. The analysis yields a significant model with desire for support as the single predictor for intention to use ($\beta = .590$, $T = 9.705$, $p < .001$). The share of explained variance is $r^2 = .345$, thus the desire for support explains about 35% of intention to use's variance. Consequently, participants reporting higher desire for support also report a higher intention to adopt life-logging technologies.

Second, we analyzed which user factors relate to the intention to adopt life-logging technologies. For this, we calculated a second stepwise multiple linear regression with the user factors *age*, *gender*, *SET*, *VIT*, *MPAM*, *PC*, and *NfP* as independent variables. This analysis yields a significant model with two factors and an explained variance of $r^2 = .058$. The model predicts that *MPAM* is strongly and positively ($\beta = .196$, $T = 2.608$, $p = .008 < .05$) and *PC* less strong and negatively ($\beta = -.152$, $T = -2.065$, $p = .040 < .05$) related to the intention to adopt life-logging technologies. Hence, participants with a higher *motivation towards physical activities* are more likely to adopt life-logging technologies, whereas a higher *PC* acts as a barrier for the adoption. Within this sample and our limited sample size of 178 participants, no other factors, such as *age*, *SET*, or *gender*, influence the intention to use.

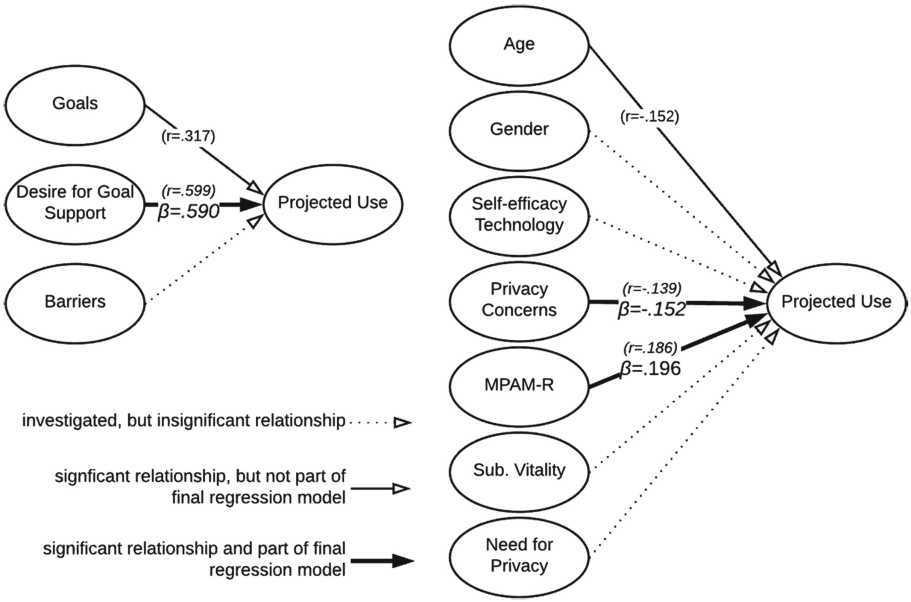



**Fig. 4.** Determinants for projected use by non-users of life-logging technologies. Left: goals, desire for support and barriers ($r^2 = .345$). Right: based on user factors ($r^2 = .058$).

## 5 Discussion, Limitations, and Outlook

The study analyzed different aspects that might influence the acceptance of life-logging technologies. To do so, user factors, e.g., demographics, privacy concerns, motivation to be physically active, etc., as well as different experience groups were compared as to the possible goals they might pursue and for whose achievement the use of CHIT might be beneficial. Furthermore, we compared the evaluation of different barriers preventing the use of life-logging devices.

Our results show that users and non-users differ in many instances in their evaluation of barriers to life-logging but also in goals they deem worthy of pursuing. While non-users are mostly opposed to arduous usage, forced sharing of their data, expenses, and malfunctions, users in turn are more concerned with the technology's appearance and battery runtime.

In accordance to previous studies, e.g., [22], concerns for one's privacy are once more the main obstacle to the use of fitness-apps and trackers. However, even though information privacy concerns or possible violations factored into 3 of 13 items, it is interesting to note that the barriers did not have a significant influence on the projected future use of life-logging technologies. Here, the goals and wish for support in achieving said goals play a much more important role.

Another interesting finding was that, although we analyzed the acceptance of a technology, specific as it might be, self-efficacy when interacting with technology did not influence the projected use. Neither did gender, both of which have been important factors in previous acceptance studies. However, other typical acceptance factors, such as perceived usefulness or effort expectancy, are still very important in the evaluation of possible barriers to the use of life-logging devices.

Even though our sample reported to be rather active on average, see Table 2, more activity was an important goal and already sufficient activity a medium important barrier. Nevertheless, this might also be a misjudgment as the standard deviations surpass the average means considerably.

While the present article analyzes the potential goals that might be supported with life-logging technologies and the perceived barriers potential users might encounter, future work should address the representativeness of the sample. Although the current study contains younger and older participants as well as users and non-users of life-logging technologies, the validity of the presented findings might be improved by using a representative sample of the population.

Furthermore, the study was conducted in Germany where life-logging is still not widely used, e.g., [32]. Other countries with more exposure to activity trackers or higher numbers of users might encounter different perceived barriers but also different goals that might be supported by life-logging devices or apps that, so far, have not reached the German usage context.

**Acknowledgements.** Parts of this work have been funded by the German Ministry of Education and Research (BMBF) under project No. KIS1DSD045 "myneData" and V5JPI004 "PAAL."

## References

1. Krug, S., Jordan, S., Mensink, G.B.M., Müters, S., Finger, J., Lampert, T.: Physical activity. Results of the German health interview and examination survey for adults (DEGS1). Bundesgesundheitsblatt **56**, 765–771 (2013)
2. Hadgraft, N.T., Lynch, B.M., Clark, B.K., Healy, G.N., Owen, N., Dunstan, D.W.: Excessive sitting at work and at home: correlates of occupational sitting and TV viewing time in working adults. BMC Public Health **15**, 899 (2015)

3. Warburton, D.E.R., Nicol, C.W., Bredin, S.S.D.: Health benefits of physical activity: the evidence. CMAJ **174**, 801–809 (2006)
4. Lee, I.-M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., Katzmarzyk, P.T., Alkandari, J. R., Bo Andersen, L., Bauman, A., Brownson, R.C., Bull, F.C., Craig, C.L., Ekelund, U., Goenka, S., Guthold, R., Hallal, P.C., Haskell, W.L., Heath, G.W., Inoue, S., Kahlmeier, S., Kohl III, H.W., Lambert, V., Leetongin, G., Loos, R., Marcus, B., Martin, B., Owen, N., Parra, D.C., Pratt, M., Ogilvie, D., Reis, R.S., Sallis, J.F., Lucia Sarmiento, O., Wells, J.C.: Impact of physical inactivity on the world's major non- communicable diseases. Lancet **380**, 219–229 (2012)
5. Rühle, J.: The Mobile Health Market in Germany, Berlin (2017)
6. Rasche, P., Wille, M., Bröhl, C., Theis, S., Schäfer, K., Knobe, M., Mertens, A.: Prevalence of health app use among older adults in Germany: national survey. JMIR mHealth uHealth **6**, e26 (2018)
7. Lupton, D., Smith, G.J.D.: "A much better person": the agential capacities of self-tracking practices. In: Ajana, B. (ed.) Metric Culture: Ontologies of Self-Tracking Practices. Emerald Publishing, London (2018)
8. Lupton, D.: Understanding the human machine. Technol. Soc. Mag. IEEE **32**, 25–30 (2013)
9. Bravata, D.M., Smith-Spangler, C., Sundaram, V., Gienger, A.L., Lin, N., Lewis, R., Stave, C.D., Olkin, I., Sirard, J.R.: Using pedometers to increase physical activity and improve health: a systematic review. JAMA **298**, 2296–2304 (2007)
10. Brauner, P., van Heek, J., Ziefle, M.: Age, gender, and technology attitude as factors for acceptance of smart interactive textiles in home environments. In: Proceedings of the 3rd International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2017), pp. 13–24. SCITEPRESS – Science and Technology Publications, Porto, Portugal (2017)
11. Himmel, S., Ziefle, M.: Smart home medical technologies: users' requirements for conditional acceptance. i-com **15**, 39–50 (2016)
12. van Heek, J., Arning, K., Ziefle, M.: The surveillance society: which factors form public acceptance of surveillance technologies? In: Helfert, M., Klein, C., Donnellan, B., Gusikhin, O. (eds.) Smart Cities, Green Technologies, and Intelligent Transport Systems, pp. 170–191. Springer, Cham (2017)
13. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. **35**, 982–1003 (1989)
14. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison-Wesley Publishing Company Inc., Reading (1975)
15. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**, 179–211 (1991)
16. Bansal, G., Zahedi, F.M., Gefen, D.: The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis. Support Syst. **49**, 138–150 (2010)
17. Motti, V.G., Caine, K.: Users' privacy concerns about wearables. In: Brenner, M., et al. (eds.) International Financial Cryptography Association FC 2015 Workshops, pp. 231–244 (2015)
18. Wieneke, A., Lehrer, C., Zeder, R., Jung, R.: Privacy-related decision-making in the context of wearable use. In: PACIS 2016 Proceedings Paper 67 (2016)
19. Clawson, J., Pater, J.A., Miller, A.D., Mynatt, E.D., Mamykina, L.: No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In: UbiComp 2015, pp. 647–658 (2015)

20. Lazar, A., Koehler, C., Tanenbaum, J., Nguyen, D.H.: Why we use and abandon smart devices. In: UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 635–646 (2015)
21. Epstein, D.A., Caraway, M., Johnston, C., Ping, A., Fogarty, J., Munson, S.A.: Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. In: CHI 2016, pp. 1109–1113 (2016)
22. Lidynia, C., Brauner, P., Ziefle, M.: A step in the right direction – understanding privacy concerns and perceived sensitivity of fitness trackers. In: Ahram, T., Falcão, C. (eds.) Advances in Intelligent Systems and Computing - Advances in Human Factors in Wearable Technologies and Game Design. Springer, Cham (2018)
23. Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A.M., Houghtaling, P., Javadikasgari, H., Desai, M.Y.: Variable accuracy of wearable heart rate monitors during aerobic exercise. Med. Sci. Sport. Exerc. **49**, 1697–1703 (2017)
24. Battenberg, A.K., Donohoe, S., Robertson, N., Schmalzried, T.P.: The accuracy of personal activity monitoring devices. Semin. Arthroplasty. **28**, 71–75 (2017)
25. Storm, F.A., Heller, B.W., Mazzà, C.: Step detection and activity recognition accuracy of seven physical activity monitors. PLoS ONE **10**, e0118723 (2015)
26. Sasaki, J.E., Hickey, A., Mavilia, M., Tedesco, J., John, D., Kozey Keadle, S., Freedson, P. S.: Validation of the fitbit wireless activity tracker(r) for prediction of energy expenditure. J. Phys. Act. Health **12**, 149–154 (2015)
27. O’Connell, S., Ólaighin, G., Kelly, L., Murphy, E., Beirne, S., Burke, N., Kilgannon, O., Quinlan, L.R.: These shoes are made for walking: sensitivity performance evaluation of commercial activity monitors under the expected conditions and circumstances required to achieve the international daily step goal of 10,000 steps. PLoS ONE **11**, e0154956 (2016)
28. Chow, J.J., Thom, J.M., Wewege, M.A., Ward, R.E., Parmenter, B.J.: Accuracy of step count measured by physical activity monitors: the effect of gait speed and anatomical placement site. Gait Posture **57**, 199–203 (2017)
29. Ryan, R.M., Frederick, C.: On energy, personality and health: subjective vitality as a dynamic reflection of well-being. J. Pers. **65**, 529–565 (1997)
30. Ryan, R.M., Fredrick, C.M., Lepes, D., Rubio, N., Sheldon, K.M.: Intrinsic motivation and exercise adherence. Int. J. Sport Psychol. **28**, 335–354 (1997)
31. Beier, G.: Kontrollüberzeugungen im Umgang mit Technik (Locus of control when interacting with technology). Rep. Psychol. **24**, 684–693 (1999)
32. Ernsting, C., Dombrowski, S.U., Oedekoven, M., O´sullivan, J.L., Kanzler, M., Kuhlmey, A., Gellert, P.: Using smartphones and health apps to change and manage health behaviors: a population-based survey. J. Med. Internet Res. **19**, e101 (2017)

# Comparison Among Standard Method, Dedicated Toolbox and Kinematic-Based Approach in Assessing Risk of Developing Upper Limb Musculoskeletal Disorders

Stefano Elio Lenzi[1(✉)], Carlo Emilio Standoli[2], Giuseppe Andreoni[2], Paolo Perego[2], and Nicola Francesco Lopomo[1]

[1] Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Brescia, via Branze 38, 25123 Brescia, Italy
{s.lenzi002,nicola.lopomo}@unibs.it

[2] Dipartimento di Design, Politecnico di Milano, via Durando 38/A, 20158 Milan, Italy
{carloemilio.standoli,giuseppe.andreoni, paolo.perego}@polimi.it

**Abstract.** OCRA (Occupational Repetitive Action) index is one of the most used method for supporting risk assessment in tasks requiring manual handling of low loads at high frequency. One of the main drawbacks of this method is that the operator analyses the activities by observing videos. This kind of procedure is inherently not objective and operator-dependent. To overcome these limitations, we developed a toolbox to support the analysis with contextual noting and wearable sensors kinematic data. Three expert operators were asked to evaluate seven videos with and without the aid of the developed toolbox. Results underlined a high inter ($R^2$ mean 0.4) and intra-operator variability (posture time percentage and technical actions (TAs) count mean errors respectively 7.44%, 4 TAs) when using the only video-based approaches. On the contrary, research outcomes showed that the introduction of wearable device allow to overcome these issues and to reduce noticeably the evaluation time (−98%).

**Keywords:** Occupational ergonomics · Wearable systems
Upper limb musculoskeletal disorders · Movement analysis

## 1 Introduction

Manual handling of loads is very common in several professions and in very different workplace and working tasks. This kind of activities implies that the workers decide - consciously and/or unconsciously and in relation to the environment itself - movements, postures and muscular strength necessary to complete the assigned task and, consequently, the level of effort to be invested in. The motor strategies or behaviours adopted by subjects in executing these actions obviously strongly affect the overall load acting on and related fatigue of musculoskeletal system. Upon the individual choice of the "optimal" motor strategies and the environmental constraints, the worker could


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 135–145, 2019.
https://doi.org/10.1007/978-3-319-94619-1_13

develop different kind of musculoskeletal pathologies and disorders that can impact also his/her daily life. Hence it is important that manual work must be designed and monitored with the aim to guarantee and protect health and well-being of all the involved employees. In this framework the analysis of postures and movement assumed by workers is critical to correctly define the level of risk to develop musculoskeletal disorders, associated to their specific working activities. Ergonomics is the reference discipline to assess and design tasks and workplaces to support the best safety and performance of the human-environment-task system. Several norms have been developed for these purposes, both for designing and for assessing the working tasks and environments (and tools).

ISO 11228 [1] is the international standard referred to manual handling in which are contained several information on how to conduct a reliable risk assessment. It is organized in three parts related to the evaluation of risk factors in lifting and carrying tasks (ISO 11228-1), push and pulling tasks (ISO 11228-2) and manual handling of low loads at high frequency tasks (ISO 11228-3). In particular, this latter part suggests the use of OCRA (Occupational Repetitive Actions) index [1, 2]. Undoubtedly this method presents several advantages such as: it takes into account several risk factors, it is applicable to jobs in which workers have to exploit different tasks (i.e. multitask jobs) and give to the analyst quantitative criteria to predict the occurrence of upper limb work-related musculoskeletal disorders. On the other hand, in most cases, operators conduct the analysis of postures and movement mainly by qualitatively observing videos acquired during the work shift of a workers' population. Thus, although the approach has a concrete scientific basis and is based on an extensive epidemiological research, the overall procedure can be affected by systematic not objective and operator-dependent (with a critical level of inter- and intra-operator variability) factors. Furthermore, the analysis could lead to misleading outcomes due to the 2D nature of the available information [3–5].

All these considerations led us to hypothesize the necessity to design novel tools and methods able to objectively and reliably quantify the level of risk to develop musculoskeletal disorders associated to working task.

In this regard, the main goal of this paper is to describe the development of an original software tool that can support and help operators in the analysis and evaluation of working task requiring manual handling of low loads at high frequencies. The method/tool is based on the integration of multimedia information with contextual noting and the acquisition of kinematic data with the use of wearable inertial sensors. Furthermore, different methods are compared with the aim to assess whether our novel approach could improve the quality and the productivity of the risk assessment with respect to the classic manual "paper-based" approach.

## 2 Materials and Methods

The proposed toolbox integrates different activities that the operator can do concerning risk assessment by means of OCRA Index. For sake of clarity, first we report the main characteristics of this index and then how we integrated them into the developed toolbox.

### 2.1 ISO 11228-3: Posture Analysis and Technical Actions Count (Checklist OCRA Index)

The third part of the ISO 11228 norm gives recommendations about the evaluation of tasks involving manual handling of low loads at high frequencies. It guides the analyst step-by-step in conducting a complete risk analysis considering all the factors that can contribute to the insurgence of work-related musculoskeletal disorders. The Standard suggests to quantify the risk of hazard using the OCRA Index.

The OCRA Index [1] (Eq. 1) is specifically defined as the ratio between the number of Actual Technical Actions (ATAs) execute by workers and the Reference Technical Actions (RTAs), that represent the ideal number of actions required to prevent the occurrence of Upper Limb Work-related Musculoskeletal Disorders (UL-WMSDs) in the execution of that task.

$$OCRA_{Index} = n_{ATA} / n_{RTA} \quad (1)$$

RTAs and ATAs must be computed for each upper limb side. In order to determine ATAs, the analyst has to compute the number of technical actions observed during the work shift related to that particular task. Then, it is possible to compute the frequency of actions per minute, through the definition of the ratio between number of actions observed and the time of the observation. The number of technical actions is computed by observing the video: the analyst takes note "on paper" about the number of estimated technical actions.

Once $n_{ATA}$ is calculated, it is possible to calculate the number of RTAs. A value for each different multiplier is computed using empirical rules described in the international standard. Then all values are multiplied and thence $n_{ATA}$ is obtained.

In particular, force and posture multipliers are really hard to be objectively determined only by observing a video. For the main purpose of this paper, only posture analysis will be take into account.

**Table 1.** Angular displacement threshold for each upper limb joint movement.

| Joint | Movement | Limits |
|---|---|---|
| Shoulder | Flexion/extension | >80° |
| Elbow | Pronation/supination (dynamic movement) | >60° |
| | Flexion/extension (dynamic movement) | >60° |
| Wrist | Palmar flexion/dorsal extension | >45° |
| | Ulnar deviation | >20° |
| | Radial deviation | >15° |

Posture multiplier is determined for each upper limb joint, by comparing the angular displacement in different anatomical planes with its threshold. The operator takes note of percentage of time in which the joint angle is greater than the threshold. Thresholds for each movement in each anatomical plane are reported in Table 1. For the elbow, the analyst has to consider only "dynamic movement", i.e. not considering absolute angles, but temporal variation. Concerning the hand, it has to be considered the percentage of time only when workers hold in their hand the object.

In particular, this work has taken into account the OCRA Checklist index method, a simplify version of the more complex OCRA. This method is the most widely used for the assessment of the overall risk in work environment.

Considering the OCRA Checklist method, it is important to underline that the percentage of time is cumulative for movements of a specific joint. A time span is considered as the time in which the worker has an incorrect posture, if at least one of the movements of the specific joint reaches an angle that is beyond the (static or dynamic) threshold.

## 2.2 Software Toolbox

The proposed toolbox was designed to follow OCRA Checklist method, giving multimedia management support and integrating the possibility of using wearable inertial sensor data. The overall toolbox was developed in Matlab (The Mathworks Inc.). As reported in Fig. 1, it is composed by four different Graphical User Interfaces (GUI). Each GUI is useful to conduct a particular analysis. The toolbox is organized as follow:

- Selection module;
- Technical Action Count module;
- Posture Analysis (Video) module;
- Posture Analysis (Inertial Sensors) module.

Once the analysis is completed, it is possible to export drawings and data in a well-organized spreadsheet file. Furthermore, the user can save in the same spreadsheet file and sheet, results of other performed analyses and trials.

### 2.2.1 Selection Module

In the first GUI, the user can choose the analysis mode: (a) count of technical action number (*Technical Actions*), evaluation of postures through video analysis (*Posture*) and automatic evaluation of postures uploading body kinematics acquired by a commercial system of wearable inertial sensors (*Smart Posture*).

### 2.2.2 Technical Action Module

This module is dedicated to the analysis of the so-called technical actions (Fig. 1). The operator has to upload a specific video recording of the working tasks from the structured video database folder (each one has been categorized with a proper label according to technical action, user, date). While the video is reproduced, the operator can select the side of the body to be analyzed and the number of technical actions, with their timing with a dedicated button. The technical actions frequency is automatically computed and displayed.

The user can also tag the instants when the worker begins to maintain an object with a single hand (static actions) and when he stops to maintain it and begin to move again (end of static action). The duration of the static action must be higher than 5 s to be considered in the computation of static action total time. Bar-plots are drawn to describe these actions, and the total and percentage of time of static actions are displayed. These data can be exported, in the most common spreadsheet file format (such as *.xls*) all the data. Corresponding pictures are automatically saved in the selected folder location.

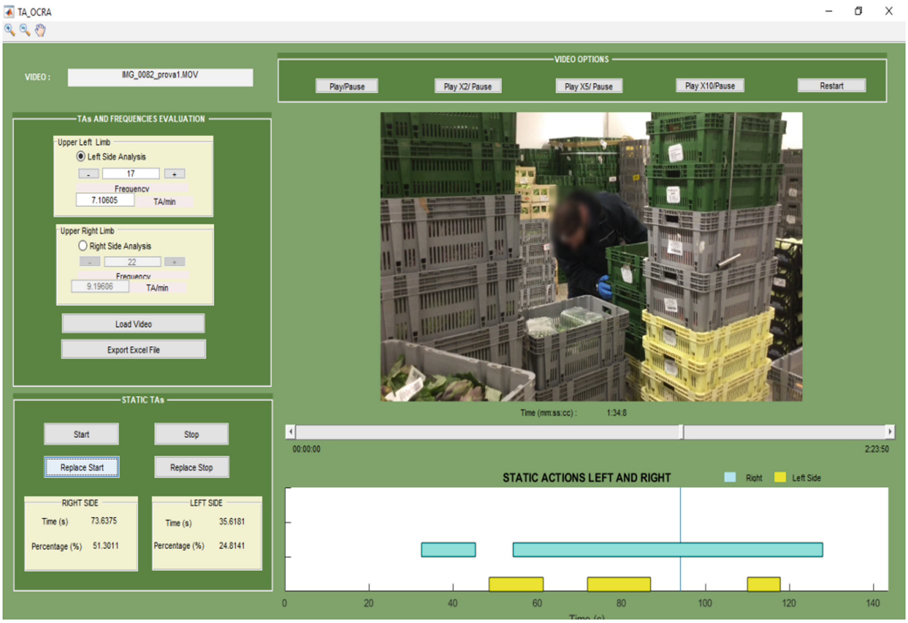



Fig. 1. Technical actions GUI.

### 2.2.3 Posture Module

Posture module (Fig. 2) was designed in order to help the user analyzing postures through the video observation. The GUI allows the operator for loading the video to be

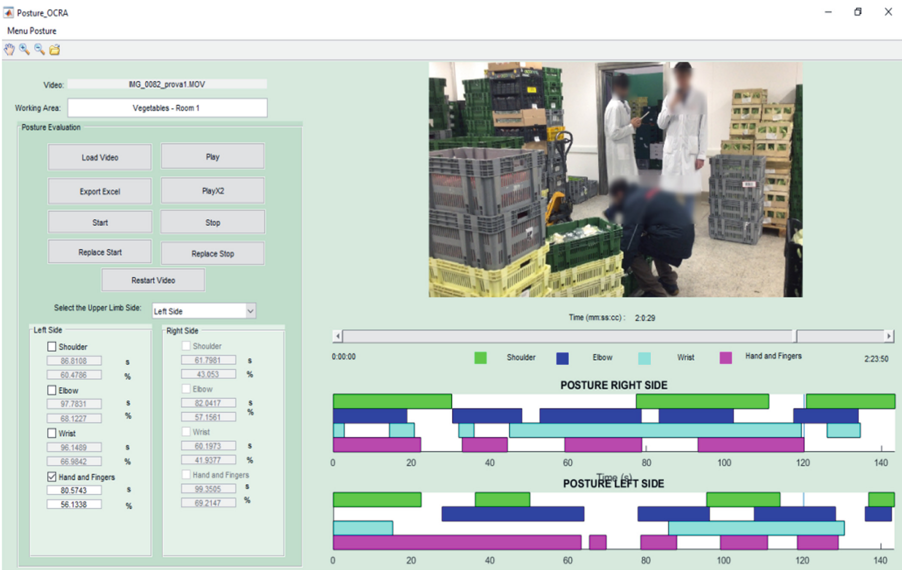



Fig. 2. Posture module

analyzed and its reproduction. Again, the operator the side of the body (left/right) and the joint he/she wants to analyze. After starting the analysis with the corresponding button, the user can mark the time whenever the worker presents an awkward posture. A bar in the lower-right graph window related to the selected side, described the joint posture and the relative percentage of time. As previously reported, data can be exported in spreadsheet file and the bar-plot can be saved too.

### 2.2.4 "Smart" Posture Module

The "Smart" posture module lets the user to import kinematic data (joint angles) acquired by means of commercial systems based on wearable inertial sensors and then it computes in an automatic way the posture evaluation and the related time and percentages of time (Fig. 3).

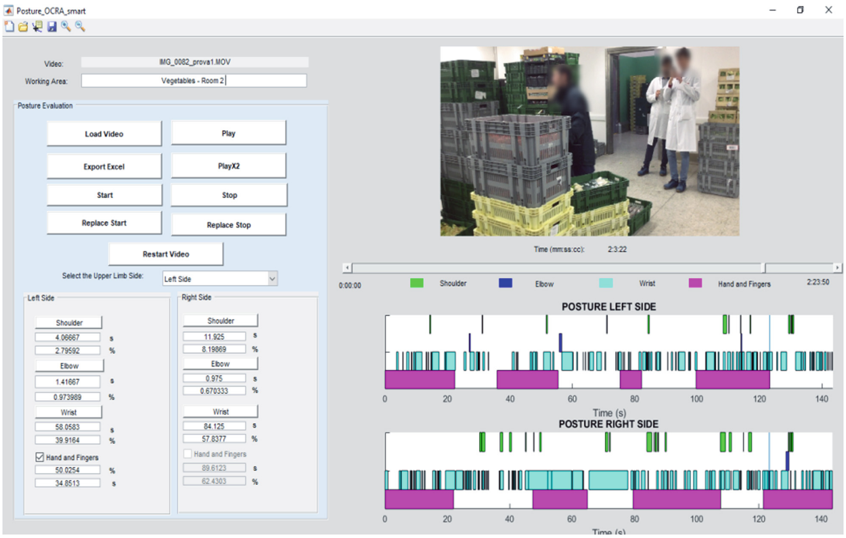



**Fig. 3.** Smart posture module

For the determination of time and percentage of time, a specific algorithm (Fig. 4) was designed, following the rules of ISO 11228-3 norm. In particular, once data have been loaded, a counter ($count_{sign}$) is initialized. This counter will take into account the discrete time of the sample that is going to be analyzed. Then the sample of the first joint angle ($i_{mov1}$) corresponding to the counter value is read. The value is compared with the relative threshold ($\text{thr}_{mov1}$) and, if the value is greater than the threshold, a counter is incremented ($\text{count}_{mov1}$). The next sample in the array of data is read and a new comparison will take place. In case the angle is no longer beyond the threshold, the algorithm starts to consider another movement of the same joint in another anatomical plane. Subsequently the flow follows the same path as seen previously. The loop can be interrupted only when the last sample has been analyzed. When this condition occurs, the percentage of time for the i[th] joint in which the worker has an awkward posture is computed using the following equation (Eq. 2):

$$t_{\%} = \left(\left(count_{mov\,i}/f_{sampl}\right) / t_{tot}\right) 100\,\% \tag{2}$$

where $count_{mov\_i}$ is the counter that counts samples in which the worker has an awkward posture, $f_{sampl}$ is the sampling frequency and $t_{tot}$ is the trial time. Regarding the elbow, the algorithm is not the same as for the other joints. In this case the algorithm must control that the value is not the same for more than 5 s.

The grab is the only percentage that cannot be automatically computed starting from kinematic data. The percentage of time in which the worker handles an object is obtained in the same way as for the posture module.

### 2.3 Experimental Protocol

The developed toolbox was preliminary tested in a real context related to large retail chains, by involving expert operators and real workers and compared to the traditional analysis in accuracy and productivity (time spent for the analysis).

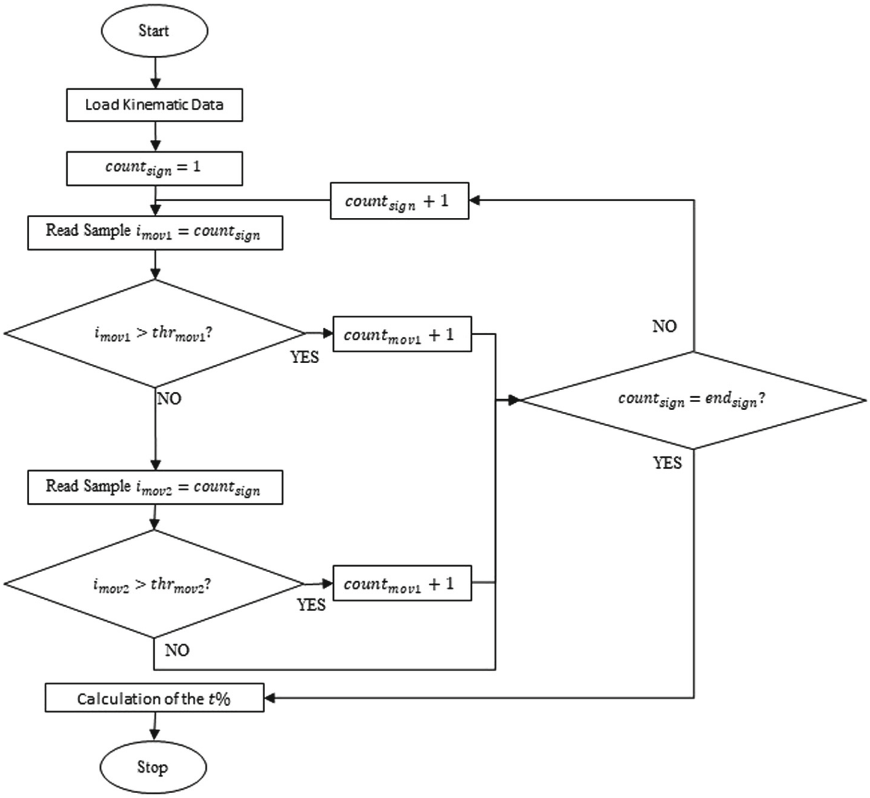



**Fig. 4.** Data and choice flow within the algorithm used to estimate percentage time of posture in the "Smart" posture module.

### 2.3.1 Participants and Study Procedures

Three expert operators were asked to evaluate seven different videos, each one showing an activity made by a worker in their working environment. All the activities recorded were characterized by repetitive task involving manual handling of low loads. Operators evaluated, for each video, both number of technical actions and postures using the Checklist OCRA method. Videos were randomly selected and showed to each operator. Participants were instructed not to share their results and not to influence each other with any kind of suggestion.

Operators evaluated videos according to the following methods: “Pen&Paper” (Method 1), video analysis supported by the toolbox (Method 2) (Technical actions and Posture) and finally by using the automatic method (Method 3), in which the posture evaluation is made automatically using the “Smart” posture module.

For each video the sequence the operators had to follow in order conduct their analysis was therefore different.

### 2.3.2 Data Acquisition and Sensors Placement Protocol

Workers, during “on field” trials, wore wearable inertial sensors. We specifically used the Notch system (Notch Interfaces Inc.), a commercial and low-cost inertial measurement units (IMU) based wearable motion capture system. Sensors were placed on the workers’ body surface as reported in the left part of Fig. 5. Sensors were attached on the body surface using cotton straps and were positioned on the following anatomical regions: stern process, left upper arm, right upper arm, left lower arm, right lower arm, left hand, right hand, pelvis.

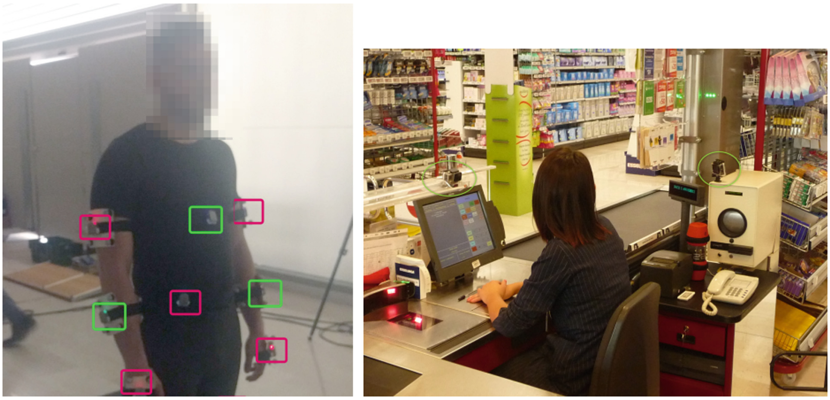


**Fig. 5.** Cameras setup (right) and sensors placement protocol (left).

Furthermore, in order to allow the simultaneous and standard video-based approach, each trial was recorded using two cameras with two distinct point of view. In this way operators can choose the best perspective from which he can properly conduct his/her analysis. Usually cameras were placed at about one meter from the workers (Fig. 5 right).

### 2.3.3 Statistical Analysis

The statistical analysis was performed with the aim of investigate inter- and intra-operator variability and errors made by using manual method (Methods 1 and 2) with respect to the automatic one (Method 3), here used as "gold standard". Inter-operator error was computed as correlation index ($R^2$ index) between results obtained by different operators using the same method. Intra-operator error was underlined computing the mean error between the two manual methods (Methods 1 and 2), when the same operator evaluates the same video. Results obtained from method 1 and method 2 by different operators, were then compared with results obtained with method 3. $R^2$ correlation was computed to underline the errors made with qualitative methods (Methods 1 and 2), with respect to the objective method (Method 3). Results for each limb are averaged between them.

Furthermore, the reduction in the overall assessment time was compared between the manual methods (Methods 1 and 2) and the automatic one (Method 3).

## 3 Results and Discussion

### 3.1 Inter-operator Variability

Results concerning the inter-operator variability analysis are reported in Table 2. The highest value reached (0.93) is for the evaluation of the shoulder made by operator 1 and operator 3 with the "Pen&Paper" method (Method 1). The minimum value obtained in this analysis is 0.00 (no correlation) for the evaluations of the elbow made by operator 1 and operator 3. It is possible to see that all the other values are randomly distributed and there is no a clear tendency in data. However, $R^2$ values for the evaluation of the shoulders with all methods showed the highest value of correlation ($R^2$ higher than 0.6).

In a general perspective, these results showed that, in a general perspective, there is a good correspondence. Furthermore, for technical action count and shoulder evaluation, operators seem to have the same metric in assessing trials. For wrist and elbow this is no longer true: this result reflects the difficulties in the determination of the joint angles from the video analysis.

**Table 2.** Inter-operator analysis results: R2 value for each method and each couple.

| Operators | Method 1 | | | | Method 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | TA | Shoulder | Elbow | Wrist | TA | Shoulder | Elbow | Wrist |
| 1 vs 2 | 0.55 | 0.77 | 0.18 | 0.11 | 0.81 | 0.71 | 0.49 | 0.06 |
| 2 vs 3 | 0.59 | 0.68 | 0.37 | 0.01 | 0.12 | 0.71 | 0.12 | 0.02 |
| 1 vs 3 | 0.59 | 0.93 | 0.00 | 0.02 | 0.42 | 0.71 | 0.39 | 0.15 |

### 3.2 Intra-operator Variability

Outcomes related to the intra-operator variability are reported in Table 3. It is possible to underline in this case how errors are always high. Highest values for technical action count are obtained for operator 3 (4.33). For postures, all percentage are not negligible too. The error become higher when considering elbow and wrist joint. The maximum error is made by operator 1 when considering the wrist evaluation.

These considerations, lead us to affirm how this kind of evaluations, that are based on the video analysis, are strongly affected by a high intra-operator variability. This means that also the same operator, analyzing the same video with two different approaches, is not able to uniquely determine results.

**Table 3.** Differences in posture and technical actions count using method 1 and 2.

| Operator | Mean error - method 1 and method 2 | | | | |
|---|---|---|---|---|---|
| | TA | Frequency (TA/min) | Shoulder | Elbow | Wrist |
| 1 | 3.93 | 3.70 | 4.43% | 8.12% | 14.38% |
| 2 | 3.29 | 3.64 | 5.22% | 8.15% | 10.15% |
| 3 | 4.33 | 10.50 | 3.24% | 5.87% | 7.39% |

### 3.3 Comparison Between Manual Methods and the "Gold Standard" and Time Reduction Analysis

Table 4 reports results of the comparison between manual methods and the automatic method base on kinematic data gathered from the IMU based movement analysis system. All $R^2$ values are very low, except when comparing wrist evaluation using methods 2 and 3. The worst results are obtained for the wrist: all the $R^2$ values are lower than 0.1.

**Table 4.** Differences in posture and technical actions count using method 1 and 2.

| Operator | Method 2–method 3 | | | Method 1–method 3 | | |
|---|---|---|---|---|---|---|
| | Shoulder | Elbow | Wrist | Shoulder | Elbow | Wrist |
| 1 | 0.14 | 0.46 | 0.05 | 0.08 | 0.06 | 0.06 |
| 2 | 0.14 | 0.09 | 0.06 | 0.10 | 0.11 | 0.00 |
| 3 | 0.02 | 0.28 | 0.01 | 0.07 | 0.04 | 0.01 |

From the analysis of these results, we can surely affirm that values obtained from the video analysis-based approach are far from the real and objective data. This is due to difficulties in evaluating movements in a 3D space. Difficulties are encountered also and above all when evaluating the wrist. This can be explained thinking at the small threshold and the presence of occlusions that can make it impossible to determine when awkward postures are reached.

Net time reduction (in both cases of about 98%) (Table 5) is a clear advantage of the automatic analysis. Data show that the mean time using method 1 is about 77 min, Companies can reduce notably time spent for risk analysis assessment. Moreover, they can monitor the real risk associated to working activities acquiring many data in a more reliable way. In this way the evaluation can be considered consistent.

**Table 5.** Mean time (min) for each method and time reduction (%) when using the automatic tool respect to the manual approach

| Trial | Time reduction (%) | | Mean time (min) | | |
|---|---|---|---|---|---|
| | Method 1 vs 3 | Method 2 vs 3 | Method 1 | Method 2 | Method 3 |
| 1 | 98.76 | 98.83 | 98.00 | 104.00 | 1.22 |
| 2 | 98.53 | 97.20 | 86.00 | 45.00 | 1.26 |
| 3 | 98.98 | 98.95 | 129.00 | 125.00 | 1.31 |
| 4 | 97.63 | 97.42 | 49.00 | 45.00 | 1.16 |
| 5 | 98.38 | 98.12 | 70.00 | 60.00 | 1.13 |
| 6 | 98.16 | 97.49 | 64.00 | 47.00 | 1.18 |
| 7 | 97.56 | 97.44 | 43.00 | 45.00 | 1.10 |
| Mean | 97.94 | 98.27 | 77.00 | 67.16 | 1.19 |

# 4 Conclusion

This work can be a good starting point to develop a complete toolbox, integrating user-friendly interface and quantitative data about worker kinematics, with the aim to help the operator conducting his risk analysis in a more efficient and objective way.

Results are promising and underlined that the analysis, although it is conducted by expert operators, may be subject to an intrinsic operator-dependent error when the evaluation is made only by observing videos. This is easy to comprehend thinking about the 3D nature of movement. Furthermore, the advantages of this kind of methodology are surely the possibility to measure the movement of workers in an easy and non-invasive way thanks to the use of the wearable inertial sensors. This kind of approach can also remove the inter and intra-operator variability due to the subjective way in which the large part of analysts in this field conduct their evaluations.

# References

1. ISO 11228:2007(E). Ergonomics - Manual handling, 01 April 2007
2. Colombini, D., Occhipinti, E.: L'Analisi e la gestione del rischio nel lavoro manuale ripetitivo: Manuale per l'uso del sistema OCRA per la gestione del rischio da sovraccarico biomeccanico in lavori semplici e complessi. Franco Angeli, Milano (2015)
3. McAtamney, L., Corlett, E.N.: RULA: a survey method for the investigation of world-related upper limb disorders. Appl. Ergon. **24**, 91–99 (1993)
4. Spielholz, P., Silverstein, B., Morgan, M., Checkoway, H., Kaufman, J.: Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. Ergonomics **44**(6), 588–613 (2001)
5. David, G.C.: Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. Occup. Med. **55**(3), 190–199 (2005)

# Evaluating an Inertial Measurement Unit Based System for After-Reach Speed Measurement in Power Press Applications

Richard F. Garnett[1(✉)], Gerard A. Davis[1], Richard F. Sesek[1], Sean Gallagher[1], Mark C. Schall Jr.[1], and Howard Chen[2]

[1] Department of Industrial and Systems Engineering, Auburn University, Auburn, AL, USA
{rfg0004, davisga, rfs0006, szg0036, mcs0084}@auburn.edu
[2] Department of Mechanical Engineering, Auburn University, Auburn, AL, USA
hzc0074@auburn.edu

**Abstract.** The objective of this study was to measure the hand speed of eighteen (18) subjects making after-reach movements from an upper palm button (UB) and lower palm button (LB) on a simulated press. Each after-reach movement was measured with a Vicon optical motion capture system and an Xsens IMU based system. A Bland-Altman analysis of the speed measured by the two technologies demonstrated a general agreement (average bias 0.19 m/s) between the measurements and a potential for using IMUs for hand-speed measures in the future. However, the computation intensity required to manipulate the Xsens data is likely too complex and time consuming for practitioners who are busy with plant activities. A simple IMU system, designed specifically for hand speed capture, could be a viable option for measuring after-reach speed in the future.

**Keywords:** Hand speed · Palm button · Machine guarding · After-reach IMU

## 1 Introduction

The OSHA 1910.217 (1971) standard established safety requirements for mechanical presses in the USA. The standard was revised in 1973 to improve machine control reliability and point-of-operation safeguarding [1]. One of the point-of-operation improvements was the development of a safe distance formula for dual palm buttons and light curtains to prevent after-reach errors. After-reach occurs when a worker attempts to reach into the press after the downward motion of the ram has been actuated. This is believed to be an inadvertent, automatic response to clear a process upset in the die [2]. The safe distance formula contains a hand speed constant based on the work of Löbl [3], who estimated that after-reach hand speed was 63 inches/second (or 1.6 m/s) [4].


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 146–157, 2019.
https://doi.org/10.1007/978-3-319-94619-1_14

Palm buttons are a two-hand-capture device intended to reduce the risk of hand injury when operating a press. They allow the machine to cycle only if both buttons are pressed at the same time, ensuring that an operator's hands cannot be in the die during the start of the downward stroke of the press. The use of two-hand palm buttons and light curtains reduces the risk of after-reach errors. After-reach is believed to occur when workers notice scrap in the die or a miss-set blank and attempt to clear it after the machine cycle has started, but before the die closes. Reaching into the die during the press's downward stroke can result in the worker's fingers, hand, or arm being in the point-of-operation when the press closes. The safe distance calculation reduces manual loading efficiency by moving the operator away from the point of operation while increasing safety. The "safe distance" calculation establishes a set-back distance that is intended to decrease the likelihood that an operator can physically reach (limited by the speed of his/her hands) into the press before it closes.

Numerous researchers have studied after-reach speed [5–8]. The majority of the researchers have reported after-reach speeds higher than the 1.6 m/s value found by Löbl in 1935. In March 1987, in response to the evidence amassed by research scientists (Pizatella, Moll, Horton, and others), The National Institute for Occupational Safety and Health (NIOSH) issued Current Intelligence Bulletin (CIB) 49 [2]. The CIB warned employers that the OSHA hand speed constant was not protective for all workers and that additional safety precautions were necessary for some workers. NIOSH did not recommend a method for evaluating press operators to determine if they were "exceeding the current OSHA hand-speed constant."

## 2 Inertial Measurement Systems

Inertial Measurement Systems (IMUs) are small, portable devices that measure acceleration, velocity, and position using miniature electro-mechanical sensors. IMUs are a product of micro-machining technology that allow the fabrication of computer board level mechanical devices. Most IMUs contain three orthogonal rate-gyroscopes and three orthogonal accelerometers, which measure angular velocity and linear acceleration [9]. Because gyroscopes are prone to drift, the system also contains three magnetometers which are designed to sense the earth's magnetic field and locate magnetic north. The output from the sensors are combined in a recursive algorithm called a Kalman Filter. This algorithm is designed to combine the data from the three types of "complementary sensors" (a technique termed sensor fusion) to minimize drift and allow the more accurate calculation of acceleration, velocity and position [10].

The design of the Kalman algorithm is a critical part of the accuracy of an IMU system. Since position data is calculated by the integration of sensor output, small errors in the accelerometer or gyroscope output builds rapidly in to larger errors in velocity (a single integrated value) and position (a double integrated value) [9]. Another source of IMU error is magnetic interference. The presence of iron (machines, rebar in floors, and so forth) in the test area can distort the earth's magnetic field and interfere with the magnetometer output, introducing error into the sensor fusion calculation [11].

A system of IMUs can be purchased in a fully integrated, commercial package that is designed for full body motion capture. For example, Xsens MVN Biomech Awinda (Xsens Technologies B.V., P.O. Box 559, 7500 AN Enschede, Netherlands) is a small, lightweight motion tracking system. It consists of seventeen, individual, wearable IMU (47 × 30 × 13 mm) sensors. The IMUs include an on-board battery and wireless communication to a proprietary motion capture software package. Xsens uses a proprietary data fusion algorithm which combines both sensor output and user entered anthropometric data. "The system is unique in its approach to estimating body segment orientation and position changes by integration of gyroscope and accelerometer signals which are continuously updated by using a biomechanical model of the human body" [10].

## 3 IMUs Versus Optical Motion Capture

IMU based MoCap systems, like Xsens, are an important innovation for human factors and ergonomics researchers since they have the potential to allow the less invasive measurement of workers on the factory floor. Traditional motion capture research has relied on optical motion capture systems (OMC). OMC systems use multiple cameras to triangulate the position of reflective markers secured to the test subject. OMC systems provide very precise position data collected at high speed (frame count 120 Hz or greater) but are not compact or portable. OMC position data can be used to calculate velocity, acceleration, joint angles and so forth and is accurate (system error is less than 2 mm during dynamic movement) [12]. Optical motion capture is considered the gold standard for segmental kinematic measurements [13].

However, OMC based research is usually confined to the research lab. OMC requires multiple cameras installed in multiple locations to view the target from multiple angles to clearly capture the movement. Lighting is critical and cameras must be rigidly mounted to minimize vibration and camera movement. The capture volume is limited and data capture is line of sight, so subjects cannot reach into blind areas (areas not seen by the cameras) or turn their body so the cameras cannot "see" the reflectors. OMC systems can be purchased as fully integrated, commercial packages. Vicon (Vicon, Oxford, 14 Minns Business Park, West Way, Oxford England, OX2 0JB) provides an integrated system that provides cameras, reflectors and software designed to capture and model the motion data.

The ease of use and potential for using IMU based motion capture in the workplace has generated a flurry of research activity to demonstrate IMU MoCap accuracy and compare it to the "gold standard" OMC. Table 1 summarizes the IMU validation effort and the finding generated:

**Table 1.** Summary of IMU accuracy research

| Study | Findings | Reference |
|---|---|---|
| Inertial measurements of upper limb motion | Development of portable IMU motion tracker to aid in home rehab of stroke victims. System demonstrated less than 5% error (arm angle, slow movements) compared to OMC | Zhou et al. [15] |
| Ambulatory measurement of arm orientation | Study of upper arm kinetics using an "ambulatory system." IMU vs OMC using RMS | Luinge et al. [16] |
| Inertial sensors for motion detection of human upper limbs | Study of portable IMU motion tracker to aid in home rehab of stroke victims. "The motion detector using the proposed kinematic model only has drifts in the measurements. Fusion of acceleration and orientation data can effectively solve the drift problem…" | Zhou and Hu [17] |
| Magnetic distortion in motion labs, implications for validating inertial magnetic sensors | Compared to OMC; Studied joint angle. Accuracy in the presence of iron was acceptable but performance deteriorated in 20–30 s | de Vries et al. [11] |
| Use of multiple wearable inertial sensors in upper limb motion tracking | IMU vs OMC measuring wrist & elbow joint angle. "Experimental results demonstrate that this new system, compared to an optical motion tracker, has RMS position errors that are normally less than 0.01 m, and RMS angle errors that are 2.5°–4.8°." | Zhou et al. [18] |
| Accuracy of inertial motion sensors in static, quasistatic, and complex dynamic motion | Tested IMU against OMC using RMSE analysis. Studied arm movement during static, quasi-static, and dynamic motion. RMS error was 1.9° to 3.5° during pendulum movement. Higher error in complex movement. Acceptable accuracy for field study | Godwin et al. [19] |
| Feasibility of using inertial sensors to assess human movement | Comparison of IMU system to electromagnetic sensors for hip measurement during walking. "We conclude that the inertial sensors studied have the potential to be used for motion analysis and clinical research." | Saber-Sheikh et al. [20] |

(*continued*)

**Table 1.** (*continued*)

| Study | Findings | Reference |
|---|---|---|
| Inertial measures of motion for clinical biomechanics: comparative assessment of accuracy under controlled conditions - effect of velocity | IMU "systems demonstrated good absolute static accuracy (mean error, 0.5°) and clinically acceptable absolute accuracy under condition of slow motions (mean error between 0.5° and 3.1°). Absolute and relative accuracy were significantly affected by velocity during sustained motions." | Lebel et al. [21] |
| Performance evaluation of a wearable inertial motion capture system for capturing physical exposure during manual material handling | Application of IMU in a work environment (lifting/material handling). Tested IMU against OMC to compare joint angle and joint velocity. "The IMU system yield peak kinematic values that differed up to 28% from OMC system." Acceptable accuracy for field use | Kim and Nussbaum [22] |
| A comparison of instrumentation methods to estimate thoracolumbar motion in field-based occupational studies | Compared to Lumbar Motion Monitor using RMSE. "Results suggest investigators should consider computing thoracolumbar trunk motion as a function of estimates from multiple IMUs using fusion algorithms…" | Schall et al. [23] |
| Accuracy and repeat-ability of an inertial measurement unit system for field-based occupational studies | "The accuracy and repeatability of an inertial measurement unit (IMU) system for directly measuring trunk angular displacement and upper arm elevation were evaluated over eight hours… Sample-to-sample root mean square differences between the IMU and OMC system ranged from 4.1° to 6.6° for the trunk and 7.2°–12.1° for the upper arm depending on the processing method | Schall et al. [24] |
| Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis | Compared to OMC; Evaluated Joint Angle using RMSE "Error remained under 5° RMSE during handling tasks, which shows potential to track workers during their daily labor." | Robert-Lachaine et al. [14] |

## 4 Research Aim

It is clear from the literature that the 1.6 m/s OSHA hand speed constant used in the “safe distance” calculation is not protective for all press operators. Multiple studies have found after-reach speeds that exceed the 1.6 m/s OSHA constant and that hand speeds vary significantly between subjects and based on palm button placement and orientation. NIOSH recommends that press operators with high hand speeds be evaluated to determine if they are “exceeding the current OSHA hand-speed constant.” These “at risk” operators would require additional safeguards.

The present study is designed to evaluate an Xsens (IMU-based motion capture) system and compare its accuracy to a Vicon (optical motion capture) system. If the Xsens accuracy is equivalent to Vicon, it would be possible for safety practitioners to conduct hand speed trials on the factory floor and identify press operators who are at risk for after-reach accidents.

## 5 Equipment

To estimate the accuracy of an IMU based system for after-reach speed measurements, subjects were fitted with two motion capture systems. The first system was an optical motion tracking manufactured by Vicon (Vicon, Oxford, 14 Minns Business Park, West Way, Oxford England, OX2 0JB). The OMC system used seven (7) Vicon T010 cameras collecting data at 120 Hz. The OMC data was processed using Vicon Nexus 1.8.5 2013 software. The second system was an Xsens Model MVN (Xsens Technologies B.V., P.O. Box 559, 7500 AN Enschede, the Netherlands) which collected data at 60 Hz. The IMU data was processed using Xsens MVN Studio BIOMECH Version 4.2.0 software.

All of the after-reach speed measurements in the literature were collected using a simulated press [4–8]. The target “hit” was captured using an electronic or photo-optic switch. The after-reach speed was determined by dividing the straight-line distance from the dual palm buttons to the target by the time measured from palm button release to target strike.

A simulated press was constructed for this experiment (Fig. 1). The fixture mimicked the geometry of commercially available power presses and included dual palm buttons at waist and shoulder height. The simulated two hand start buttons (56 cm apart) were mounted at waist height (84 cm from the floor) and shoulder height (160 cm from the floor) with a near waist height after-reach target (107 cm from the floor). These dimensions are typical of free standing, C frame presses used in industry and in the simulated press used in other experiments. The apparatus was constructed of wood (to minimize magnetic interference to the IMU system) and counterweighted to prevent tipping, with all sharp edges padded to prevent injury. The press was painted flat black to minimize reflectivity that might interfere with the OMC system (Fig. 1).

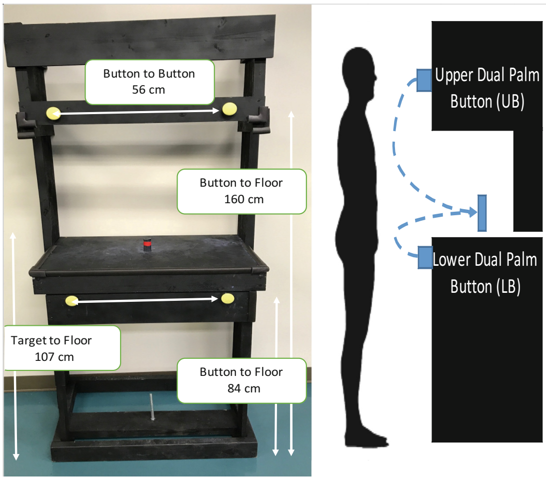



**Fig. 1.** Waist height and shoulder height dual palm button layout

# 6 Subjects

Data for this study was drawn from a larger experiment to measure after-reach speed. The criteria for experimental subjects were to be 19 to 30 years old, with no history of heart disease, stroke or breathing disorder that prevented exercise, and no medical history or injury to their back, shoulders, arms or hands that would prevent rapid movement. The experiment was conducted using twenty-seven (27) male subjects. The male subjects ranged in age from 20 to 29 years with an average age of 23.6 years old (SD = 2.97 years). On the day of the experiment, subjects were interviewed to insure they met the requirements of the study, they reviewed the experimental procedure, and signed a voluntary consent form. After signing the consent, the subject's measurements were taken (height and length measurements of their arms, legs, and arm reach). Vicon reflective markers and Xsens IMUs were then placed on the subject's shoulder, elbow, wrist, and hand. Only right hand dominant subjects were used in this experiment.

After placing the IMUs, the subject's anthropometric measurements were entered into the Xsens MVN Studio software, which allowed Xsens to establish a model of the subject's body segment lengths. The Xsens system was then calibrated with an "N-pose" (arms neutral beside the body) to establish the sensor to segment orientation for each IMU sensor [10]. The MVN Studio software rated the sensor to segment calibration quality (good, acceptable, fair, and poor); only systems calibration rated as good or acceptable were used in this study.

The study consisted of the subject reaching from the waist height palm button (LB) to the target and from the shoulder height palm button (UB) to the target (Fig. 1). The study was a randomized design with order of the conditions decided based on a coin flip. The subject was asked to practice each movement until they felt comfortable before an actual measurement was taken. The subject was given a "one-two-three-go" signal and then asked to make a rapid hand movement from the fixed starting position to the target position. The movement was repeated three times. After completing the three trials, the researcher explained the next start position, the subject was allowed to practice, and the subsequent trial was performed.

## 7 Data Collection

The majority of Xsens/Vicon comparison studies (Table 1) were position based. The x, y, and z positions for a movement were extracted from the IMU system and the OMC system and the position curves were compared using root mean squared error analysis. Accuracy results were reported as joint angles. For this study, the position data was extracted from both systems and converted to a velocity by dividing by time. This allowed the direct comparison of after-reach speed for each trial.

Several hurdles had to be overcome to develop an objective comparison between the two systems. Since the Xsens system was developed as a MoCap system, much of the data developed by the MVN Studio BIOMECH program is calculated for the center of the joints (for example the right wrist) rather than the appendages (the right hand). To allow a direct correlation with the Xsens position data for the right wrist during an after-reach event, Vicon markers were attached to the left and right sides of the wrist. The Vicon velocity data for both wrist markers was averaged to estimate the velocity for the center of the wrist.

During an after-reach movement, the operator's hand moves along a curvilinear path from the palm button (start position) to the target (Fig. 2). Velocity data for the wrist joint can be downloaded directly from Xsens. However, this velocity data is the angular velocity of the wrist as it travels along the curvilinear path to the target. This velocity is higher than the linear velocity. OSHA's hand speed constant is based on the linear velocity of the hand as it moves toward the target. To provide an "apples to apples" comparison of hand speed a Python program (Version 2.7) was used to extract the position data, approximate the linear distance from current position to the point of origin (linear distance), and calculate the average velocity of the movement over the travel path.

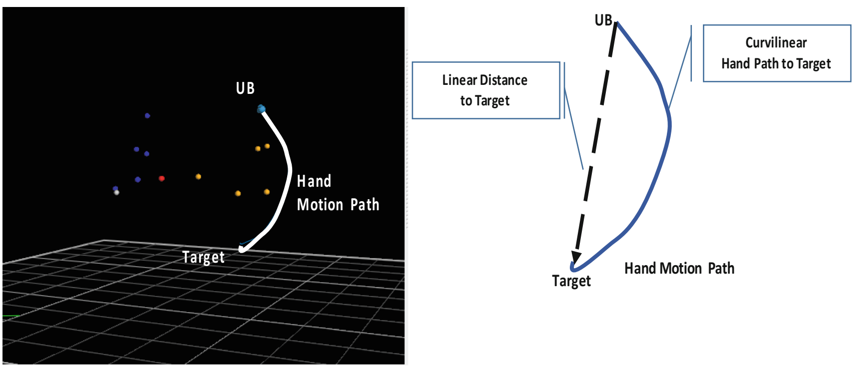



**Fig. 2.** Trace showing hand path traveled during an after-reach event (based on Vicon data)

## 8 Statistics and Analysis

The average velocity for each trial was compared using a Bland-Altman Plot. The Bland-Altman methodology was developed in the 1980's and is designed to compare two methods of measure and determine if they agree statistically [25]. If the methods "agree" then they are interchangeable and either can be used without a significant loss of accuracy [25, 26]. The Bland-Altman technique involves plotting the average of the reading taken by both devices against their difference for each trial. If the plotted points fall between the limits of agreement (LOA) for the process (two standard deviations from the mean) the instruments are in agreement. Bland and Altman have continued to refine the technique and extend their method to experiments with repeated measure in 2010 [26]. Zou suggested improvements to the method of calculating the LOA based on a one-way random effects model measures when the true experimental value varies trial to trial [27].

The study included twenty-seven (27) male subjects who attempted three trials at the waist height palm button (LB) location to the target and three trials from the shoulder height palm button (UB) location to the target. Only subjects that completed three complete trials without gaps in the position data were included in the analysis. Of the twenty-seven trials for the UB after-reach, three of the Xsens files were missing or corrupted and six of the Vicon files had gaps in the position data (subject's wrist rotated during the movement, obscuring the reflector for the right wrist marker) and were not included in the analysis. For twenty-seven runs LB after-reach trials three of the Xsens files were corrupted and six of the Vicon files had gaps in the position data (subject's wrist rotated during the movement, obscuring the reflector for the wrist marker) and were not included in the analysis.

The eighteen UB after-reach speed runs (3 trials/run) measured by Vicon averaged 1.40 m/s (SD = 0.24 m/s). This is the average speed of the left & right side of the wrist and represents the center of the wrist. The eighteen UB after-reach speed runs (3 trials/run) measured by Xsens averaged 1.22 m/s (SD = 0.23 m/s). An analysis of the normality of the difference of the two methods using the Anderson Darling Test ($H_o$ = Data follows a normal distribution) failed to reject the $H_o$ ($p_{value}$ = 0.7840). The Bland-Altman plot of the data is shown in Fig. 3. The Limits of Agreement were calculated using the method recommended by Zou. The Bland-Altman Plot for the UB after-reach (Fig. 3) indicates that all values fall inside the upper and lower limits of agreement. The difference of the values (Vicon-Xsens) indicate a bias of 0.18 m/s.

The eighteen LB after-reach speed runs (3 trials/run) measured by Vicon averaged 0.94 m/s (SD = 0.16 m/s). This is the average speed of the left & right side of the wrist (i.e. representing the center of the wrist). The eighteen LB after-reach speed runs (3 trials/run) measured by Xsens averaged 0.74 m/s (SD = 0.15 m/s). An analysis of the normality of the difference of the two methods using the Anderson Darling Test ($H_o$ = Data follows a normal distribution) failed to reject the $H_o$ ($p_{value}$ = 0.2631).

The Bland-Altman plot of the data is shown in Fig. 4. The Limits of Agreement were calculated using the method recommended by Zou [27]. The Plot for the LB after-reach shows that all plotted values fall inside the upper and lower limits of agreement. The difference of the values (Vicon-Xsens) shows a bias of 0.20 m/s. The bias between the UB (0.18 m/s) and LB (0.20 m/s) plots were different but an ANOVA of the differences ($H_o$ = values are equal) failed to reject the $H_o$ ($p_{value}$ = 0.6444).

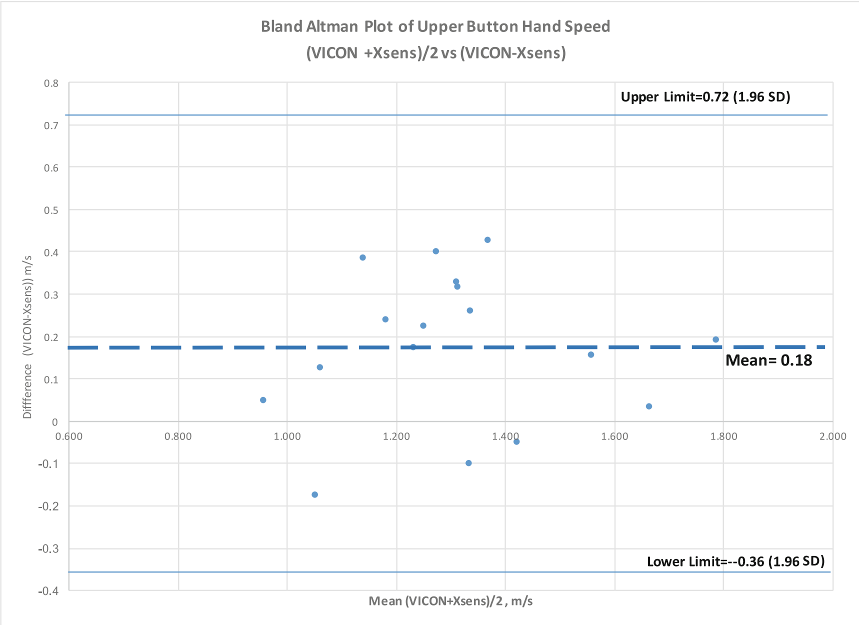



**Fig. 3.** Bland-Altman plot of UB after-reach speed, comparison of Vicon and Xsens

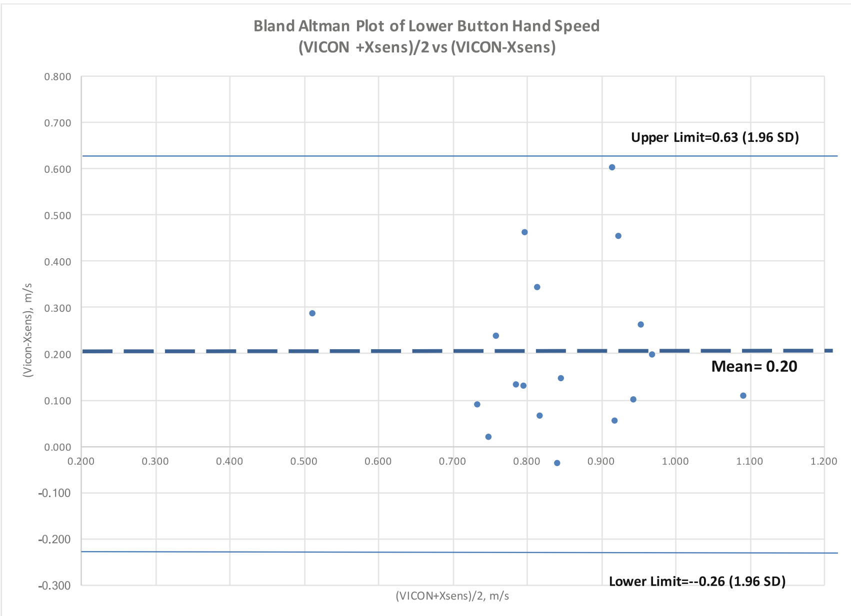



**Fig. 4.** Bland-Altman plot of LB after-reach speed, comparison of Vicon and Xsens

## 9 Discussion of Results and Conclusions

The data (UB/LB) were normally distributed and are good candidates for Bland-Altman analysis. Both plots demonstrated bias, but in general they indicate acceptable agreement between the two methods (average bias 0.19 m/s). The difference of the individual trials was calculated as "Vicon-Xsens", so the positive value of both biases indicates that the Xsens measurements are consistently lower than the Vicon measurements. This can be corrected by adding the bias value to the Xsens speed measurement.

The general agreement between the two speed measurement methods show a potential for using IMUs for speed measure in the future. However, the use of an Xsens system for this type of measure seems impractical. The Xsens software is designed to track joint velocity rather than finger or hand velocity. The computational intensive procedures required to download the joint position data, manipulate it to hand position, and calculate the linear velocity from this data may be too complex and time consuming for typical safety practitioners who are busy with day to day plant activities.

However, the IMU system exhibits reasonable accuracy for ballistic hand speed measurement. A simple IMU system, designed specifically for hand speed capture, could be a viable application for measuring after-reach speed in the future. The operation of IMUs in a ferromagnetic and electromagnetic rich environment, like that found on the factory floor, must be evaluated before the technology may be considered for adoption in industry.

**Acknowledgement.** This publication was partially supported by Grant # 2T420H008436 from NIOSH. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Use of trade names is for identification only and does not constitute endorsement by the Public Health Service or by the U.S. Department of Health and Human Service.

## References

1. Baldwin, D.M.: The power press controversy- a status report. Job Saf. Health (OSHA) **4**(5), 12–18 (1976)
2. NIOSH: CDC - NIOSH Publications and Products - Injuries and Amputations Resulting from Work with Mechanical Power Presses (87-107), Current Intelligence Bulletin 49 (2017)
3. OSHA Etool Website, CFR 1910.217 Appendix I. https://www.osha.gov/SLTC/etools/machineguarding/appendices/appendix_i.html
4. Lobl, O.: On the Two-Hand Insertion on Eccentric Presses-A Study of the Limits of the Protective Effect. Reichsarbeitsblatt (Berlin), 20 (part 111) (1935)
5. Pizatella, T.J., Moll, M.B.: Simulation of the after-reach hazard on power presses using dual palm button actuation. Hum. Factors **29**(1), 9–18 (1987)
6. Pizatella, T.J., Etherton, J.R., Jensen, R., Oppold, J.A.: Investigation of the after-reach hazard in two-hand controlled power press operations. Scand. J. Work Environ. Health **9**(2), 194–200 (1983)
7. Horton, J.T., Pizatella, T.J., Plummer, R.W.: The effect of palm button location on hand reach speed for power press operations. In: Trends in Ergonomics/Human Factors III. Elsevier Science Publishers B.V. (1986)

8. Jensen, R., Stobbe, T.: Safe distance for machinery actuators: is after-reach speed a constant?. In: Advances in Safety Management and Human Factors, pp. 321–331. Springer (2016)
9. Woodman, O.J.: An introduction to inertial navigation (No. Technical report 696). University of Cambridge, Computer Laboratory (2007)
10. Roetenberg, D., Luinge, H., Slycke, P.: Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors, vol. 3 (2009)
11. de Vries, W.H.K., Veeger, H.E.J., Baten, C.T.M., van der Helm, F.C.T.: Magnetic distortion in motion labs, implications for validating inertial magnetic sensors. Gait Posture **29**(4), 535–541 (2009)
12. Merriaux, P., Dupuis, Y., Boutteau, R., Vasseur, P., Savatier, X.: A study of Vicon system positioning performance. Sensors **17**(7), 1591 (2017)
13. Ceseracciu, E., Sawacha, Z., Cobelli, C.: Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept. PLoS ONE **9**(3), e87640 (2014)
14. Robert-Lachaine, X., Mecheri, H., Larue, C., Plamondon, A.: Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. Med. Biol. Eng. Comput. **55**(4), 609–619 (2017). https://doi.org/10.1007/s11517-016-1537-2
15. Zhou, H., Hu, H., Tao, Y.: Inertial measurements of upper limb motion. Med. Biol. Eng. Comput. **44**(6), 479–487 (2006)
16. Luinge, H.J., Veltink, P.H., Baten, C.T.M.: Ambulatory measurement of arm orientation. J. Biomech. **40**(1), 78–85 (2007)
17. Zhou, H., Hu, H.: Inertial sensors for motion detection of human upper limbs. Sens. Rev. **27** (2), 151–158 (2007)
18. Zhou, H., Stone, T., Hu, H., Harris, N.: Use of multiple wearable inertial sensors in upper limb motion tracking. Med. Eng. Phys. **30**(1), 123–133 (2008)
19. Godwin, A., Agnew, M., Stevenson, J.: accuracy of inertial motion sensors in static, quasistatic, and complex dynamic motion. J. Biomech. Eng. **131**(11), 114501 (2009)
20. Saber-Sheikh, K., Bryant, E.C., Glazzard, C., Hamel, A., Lee, R.Y.W.: Feasibility of using inertial sensors to assess human movement. Man. Ther. **15**(1), 122–125 (2010)
21. Lebel, K., Boissy, P., Hamel, M., Duval, C.: Inertial measures of motion for clinical biomechanics: comparative assessment of accuracy under controlled conditions - effect of velocity. PLoS ONE **8**(11), e79945 (2013)
22. Kim, S., Nussbaum, M.A.: Performance evaluation of a wearable inertial motion capture system for capturing physical exposure during manual material handling. Ergonomics **56**, 314–326 (2013)
23. Schall, M.C., Fethke, N.B., Chen, H., Gerr, F.: A comparison of instrumentation methods to estimate thoracolumbar motion in field-based occupational studies. Appl. Ergon. **48**, 224–231 (2015)
24. Schall, M.C., Fethke, N.B., Chen, H., Oyama, S., Douphrate, D.I.: Accuracy and repeatability of an inertial measurement unit system for field-based occupational studies. Ergonomics **59**(4), 591–602 (2016)
25. Altman, D.G., Bland, J.M.: Measurement in medicine: the analysis of method comparison studies. J. R. Stat. Soc. Ser. D (Stat.) **32**(3), 307–317 (1983)
26. Bland, J.M., Altman, D.G.: Statistical methods for assessing agreement between two methods of clinical measurement. Int. J. Nurs. Stud. **47**(8), 307–310 (2010)
27. Zou, G.: Confidence interval estimation for the bland-altman limits of agreement with multiple observations per individual. Stat. Methods Med. Res. **22**(6), 630–642 (2013)

# Smart Clothing Design Issues in Military Applications

Sofia Scataglini[1,2(✉)], Giuseppe Andreoni[3], and Johan Gallant[1]

[1] Royal Military Academy, 1000 Brussels, Belgium
sofia.scataglini@rma.ac.be
[2] Military Hospital Queen Astrid, 1120 Brussels, Belgium
[3] Department of Design, Politecnico di Milano, 20158 Milan, Italy

**Abstract.** Smart clothes development history started in the military field and this still remains a main application field. A soldier is like a high-performance athlete, where monitoring of physical and physiological capabilities of primary importance. Wearable systems and smart clothes can answer this need appropriately. Smart cloth represents a "second skin" that has a close, "intimate" relation with the human body. The relation is physiological, psychological, biomechanical and ergonomical. Effectiveness of functional wear is based on the integration of all these considerations into the design of a smart clothing system. The design of smart cloth is crucial to obtain the best results. Identifying all the steps involved in the co-design workflow can prevent a decrease in wearer's performance ensuring a more successful design. This paper presents all the steps involved in the workflow for the design of a proposed solution of a smart garment for monitoring soldier's performance.

**Keywords:** Human factors · Smart clothing · Wearable monitoring Military

## 1 Introduction

Smart clothing or "intelligent textile" represents the new class of wearable textile design era 2.0 with interactive technologies, intended to be attractive, comfortable and 'fit for purpose' for the identified user. The fields of applications are very different such as healthcare, fitness, sport [1], lifestyle, space exploration, public safety, and military [2].

The first idea to embed sensors into the garments in the military field was by researchers of the Georgia Institute of Technology on a US Defense Advanced Research Project Agency grant [3]. The main goal was to monitor the status of the soldier and to reveal eventual injuries and their influence on his/her health. Scataglini et al. [4] classified the current main applications of smart clothing in military field in health monitoring, environmental safety monitoring, stress management and empowering human function.

Sensing clothing offers the unique opportunity to implement a non-intrusive monitoring, i.e. they represent an extraordinary tool for observing and analysing the complex Human-Machine-Environment system in specific tasks (Fig. 1).


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 158–168, 2019.
https://doi.org/10.1007/978-3-319-94619-1_15

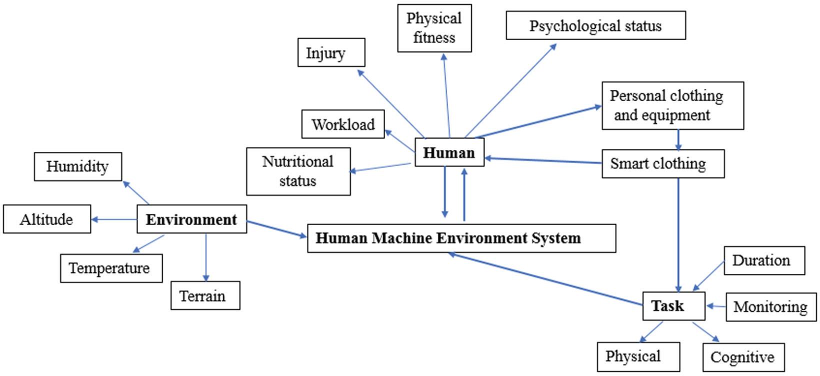



**Fig. 1.** Human-product-task system performance and the main parameters to be monitored specifically in the military applications.

The main requirements that a smart clothing is called to achieve are functionality, usability, monitoring duration, wearability, maintainability and connectivity [5], but comfort is another key and transversal factor (also among the previous features) to be considered. The impact of comfort on wearable monitoring technologies has been recognized as an important aspect of its design. Wearing an uncomfortable system compromises the soldier's ability to do his/her job [6].

The functional design process, as well as the knowledge and intuition about the body interface gained from the study of functional design methodologies can help to broaden the scope of interdisciplinary variables considered in the design of wearable technology, and thereby produce a more successful design. The design process should begin through the anthropometric data retrieval and analysis of the user, and the identification of the user's needs. Anthropometry, or measurement of the body, is a key of the clothing design and the placement of smart textiles around the body. Volume, shape, weight, and adherence to the body of wearable devices must be designed to not affect or interfere with natural movements. Design must also consider the wearability in situation-specific movements required for the accomplishment of a task. In more extreme situations, where weight is a crucial factor as well as volume, stationary or dynamic balance should be preserved and not modified.

Finally, but not exhaustively, other architectural requirements are the connectivity between the textile sensor/component and the electronic part and the connectivity towards the external world.

Once these design criteria have been established, the initial aesthetic design is created within the framework of the user's needs.

The purpose of this paper is to present the methodological approach and the definition of requirements and specifications for the design of smart clothing for military applications, using Belgian soldiers as case study.

## 2 Methods

Wearable have a close, even intimate, relation with the human body. This relation is physical, physiological, and functional. For this reason, the design of "intelligent" garments involves two different macro areas: design issues and technological issues (Fig. 2). The first one comprehend specifically physical and design related factors such as anthropometry, gender issues, body positions, wearability, elasticity and adherence of the body fixing element or of the garments. The technological issues comprehend the requirements related to the technology as sensing and processing, data transmission and power supply [7]. Design and technological issues are the two main macro areas involved in the process together with the esthetical one.

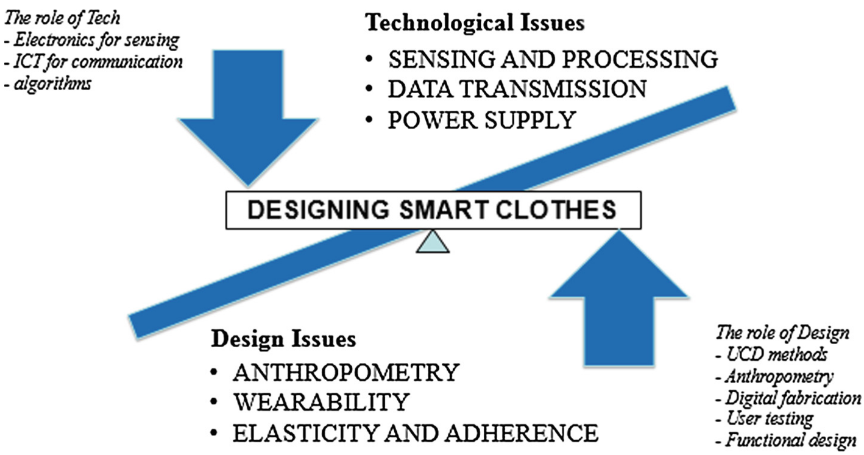



**Fig. 2.** The two macro-areas of the design process

The design thinking process begins at the empathise stage with the analysis of the anticipated user and the identification of the end-user needs (Fig. 3).

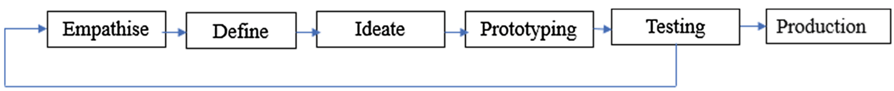



**Fig. 3.** Design thinking process.

### 2.1 Empathise: User Analysis

This step consists in the observation of the user by questionnaire. An user experience questionnaire 2.1.1 was designed to investigate the importance of the following features in designing smart clothing focus on the military field.

### 2.1.1 Questionnaire

Please rate the following variables on a scale. From 1–5 as to the importance they play in your perception of comfort and use.

5 very important
4 of some importance
3 of moderate importance
2 of little importance
1 of no importance

Place a number from 1 to 5 that indicates the rating you would give that variable.

1. Ease of movement
2. Texture
3. Weight of the fabric
4. Color
5. Fit
6. Thermal
7. Care of the garment
8. Self-confidence in appearance
9. Transmission and data storage
10. Easy to clean
11. Functionality and reliability
12. Ease of use
13. Proxemics (as human perception in space)
14. Use for sport activity
15. Use for military training
16. Do you like the idea to have it as military equipment? If the answer is yes, do you have any suggestions?

The empathy phase highlighted that the smart clothing should be a smart shirt addressed to the soldier in military training and can be included as part of the military equipment.

In fact, the soldier is like a high-performance athlete, but he/she must perform at a lower physical level for an extended period and career. The physical efforts of the tactical athlete cannot be picked out or periodized like for sports athletes. The prediction of events is highly uncertain in the military context. Musculoskeletal injuries are the leading cause for the medical profile during training and military operations. Big efforts are made in the screening and prediction of these kinds of injuries. It is very clear that the multifactorial ground of these injuries makes them hard to predict.

The performance and underperformance of a tactical athlete do not solely depend on physical factors but also on mental fitness, nutrition, rest & recovery and so on. The (re)occurrence of musculoskeletal injuries follows the same path. Smart technology can allow collecting many data on the physical load on the tactical athlete, the physiological response to it, the behavior of the individual, etc.…

This technology enables the military rehabilitation specialist to monitor performance of the patients defining specific protocols based on algorithms and end-user interface adapted to the user's needs.

Body-worn health monitoring systems measure physiological signals in real time to track individuals' physical condition and performance.

Combining data from established human monitoring technology with environmental and performance information, provides detailed live feedback from military rehabilitation training, monitors their well-being and records changes in ability over time.

## 2.2 Ideate: The Concept

Designing smart clothes requires as first task the identification of measuring parameter requirements, whose first need is determining sensor placement onto the body and their wearability. For sensor positioning, according to [8], we need to consider areas that are relatively the same size across adults, areas that are larger as surface areas and finally areas that have low movements (Fig. 4).

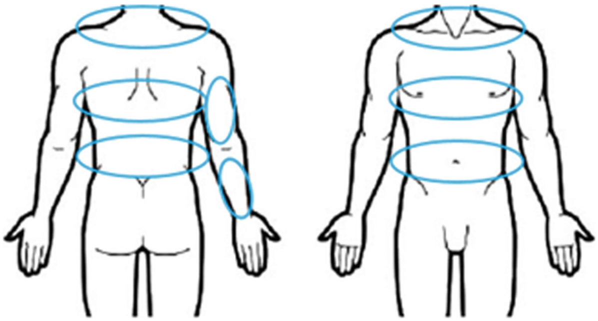


**Fig. 4.** Areas of the upper body parts for the sensor placement according with Gemperle et al. [8] criteria.

Considering the Belgian soldier equipment, we eliminated the neck areas and the arms areas. We chose the chest areas due to the equipment interaction and the shooting movements (Fig. 5).

Washability [9] and sterilization or disinfection should be considered if applicable in special applications like the military. When protection is a driving force, other factors must be taken into consideration. In addition to fabric durability, air permeability, pore size and moisture transport, factors such as dust infiltration, dirt, and sand resistance, are expected to play a vital role in both, comfort and protection.

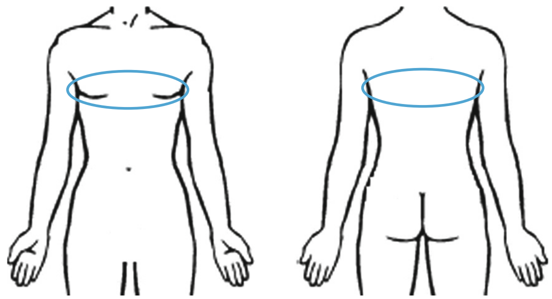


**Fig. 5.** Areas of the upper body parts for the sensor placement of our smart shirt as an unobtrusive placement.

Elasticity also plays an important aspects permitting a close contact between the sensor and the skin. Elasticity is obtained by the design impose criteria that preserve comfort and wearability of the smart clothes. High stretch fabrics provide mobility, tight-fitting and enhanced performance. At the same time the adherence should not to be invasive for posture and movements or create a thermal discomfort. According to these criteria a smart shirt was designed with stretch fabrics and cut close to the body for enhanced performance. The observation of the upper body soldier equipment (such as the rucksack, vest and helmet combined with motor task (jumping, climbing, shooting, running)) gave the idea of the first constraints and requirements in the design of smart garment for soldier. The smart cloth was designed to be used as underwear garment. This means that adding equipment on it should not create inter-ference with the sensing technology embedded in the cloth. Ruck-sack strap position should be tested in the ideate steps as while impact protection and weapon interactions.

The intelligent garment is a washable smart shirt based on the wearable 2.0. is intended for monitoring the heart rate activity and the body acceleration (Fig. 6).

The base fabric is complemented with textrodes textrode (textile electrodes) that are constituted of electrically conductive yarns that are in contact with the skin. The main advantage of using the textrode is the non invasivity in terms of skin irritation for long monitoring due to the unnecessary use of the electrogel. They are embedded into the clothes and positioned on the 10th rib at the distance of 15 cm between each other enabling transthoracic electrical bioimpedence measurements [10].

We distinguish metal yarns (stainless steel, copper or silver mixed with natural or synthetic fibers) and yarns containing electro-conductive fibers (polymeric or carbon coated threads). In that case, the textrode does not required electro conductive gel but sweat (perspiration). As well, strain gage textile sensors are used to analyse breathing parameters. Mechanically, they elongate in relation to body motion structures [5].

Data connection between the textile/component is facilitated by two fasteners (nickel free material). The device is supported by a mobile app designed at POLIMI for collecting and managing data related to 3D accelerations (Anterior-Posterior, Medio-Lateral and Vertical) and the ECG potential in real-time.

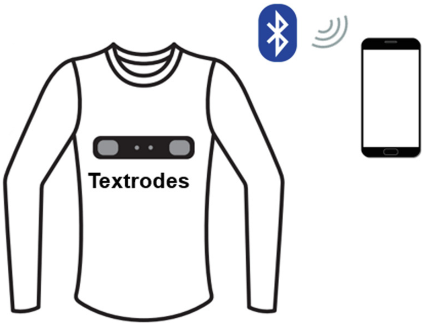



**Fig. 6.** The smart system.

## 2.3 Prototyping: Garment Design

Once these criteria have been established, the initial esthetic design is created within the framework of the user's needs.

A co-design workflow can be used to define the functional cloth.

Iterative studies are evaluated and re-evaluated based on physiological, ergonomical and biomechanical monitoring of the wearer's performance. Therefore, alternative solutions are generated for each decision. Iterative co-design steps (Fig. 7) were used to influence the modifications made from the first prototype, and the design process began again [11]. This ensured that corrections had been made before the design was finalized.

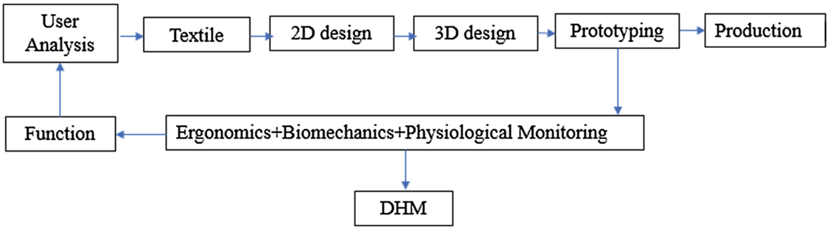



**Fig. 7.** Garment co-design workflow.

Garment can be design manually drafted (traditional) or computer aided drafting. Both require the illustration of the flat drawing, the patter making and grading, and the process sheet making.

Garments were designed using the traditional 2D flattened pattern approach.

Traditional pattern making refers that pattern makers draw garment constructions lines on a paper. Anthropometry plays an important role in pattern grading.

Starting from an anthropometric approach collected on 1615 soldiers we defined the "average anthropometric soldier measure" [12]. To obtain the single segment length we used Drillis and Contini [13] that expresses the segment length as a function of the body height (Fig. 8).

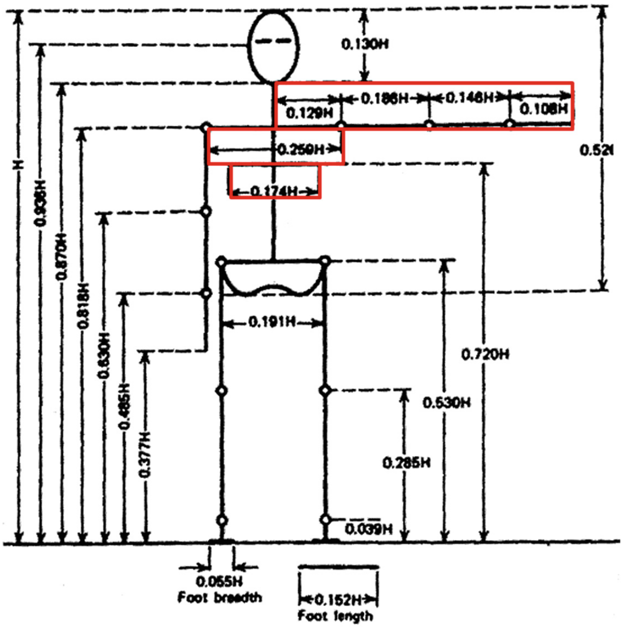



**Fig. 8.** Body segment lengths expressed as proportion of body height (H) by Drillis and Contini and the identification of body segment lengths related to the thorax (marked in red) expressed as proportion of body height (H).

Thanks to this calculation, we were able to define the horizontal (width at chest and waist) and the vertical measurements (sleeves length and total length) of the smart shirt.

In term of prototyping, three smart cloths were realized [14].

The first one was inspired by body mapping of the sweating in male athletes [15]. Qualitative study on a soldier population reveals that the shirt was not comfortable because of a limited movement at the armpit (Fig. 9).

Adhesives and bonding films were used to join the fabrics together creating the 3D shell ensuring performance and reliability with excellent bond strength and washability.

A second one solved that limit thanks to the use of stretch fabric at the armpit permitting comfort and flexibility (Fig. 10).

**Fig. 9.** The first prototype-front

That feature allows us to extend the use of the smart shirt in sports such as climbing, rowing, archery, basketball, and volley-ball where a high range of motion at the armpit is requested.

The second prototype presents a hidden pocket able to transform the smart cloth by inserting a pad of smart material for shock absorption and impact protection during activities like shooting.

**Fig. 10.** The second prototype-front.

## 2.4 Testing

Comfort includes the physiological and psychological aspects on one hands, and the biomechanical and ergonomical aspects on the other. More attention should to be paid to understanding ergonomic issues, heat stress and the clothing. The degree of thermophysiological comfort is defined by the thermophysiological characteristics of the textile as well the range of the motion while performing a task.

The thermophysiological can be measured using a thermal imaging technique. While the range of motion can be measured using an inertial motion capture system.

Nowadays Digital Human Modeling (DHM) offers innovative possibilities to integrate physiological, kinematic and kinetics data providing a data representation of the virtual soldier wearing a smart cloth during a motor task (Fig. 11).

The goal is to identify and quantify the main critical factors influencing the performance in terms of effort and correct task completion and to evaluate the influence of the equipment design on this performance.

**Fig. 11.** The DHM with the second prototype.

Iterative studies on the co-design workflow permitted us to redefine the final functional garment realizing a third prototype, a vest that can be used for monitoring soldier's performance in terms of training, injuries, and psychological status monitoring inside and outside the water.

This concept aims at reducing the number of patterns and consequently the number of joins of the fabrics permitting more comfort and flexibility. The vest is a breathable fitting garment that gives freedom of movement.

## 3 Conclusion

In this paper, we have presented some issues related to the design smart clothing for military applications, using Belgian soldiers as case study. Smart t-shirt monitoring capabilities can be extended from the lab to the field permitting a continuous physiological monitoring (HR and HRV) and the wearer's gesture detection.

The smart technology equipment is composed by:

- A wearable device (waterproof package) measuring and recording motion data (3D accelerations) and physiological data (heart rate variability).

- A smart shirt (long sleeves) with two textrodes embedded into the cloth. The new shirt presents a hidden pocket inserted on the right shoulder. This gives the possibility to transform the shirt inserting a pad of smart material for shock absorption and impact protection during activities like shooting.
- A smart vest with two textrodes embedded into the cloth.
- A mobile app for real-time processing and visualization.

**Acknowledgement.** The authors would like to thank you all the subjects that took parts in the study for their patience, commitments, and motivated collaboration.

## References

1. Perego, P., Moltani, A., Andreoni, G.: Sport monitoring with smart wearable system. Stud. Health Technol. Inf. **177**, 224–228 (2012)
2. Sahin, O., Kayacan, O., Bulgun, E.Y.: Smart textile for the soldiers of the future. Defence Sci. J. **55**, 195–205 (2005)
3. Bonato, P.: Advances in wearable technology and applications in physical medicine and rehabilitation. J. Neuroeng. Rehabil. **2**(1), 2 (2005)
4. Scataglini, S., Andreoni, G., Gallant, J.: A review of smart clothing in military. In: Proceeding WearSys@MobiSys 2015, pp. 53–54 (2015)
5. Gilsoo, C.: Smart Clothing: Technology and Applications. CRC Press, Boca Raton (2009)
6. Tharion, W.J., Buller, M.J., Karis, A.J., Muller, S.P.: Acceptability of a wearable vital sign detection system. In: Proceedings of the Human Factors and Ergonomics Society (2007)
7. Andreoni, G., Standoli, C.M., Perego, P.: Sensorized garment for biomedical monitoring. Design issues. In: International Conference of Sensors and Applications (2015)
8. Gemperle, F., Kasaback, C., Stivoric, J., Bauer, M., Martin, R.: Design for wearability. In: Proceedings of the 2nd IEEE International Symposium on Wearable Computers (1988)
9. Tao, X., Koncar, V., Huang, T.-H, Shen, C.L., Ko, Y.C., Jou, G.T.: How to make reliable, washable, and wearable textronic devices. In: Chung, H.-J., Kim, T. (eds.) Sensors, Basel, Switzerland (2017)
10. Scataglini, S., Truyen, E., Perego, P., Gallant, J., Tiggelen, D.V., Andreoni, G.: Smart clothing for heart rate variability measures in military. HBIM J. **1**, 74 (2017)
11. Scataglini, S., Truyen, E., Perego, P., Gallant, J., Tiggelen, D.V., Andreoni, G.: Smart clothing for human performance evaluation: biomechanics and design concepts evolution. In: 5th International Digital Human Modeling Symposium, Germany, Bonn (2017)
12. Scataglini, S., Andreoni, G., Truyen, E., Warnimont, L., Gallant, J., Tiggelen, D.V.: Design of smart clothing for Belgian soldiers through a preliminary anthropometric approach. In: Proceedings 4th DHM Digital Human Modeling, Montréal, Québec, Canada, 15–17 June (2016)
13. Drillis, R., Contini, R.: Body Segment Parameters, Report 1166-03. Office of Vocational Rehabilitation, New York (1966)
14. Scataglini, S.: Ergonomics of gesture: effect of body posture and load on human performance, Joint Ph.D. Politecnico di Milano and Belgium Royal Military Academy (2017). (https://www.politesi.polimi.it/handle/10589/136840)
15. Smith, C.J., Havenith, G.: Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia. Eur. J. Appl. Physiol. **111**(7), 1391–1404 (2011)

# Design of Intelligent Obstacle Avoidance Gloves Based on Tactile Channel

Tongtong Zhang(✉) and Lei Zhou

School of Mechanical Engineering, Southeast University, Nanjing, China
{zhangtongtong, zhoulei}@seu.edu.cn

**Abstract.** Nowadays there is a global concern with driving safety issues. This article proposes a system based on tactile cues, designed to help drivers find obstacles while driving and achieve safe driving. The feature of this system compared to other systems is that it uses a tactile perception and reminding method that has not been used in the previous designs of the same type to build a simple tactile display to prompt the user. In addition, a user use test was performed for the vibration mode proposed in this issue, and the accuracy and naturalness of the system were evaluated by the detection of the user's use of sensations. At the same time, the design is completely independent of the vehicle's system. It can be installed not only on bicycles and motor vehicles, but also on special vehicles such as disabled bicycles and urban sweepers.

**Keywords:** Arduino · Vehicle · Obstacle avoidance · Electronic system

## 1 Introduction

Volvo proposes Eye Car technology, the working principle is that the position of the eyes of the driver is first read by the eye position sensor, and then the motor is automatically used to lift the seat to the optimal height position, thereby providing the driver with the best sight to grasp the road conditions. At the same time, the motor will automatically adjust the steering wheel, pedal, center console and even floor height to provide the most comfortable driving position [1]. Finally, the B pillar of the car was redesigned and removed from the driver's eyes [2]. The core of Ford proposes CamCar technology is to use micro-cameras to detect road conditions. After computer processing, the measured information is presented to the driver from a central main display and two additional side displays. The displayed image will be transformed according to real-time conditions, providing the driver with information on the lateral and visual dead angles of the vehicle [3]. Mazda proposes SensorCar technology, sensors in the rear bumper monitor the traffic flow at the rear of the vehicle. When a rear-end collision is about to occur, the warning system will activate the electric seat belt pretensioner, automatically tighten the seat belt, the warning icon on the instrument panel will light, and the rear speaker will alarm. Therefore, it plays a very good protection and early warning effect [3]. Mercedes-Benz proposes preventive fatigue prevention technology. If the driver's operation is judged by the system to become obstructed by rapid stability, and the vehicle has been driving continuously for more than two hours, the icon of the small coffee cup in the center of the instrument panel will light up. This reminds


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 169–179, 2019.
https://doi.org/10.1007/978-3-319-94619-1_16

the drivers that they need to rest and this icon will disappear after the parking flameout [3–7]. Tian Zhihong (2007) used ultrasonic distance measurement principles to design an automatic wheelchair capable of intelligent obstacle avoidance. Multiple sets of ultrasonic ranging sensors were installed in various directions of the wheelchair, and the real-time distance value measured by the ultrasonic ranging sensor was processed. In addition, a certain obstacle avoidance strategy was added to achieve the function of helping people with mobility difficulties to avoid obstacles intelligently [8]. The Ricardo Queiros and Francisco Correa Alegria (2001) of the University of Lisbon applied the cross-correlation and sine generator technology to the ultrasonic ranging system, which further improved the resolution of the ultrasonic ranging system and drastically advanced the ultrasonic ranging technology. The design mainly used two methods to improve the resolution of ultrasonic distance measurement in the air. First, ultrasonic transmission and reception signals and cross-correlation technology are combined to detect the propagation time of ultrasonic waves in the air [9]. Second, the sine generator and other transmitted and received signals are used to detect the phase shift generated during propagation of the ultrasonic wave in the air, and the sine generator can effectively eliminate the influence of the phase shift, so that a very high ultrasonic ranging resolution can be obtained [10]. Joseph C. Jackson et al. of the Imperial College of London jointly published a book describing in detail the principles and implementation of ultrasonic distance measurement using the transit time method. After research, they eventually applied transit time ultrasonic ranging technology to mobile robots to avoid obstacles, internal structure detection and medical imaging, etc. [11]. Infrared distance measurement under good weather conditions, the actual atmospheric extinction coefficient value can be obtained by actual measurement or calculated by software such as Lowtran, Hitran, etc. Based on the actual parameters of the atmosphere, and a farther measurement distance can be obtained with good distance resolution [12]. The disadvantage of this method is that it needs precise atmospheric conditions parameters, that is, it has poor generality. it needs to change the corresponding parameters every time. These parameters will fluctuate greatly with changes in the weather, and these parameters are not easily measurable, and are difficult to apply in electronic systems in a short time [13]. The advantages of infrared ranging sensors are that they are cheap, easy to use, safe and practical; and infrared rays cannot penetrate obstacles, so no signal crosstalk occurs [14].

This paper summarizes the status quo of domestic bicycle safety, as well as the existing safety assisted driving systems on the market. Based on the Arduino, a new traffic safety warning system consisting of three subsystems for measurement, calculation and implementation was built to constantly detect nearby obstacles and remind cyclists to avoid obstacles. In this project, ultrasonic sensor probes are used to detect surrounding obstructions in real time. Through the calculation and processing of the Arduino development board, a completely different way from the previous technology, the vibration sense haptic gloves is used to warn the user of the presence of obstacles. This enables effective and rapid delivery of information without affecting driving.

## 2 Overall System Design

To complete the basic functions of the system, summarize the existing systems and summarize the three subsystems necessary for the system to complete the functions.

The sensor usually completes the main functions, is responsible for continuously collecting the surrounding road conditions and transmitting it to the computing system, and can also be supplemented by a simple circuit to convert the electrical signals transmitted by the sensors to the computing system. One or more microprocessors working together to collect and process the electrical signals sent by the detection system, and output electrical signals to the execution system so that the execution system can complete the corresponding tasks. Composed of LED, buzzer, motor and other original parts, responsible for the implementation of the information output by the computing system, the implementation results are perceived by the user. According to the functions of these three subsystems, the system workflow is designed, as is shown in Fig. 1.

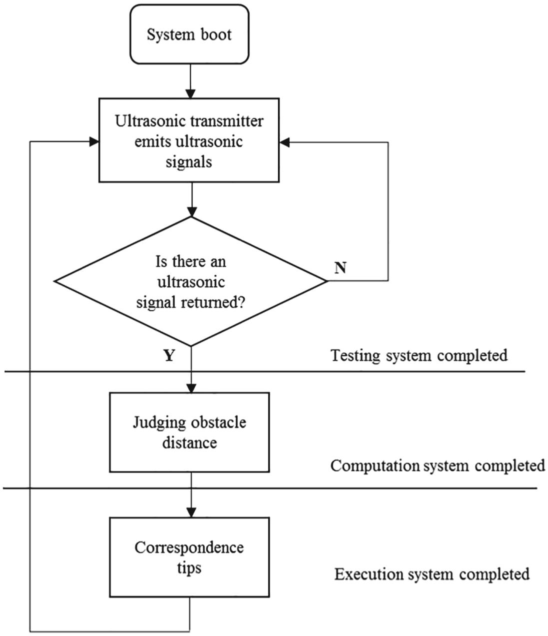



**Fig. 1.** System workflow

## 3 Detection System Design

This chapter classifies and enumerates various types of obstacles encountered during the course of driving a bicycle, analyzes its characteristics one by one, and proposes a theoretical identification method.

Since multiple sensors are mounted on the same plane, the position of the sensor can be determined by two-dimensional coordinates after determining the direction in which the sensor is facing. Having determined the two-dimensional coordinates of each sensor, the type of the obstacle can be initially determined by the time difference of the data received by each sensor. However, this method can only determine the volume and speed of obstacles, so the characteristics of the obstacles are mainly analyzed in both volume and speed. Common driving obstacles fall into three categories, as shown in Fig. 2.

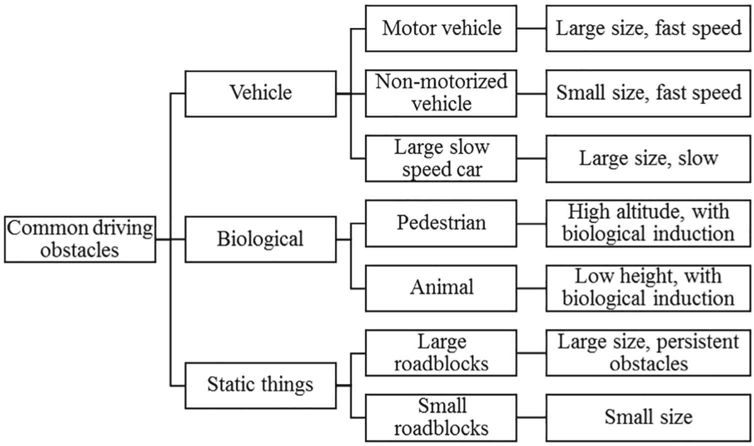



**Fig. 2.** Common driving obstacles and basic features

When there is a motor vehicle approaching from the rear, a plurality of ultrasonic sensors installed behind the bicycle detect obstacles at the same time, and a plurality of sensors detect that the distance is getting closer and closer, and the approaching speeds are substantially equal, and the presence of the vehicle is identified. When passing on both sides, the sensor recognizes that the order of the motor vehicle is from the back to the front, and the vehicle can be identified by setting a certain delay time range after determining the distance of the sensor. Non-motorized vehicles can only be detected by one sensor. When approaching from the rear, only one or a plurality of vertically mounted sensors detect the obstacles getting closer and closer, and the approaching speeds are substantially equal, identifying the existence of a non-motor vehicle. When passing from both sides, the rear sensor first recognizes the presence of an obstacle, and then the obstacle leaves the detection range. After a short time, the front sensor recognizes the presence of an obstacle, and the distance from the obstacle identified by the rear sensor is basically the same. It can be determined as a non-motor vehicle. Large-scale slow-moving cars usually emit loud sounds that are easily noticed by the rider and are therefore not considered.

The method of detecting biological and stationary objects is similar to the method of detecting the vehicle.

Through the above methods, it is possible to complete the discrimination and recognition of common driving obstacles by the sensors. Through the calculation and comparison of the calculation system, a corresponding prompt for the rider can be given.

According to the HC-SR04 Ultrasonic Ranging Module Instruction Manual, first the module's trigger signal input and recall signal output need to be defined (for example, the rear-side sensor, the same below):

```
int EchoB=13;
int TrigB=12;
```

Set port status on Arduino:

```
pinMode(EchoB, INPUT);
pinMode(TrigB, OUTPUT);
```

10 microsecond high signal to the TRIG pin:

```
digitalWrite(TrigB,LOW);
delayMicroseconds(2);
digitalWrite(TrigB,HIGH);
delayMicroseconds(10);
digitalWrite(TrigB,LOW);
```

Record ECHO pin high time with cmB:

```
cmB=pulseIn(EchoB,HIGH);
```

Convert time to distance:

```
cmB=cmB/58;
```

The value recorded by cmB is the distance from the rear obstacle to the sensor in centimeters.

## 4 Executive System Design

The design flow of the execution subsystem system is a cyclic process that proposes the model, implements the model, determines the user and finally determines the overall scheme of the system, as shown in Fig. 3. This mode requires that a specific vibration mode scheme be put forward, and the scheme should be embodied as an entity, and then the actual user use test should be conducted to evaluate the accuracy and naturalness of the vibration mode.

### 4.1 Execution System Design

The difference in the intensity of the vibration, the time of the vibration and the area of the vibration allows the user to recognize the vibration generated by different distances. In the process of riding, enough attention is needed to focus on the visual and auditory. The rider's attention to touch is not too high, and people's perception of vibration intensity often requires higher attention. Therefore, this project does not choose the vibration intensity as the basis for distinguishing vibrations. The time of vibration is relatively easy to perceive. Although everyone has a different perception of time, most people have the ability to compare the vibration time, so the length of the vibration time can be used as one of the elements of the oscillator vibration mode design. The vibrating area is the easiest to discriminate and will also be used as an element of the vibration pattern design.

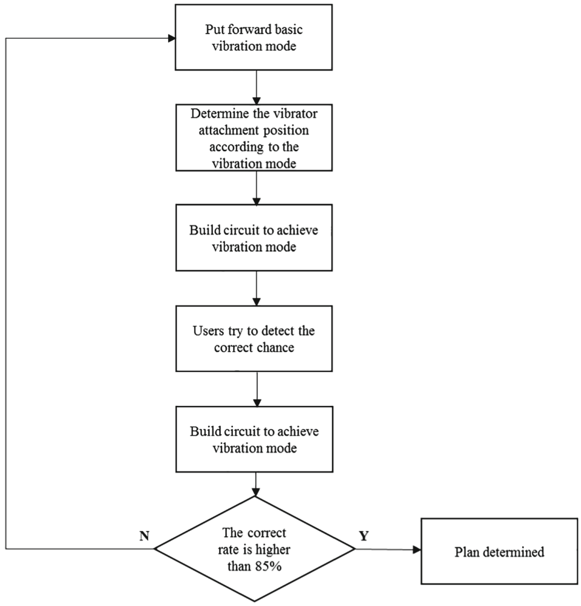



**Fig. 3.** Implementation system design flow

In order to detect the direction and the distance of obstacles, at least three vibrators are required, representing the left, rear and right directions, the forward direction is the normal direction of the rider's vision and is therefore not considered. Three different arrangements of vibrators were tried, as is shown in Fig. 4. In the first arrangement, the right hand little finger serves as the reference direction on the right, and vibrators are placed on the thumb, the little finger and the wrist respectively. Although this oscillator arrangement mode can clearly indicate the direction of obstacles, the prompts for different distances can only be distinguished by multiple vibrations. This method of presentation is not desirable. The second arrangement mode places two vibrators in place of each vibrator, and the vibratory tactile sensation can be achieved by controlling the number of vibrator elements. However, after the actual completion of this method, it was found that the difference in the sense of vibration was not obvious, and the correct rate of the prompt was very low. Therefore, this scheme was abandoned. The third scheme adds a large number of vibrators and places them in different areas. Through the sequential vibration of different areas, the user is prompted to locate the obstacles. The specific tips are: (1) For far-distance obstacles, only the fingertips corresponding to the obstacle orientation trigger short vibrations. (2) For near-obstacles, intermittently vibrate the fingers of the finger corresponding to the direction of the obstacle. (3) For the most recent obstacles, start with the finger pads that correspond to the relative directions of the obstacles, vibrate one by one in the horizontal or vertical sequence, and finally the fingertips that correspond to the direction of the obstacles vibrate for the longest time.

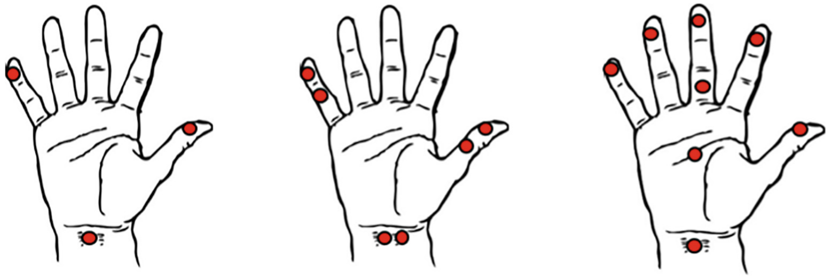


**Fig. 4.** Vibrators arrangement modes to show distance differences

Obstacle avoidance gloves based on tactile channel was designed, as is shown in Fig. 5.

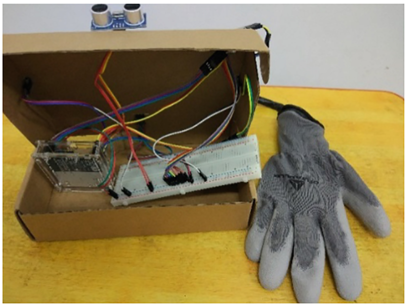
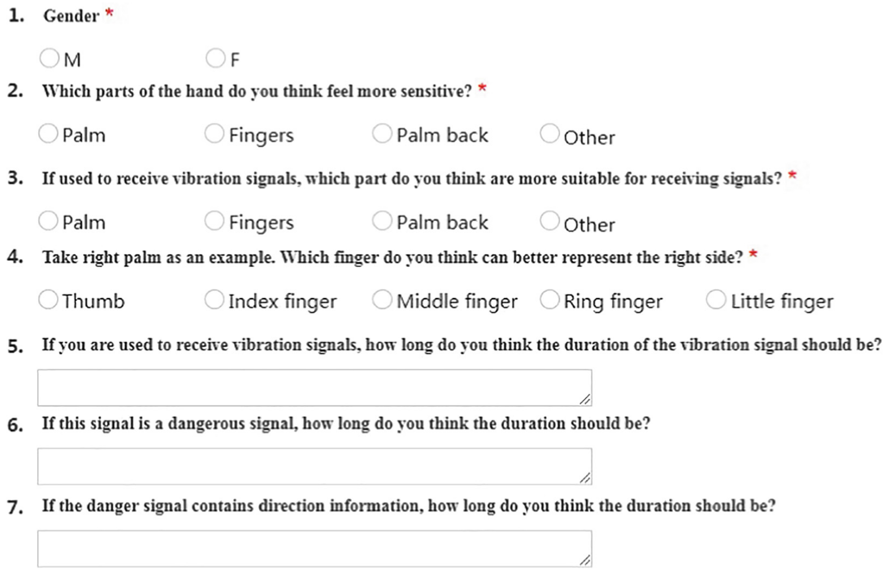


**Fig. 5.** Obstacle avoidance gloves based on tactile channel

### 4.2 Execution System Effect Test

The ultimate goal of this system is to create a hardware device that can be widely used. The device uses a tactile language, more precisely a vibration alert language. Different from visual and auditory information, tactile language is often just a kind of information that helps people to understand things. Few healthy people use touch as the main source of information. Therefore, the vibration operation mode of the device needs to be used and adapted by most users. Instead of allowing users to adapt to the device's prompt language, it is better to let the device actively adapt to the user's way of thinking and thinking habits. So if the system wants to design a set of vibration prompting language suitable for most people, it will require a certain amount of sample research.

In order to facilitate data recording and statistics, the survey was conducted using questionnaires and live-action tests. The designed questionnaire is as shown in Fig. 6:



1. **Gender** *

M F

2. **Which parts of the hand do you think feel more sensitive?** *

Palm Fingers Palm back Other

3. **If used to receive vibration signals, which part do you think are more suitable for receiving signals?** *

Palm Fingers Palm back Other

4. **Take right palm as an example. Which finger do you think can better represent the right side?** *

Thumb Index finger Middle finger Ring finger Little finger

5. **If you are used to receive vibration signals, how long do you think the duration of the vibration signal should be?**

6. **If this signal is a dangerous signal, how long do you think the duration should be?**

7. **If the danger signal contains direction information, how long do you think the duration should be?**

**Fig. 6.** Questionnaire design

The number of participants is 59, including 32 men and 27 women. The result of the questionnaire shows that the vast majority of people thought that the fingers and the palm feel the most sensitive, so they are more suitable for receiving signals. In question 4, 56 people thought that little finger can better represent the right side. For questions 5, 6 and 7, since the survey result is a time range, the average of the minimum and maximum values of the interval is taken as the minimum and maximum values of the survey result range, respectively. The result time ranges are as follow. The result of question 5 is 1.217 s and 4.443 s, question 6 is 0.381 s and 1.239 s, question 7 is 0.515 s and 1.937 s.

According to the contents of the questionnaire, the vibratory alert method described in Sect 4.2 was set up. Three types of vibrations were performed in an out-of-order manner. The following three questions were set for the subjects: (1) Which direction do you think the prompt indicates the risk is coming from? (2) Which vibration do you think is the strongest? Which time is the weakest? (3) Can you roughly describe the sequence of vibrations?

Test results are expressed as percentages. For question 1, 100% of the participants can accurately describe the direction of the obstacle. For question 2, 100% of the participants can judge the strength of the vibration method. For question 3, 13 people (about 22%) can accurately tell the vibration sequence of each vibrator, 43 people (about 73%) can describe which area is vibrating first and which area is vibrating later. Three people (about 5%) can only tell which vibrator is the most vibrated.

Based on the above experimental results, it is determined that the installation position of the vibrator is located in the palm and the finger pulp, the reference direction with the right hand little finger as the right, and the vibration period is 0.6 s. At the same time, it is proved that the vibration mode originally set by the system can be accepted by most people. The three levels of vibration intensity and sequence can also be recognized by most people. However, due to the small total sample size, there may be some errors.

### 4.3 Execution System Code Design

Define the detection distance, which is bounded by 5 cm and 10 cm:

```
double Length1,Length2,Length3;
Length1=5;
Length2=10;
Length3=20;
```

Compare the obstacle distances in the four directions and copy the minimum value to cm:

```
if(cmB<cmF){cm1=cmB;};
if(cmF<cmB){cm1=cmF;};
if(cmL<cmR){cm2=cmL;};
if(cmR<cmL){cm2=cmR;};
if(cm1<cm2){cm=cm1;};
if(cm2<cm1){cm=cm2;};
```

If the rear obstacle is the shortest and cmB is the smallest, enter the execute statement using the if statement:

```
(cmB<cmF && cmB<cmL && cmB<cmR)
```

Determine the distance range of the rear obstacle. If it is between length2 and length3, that is, more than 10 cm and less than 20 cm, then,in a time of 0.6 s, vibrate the vibrator at the wrist twice for 50 ms each, and the interval between the two shocks is 250 ms.

```
if(cmB>Length2 and cmB<Length3){
 digitalWrite(Long4,HIGH);
 delay(50);
 digitalWrite(Long4,LOW);
 delay(250);
 digitalWrite(Long4,HIGH);
 delay(50);
 digitalWrite(Long4,LOW);
 delay(250);};
```

If it is between length1 and length2, that is, greater than 5 cm and less than 10 cm, shake the vibrator at the wrist twice within 0.6 s, the first 50 ms, the second 250 ms, interval between 250 ms:

```
if(cmB<Length2 and cmB>Length1){
 digitalWrite(Long4,HIGH);
 delay(50);
 digitalWrite(Long4,LOW);
 delay(250);

 digitalWrite(Long4,HIGH);
 delay(250);
 digitalWrite(Long4,LOW);
 delay(50);};
```

If it is within length1, that is, less than 5 cm, vibrate the vibrators at the middle fingertip, middle finger root, palm, and wrist in order of 0.6 s, vibrating the first three times for 100 ms and the fourth for 300 ms:

```
if(cmB<Length1){
 digitalWrite(Long1,HIGH);
 delay(100);
 digitalWrite(Long1,LOW);

 digitalWrite(Long2,HIGH);
 delay(100);
 digitalWrite(Long2,LOW);

 digitalWrite(Long3,HIGH);
 delay(100);
 digitalWrite(Long3,LOW);

 digitalWrite(Long4,HIGH);
 delay(300);
 digitalWrite(Long4,LOW);};
```

Finally close all vibrators, end the statement, and proceed to the next cycle:

```
digitalWrite(Long1,LOW);
digitalWrite(Long2,LOW);
digitalWrite(Long3,LOW);
digitalWrite(Long4,LOW);
```

## 5 Conclusion

The safe driving intelligent system will encounter various obstacles during the operation, find obstacles beyond the horizon of the cyclist, and can get the exact distance between the sensor and the obstacle, which will provide great help for the safe driving of the vehicle. Therefore, this paper designed and developed an Arduino-based intelligent obstacle avoidance prompting system to implement automatic detection obstacles and humanized prompts. The system involves hardware and software development. After a comparative analysis, it is found that the Arduino platform is easy to use, low cost, and excellent in performance, and is in line with market trends. Therefore, the design and development of this system is based on the Arduino open source platform.

**Acknowledgement.** This paper is supported by Natural Science fund of Jiangsu Province (No. BK20150636).

## References

1. Zhang, H., Han, W.: Development of automobile safety technology. Heavy Veh. (2), 23–24 (2002). https://doi.org/10.3969/j.issn.1007-211x.2002.02.009. (in Chinese)
2. Huo, Q.: Looking at the world in the eye - volvo's vision monitoring and regulation system. World Car **10** (2008). (in Chinese)
3. Zhang, L.: Exploration of new measures for driving safety of vehicles. (21) (2012). (in Chinese)
4. Schoeneburg, R., Breitling, T.: Enhancement of active & passive safety by future pre-safe systems. In: Proceedings of the 19th International Technical Conference (2005)
5. Coelingh, E., Jakobsson, L., Lind, H., et al.: Collision Warning With Auto Brake - A Real Life Safety Perspective. U.S. Department of Transportation National Highway Traffic Safety Administration (2007)
6. Aoki, H., Aga, M., Miichi, Y., et al.: Safety Impact Methodology (SIM) for effectiveness estimation of a Pre-Collision System (PCS) by utilizing driving simulator test and EDR data analysis. SAE Technical Paper, 01-1003 (2010)
7. Ma, R., Kaber, D.B.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. Int. J. Ind. Ergon. **35**(10), 939–953 (2005)
8. Wang, G.: Study on ultrasonic distance sensor. Heilongjiang University (2014). (in Chinese)
9. Ohga, J., Takei, T., Moriyama, N.: Wideband piezoelectric loudspeakers using a tuck shaped PVDF bimorp. IEEE Trans. Dielectr. Electr. Insul. **4**(17), 23–56 (2010)
10. Jiang, B., Yang, Z.G., Liu, X.L., et al.: Study of a vertical drive ultrasonic feeder. Opt. Precis. Eng. **16**(6), 1082–1086 (2008)
11. Wang, Y.: Exploration of ultrasonic principles and modern applications. J. Guizhou Univ. (Natl. Sci.) **11**(7), 56–78 (2005). (in Chinese)
12. Lu, Y., Shi, J., Ling, Y., et al.: Infrared passive location research. Infrared Laser Eng. **30**(6), 405–409 (2001). (in Chinese)
13. Fu, X., Wu, D.: Single-station 3D infrared passive ranging algorithm with angle measurement only. J. Ordnance **29**(10), 1188–1191 (2008). (in Chinese)
14. Liang, S., Xu, L.: Research and design of multi-sensor ranging system for mobile robots. Comput. Appl. **28**, 340–343 (2008). (in Chinese)

# Study on Thermal Comfort of Virtual Reality Headsets

Zihao Wang(✉), Ke Chen, and Renke He

School of Design, Hunan University, Changsha 410000, China
wangzihao0716@gmail.com, ckechen@hnu.edu.cn,
renke8@hotmail.com

**Abstract.** This study aimed to investigate the thermal properties and subjective comfort of Virtual Reality (VR) headsets. Three VR products were selected. Twenty-seven university students were invited to experience the VR for 45 min. Each participant wore all VR headsets in three different days. The microclimate temperature and relative humidity were measured by thermocouples. Subjective thermal discomfort was evaluated using a 10-point visual analog scale. The results shown that the average microclimate temperature of the three VR headsets was increasing, and humidity increased at the beginning and then decreased. The subjective thermal discomfort of the three VR headsets increased with duration of use. The differences of average microclimate temperature, humidity and subjective thermal discomfort were significant among three headsets ($p < 0.05$). SONY PlayStation had the best thermal appreciation, followed by OCULUS Rift and GOOGLE Daydream.

**Keywords:** Virtual Reality · Head-mounted display · Thermal comfort
Microclimate

## 1 Introduction

Virtual Reality (VR) systems have been widely applied in industry, education, public and domestic settings. Head-mounted display (HMD) is used for providing image for many VR headsets [1]. Previous study indicates the head is the most sensitive region for whole body thermal comfort, and the thermal comfort of headgears directly determines its user experience [2–5], thus improving thermal comfort of headgear is likely to enhance the willingness to wear headgear [6]. A few studies related to thermal comfort of VR headsets mentioned that head-mounted displays (HMD)s are often of enclosed design and generate a considerable amount of heat in powering the displays [7], but no study has been empirically conducted to evaluate the thermal comfort of VR headsets, it is therefore important to conduct the study.

Methods used in evaluating thermal comfort of headgears can be categorized into numerical, biophysical, combined numerical with biophysical, and user trials [6]. Temperature and relative humidity within the microclimate, which is a relatively enclosed air layer between human skin and the HMDs, were the most frequently used indicators to evaluate thermal properties of industrial head-worn products, such as


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 180–186, 2019.
https://doi.org/10.1007/978-3-319-94619-1_17

bicycle helmets [3], motorcycle helmets [8], cricket helmets [9], protective garments [10], ski-boot liner [11] and back protectors [12].

This study aimed to investigate the thermal properties (microclimate temperature and relative humidity) and subjective discomfort of VR headsets. Through experiment, physiological parameters were integrated with subjective thermal sensations collected from a questionnaire, to have an overall understanding of the thermal comfort of the VR headsets.

## 2 Method

### 2.1 Materials

In the experiment, three VR products available on the market were selected: SONY PlayStation VR, OCULUS Rift Consumer Version 1, GOOGLE Daydream, as shown in Fig. 1. Each VR product has its own platform system, whereas same type of videogames were chosen in the test to ensure the amount of physical activities are the same among different platforms.

(1) SONY PlayStation VR (2) OCULUS Rift Consumer Version 1 (3) GOOGLE Daydream.

**Fig. 1.** Three VR headsets used in this experiment

HMDs were differed with other headgears in that the display itself will be heated, so the temperature and relative humidity between the display and human face will change when using. The thermocouples (iButton Hygrochron Temperature and Humidity Logger - DS1923, as shown in Fig. 2) were used to measure temperature and relative humidity in the microclimate with time interval of 30 s. Three locations were identified to place sensors inside the microclimate. The first thermocouple was located upon the right cheekbone of the face, the second sensor was located in the center of eyebrows, and the third sensor was located inside the left inner wall of the HMD cradle. The placement of all three sensors did not interfere with the VR experience process.

The experiment was conducted in a controlled environment. Laboratory average air temperature was 22.92 ± 1.38 °C, and the relative humidity was 38.56 ± 7.43%. The temperature was controlled through an air conditioning. The ambient temperature and the relative humidity were measured using Testo 610 Thermo hygrometer.

SPSS were used for data input and data analysis. Repeated measures ANOVA was

**Fig. 2.** The temperature and humidity sensors used in this experiment

performed to determine whether there were statistical differences of mean temperature, humidity and subjective thermal discomfort among three different VR headsets. The alpha level ($p < 0.05$) was set.

### 2.2 Participants

A number of 27 university students (17 women and 10 men) with average age of 22.78 (standard deviation of 2.28) participated in the experiment. All of the participants were in good health condition and were informed the procedure and purpose of the study. Their participation was voluntary. Individual demographic information (e.g. gender, age) was collected. Each participant experienced all of the three VR headsets in three different days at the same time of day.

### 2.3 Procedure

The experiment employed a within-group design. Each participant was wearing the three VR headsets randomly in three different days. In order to investigate the performance of thermal properties of VR headsets during the course of using as well as thermal comfort, the experiment time lasted for 45 min. Before experiencing VR, participants were given 10 min to acclimatize the environment, which is independent from the activities simulated [11, 12]. During the acclimatization phase, written consent was obtained from participants and thermocouples were placed on the participant's face and inside the HMD. Participants wore the VR headset properly with the help of experimenters. During 45 min VR experiencing, participants were stayed within a specified area, and the thermal couple record the microclimate temperature and relative humidity every 30 s. A self-rated subjective question was asked at 10th, 20th, 30th, and 45th min of the experiment. The subjective question was "*would you please report your thermal comfort of your head region at this moment*?" The responses were based on a 10-point visual analog scale, in which "0" indicates no discomfort and "10" indicates extremely discomfort.

## 3 Result and Discussion

### 3.1 Microclimate Temperature

The average microclimate temperature of the three VR headsets was 30.02 °C during the 45 min of VR experiencing. The average temperature of microclimate was increasing from 24.79 °C to 32.90 °C, with a rise of 8.11 °C in 45 min, and the result of recorded temperature is shown in Table 1, GOOGLE had the highest average microclimate temperature (30.39 °C) in 45 min, followed by OCULUS (30.14 °C), and SONY (29.55 °C). The difference between the three devices in average microclimate temperature was significant ($p < 0.05$). Post-hoc analysis showed the difference between GOOGLE and OCULUS was not significant.

**Table 1.** Experiment result of microclimate temperature (°C)

| | 1st min | 10th min | 20th min | 30th min | 45th min | Average |
|---|---|---|---|---|---|---|
| SONY | 25.23 | 28.69 | 30.37 | 31.39 | 32.06 | 29.55 |
| OCULUS | 24.94 | 28.83 | 31.03 | 32.42 | 33.50 | 30.14 |
| GOOGLE | 24.66 | 29.85 | 31.73 | 32.59 | 33.14 | 30.39 |
| Average | 24.94 | 29.12 | 31.04 | 32.13 | 32.90 | 30.02 |

Previous research on headgear thermal property has summarized the range of change in microclimate temperature [6], the results in this study was consistent within the range of 26 °C to 36.5 °C. Literature indicated forehead has higher sweating rate compared with the regions of temple, rear or vertex [13] and all VR headsets covers eyes and forehead, so a rising temperature in the forehead microclimate might result in thermal discomfort. Studies also found that discomfort is relating to the rapid increase of temperature in head area when wearing headgears [14], and it's also happened when wearing a helmet lacking of heat dissipation [15]. In this experiment, a quick change of microclimate temperature might cause heat discomfort for all users. Moreover, previous study also proved that users could experience thermal discomfort when temperature is beyond a narrow range of ±1 °C [16]. Increment in microclimate temperature when using VR headsets in this study could cause thermal discomfort for participants.

### 3.2 Microclimate Relative Humidity

The average relative humidity of the three headsets was 48.87% during 45 min of experiment. The average initial relative humidity was 45.14%, and it was rapidly rising to a peak point of 53.38% at the 7th min, and then started to relatively slowly decline. The relative humidity at different test times is shown in Table 2. At the 45th min, the average relative humidity was 46.67%, which was increased by only 1.53% compared with baseline. The changes in the relative humidity seemed to play a secondary role in influencing thermal comfort in addition to microclimate temperature increase. The average relative humidity differences between the three VR devices were significant ($p < 0.05$): OCULUS (58.14%) > SONY (44.62%) > GOOGLE (43.84%), while there was no statistical difference between SONY and GOOGLE in average relative humidity.

**Table 2.** Experiment result of microclimate relative humidity (%)

| | 1st min | 10th min | 20th min | 30th min | 45th min | Average |
|---|---|---|---|---|---|---|
| SONY | 44.61 | 50.14 | 44.91 | 43.29 | 40.12 | 44.62 |
| OCULUS | 47.01 | 62.52 | 62.49 | 60.78 | 57.92 | 58.14 |
| GOOGLE | 43.79 | 46.94 | 44.17 | 42.34 | 41.98 | 43.84 |
| Average | 45.14 | 53.20 | 50.52 | 48.80 | 46.67 | 48.87 |

### 3.3 Subjective Discomfort Evaluation

The subjective thermal discomfort of the three VR devices increased with time (Table 3). The mean subjective discomfort score was 2.81 with slight discomfort according to the scale (from 0 - no discomfort to 10 - extremely discomfort), and the differences of subjective thermal discomfort were significant among three VR headset ($p < 0.05$): GOOGLE (3.32) > OCULUS (2.79) > SONY (2.31). Results of post-hoc comparison indicated that GOOGLE was much uncomfortable than SONY, but there were no significant differences between GOOGLE and OCULUS, neither OCULUS nor SONY.

The thermal discomfort level of all VR headsets was acceptable in the current study. Previous literatures indicate that unfavorable thermal sensation might result in not wearing protective helmets when it was recommended to [17, 18]. Thus, in order to provide better user experience, the thermal properties of HMDs need to be improved.

**Table 3.** Experiment result of subjective discomfort evaluation

| | 10th min | 20th min | 30th min | 45th min | Average |
|---|---|---|---|---|---|
| SONY | 1.44 | 1.85 | 2.56 | 3.37 | 2.31 |
| OCULUS | 1.78 | 2.44 | 3.22 | 3.70 | 2.79 |
| GOOGLE | 2.04 | 2.89 | 3.70 | 4.63 | 3.32 |
| Average | 1.76 | 2.39 | 3.16 | 3.9 | 2.81 |

Comparing with SONY and OCULUS with T-shape head strap, GOOGLE has a single headband to be fastened when wearing. In order to prevent the device from falling off the face during use, the headband pressure of GOOGLE would cause facial discomfort. Moreover, the internal material for GOOGLE and OCULUS is composed of cotton. The shell of GOOGLE is made from cotton as well. Cotton has lower thermal conductivity and impedes heat dissipation, which may be the reason that microclimate temperature for OCULUS and GOOGLE were higher than SONY. SONY had the lowest microclimate temperature and highest subjective thermal comfort rating compared with OCULUS and GOOGLE. It is suggested that higher microclimate temperature and higher relative humidity might be the main causes of subjective thermal discomfort when using VR headsets.

## 4 Conclusion

During the 45 min of VR experiencing, the average microclimate temperature of all three headsets increased by 8.11 °C, which might cause thermal discomfort according to a previous study on helmet [14].

Humidity of all three devices increased at the beginning and then decreased, and it has increased by 1.53% within 45 min, which seemed to play a secondary role in thermal discomfort feeling.

Lastly, subjective evaluation of the thermal discomfort showed that there are differences among three VR headsets. The device with better thermal appreciation was SONY PlayStation VR, followed by OCULUS Rift Consumer Version 1 and GOOGLE Daydream.

This study investigated the thermal properties of three VR headsets in terms of microclimate temperature and relative humidity, as well as subjective thermal sensation. However, independent variables such as materials, types of wearing and ventilation design were not controlled, reasons leading to the experimental result and differences between three headsets could not be concluded. A further research with controlled condition is needed to figure out how material, and ventilation design could impact on the thermal properties.

## References


1. Sharples, S., Cobb, S., Moody, A., Wilson, J.R.: Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. Displays **29**(2), 58–69 (2008)
2. Bogerd, C.P., Brühwiler, P.A.: The role of head tilt, hair and wind speed on forced convective heat loss through full-face motorcycle helmets: a thermal manikin study. Int. J. Ind. Ergonomics **38**(3–4), 346–353 (2008)
3. Brühwiler, P.: Radiant heat transfer of bicycle helmets and visors. J. Sports Sci. **26**(10), 1025–1031 (2008)
4. Buyan, M., Brühwiler, P.A., Azens, A., Gustavsson, G., Karmhag, R., Granqvist, C.G.: Facial warming and tinted helmet visors. Int. J. Ind. Ergonomics **36**(1), 11–16 (2006)
5. Ishigaki, T., Fujishiro, H., Tsujita, J., En, Y., Yamato, M., Nakano, S., Hori, S.: Relationship between helmet temperature and tympanic temperature during American football practice. Japan. J. Phys. Fitness Sports Med. **50**(3), 333–338 (2001)
6. Bogerd, C.P., Aerts, J.M., Annaheim, S., Bröde, P., De Bruyne, G., Flouris, A.D., Rossi, R. M.: A review on ergonomics of headgear: thermal effects. Int. J. Ind. Ergonomics **45**, 1–12 (2015)

7. Costello, P.J.: Health and safety issues associated with virtual reality: a review of current literature, pp. 1–23. Advisory Group on Computer Graphics (1997)
8. Brühwiler, P.A.: Role of the visor in forced convective heat loss with bicycle helmets. Int. J. Ind. Ergonomics **39**(1), 255–259 (2009)
9. Pang, T.Y., Subic, A., Takla, M.: A comparative experimental study of the thermal properties of cricket helmets. Int. J. Ind. Ergonomics **43**(2), 161–169 (2013)
10. Wardiningsih, W., Troynikov, O., Nawaz, N., Watson, C.: Influence of wearing impact protective garment on thermophysiological comfort of the wearer. Proc. Eng. **72**, 551–556 (2014)
11. Colonna, M., Moncalero, M., Nicotra, M., Pezzoli, A., Fabbri, E., Bortolan, L., Schena, F.: Thermal behaviour of ski-boot liners: effect of materials on thermal comfort in real and simulated skiing conditions. Procedia Engineering **72**, 386–391 (2014)
12. Dotti, F., Ferri, A., Moncalero, M., Colonna, M.: Thermo-physiological comfort of soft-shell back protectors under controlled environmental conditions. Appl. Ergonomics **56**, 144–152 (2016)
13. Smith, C.J., Havenith, G.: Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia. Eur. J. Appl. Physiol. **111**(7), 1391–1404 (2011)
14. Van Brecht, A., Nuyttens, D., Aerts, J.M., Quanten, S., De Bruyne, G., Berckmans, D.: Quantification of ventilation characteristics of a helmet. Appl. Ergonomics **39**(3), 332–341 (2008)
15. Neave, N., Emmett, J., Moss, M., Ayton, R., Scholey, A., Wesnes, K.: The effects of protective helmet use on physiology and cognition in young cricketers. Appl. Cogn. Psychol. **18**(9), 1181–1193 (2004)
16. Fanger, P.O., Hojbjerre, J., Thomsen, J.O.: Man's preferred ambient temperature during the day. Archives des sciences physiologiques **27**(4), 395–402 (1973)
17. Orsi, C., Stendardo, A., Marinoni, A., Gilchrist, M.D., Otte, D., Chliaoutakis, J., Morandi, A.: Motorcycle riders' perception of helmet use: complaints and dissatisfaction. Accid. Anal. Prev. **44**(1), 111–117 (2012)
18. Papadakaki, M., Tzamalouka, G., Orsi, C., Kritikos, A., Morandi, A., Gnardellis, C., Chliaoutakis, J.: Barriers and facilitators of helmet use in a Greek sample of motorcycle riders: which evidence? Transp. Res. Part F: Traffic Psychol. Behav. **18**, 189–198 (2013)

# Digitization of Manufacturing Companies: Employee Acceptance Towards Mobile and Wearable Devices

Laura Merhar[1(✉)], Christoph Berger[1], Stefan Braunreuther[1,2], and Gunther Reinhart[1]

[1] Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Provinostr. 52, 86153 Augsburg, Germany
laura.merhar@igcv.fraunhofer.de
[2] Hochschule Augsburg University of Applied Sciences, An der Hochschule 1, 86161 Augsburg, Germany

**Abstract.** In the production environment, digital assistance systems become more important as an interface between human and machines. Considering implementation processes of new technologies, like mobile and wearable devices, employee acceptance is crucial. The research objective is to describe the current level of acceptance of production staff and to reveal the influencing factors for technology acceptance towards mobile and wearable devices in manufacturing environment. The impact of the degree of familiarity with the medium is of particular interest, because of the first-time implementation of devices that are already conversant from private surroundings, such as smartphones. The research method features an explorative, qualitative interview approach to examine complex attitude and decision-making processes in a theory-generating way. The qualitative content analysis implies differences between familiar and fairly unknown devices. Further results reveal several influencing factors such as robustness, wearing comfort, customizability, support of management or data protection that should be considered in implementation strategies.

**Keywords:** Industry 4.0 · Human factors · Technology acceptance
Mobile devices · Wearable devices · Digital assistance systems
Production industry · Employee participation

## 1 Introduction

The digital shift is ubiquitous and changes global markets. In an economic or industrial context, the digital transformation is often covered under the term Industry 4.0. Industry 4.0 comprises enhancements in industrial production systems in the context of digitization. At its core, there is the real-time capable and smart connection of all players including machines, workers as well as information and communication technology (ICT) [1]. Thereby Industry 4.0 does not only go beyond the manufacturing process in particular, but rather affects all sectors 'from order management, research and development, manufacturing, commissioning, delivery up to the utilization and the


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 187–197, 2019.
https://doi.org/10.1007/978-3-319-94619-1_18

recycling of produced goods' [1]. The overall goal of all efforts summarized under the term Industry 4.0 is to overcome current challenges in the changing working world. Companies are under pressure to maintain their competitive position in globalizing and fast-changing markets. Product cycles are becoming shorter and shorter, the innovation pressure and the demand for individualized products are increasing, which results in decreasing lot sizes. In some countries demographic change leads to an aging society which causes companies to adjust work tasks to less powerful employees [2–4]. Those changes imply an increasing complexity that shall be coped with the potentials of digitization. Mobile and wearable digital assistance systems are an important technological development that helps handling the upcoming level of complexity in the fast-changing working world [5, 6].

# 2 Theoretical Framework

## 2.1 Digital Assistance Systems in Production and Private Sectors

Digital assistance systems can be defined as technologies that support humans in performing certain tasks. They can provide physical support, such as human-robot-interaction or exoskeletons, or cognitive support, such as digital information systems [7, 8]. This paper focuses on cognitive systems for informational assistance. The aforementioned challenges in the production environment confront employees with changing information and work tasks. Digital assistance systems offer new ways of technical assistance, they are location-independent, provide hands-free communication and real-time information to handle different upcoming challenges. They contain a wide range of *technologies*, like virtual or augmented reality, *devices*, like tablets or data glasses and *use cases*, like information systems for manual assembly. The further include *input modalities*, like voice or gesture control, and different possibilities of *information presentation* [5, 9–11]. This paper focuses on mobile and wearable devices from the consumer sector due to their increasing use in the production environment as well as their lower costs compared to customized devices. In particular, smartphones, tablets, data glasses and smart watches are chosen to be the objects of investigation due to their multiple features and thus their great importance for future production [11].

In the private sector in Germany, digitization has been going on for many years and especially tablets and smartphones are widely spread among the population [12]. Those devices now enter the production environment, which characterizes the uniqueness of the current situation and may have an influence on the workers perceptions. Especially small and medium sized enterprises (SME) are at the beginning of digitization processes, which is why they are central in this research [13].

## 2.2 Human Factors in Production and the Role of Acceptance

Technological innovations in companies mostly come along with social innovations [14]. The term socio-technological system describes the cooperation of humans and machines to collaboratively carry out a task. A central purpose in order to develop such working systems in the context of Industry 4.0 is to reconcile the factors human, technology and organization [15, 16]. Despite of a prevailing technocratic view, the opinion that human beings play a significant role in future smart factories becomes prevalent [17]. Nevertheless, operational and technical decisions are often made merely on the basis of efficiency, productivity or competition, whereas the role of the human and his sensitivities should also be considered as one of the most essential factors for decision-making in technology innovation processes [16, 18]. Especially two aspects are pointed out: The technology – in this case the digital assistance systems – should be developed based on the workers requirements and primarily acceptance-encouraging.

Acceptance towards a technology can be seen as an important condition to sustain worker satisfaction and to enable an efficient and trouble-free usage. Otherwise, lacking acceptance may lead to barriers in willingness or abilities of the employees or in the worst case to avoidance. Change is not always perceived positive because technical transformations affect one's habituated behavior and request adaption to new situations and processes [18, 19]. Acceptance is a key concept in many research areas. An established view is the differentiation between attitude-acceptance and behavior-acceptance. Attitude-acceptance includes the positive attitude and valuation of a subject matter and therefore cognitive and affective components. Behavior-acceptance means the performed action itself, which is visible and includes attitude-acceptance [20]. In this study, attitude-acceptance is focused because the research subject covers employees in companies that not have implemented assistance systems yet, so the attitude and level of acceptance is located before the actual use. Acceptance towards technologies is usually investigated under the term technology acceptance.

### 2.3 Technology Acceptance

One of the most popular, widespread and evolved models is the Technology Acceptance Model (TAM) from Davis [21, 22]. It is based on the Theory of Reasoned Action (TRA) [23] and the Theory of Planned Behavior (TPB) [24]. The TAM implies that the *perceived ease of use* as well as the *perceived usefulness* affect the *attitude toward use* and the *intention to use*. The model is approved in numerous studies but criticized because of excluding further important influencing factors [15, 25]. Further developments of the model – TAM 2 and TAM 3 – include determinants of perceived ease of use, like *image* or *job relevance*, and of perceived ease of use, like *computer anxiety* or *objective usability* [26–28].

The Unified Theory of Acceptance and Use of Technology (UTAUT) comprises eight acceptance models, including TAM, and has four key constructs: *performance expectancy*, *effort expectancy*, *social influence* and *facilitating conditions*. *Gender*, *age*, *experience*, and *voluntariness of use* are presumed to moderate the impact of the four key constructs on intention to use and behavior. The four key constructs include various factors or sub constructs that cover further important influencing factors on technology acceptance [29].

Furthermore, several models and constructs are commonly applied in technology acceptance research. For example the Task Technology Fit Model (TTFM) is of interest for this study. Therefore, the accordance of task demands, technology functionalities and individual user skills – *the task-technology-fit* – affects acceptance and behavior [30].

Popular and familiar approaches are technology appropriation theories and considerations to innovation barriers. The Model of Technology Appropriation (MTA) describes technology adoption as the relation of people's desire, the capabilities and implications of technology as well as the situations of usage. It includes inter alia appropriation criteria, like *leisure use*, and disappropriation criteria, like *unusability* [31]. In the German research environment, acceptance and innovation barriers complement the discussion. These are for example motivational or ability barriers, lacking employee participation, lacking know-how or missing recognition of the technology's potentials [32]. In the course of Industry 4.0 further factors may play a role, for example concerns about work safety, data-security or the preservation of work-related values and meanings.

## 3 Research Issue

In conclusion, technologies like digital assistance systems for the production environment should be developed and implemented under the principle of maximizing employee acceptance. To do so, research needs to be conducted for a deeper understanding. Whereas technology acceptance in general is a relatively well studied field, especially in market analysis for the consumer goods sector, there is a lack of research concerning the level and influences of worker acceptance of mobile and wearable devices in the production environment.

Therefore, the research objective of this study is to describe the current level of acceptance of the production staff and to reveal the influencing factors on technology acceptance towards mobile and wearable devices in the manufacturing environment. The implementation process of digital assistance systems has only just begun, which underlines the importance of investigating technology acceptance at an early stage. The impact of the degree of familiarity with the medium is of particular interest, because of the first-time implementation of devices that are already conversant from private surroundings, such as smartphones and tablets.

## 4 Methodology

### 4.1 Design

While many researchers choose a quantitative approach to examine technology acceptance by means of the aforementioned models, this study uses a qualitative approach for two reasons: First, quantitative acceptance models are criticized for not being able to explain complex decision making and attitude processes that exceed the rather inflexible models [33]. Second, the consideration of the use of tablets or smart watches in the production environment is a fairly new situation, which demands a flexible research approach. Qualitative and explorative social research provides methods that are open and flexible for new circumstances. Moreover, they are able to reveal complex subjective thought and attitude patterns as well as probably unexpected results [33]. Therefore, the research method of this study features an explorative, qualitative interview approach that is oriented towards the theory-generating Problem-Centered Interview from Witzel [34]. This approach focuses on a combination of induction and deduction and includes theoretical prior knowledge, in this case especially the aforementioned acceptance models, as well as the required openness for explorative research [34].

### 4.2 Interview Guideline and Pretest

Core element of this interview study is the developed problem-centered interview guideline, which serves a structuring purpose. After an information part about the data acquisition and usage, the interview starts with a problem-orientated introductory example. It motivates the participant to narrate and it helps to settle in the interview situation. The main phase includes several open and narrative-generating questions with more specific sub-questions that are aligned with established models, previous research and assumptions from scientific literature about technology acceptance (see Sect. 2.3) while providing the necessary openness for qualitative research. In the final phase, the interviewee has the possibility to add his thoughts and is asked to draw a conclusion.

The interview guideline was pretested with one interviewee that additionally gave feedback regarding comprehensibility. Afterwards, the guideline has been slightly adapted to a more natural course of conversation. The pretest is included in the analysis.

### 4.3 Procedure

The five interviewees were selected on the basis of different criteria to cover various fields of activity: All of them work in the production environment of small or medium sized companies in Germany but in different levels and areas, such as assembly or quality assurance. Unfortunately, no female participants were found for the study, which is why gender had to be excluded as a possible influence. The theoretical saturation was reached after five interviews.

The interviews were conducted at the end of the year 2016. At the time of the interviews the age of the participants was between 28 and 49 with an average of 35, 4 years and they worked in assembly, manufacturing or inspection. The interview duration was between 29 and 68 min with an average of 51 min. All interviews were recorded and fully transcribed.

### 4.4 Data Analysis

The analysis of the transcribed interviews was executed in accordance with Qualitative Content Analysis from Mayring [35]. This method for text analysis reduces the content up to the relevant passages in order to build a systematic category system that constitutes a reflection of the base material.

Deductive categories from the aforementioned acceptance models and theories formed a preliminary category system for the initial analysis. In the course of the analysis, this category system was supplemented, extended and modified gradually (inductive category development). MAXQDA was used as data analysis software. The result of the analysis, the final category system, is explained in the following chapter.

# 5 Results

The following chapter outlines the results focused on the comparison of mobile and wearable devices. First, the level of acceptance is described (Fig. 1). Second, various influencing factors on acceptance are explained (Fig. 2).

## 5.1 Level of Acceptance

*Awareness of Change.* The participants are aware of the necessity of a change. They recognize current challenges and support efforts of their management towards digitization in general.

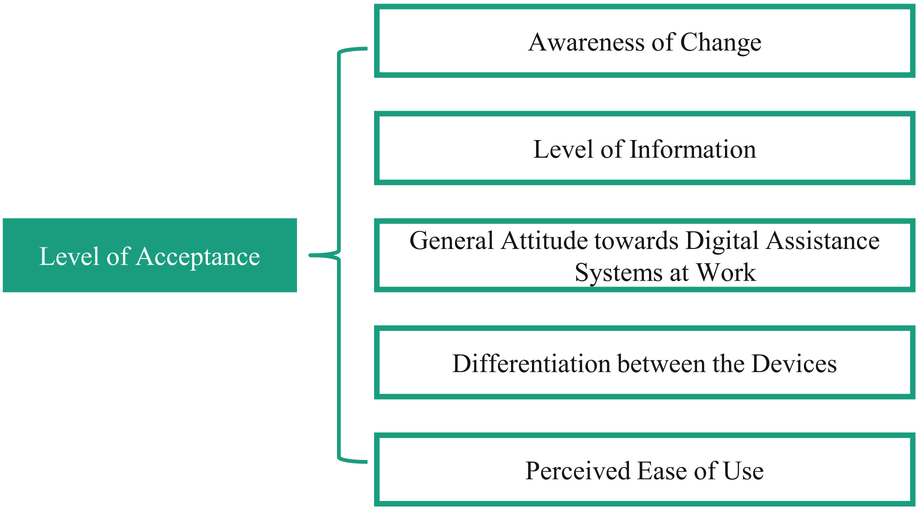



**Fig. 1.** Main categories of level of acceptance.

*Level of Information.* Most participants do not feel well informed about the usage and benefits of digital assistance systems in production contexts. One participant assesses his own level of information slightly higher based on his usage experience with digital devices in the private sector, although he does not have a clear vision of specific use cases.

*General Attitude Towards Digital Assistance Systems at Work.* Overall the attitude towards digital devices at work is characterized by optimism and openness.

*Differentiation Between the Devices.* The interviews show major differences between mobile and wearable devices. Tablets are mainly favored because of usage experience from the private sector, as supposed before. Furthermore, it is assumed that tablets are equipped with the best display options with a simultaneously high level of mobility as well as possibilities of a flexible installation at a working station. Smartphones are on the second place because of high mobility but fewer display options due to a smaller screen size. Wearable devices score poorer in general due to skepticism towards technology performance and wearing comfort. Data glasses get rejected in case of long wearing periods, frequent information in the field of vision or for wearers of glasses. Smart watches are considered to disturb manual activities or to damage sensitive components. In contrast to tablets and smartphones, the participants only have very little or no usage experience from the private sector regarding wearables.

*Perceived Ease of Use.* For devices that were already tried at home – mainly tablets and smartphones – the level of difficulty is considered to be low. In general, the perceived ease of use is higher if trainings and an extensive induction phases are provided.

In conclusion, prior experiences seem to have a relevant impact on the current level of acceptance before the implementation of digital assistance systems. Therefore, mobile devices show a higher acceptance level in the context of the study than wearable devices with a fairly low level.

### 5.2 Influencing Factors on Acceptance

*Perceived Usefulness.* Digital assistance systems are considered useful, for example if new tasks that are not solvable through practical knowledge, have to be carried out. The devices shall grant benefits for the worker, for example higher productivity, less reject

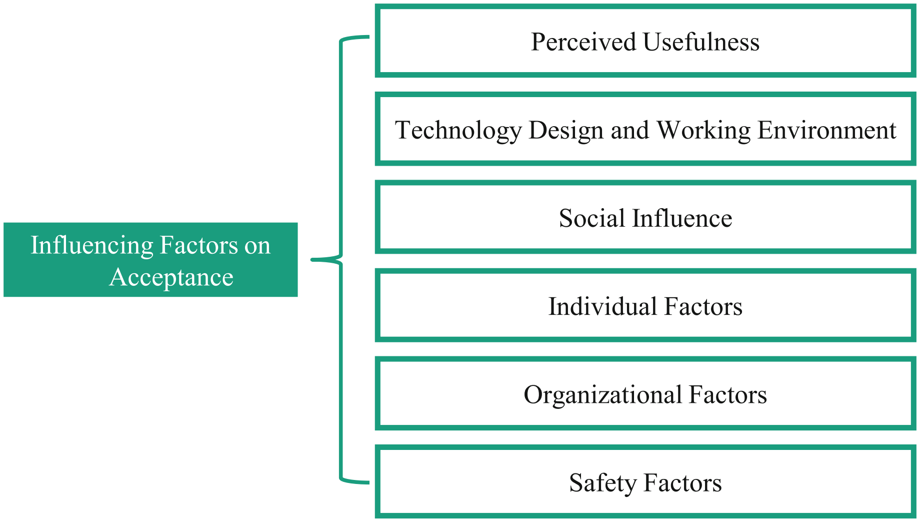



**Fig. 2.** Main categories of influencing factors on acceptance.

parts, more information, shorter distances, an improved flow of information and easier coordination between different areas, targeted assistance, less documentation effort – overall a saving of time, fewer errors, support and communication. Negative influences would be an increasing dependency on technology and a loss of significance of practical knowledge.

*Technology Design and Working Environment.* Digital assistance systems should be designed primarily easy in relation to structure and intuitive operation. Further important aspects are trouble-free functionality, reliability, up-to-dateness of the device and software, performance and rapid data-transfer. In return, malfunctions or slowness are significant hindering factors. Moreover, the devices should be robust against surrounding threats like oil and dust or against dropping. Further designing aspects are device and display size including clarity. Regarding clarity, the participants are skeptical towards data glasses and smart watches. They are concerned that wearables might be hindering, disturbing or like a foreign object that they do not want to wear close to the body all the time. At least, several factors related to information provision are important: information should be found and structured easily and a help feature should provide additional information on demand. At best, the device is customizable to the users' individual requests.

*Social Influence.* Social factors include support from the management, the influence of colleagues and the atmosphere in the group as well as external effects like a positive image of being an innovative company.

*Individual Factors.* The most important impact in this section is private experience. The participants are familiarized with the handling of smartphones and tablets, whereas no one has got or tried wearables in-depth. Some influencing factors correlate according to the participants with private experience. They consider the devices to be less difficult in handling, more work-simplifying and the usability to be easier. In general the participants are much more open towards familiar devices. Negative factors in the individual sector might be the fear of rationalization and again a loss of significance of prior working experience.

*Organizational Factors.* As previous studies have shown, information, participation and training are very important acceptance factors for the participants. As further factors, they mention early and consistent information and participation, justification for the project from the management and testing procedures on-the-job. Another conducive aspect is the equal distribution of the devices. Inversely, envy among colleagues is stated as a hindering factor for acceptance.

*Safety Factors.* Data safety, data protection and work safety are mentioned as influencing factors. The participants highlight that sensitive data must not leak out. Regarding data protection the opinions differ. Some participants fear a data misuse of the management in form of performance assessment. Others would either like to use their data for their benefit, for example in salary negotiations, or do not care about data usage.

In conclusion, many factors from technical, organizational and social areas seem to have an impact on acceptance towards digital assistance systems. Some of these are already known from scientific literature, but also several context-specific factors beyond the acceptance models were identified, like information design or safety aspects.

## 6 Conclusion and Outlook

The results show that there is a difference between mobile and wearable devices concerning the current level of acceptance of digital assistance systems in the production environment. Whereas the participants of the study largely approve tablets and smartphones, they are rather skeptical towards smart watches or data glasses. Moreover, the results reveal several influencing factors that might have a significant impact on technology acceptance in the specific context.

Due to an explorative, qualitative research approach with few participants the study is limited regarding statements about frequency and distribution. Therefore, generally valid conclusions cannot be drawn without further research.

First cautious assertions include conclusions for implementation strategies. Mobile devices that are widespread in the private sector, like tablets and smartphones, have the advantage that they are familiar regarding operation and as well presumably already accepted. If companies want to establish wearable devices, they should start with a detailed and extensive introduction to the device and its features, an accentuation of the advantages and comprehensive testing possibilities. Regarding the device, different factors should be considered, for example display size, surrounding conditions or device-specific acceptance barriers.

In conclusion, the study provides findings that could not have been revealed with a quantitative approach. The next step is to revise the results of this study with a larger sample to generate universal statements as a basis for future development and successful implementation of digital assistance systems.

**Acknowledgement.** This publication is part of the “SynDiQuAss” project for excellence in manual assembly in the context of Industry 4.0 funded by the German Ministry of Education and Research (BMBF) and the European Social Fund (ESF) (grant no. 02L15A281).

## References

1. Neugebauer, R., Hippmann, S., Leis, M., Landherr, M.: Industrie 4.0 – from the perspective of applied research. Proc. CIRP **57**, 2–7 (2016)
2. Abele, E., Reinhart, G.: Zukunft der Produktion. Carl Hanser, München (2011)
3. Lee, J.: Smart factory systems. Inf. Spektrum **38**(3), 230–235 (2015)
4. Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K.: Cyber-physical systems in manufacturing. CIRP Ann. Manufact. Technol. **65**(2), 621–641 (2016)
5. Hold, P., Erol, S., Reisinger, G., Sihn, W.: Planning and evaluation of digital assistance systems. Proc. Manufact. **9**, 143–150 (2017)
6. Prinz, C., Morlock, F., Freith, S., Kreggenfeld, N., Kreimeier, D., Kuhlenkötter, B.: Learning factory modules for smart factories in industrie 4.0. Proc. CIRP **54**, 113–118 (2016)

7. Yang, X., Plewe, D.A.: Assistance systems in manufacturing: a systematic review. In: Schlick, C., Trzcieliński, S. (eds.) Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future, vol. 490, pp. 279–289. Springer, Cham (2016)
8. Reinhart, G.: Handbuch Industrie 4.0. Carl Hanser, München (2017)
9. Merkel, L., Berger, C., Schultz, C., Braunreuther, S., Reinhart, G.: Application-specific design of assistance systems for manual work in production. In: 2017 IEEE International Conference on Industrial Engineering and Engineering Management, Singapore (2017)
10. Nelles, J., Kuz, S., Mertens, A., Schlick, C.M.: Human-centered design of assistance systems for production planning and control: the role of the human in industry 4.0. In: 2016 IEEE International Conference on Industrial Technology (ICIT), Taipei, pp. 2099–2104 (2016)
11. Gorecky, D., Schmitt, M., Loskyll, M., Zühlke, D.: Human-machine-interaction in the industry 4.0 era. In: 12th IEEE International Conference on Industrial Informatics (INDIN), Porto Alegre, pp. 289–294 (2014)
12. Koch, W., Frees, B.: ARD/ZDF-Onlinestudie 2017: Neun von zehn Deutschen online. Media Perspektiven **9**, 434–446 (2017)
13. Bley, K., Leyh, C., Schäffer, T.: Digitization of German enterprises in the production sector – do they know how "digitized" they are? In: Proceedings of the 22nd Americas Conference on Information Systems (AMCIS) (2016)
14. Georg, A., Hellinger, A.: Soziale und technische Innovationen in der Industrie 4.0 gestalten. In: Schlick, C.M. (ed.) Arbeit in der digitalisierten Welt, pp. 57–65. Campus, Frankfurt/New York (2015)
15. Legris, P., Ingham, J., Collerette, P.: Why do people use information technology? A critical review of the technology acceptance model. Inf. Manage. **40**(3), 191–204 (2003)
16. Deuse, J., Busch, F., Weisner, K., Steffen, M.: Differenzielle Arbeitsgestaltung durch hybride Automatisierung. In: Schlick, C.M. (ed.) Arbeit in der digitalisierten Welt, pp. 235–245. Campus, Frankfurt/New York (2015)
17. Vaidya, S., Ambad, P., Bhosle, S.: Industry 4.0 – a glimpse. Proc. Manufact. **20**, 233–238 (2018)
18. Iverson, R.D.: Employee acceptance of organizational change: the role of organizational commitment. Int. J. Hum. Resour. Manage. **7**(1), 122–149 (1996)
19. Lines, B.C., Sullivan, K.T., Smithwick, J.B., Mischung, J.: Overcoming resistance to change in engineering and construction: change management factors for owner organizations. Int. J. Proj. Manage. **33**, 1170–1179 (2015)
20. Arning, K., Ziefle, M.: Understanding age differences in PDA acceptance and performance. Comput. Hum. Behav. **23**, 2904–2927 (2007)
21. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**(3), 319–340 (1989)
22. Lee, Y., Kozar, K.A., Larsen, K.R.T.: The technology acceptance model: past, present, and future. Commun. Assoc. Inf. Syst. **12**(1), 752–780 (2003)
23. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Addison-Wesley, Reading (1975)
24. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**(2), 179–211 (1991)
25. King, W.R., He, J.: A meta-analysis of the technology acceptance model. Inf. Manage. **43**, 740–755 (2006)
26. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage. Sci. **46**(2), 186–204 (2000)
27. Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. Decis. Sci. **39**(2), 273–315 (2008)

28. Marangunic, N., Granic, A.: Technology acceptance model: a literature review from 1986 to 2013. Univ. Access Inf. Soc. **14**(1), 81–95 (2015)
29. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **37**(3), 425–478 (2003)
30. Goodhue, D.L., Thompson, R.L.: Task-technology fit and individual performance. MIS Q. **19**(2), 213–236 (1995)
31. Carroll, J., Howard, S., Vetere, F., Peck, J., Murphy, J.: Just what do the youth of today want? technology appropriation by young people. In: Proceedings of the 35th System Sciences HICSS, pp. 1777–1785 (2002)
32. Anstadt, U.: Determinanten der individuellen Akzeptanz bei Einführung neuer Technologien: Eine empirische arbeitswissenschaftliche Studie am Beispiel von CNC-Werkzeugmaschinen und Industrierobotern. Peter Lang, Frankfurt (1994)
33. Vogelsang, K., Steinhüser, M., Hoppe, U.: A qualitative approach to examine technology acceptance. In: International Conference on Information Systems: Reshaping Society Through Information Systems Design (2013)
34. Witzel, A.: The problem-centered interview. Forum Qual. Soc. Res. **1**(1), Article 22 (2000). http://nbn-resolving.de/urn:nbn:de:0114-fqs0001228
35. Mayring, P.: Qualitative content analysis. Forum Qual. Soc. Res., **1**(2), Article 20 (2000). http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204

# Determination of Cognitive Assistance Functions for Manual Assembly Systems

Lukas Merkel[1](✉), Christoph Berger[1], Stefan Braunreuther[1,2], and Gunther Reinhart[1]

[1] Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Provinostr. 52, 86153 Augsburg, Germany
lukas.merkel@igcv.fraunhofer.de
[2] Hochschule Augsburg University of Applied Sciences, An der Hochschule 1, 86161 Augsburg, Germany

**Abstract.** Since a growing number of variants increase complexity in today's production systems, higher flexibility is needed. However, automated production systems are often not economical in high-variant production scenarios. Therefore, human flexibility plays an important role, especially for assembly tasks. In order to increase human flexibility in manual assembly a variety of assistance systems providing cognitive support for individual workers has been developed in recent years. Cognitive assistance systems can support assembly workers by providing, processing or collecting information. This paper presents an approach to determine cognitive assistance functions in manual assembly. The need for different assistance functions is investigated in order to make a needs-based selection. The results can then be matched with suitable technologies to design an assistance system. An application of this approach is shown for a manual assembly system in the learning factory for cyber-physical production systems in Augsburg, Germany.

**Keywords:** Cognitive assistance · Manual assembly
Technology management · Industry 4.0 · Human factors

## 1 Introduction

In order to be successful in saturated markets, more companies are offering customer-individual products [1]. Besides production in lot size one, volatile demand [2] increases complexity in production. Furthermore, a rising product complexity [3] and shorter product lifecycles [4] are challenges producing companies face today. Rising complexity in production not only increases quality costs [3] but also requires a broad qualification of employees [2]. In order to master rising complexity, production systems need to become more flexible.

Depending on the industrial sector, assembly is responsible for up to 70% [5] of total production costs. Since manual work is the most flexible factor of production [6], flexibility in production is especially enabled by manual assembly. Manual assembly workers can be supported by assistance systems in perceiving assembly tasks, making


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 198–207, 2019.
https://doi.org/10.1007/978-3-319-94619-1_19

decisions and executing assembly tasks [1]. Physical assistance systems can be applied to improve ergonomics in handling heavy components. Cognitive assistance systems can support manual assembly workers by providing, processing or collecting information. Due to rising complexity in manual assembly, the importance of correct and relevant information in manual assembly is rising [7]. Therefore, cognitive assistance systems play an important role to increase flexibility in manual assembly systems. Furthermore, the use of cognitive assistance systems can result in higher efficiency and error-free assembly [8].

## 2 Complexity Evaluation in Manual Assembly

In order to design a needs-based cognitive assistance system in manual assembly, it is necessary to analyze the required assembly tasks. Several models have already been developed for measuring complexity in assembly. Samy et al. [9] are using detailed geometrical information to measure product assembly complexity. Zeltzer et al. [10] identified 11 complexity drivers in assembly to quantify complexity for mixed-model assembly workstations. These complexity drivers cover several technical aspects from the analyzed assembly operations. These models can be used to reduce complexity during the design-phase of a product.

In order to develop a needs-based assistance system for a given product, worker-specific information has to be considered as well. Zäh et al. [11] are using a temporal, a cognitive and a knowledge-based factor in order to measure complexity of manual assembly operations. Besides product complexity Claeys et al. [12] are also considering operational complexity of the workstation as well as operator-specific information. Hold et al. [13] are using MTM to structure assembly tasks and also consider human error probability for each task group.

Several existing models allow a very detailed analysis and quantification of assembly operations. Since product life-cycles are shortening [4] and the investment of cognitive assistance system has to be amortized, the lifecycle of a cognitive assistance system is often longer than product lifecycles. Therefore, it is necessary to consider product changes during the lifecycle of the assistance system. Furthermore, several models for measuring assembly complexity focus on detailed technical aspects and don’t consider legal aspects, such as documentation tasks.

## 3 Determination of Cognitive Assistance Functions

This section introduces a method for needs-based determination of cognitive assistance functions for a given manual assembly process. It can be applied for workshop assembly scenarios which are not equipped with cognitive assistance systems. In this method the complexity of an assembly task is evaluated in order to select suitable assistance functions. In order to include all relevant aspects all stakeholders (e.g. industrial engineering, assembly workers, production IT) should be involved in the evaluation process.

### 3.1 Assembly Process Analysis and Complexity Evaluation

In order to design an application-specific assistance system, the assembly process shall be analyzed from worker-centric perspective. Assembly processes is therefore divided into different tasks, which are defined in the work plan. For each assembly task an assembly worker has to perceive information [11], execute the assembly task and document the performed assembly. In order to introduce a needs-based assistance system the complexity of an assembly task needs to be evaluated in different dimensions. The goal of this method is to develop an applicable assessment of manual assembly that focusses on key aspects in order to limit the evaluation effort.

Therefore, for each assembly task three dimensions of complexity shall be analyzed:

- *Perception Complexity (PC)* aims to evaluate the cognitive workload an assembly worker before the physical assembly execution.
- *Execution Complexity (EC)* aims to evaluate the risk of an assembly error during the assembly.
- *Documentation Complexity (DC)* aims to evaluate the necessity and effort of documentation during the assembly task.

#### 3.1.1 Perception Complexity

During several workshops with production engineers from different industrial sectors three main complexity drivers were identified that increase the cognitive workload during the perception of an assembly task. They include worker-specific information [12], variant-specific information [10] as well as influences from product changes. In order to evaluate the Perception Complexity of an assembly task, the following three components are introduced:

- The *Employee Qualification Factor (EQF)* describes if assembly workers have the necessary qualification for the performed assembly task.
- The *Customization Level (CL)* describes the influence of product customization on the analyzed assembly task.
- The *Change Frequency (CF)* describes the time intervals the assembly task changes during product changes or the introduction of new products.

In order to achieve a compromise between accuracy and practicability, the three components are evaluated with integer values on a scale from 1 to 10. The three complexity components are multiplied and weighted equally to quantify the Perception Complexity:

$$\mathrm{PC} = \mathrm{EQF} \cdot \mathrm{CL} \cdot \mathrm{CF}. \tag{1}$$

Table 1 shows typical scenarios for the lowest and highest value for the three Perception Complexity components. To assess the complexity for a given assembly task, an intermediate value within the given limits has to be chosen.

**Table 1.** Perception complexity components with lowest and highest value.

| Perception Complexity component | Lower limit: 1 | Upper limit: 10 |
|---|---|---|
| Employee qualification factor | All employees are trained for the assembly task | The assembly task is regularly performed with new employees that lack necessary qualification |
| Customization level | The assembly task is identical in every order | The assembly task is highly dependent on order-specific information |
| Change frequency | The assembly task is not influenced by product changes | The assembly task is frequently changing due to product changes |

### 3.1.2 Execution Complexity

In order to quantify the Execution Complexity of an assembly task, the risk of an assembly error shall be evaluated. In risk assessment applications a *risk priority number* [14] is commonly used for evaluation. This method shall be applied for the Execution Complexity in manual assembly:

- *Severity (S)* describes the impact of an assembly error on product quality and human safety.
- *Occurrence (O)* describes the frequency of assembly errors in the analyzed assembly step.
- *Detection (D)* describes the probability an assembly error is detected before the final assembly step is completed.

According to the risk priority number, the three components are evaluated with integer values on a scale from 1 to 10. In order to quantify the Execution Complexity, the three complexity components are multiplied [14]:

$$\mathrm{EC} = \mathrm{S} \cdot \mathrm{O} \cdot \mathrm{D}. \tag{2}$$

Table 2 shows typical scenarios for the lowest and highest value for the three Execution Complexity components.

**Table 2.** Execution complexity components with lowest and highest value.

| Execution complexity | Lower limit: 1 | Upper limit: 10 |
|---|---|---|
| Severity | An error in the assembly has no effect on the product quality | An error in the assembly task always has an hazardous effect |
| Occurrence | History shows no failures in the analyzed assembly task | Failures in the analyzed assembly task are almost certain |
| Detection | Every error in the assembly task is directly detected at the next process step | Every error in the assembly task cannot be detected in all following process steps |

### 3.1.3 Documentation Complexity

In order to evaluate the Documentation Complexity, the reason for documentation and its effort are analyzed. Moreover, the reason for documentation can be divided into internal and external reasons factors. In order to evaluate the Documentation Complexity of an assembly task, the following three components are introduced:

- *Documentation Necessity (DN)* describes the importance of documentation caused by external factors, such as legal aspects or customer demands.
- *Internal Information Value (IIV)* describes if a documentation of the assembly step is important for internal assembly data evaluations such as productivity indicators or traceability of components.
- *Documentation Effort (DE)* describes the effort of manual documentation without digital assistance.

In order to achieve a compromise between accuracy and practicability, the three components are evaluated with integer values on a scale from 1 to 10. The three complexity components are multiplied and weighted equally to quantify the Documentation Complexity:

$$EC = DN \cdot IIV \cdot DE. \tag{3}$$

Table 3 shows typical scenarios for the lowest and highest value for the three Execution Complexity components.

**Table 3.** Documentation complexity components with lowest and highest value.

| Documentation complexity | Lower limit: 1 | Upper limit: 10 |
|---|---|---|
| Documentation necessity | No external party requires a documentation of the essential task | Documentation of every aspect of the assembly task is required |
| Internal information value | A documentation of the assembly task creates no internal benefits | A documentation of the assembly creates a high benefit for internal evaluations |
| Documentation effort | The documentation effort of the assembly task is negligible | The documentation effort surpasses the value-adding assembly time by many times |

## 3.2 Cognitive Assistance Functions

### 3.2.1 Framework

A cognitive assistance system is able to exchange information with the objects *assembly worker*, *IT systems* and *product* [15]. Figure 1 shows an illustration if the information flow framework.

Communication with the assembly worker consists of input given by the worker as well as information provided for the worker by the assistance system. An assistance system can also communicate with existing IT Systems such as enterprise resource

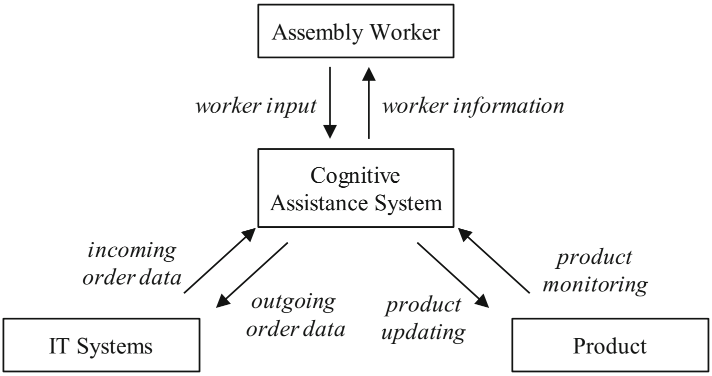



**Fig. 1.** Information flow framework for cognitive assistance systems in manual assembly [15].

planning systems or manufacturing execution systems. This communication is used to receive order information as well as to document an accomplished assembly. Furthermore, an assistance system can also read and write data from a product if it or the workpiece carrier is equipped with rewritable digital memory. An information flow with the product can also be realized by intelligent tools that are part of the assistance system.

Assuming in each assistance function an assistance system always has one data input and one data output, there are two possible outputs for every three inputs. In total, there are six possible assistance functions for the communication with assembly worker, IT systems and product. Figure 2 shows the information flow of the six assistance functions and the corresponding description. The six assistance functions are to be seen as a toolbox and can be combined as necessary for a given use case. The following section connects these six assistance functions with the evaluated complexity dimensions and gives a brief explanation how the six assistance functions can support manual assembly tasks.

### 3.2.2 Selection of Assistance Functions

In order to make a selection of the introduced assistance function, they are mapped in Table 4 with the three complexity dimensions. A high Perception Complexity indicated the need for *worker information* in order to provide the relevant information for the assembly task at the right time. A high Execution Complexity can be handled by the assistance functions *quality assurance* and product *manipulation*. In *quality assurance* the assistance system supervises the assembly process and only gives feedback if an error was made. A *product manipulation* can be realized in order to avoid manual mistakes, e.g. by introducing a torque-controlled screwdriver. Documentation can be assisted in several ways. A *manual documentation* can be integrated easily by using keyboard or touchscreen input. *Automatic documentation* can be realized using a camera at an assembly workstation. A *product documentation* is only possible if the product is equipped with digital memory, e.g. radio-frequency identification (RFID) transponders.

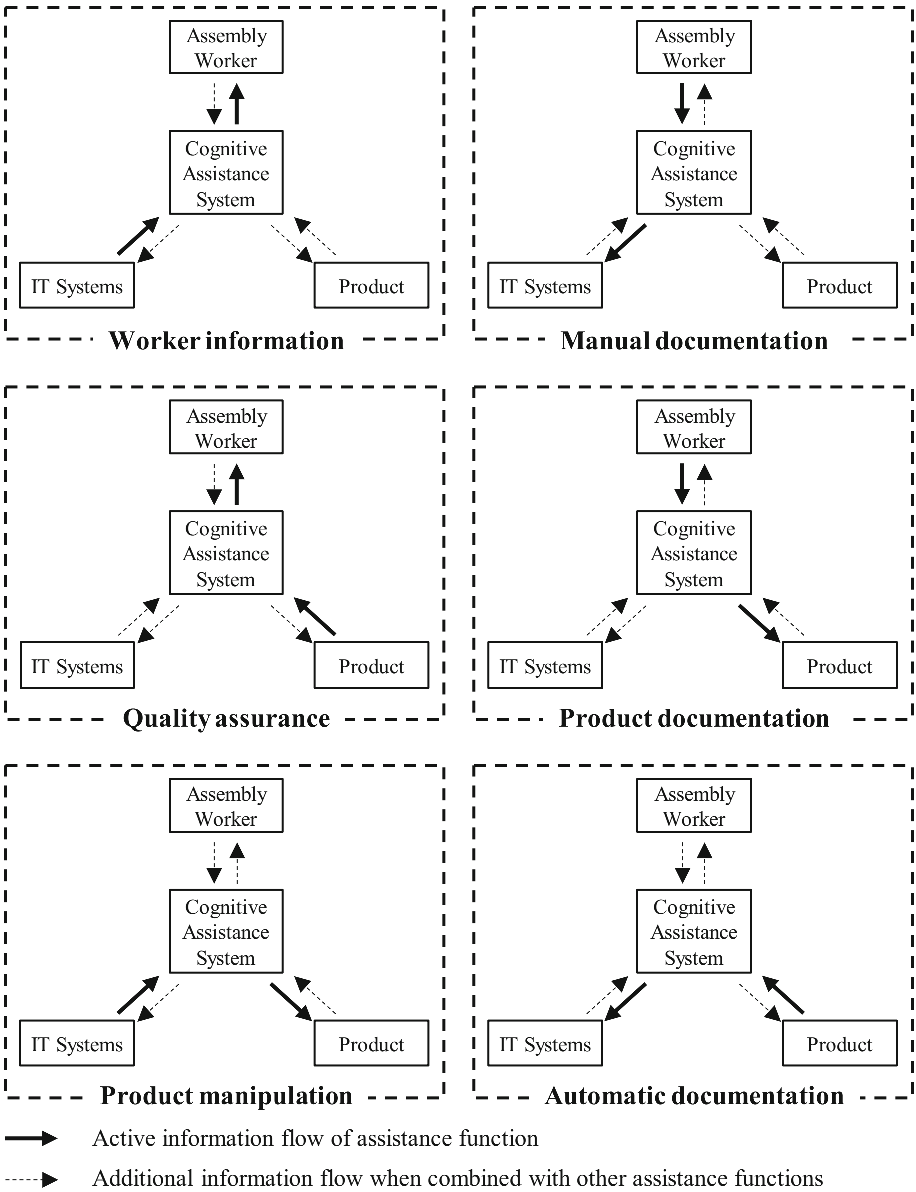



**Fig. 2.** Information flow of six developed assistance functions.

**Table 4.** Mapping of complexity dimensions and assistance functions.

| Assistance function | Perception complexity | Execution complexity | Documentation complexity |
|---|---|---|---|
| Worker information | ● | | |
| Manual documentation | | | ● |
| Quality assurance | | ● | |
| Product documentation | | | ● |
| Product manipulation | | ● | |
| Automatic documentation | | | ● |

## 4 Application in the Learning Factory for Cyber-Physical Production Systems

This section shows an application of the method in the Learning Factory for Cyber-Physical Production Systems. The learning factory consists of six assembly stations for remote-controlled cars [16]. It is organized as a flexible workshop assembly with dedicated parts at each workstation. Depending on the customer's product configuration every product is only transported to the relevant workstations.

The method is applied on workstation #4 'Electronics' of an assembly system for remote-controlled cars. In workstation #4 there are five assembly tasks to be executed. At first a battery is installed for which a customer can choose from three different options. The installation of a wrong battery cannot be detected after that chassis has been put on. Then, one of two antenna sleeves is inserted. In contrast to the battery, a wrong antenna sleeve can still be detected when the chassis has been put on. If demanded by the customer, the car is equipped with a light. For every car the speed controller is connected and an electronic test is performed. The final two assembly tasks are identical in every order but essential for the functionality. Table 5 shows the evaluation of assembly tasks at workstation #4 which was performed by developers of the learning factory.

**Table 5.** Evaluation of workstation #4 in the learning factory for cyber-physical assistance systems.

| Assembly tasks | PC | | | EC | | | DC | | |
|---|---|---|---|---|---|---|---|---|---|
| | EQF | CL | CF | S | O | D | DN | IIV | DE |
| Battery installation | 6 | 8 | 6 | 3 | 3 | 2 | 1 | 1 | 4 |
| | PC = 288 | | | EC = 18 | | | DC = 4 | | |
| Insertion of antenna sleeve | 6 | 6 | 2 | 3 | 3 | 5 | 1 | 1 | 4 |
| | PC = 72 | | | EC = 45 | | | DC = 4 | | |
| Light installation | 6 | 4 | 6 | 3 | 2 | 3 | 1 | 1 | 2 |
| | PC = 144 | | | EC = 18 | | | DC = 2 | | |
| Speed controller connection | 4 | 1 | 1 | 7 | 4 | 10 | 1 | 1 | 2 |
| | PC = 4 | | | EC = 280 | | | DC = 2 | | |
| Electronic test | 6 | 1 | 1 | 7 | 1 | 10 | 5 | 1 | 2 |
| | PC = 6 | | | EC = 70 | | | DC = 10 | | |

The evaluation of the Perception Complexity results in high values for the battery installation and light installation. As a consequence, *worker information* has been applied by introducing tablets with order-specific information. In addition, the implementation of a pick-by-light system would be possible.

The Execution Complexity results in high values for the speed controller connection. However, this connection is already ensured by the electronic test. A possible quality assurance system would require a connection of the speed controller with the assistance system. Then a quality assurance procedure could be integrated into the assembly workflow.

A documentation of the installed components is currently not required. However, it is planned to implement a tablet-based documentation at a later station in order to perform additional quality checks.

Besides the evaluation of workstation #4, the presented approach is applied at every workstation of the learning factory. If new product variants are introduced, the evaluation will be reviewed in order to identify chances for the need of assistance.

## 5 Summary and Outlook

Producing companies face several challenges today. Cognitive assistance system can help to master rising complexity in manual assembly systems. In order to realize an application-specific and needs-based design of cognitive assistance systems a method for complexity evaluation and a framework for cognitive assistance functions have been introduced. Complexity is evaluated for each assembly task in dimensions of perception, execution and documentation. For each complexity dimension the most important influences are evaluated in order to make a needs-based selection of assistance functions. The developed framework for assistance functions introduces six directions of information flow that can be used when applying a cognitive assistance system at a manual assembly station.

In the future, the method has to be evaluated in different industrial applications. The method has to be extended in order to select technologies and components to configure to realize a cognitive assistance system. Furthermore, an economic evaluation has to be introduced in order to identify the appropriate degree of assistance during the assembly process.

## References

1. Reinhart, G.: Handbuch Industrie 4.0. Carl Hanser, München (2017)
2. Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., Schlund, S.: Produktionsarbeit der Zukunft - Industrie 4.0. Fraunhofer Verlag, Stuttgart (2013)
3. Schuh, G.: Produktkomplexität managen. Carl Hanser, München (2005)
4. Abele, E., Reinhart, G.: Zukunft der Produktion. Carl Hanser, München (2011)
5. Gairola, A.: Montagegerechtes Konstruieren. Ein Beitrag zur Konstruktionsmethodik. TH Darmstadt (1981)

6. Eversheim, W.: Organisation in der Produktionstechnik Band 1. VDI-Verlag, Düsseldorf (1996)
7. Lang, S.: Durchgängige Mitarbeiterinformation zur Steigerung von Effizienz und Prozesssicherheit in der Produktion. Meisenbach Verlag, Bamberg (2007)
8. Günthner, W., Klenk, E., Tenerowicz-Wirth, P.: Adaptive Logistiksysteme als Wegbereiter der Industrie 4.0. In: Bauernhansl, T., ten Homel, M., Vogel-Heuser, B. (eds.) Industrie 4.0 in Produktion, Automatisierung und Logistik, pp. 297–323. Springer, Wiesbaden (2014)
9. Samy, S.N., ElMaraghy, H.: A model for measuring products assembly complexity. Int. J. Comput. Integr. Manuf. **23**(11), 1015–1027 (2010)
10. Zeltzer, L., Limère, V., Van Landeghem, H., Aghezzaf, E., Stahre, J.: Measuring complexity in mixed-model assembly workstations. Int. J. Prod. Res. **51**(15), 4630–4643 (2013)
11. Zaeh, M.F., Wiesbeck, M., Stork, S., Schuböh, A.: A multi-dimensional measure for determining the complexity of manual assembly operations. Prod. Eng. Res. Dev. **3**, 489–496 (2009)
12. Claeys, A., Hoedt, S., Soete, N., Van Landeghem, H., Cottyn, J.: Framework for evaluating cognitive support in mixed model assembly systems. IFAC-PapersOnLine **48**(3), 924–929 (2015)
13. Hold, P., Erold, S., Reisinger, G., Sihn, W.: Planning and evaluation of digital assistance systems. Procedia Manuf. **9**, 143–150 (2017)
14. Wilson, R., Hill, A.V.: The Operations Management Complete Toolbox. Pearson Education, London (2013)
15. Merkel, L., Berger, C., Schultz, C., Braunreuther, S., Reinhart, G.: Application-specific design of assistance systems for manual work in production. In: IEEE International Conference on Industrial Engineering and Engineering Management, Singapore (2017)
16. Merkel, L., Atug, J., Merhar, L., Schultz, C., Braunreuther, S., Reinhart, G.: Teaching smart production: an insight into the learning factory for cyber-physical production systems (LVP). Procedia Manuf. **9**, 269–274 (2017)

# Interpersonal Resonance: Developing Interpersonal Biofeedback for the Promotion of Empathy and Social Entrainment

John M. Tennant(✉), Simon Cook, Mihnea C. Moldoveanu, Jordan B. Peterson, and William A. Cunningham

University of Toronto, Toronto, ON, Canada
john.tennant@mail.utoronto.ca

**Abstract.** We propose an interpersonal biofeedback technology which uses music-like stimuli to convey a user's physiological information to another observer. It is argued that this interpersonal biofeedback may facilitate empathy and interpersonal entrainment. We argue that music is an optimal carrier for biofeedback because it naturally regulates psychophysiology, is cross-modally associated with emotion, and can be attended to peripherally. We propose a research study to investigate the effects of interpersonal biofeedback on emotional mindreading.

**Keywords:** Interpersonal synchrony · Biofeedback · Empathy
Entrainment

## 1 Introduction

How do humans understand each other? Undoubtedly, this process requires a great deal of unconscious expression and processing of nonverbal cues: people unconsciously express emotions through actions and expressions, and mirror and entrain with others to establish shared psychophysical context [1]. They also attend to these nonverbal cues use them to make cognitive inferences about others' mental states [2]. These two processes together contribute to empathy, the ability to generate an understanding of the meaning and intent behind another person's actions. What if technology could facilitate this process in computer-mediated interaction? In this paper we outline the development of an interpersonal biofeedback system designed to facilitate mental inference between users. We explore theoretical and practical concerns in the development of such an technology, and propose applications in psychological research.

## 2 Designing Interpersonal Biofeedback

Traditionally, biofeedback involves presenting a user with data about their own biological processes. Biofeedback training helps users better understand, predict, or control those processes. However, little research has been done on exposing users to another person's biosignals, which is our primary concern. A single study serves as


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 208–214, 2019.
https://doi.org/10.1007/978-3-319-94619-1_20

evidence to the plausibility of this approach. Researchers showed that listening to another person's heartbeat while conversing has been shown to produce feelings of intimacy similar to eye contact [3]. We hypothesize that an interpersonal biofeedback technology may be able promote empathy through two distinct but compatible mechanisms of action: (1) by facilitating psychophysiological entrainment between the user and the target (entrainment approach), and (2) by providing users with relevant information for making mental inference (social information-processing approach). In this paper, we propose an interpersonal biofeedback technology where users are given perceptual access to another person's nonverbal signals as quasi-musical auditory stimuli. In order to maximize the effectiveness of the biofeedback intervention, the system will be designed to optimize both possible mechanisms of action. Each foregrounds different considerations which inform the major decisions in the design of a biofeedback system: appropriate selection of signals to encode (the message), and the method of encoding (the medium).

### 2.1 Interpersonal Synchrony

People tend to unconsciously mirror, and fall into rhythms with other people, especially when performing similar tasks. Interpersonal synchrony researchers have studied emergent synchrony in pairs and groups of interacting participants. Researchers have found that overt motor behaviour such as rocking [4], tapping [5], speaking [6], and posturing [7] entrain between interacting participants. These synchronization processes have implications for social cognition. When pairs and groups of interacting people become entrained, they may find each other easier to understand and cooperate with. Indeed, researchers have found that induced or incidental motoric entrainment has positive effects on affiliation [8], empathy [5], cooperation [9], and altruism [10, 11]. Deficits in the ability to entrain with a partner have been implicated in social disorders such as Autism Spectrum Disorder [5].

More covert or difficult-to-perceive phenomena such as breathing [12], heart rate [13], skin conductance [14], and EEG activity [15] also show entrainment effects. But the causal role of these entrainment processes is not known. These covert synchrony processes have proven more difficult to study due to the difficulty of directly manipulating physiological or neurological synchrony. Thus research on covert interpersonal synchrony has been limited to correlational and observational study. So while researchers have been able show that entrainment correlates with feelings of intimacy [14], they have not been able to manipulate physiological synchrony directly in order to attribute to it a causal role in social cognition. Fortunately, many of the processes which show entrainment effects between people also show entrainment effects to repetitive environmental stimuli. The physiological rhythms of heart rate [16, 17], breathing [18], brainwaves [19], and motoric behaviour (e.g. dancing, foot tapping) [20] all show entrainment to both social and repetitive sensory stimuli (stimuli such as flashing lights, repetitive sounds (including music) and tactile sensations [21]). This suggests that an interpersonal biofeedback system which represents these biorhythms as entrainment-ready repetitive stimuli may allow us experimentally manipulate physiological synchrony, and consequently allow us to explore physiological synchrony's causal role in social cognition.

### 2.2 Social Information Processing

Given access to informative cues about another person's mental state, users should be able to improve their ability to make mental inferences, regardless of whether they experience physiological synchrony with the target. Learning and conditioning research shows that organisms readily learn and use (1) stable, (2) informative, and (3) relevant stimulus pairings [22] in problem-solving contexts. Accordingly, for users to make effective use of the interpersonal biofeedback system, this research suggests that encoded cues should (1) reliably and (2) informatively covary with mental states, and should be (3) non-redundant, and this difficult to perceive under normal circumstances. Cues such as facial expressions, vocal prosody, and body language covary with mental states [22–25] (indeed they likely serve a communicative function [26]). But, while informative, these overt expressions of emotion are so easy to perceive in normal circumstances that biofeedback information about them would be redundant. In contrast, physiological signals like heartbeat, breathing, skin conductance also covary with mental states [27], but these are difficult to perceive under normal circumstances. They are therefore good candidates for inclusion in the biofeedback system.

### 2.3 Interim Summary

Both interpersonal synchrony research and the social-information processing approach suggest that making physiological signals perceptually available to users may improve empathy. In addition to the content of the biofeedback system, it is important to consider the medium (i.e. the encoding scheme) used to present the data to the user. We propose using quasi-musical auditory stimuli to convey this information.

### 2.4 Musical Auditory Biofeedback

Music is an ideal carrier for interpersonal biofeedback data. It is one of the most researched targets for environmental entrainment of both motor action and biorhythms [28], and it automatically regulates psychophysiology [29]. Thus, music shows promise as an entrainment-ready stimulus for facilitating interpersonal synchrony.

Humans have highly sophisticated musical-cognitive abilities, and can process and entrain to many layers of complex sensation. Hence, use of musical/auditory signaling will allow for the simultaneous encoding of multiple channels of biofeedback data, into a stimulus which is easily discriminable, and readily parsed. More generally, auditory stimuli will not take up attention or space in the visual domain, ensuring that focal tasks requiring visual attention will be minimally hindered by the use of biofeedback.

Finally, music seems to have an intuitive relationship to the communication of mental states. People speak cross-culturally about the emotions or mental states conveyed by music [30], and musical biofeedback has demonstrated efficacy as a carrier for biofeedback information in individuals [31]. Music is therefore an intuitive choice for representing psychophysical states in interpersonal biofeedback. Researchers have also found that certain musical parameters are cross-modally congruent with patterns of motion of on-screen objects and certain emotional states [32]. We can make use of these pairings in designing our biofeedback technology to ensure that the stimulus decoding is as intuitive as possible.

# 3 Proposed Biofeedback System

We propose an interpersonal biofeedback technology which conveys a user's physiological information to another observer in the form of quasi-musical stimuli. It is hypothesized that by increasing users' perceptual access to relevant physiological signals in another person, we will be able to enhance their ability to discern the other person's mental state. This technology will take advantage of two potentially independent, but mutually compatible processes in empathy. First, if the encoded physiological signals are indicative of the user's mental state, then learning & conditioning research suggests that an observer will be able to learn to use the biofeedback signal to make correct mental inferences. Second, interpersonal synchrony research suggests that the increased perceptual access will facilitate entrainment of physiological processes between the pairs of users, and that this entrainment will facilitate correct mental inference. We hypothesize that increasing perceptual access to physiological processes such as breathing, heart rate, and skin conductance will promote correct inferences through one or both of these processes. We will then be able to test if degree of entrainment predicts accuracy of mental inference judgements.

# 4 Research Utility

## 4.1 Current Research

Our current research is making use of the biofeedback system as follows:

Phase 1: Participants' psychophysiology (breathing, heart-rate & skin conductance) is recorded while they view a series of emotional video clips. This is similar to manipulations by Davis et al. [33, 34]). All participants are video recorded throughout. After each emotional video clip, participants log their emotional experience using a questionnaire and journal.

Phase 2: New participants watch all the emotional video clips shown to participants in phase 1. They are then shown video recordings of participants from phase 1. These video recordings show the target (with no sound) watching one of the emotional video clips. The content on the screen is hidden from view, however, so phase-2 participants cannot tell directly which clip the participant in the recording is being shown. Some participants hear the target's biofeedback through headphones, others receive none or sham (biofeedback from a different target). Phase-2 participants' physiology is recorded throughout. After watching each video recording, participants report on the emotions experienced by the subject in the video, and try to guess which video the subject was watching. We then examine (1) if hearing biofeedback from the target improves the accuracy of emotional judgements, (2) if participants show patterns of physiological synchrony with the target (and/or with the sham biofeedback), and (3) if such synchrony improves performance on emotional empathy.

Two important limitations of this research are as follows: (1) Effective use of the interpersonal biofeedback system might require extended training which is not included in this research paradigm, (2) interesting effects of the system might be found in contexts where two people use the system simultaneously. In this situation, the mutual entrainment and interpersonal dynamics of experiencing a real interaction partner's biorhythms may more effectively engage social-cognitive processes and intuitive learning of the encoding system. Future research should investigate these exciting prospects.

## 5 Conclusion

Due to the limitations of previous research on covert interpersonal synchrony, researchers have not been able to assess the causal role of covert synchrony in social-cognitive processes such as empathy and trust. Although behavioural research has supported the hypothesis that embodied simulation of another person aids in empathy and promotes intimacy, the covert synchrony research has not been able to follow suit. Skepticism about the causal role of covert synchrony in producing these effects is warranted, as there are no clear mechanism for people to synchronize their biorhythms. The approach outlined here provides a means for clarifying this dispute. If the entrainment intervention is effective in manipulating interpersonal synchrony, then researchers will be able to examine the effects of covert synchrony much more directly. If successful manipulations of synchrony show social-cognitive effects on empathy/empathy, then this will be strong support for the causal role of covert synchrony in social cognition. On the contrary, if increases in covert synchrony are not associated with these effects, it would support a more cognitive theory of empathy which ascribes greater importance to attention to social cues and mental inference. Each of these results would be valuable. Significant or null findings resulting from this research could inform the direction of future social-cognitive theory and research programs. Whether the primary mechanism is entrainment, or attention to social cues, the interpersonal biofeedback methodology developed here may contribute to the development of interventions and technologies for enhancing social cognition in broader contexts.

## References


1. Zaki, J.: Empathy: a motivated account. Psychol. Bull. **140**, 1608–1647 (2014)
2. Zaki, J., Ochsner, K.: You, me, and my brain: self and other representations in social cognitive neuroscience. In: Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind, vol. 26, p. 48 (2011)
3. Janssen, J.H., Bailenson, J.N., IJsselsteijn, W.A., Westerink, J.H.D.M.: Intimate heartbeats: opportunities for affective communication technology. IEEE Trans. Affect. Comput. **1**, 72–80 (2010)
4. Richardson, M.J., Marsh, K.L., Isenhower, R.W., Goodman, J.R.L., Schmidt, R.C.: Rocking together: dynamics of intentional and unintentional interpersonal coordination. Hum. Mov. Sci. **26**, 867–891 (2007)

5. Koehne, S., Hatri, A., Cacioppo, J.T., Dziobek, I.: Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. Cognition **146**, 8–15 (2016)
6. Kinsbourne, M., Marcel, K., Scott Jordan, J.: Embodied anticipation: a neurodevelopmental interpretation. Discourse Process. **46**, 103–126 (2009)
7. Schmidt, R.C., Morr, S., Fitzpatrick, P., Richardson, M.J.: Measuring the dynamics of interactional synchrony. J. Nonverbal Behav. **36**, 263–279 (2012)
8. Hove, M.J., Risen, J.L.: It's all in the timing: interpersonal synchrony increases affiliation. Soc. Cogn. **27**, 949–960 (2009)
9. Marsh, K.L., Richardson, M.J., Schmidt, R.C.: Social connection through joint action and interpersonal coordination. Top. Cogn. Sci. **1**, 320–339 (2009)
10. Valdesolo, P., DeSteno, D.: Synchrony and the social tuning of compassion. Emotion **11**, 262–266 (2011)
11. Trainor, L.J., Cirelli, L.: Rhythm and interpersonal synchrony in early social development. Ann. N. Y. Acad. Sci. **1337**, 45–52 (2015)
12. Codrons, E., Bernardi, N.F., Vandoni, M., Bernardi, L.: Spontaneous group synchronization of movements and respiratory rhythms. PLoS One **9**, e107538 (2014)
13. Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., Louzoun, Y.: Mother and infant coordinate heart rhythms through episodes of interaction synchrony. Infant Behav. Dev. **34**, 569–577 (2011)
14. Slovák, P., Tennent, P., Reeves, S., Fitzpatrick, G.: Exploring skin conductance synchronisation in everyday interactions. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction Fun, Fast, Foundational - NordiCHI 2014 (2014)
15. Yun, K., Watanabe, K., Shimojo, S.: Interpersonal body and neural synchronization as a marker of implicit social interaction. Sci. Rep. **2**, 959 (2012)
16. Bernardi, L., Porta, C., Sleight, P.: Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence. Heart **92**, 445–452 (2006)
17. Fukumoto, M., Nomura, S., Sawai, M., Imai, J.-I., Nagashima, T.: Investigation of synchronization between musical beat and heartbeat with cardio-music synchrogram. Nonlinear Theory Appl. IEICE **1**, 146–152 (2010)
18. Haas, F., Distenfeld, S., Axen, K.: Effects of perceived musical rhythm on respiratory pattern. J. Appl. Physiol. **61**, 1185–1191 (1986)
19. Zhuang, T., Zhao, H., Tang, Z.: A study of brainwave entrainment based on EEG brain dynamics. Comput. Inf. Sci. 2, (2009)
20. Trainor, L.J., Gao, X., Lei, J.-J., Lehtovaara, K., Harris, L.R.: The primal role of the vestibular system in determining musical rhythm. Cortex **45**, 35–43 (2009)
21. Huang, T.L., Charyton, C.: A comprehensive review of the psychological effects of brainwave entrainment. Altern. Ther. Health Med. **14**, 38–50 (2008)
22. Rumbaugh, D.M., King, J.E., Beran, M.J., Washburn, D.A., Gould, K.L.: A salience theory of learning and behavior: with perspectives on neurobiology and cognition. Int. J. Primatol. **28**, 973–996 (2007)
23. Russell, J.A.: Core affect and the psychological construction of emotion. Psychol. Rev. **110**, 145–172 (2003)
24. Ekman, P.: Emotional and conversational nonverbal signals. In: Language, Knowledge, and Representation, pp. 39–50. Springer, Heidelberg (2004)
25. Aviezer, H., Trope, Y., Todorov, A.: Body cues, not facial expressions, discriminate between intense positive and negative emotions. Science **338**, 1225–1229 (2012)
26. Williams, W.C.: Social goals shape the facial expression of emotion: subjective, behavioral, and neural bases of communicating feelings (2017)
27. Oehme, A., Herbon, A., Kupschick, S.: Physiological Correlates of Emotions (2007)

28. Trost, W., Vuilleumier, P.: Rhythmic entrainment as a mechanism for emotion induction by music. In: Multidisciplinary Perspectives on Musical Arousal, Expression, and Social Control, pp. 213–225 (2013)
29. van der Zwaag, M.D., Westerink, J.H., van den Broek, E.L.: Emotional and psychophysiological responses to tempo, mode, and percussiveness. Music Sci. **15**, 250–269 (2011)
30. Cross, I., Woodruff, G.E.: Music as a communicative medium. Prehist. Lang. **1**, 113–144 (2009)
31. Vidyarthi, J., Riecke, B.E., Gromala, D.: Sonic cradle. In: Proceedings of the Designing Interactive Systems Conference on - DIS 2012 (2012)
32. Sievers, B., Polansky, L., Casey, M., Wheatley, T.: Music and movement share a dynamic structure that supports universal expressions of emotion. Proc. Natl. Acad. Sci. U. S. A. **110**, 70–75 (2013)
33. Davis, J.I., Senghas, A., Ochsner, K.N.: How does facial feedback modulate emotional experience? J. Res. Pers. **43**, 822–829 (2009)
34. Davis, J.I., Senghas, A., Brandt, F., Ochsner, K.N.: The effects of BOTOX injections on emotional experience. Emotion **10**, 433 (2010)

# IMU-Based Motion Capture Wearable System for Ergonomic Assessment in Industrial Environment

Francesco Caputo[1(✉)], Alessandro Greco[1], Egidio D'Amato[1], Immacolata Notaro[1], and Stefania Spada[2]

[1] Department of Engineering, University of Campania Luigi Vanvitelli, via Roma 29, 81031 Aversa, CE, Italy
{francesco.caputo,alessandro.greco, immacolata.notaro}@unicampania.it, egidio.damato@unina.it

[2] FCA Italy – EMEA Manufacturing Planning & Control – Ergonomics, Gate 16, Corso Settembrini 53, 10135 Turin, Italy
Stefania.spada@fcagroup.com

**Abstract.** The study of human factors is fundamental for the human-centered design of Smart Workplaces. IIoT (Industrial Internet of Things) technologies, mainly wearable devices, are becoming necessary to acquire data, whose analysis will be used to make decision in a smart way. For industrial applications, motion-tracking systems are strongly developing, being not invasive and able to acquire high amounts of data related to human motion in order to evaluate the ergonomic indexes in an objective way, as well as suggested by standards. For these reasons, a modular inertial motion capture system has been developed at the Department of Engineering of the University of Campania Luigi Vanvitelli. By using low cost Inertial Measurement Units – IMU and sensor fusion algorithms based on Extended Kalman filtering, the system is able to estimate the orientation of each body segment, the posture angles trends and the gait recognition during a working activity in industrial environment. From acquired data it is possible to develop further algorithms to online asses ergonomic indexes according to methods suggested by international standards (i.e. EAWS, OCRA, OWAS). In this paper, the overall ergonomic assessment tool is presented, with an extensive result campaign in automotive assembly lines of Fiat Chrysler Automobiles to prove the effectiveness of the system in an industrial scenario.

**Keywords:** Motion capture · Wearable device · IMU · Industrial ergonomics Industrial environment

## 1 Introduction

The use of wearable devices in industrial environment, during the last years, continues to capture high levels of interest, as companies across a wide variety of industrial sectors conduct several trials with different set of devices.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 215–225, 2019.
https://doi.org/10.1007/978-3-319-94619-1_21

The main manufacturing companies are realizing that to be competitive in the market, especially during the Industry 4.0 era, the introduction of new technologies directly in the factory is needed, in order to make decision in the best way, optimizing production lines and making workplaces safe for workers.

Within manufacturing industries, in particular the automotive ones, ergonomics can be considered a very important aspects for the continuous screening about the working conditions, as well as a fundamental variable in the design of workplaces.

Fundamental analyses concern the working postures assumed by the workers. Posture angles become fundamental for studying this ergonomic factor and they can be provided by using wearable inertial motion capture system, because of their low invasiveness in a manufacturing plant, allowing the worker to normally continue its activity.

Indeed, as MEMS (Micro Electro-Mechanical Systems) inertial sensors are compact and light, they have been a popular choice for applications such as motion tracking or animation. Some interesting uses are described in [1–4].

As well as in [5–7], in this paper multiple micro Inertial Measurement Units (IMUs) are involved to analyze human postures. A Kalman filter is used to compute the estimation of the attitude for each IMU, by combining a series of measurements affected by noise.

This research is aimed to propose a body motion tracking system, already partially presented in [8, 9], composed by several IMUs, developed at the Department of Engineering of the University of Campania *Luigi Vanvitelli* in collaboration with LinUp S.r.l.

A test case is presented in order to prove the reliability of the system and how it allows to automatically evaluating ergonomic indexes by analyzing working postures in detail.

## 2 Modular Body Motion Tracking System

The core of the proposed motion tracking system is an IMU composed by a triaxial accelerometer, a triaxial gyroscope and a triaxial magnetometer. These inertial sensors represent the minimum set of measures to estimate the attitude of a rigid body system. In our laboratory, a low cost IMU has been developed based on a STM32F103 CPU and the MPU6050 (accelerometer and gyroscope) and the HMC5883L (magnetometer) as MEMS sensors.

Considering the human body as a system made of several rigid segments, connected by joints, each IMU can be mounted on a single part to estimate its pose. By concurrently monitoring every IMU it is possible to reconstruct the human posture.

Consider $\varepsilon = X_E Y_E Z_E$ as the earth fixed reference frame, coincident with the pelvis reference system at the initial time, such that the origin is in the center of mass of the body, the $Z$ axis is parallel to the gravity vector, the $X$ axis points to the front of the body and the $Y$ axis creates a left-handed reference system with the other two axes (Fig. 1).

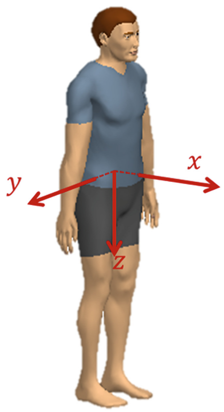



**Fig. 1.** Human body reference system.

Each segment is defined through a local frame $\mathcal{B}$ in agreement with its orientation and overlapped to the earth frame at the initial time. To avoid singularities, the orientation of each segment can be determined using a quaternion based orientation for each segment.

The transformation of an arbitrary vector $\boldsymbol{x}$ between the fixed frame ($\varepsilon$) and the local frame ($\mathcal{B}$) can be written as follows:

$$\boldsymbol{x}^b = \boldsymbol{C}_{BE}(q(t))\boldsymbol{x}^E, \tag{1}$$

where $\boldsymbol{q} = [q_1, q_2, q_3, q_4]$ is the quaternion vector and $\boldsymbol{C}_{BE}$ is the rotation matrix defined as follows:

$$\boldsymbol{C}_{BE}(\boldsymbol{q}) = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_0^2 & 2(q_1q_2 - q_3q_0) & 2(q_1q_3 + q_2q_0) \\ 2(q_1q_2 + q_3q_0) & -q_1^2 + q_2^2 - q_3^2 + q_0^2 & 2(q_2q_3 - q_1q_0) \\ 2(q_1q_3 - q_2q_0) & 2(q_2q_3 + q_1q_0) & -q_1^2 - q_2^2 + q_3^2 + q_0^2 \end{bmatrix} \tag{2}$$

The minimum set of segments, considered in our tests, is composed by: pelvis, trunk, right (left) arm, right (left) forearm, right (left) upper leg and right (left) lower leg.

Furthermore, segments have been grouped in set of 4, to create general purpose modules with the same software that independently estimate global and local orientation of each segment in the own part. This scheme has been useful to limit acquisition burden on the central CPU, represented by a Raspberry Pi board, one for each group. Figure 2 shows the main components of the hardware (HW).

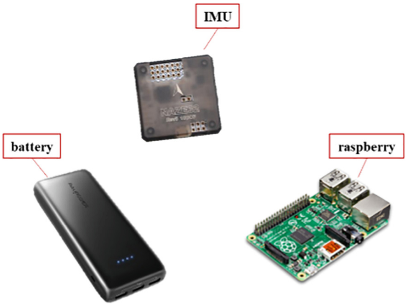



**Fig. 2.** Components of the HW.

The upper limb pose is estimated by using two modules. Each one measures the orientation of pelvis, trunk, right (left) arm and right (left) forearm. Pelvis and trunk measures are redundant. Lower parts are made of: pelvis, upper leg, lower leg. Such partition makes it possible to create general purpose modules that independently estimate global and local orientation of each segment in the own part.

The attitude estimation relies on an Extended Kalman Filter based on the following kinematic model:

$$\begin{cases} \dot{\boldsymbol{q}} = \frac{1}{2}\boldsymbol{M}\boldsymbol{q} \\ \dot{\boldsymbol{b}} = -\frac{1}{\tau}\boldsymbol{b} \end{cases} \tag{3}$$

$$\boldsymbol{M} = \begin{bmatrix} 0 & -(p-b_p) & -(q-b_q) & -(r-b_r) \\ (p-b_p) & 0 & (r-b_r) & -(q-b_q) \\ (q-b_q) & -(r-b_r) & 0 & (p-b_p) \\ (r-b_r) & (q-b_q) & -(p-b_p) & 0 \end{bmatrix} \tag{4}$$

$$\begin{cases} \boldsymbol{a} = \boldsymbol{C}_{BE}\boldsymbol{g} \\ \boldsymbol{B} = \boldsymbol{C}_{BE}\boldsymbol{B}_E, \end{cases} \tag{5}$$

where $\boldsymbol{b} = [b_p, b_q, b_r]^T$ is the gyroscope bias, $\boldsymbol{\omega} = [p, q, r]^T$ the angular velocity, $\boldsymbol{a}$ and $\boldsymbol{B}$ are respectively the acceleration and the magnetic field vector in the local reference system, and $\boldsymbol{g}$ and $\boldsymbol{B}_E$, are gravity vector and earth magnetic field in the fixed reference system. In the state space, the state vector is $= [\boldsymbol{q}^T, \boldsymbol{b}^T]^T$, the output vector is $\boldsymbol{y} = [\boldsymbol{a}^T, \boldsymbol{B}^T]^T$ and the input vector is $\boldsymbol{u} = \boldsymbol{\omega}$.

The output for each segment is based on the Tait-Bryan angles that describe a rotation around the $z$ axis ($\psi$ yaw angle), a rotation around the $y$ axis ($\theta$ pitch angle) and a rotation around the $x$ axis ($\phi$ roll angle).

Figure 3 shows the motion tracking system in upper-body configuration.

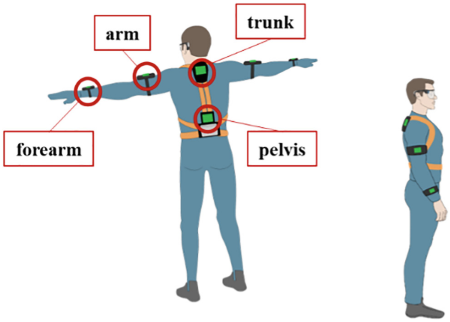



**Fig. 3.** Body motion tracking system in upper-body configuration.

## 3 Test Case

A wide campaign of experimental tests in several assembly lines of FCA plants has been conducted, in order to test the reliability of the system in an industrial environment.

Below, a test case related to an activity carried-out at the assembly line of FCA plant of Melfi is described. The task under examination is the “central-tunnel assembly” of Jeep Renegade and Fiat 500 X.

Table 1 shows a schematic description of the task.

**Table 1.** Central – tunnel assembly tasks, left side of the car.

| No | Task |
|---|---|
| 1 | Pick dashboard panel from the cart |
| 2 | Walk to the car and place the panel on the car floor |
| 3 | Pick and place central-tunnel |
| 4 | Pick and place dashboard panel |
| 5 | Pick screwdriver from the cart |
| 6 | Perform screwing no 1 |
| 7 | Perform screwing no 2 |
| 8 | Make cables connection at the rear of the central-tunnel |

Inertial data about this activity have been acquired by using the only upper-body configuration.

The algorithm provides several data per each segment: attitude quaternions, attitude Euler angles and posture angles. All of these information are enough for a detailed analysis of human motion during the working activities.

In particular, figures below show the trends over the time of the main posture angles about pelvis, trunk, arms and forearms.

Figure 4 shows the trend about the rigid rotation of the pelvis, during a working cycle. In the figure are also underlined the time intervals of the sub-tasks of Table 1.

Pelvis rotation indicates the orientation of the IMU located on the pelvis with regard to the initial orientation of the magnetic field vector.

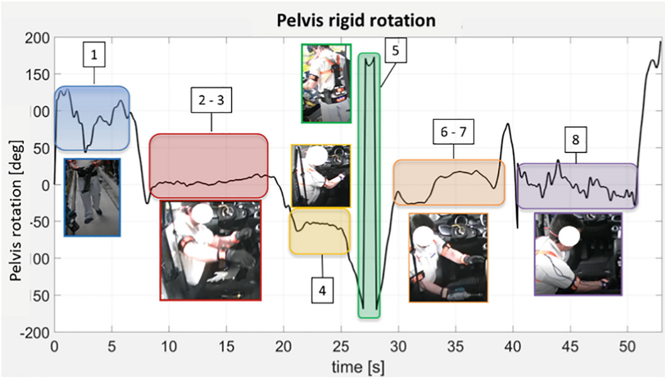



**Fig. 4.** Pelvis rigid rotation and time interval of tasks.

**Table 2.** Acquisition frames from a working cycle: observational evaluation of posture angles.

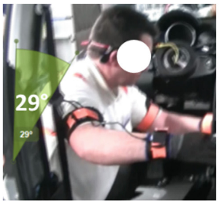
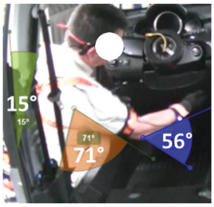
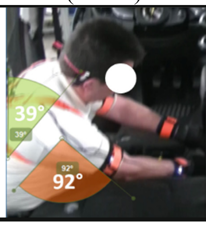
| Acquisition time = 15 s (task 3) | | Acquisition time = 24 s (task 4) | | Acquisition time = 44 s (task 8) | |
|---|---|---|---|---|---|
|  | |  | |  | |
| Main posture angles | | Main posture angles | | Main posture angles | |
| Trunk flexion fw | 29° | Trunk flexion fw | 15° | Trunk flexion fw | 39° |
| Arm elevation | ~90° | Arm elevation | 71° | Arm elevation | 92° |
| Elbow flexion | ~90° | Elbow flexion | 56° | Elbow flexion | ~10° |

Table 2 shows three frames related to task 3, 4 and 8 of Table 1, with the observational evaluation of three main human body posture angles: trunk flection forward, arm elevation and elbow flection.

Figures 5, 6 and 7 show the trends over the time about trunk flection forward, left arm elevation and left elbow flection. The times of frames of Table 2 are circled in black.

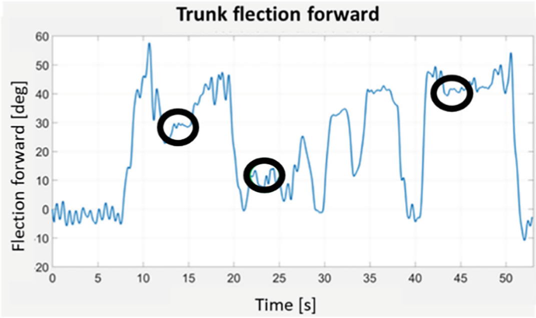



**Fig. 5.** Trunk flection forward trend in a working cycle.

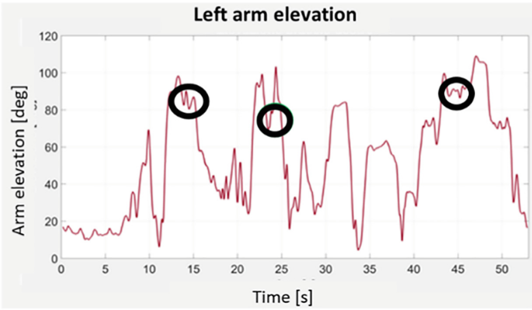



**Fig. 6.** Left arm trend in a working cycle.

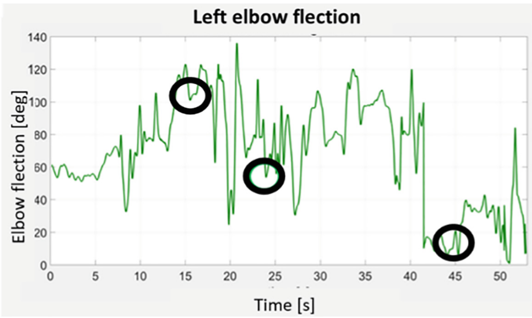



**Fig. 7.** Left elbow flection trend in a working cycle.

Table 3 shows the posture angles values, related to the three same frames of Table 2, evaluated by processing data acquired by the motion capture system. Values are coherent with those ones shown in Table 2.

By combining data related to attitude quaternions and posture angles, an algorithm has been coded, by using MATLAB® programming language, to automatically assign

**Table 3.** Acquisition frames from a working cycle: experimental evaluation of posture angles.

| Acquisition time = 15 s (task 3) | | Acquisition time = 24 s (task 4) | | Acquisition time = 44 s (task 8) | |
|---|---|---|---|---|---|
| Main posture angles | | Main posture angles | | Main posture angles | |
| Trunk flexion fw | 28.64° | Trunk flexion fw | 14.04° | Trunk flexion fw | 41.3° |
| Arm elevation | 86.8° | Arm elevation | 70.48° | Arm elevation | 90.9° |
| Elbow flexion | 99.1° | Elbow flexion | 52.68° | Elbow flexion | 9.42° |

the working posture points according to European Assembly Work-Sheet (EAWS) [1]. The EAWS is a recently developed 1st level screening method for biomechanical overloads. It is composed by four sections. It is particularly suitable for the study of repetitive work activities. Section 1 of EAWS evaluates the working postures and movements with low additional physical effort. It considers static working postures, with a duration of at least 4 consecutive seconds.

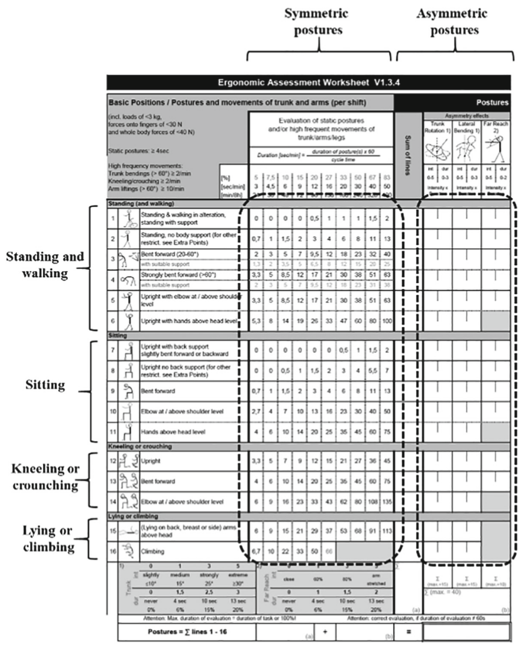



**Fig. 8.** Section 1 of EAWS.

Figure 8 shows the Sect. 1 of the EAWS. This section is able to evaluate the “working postures and movements with low additional physical effort”. It considers both symmetric (bent forward, elbow at shoulder lever and hands above head level) and asymmetric postures (trunk lateral flection, trunk torsion and far reach) in four conditions: standing posture, sitting posture, kneeling or crunching posture and lying or climbing posture.

The total score is done by adding both symmetric and asymmetric points.

Table 4 shows the duration of the static postures characterizing the analyzed activity.

**Table 4.** Duration of static postures for central-tunnel assembly.

| Static Postures | | | Total time [*s/cycle*] | Time to 60 s [*s/min*] |
|---|---|---|---|---|
| Symmetric | Standing and walking | | 27,24 s (+9.1 s balancing difference) | 36,34 *s/min* |
| | Bent forward (20°-60°) | | 25,66 s | 25,66 *s/min* |
| Asymmetric | - | | - | - |

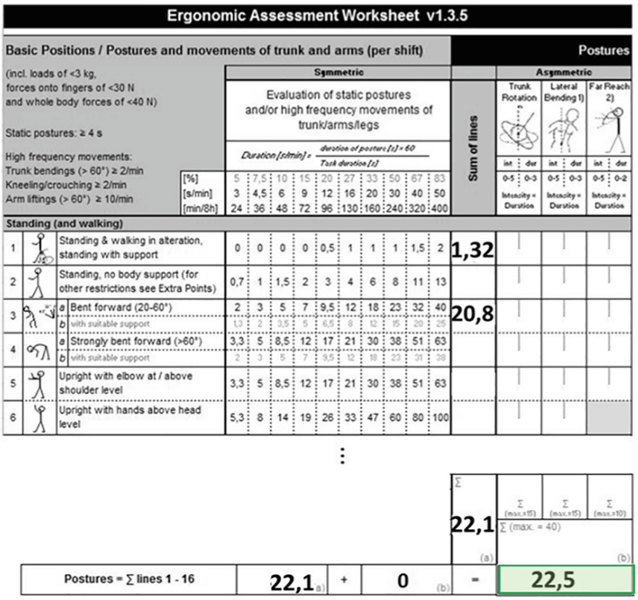



**Fig. 9.** EAWS Sect. 1 score for central-tunnel assembly.

Being the central-tunnel assembly carried-out in standing posture, Fig. 9 shows the Sect. 1 of EAWS checklist and its score, evaluated with experimental data provided by the motion tracking system. The total score, considering the risk area shown in Fig. 10, is within the low risk area and the activity can be considered safe concerning working postures.

| EAWS Risk Index | |
|---|---|
| 0÷25 | Low Risk |
| 26÷50 | Medium Risk |
| >50 | High Risk |

**Fig. 10.** Risk areas for EAWS.

## 4 Conclusions

The present research is aimed to propose a homemade inertial motion tracking system for industrial application, in order to acquire data for human motion analysis and ergonomic indexes evaluation directly on the assembly line.

The system is composed by four independent modules in full-body configuration, with each module composed by 4 IMUs and a raspberry device for data registration.

To support the reliability of the system in an industrial environment, a test case, regarding a working activity in a FCA plant assembly line, has been proposed. The high number of acquired and processed data, by using the system in upper-body configuration, gave the chance to perform a detailed analysis of the human motion and a fast and automatic evaluation of the working postures according to the desired ergonomic method.

**Acknowledgments.** The authors would like to acknowledge the FCA – Fiat Chrysler Automobiles, EMEA Manufacturing Planning & Control – Ergonomics, and the LinUp S.r.l. for supporting the research work on which this paper is based.

## References

1. Veltink, P., Bussmann, H., de Vries, W., Martens, W., van Lummel, R.: Detection of static and dynamic activities using uniaxial accelerometers. IEEE Trans. Rehabil. Eng. **4**, 375–385 (1996)
2. Lyons, G., Culhane, K., Hilton, D., Grace, P., Lyons, D.: A description of an accelerometer-based mobility monitoring technique. Med. Eng. Phys. **27**, 497–504 (2005)
3. Mayagoitia, R., Nene, A., Veltink, P.: Accelerometer and rate gyroscope measurement of the kinematics: an inexpensive alternative to optical motion analysis systems. J. Biomech. **35**, 537–542 (2002)
4. Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Bula, C., Robert, P.: Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. IEEE Trans. Biomed. Eng. **50**, 711–723 (2003)
5. Roetenberg, D., Luinge, H., Slycke, P.: Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors (2009)
6. Zhou, H., Stone, T., Hu, H., Harris, N.: Use of multiple wearable inertial sensors in upper limb motion tracking. Med. Eng. Phys. **30**(1), 123–133 (2008)
7. Yun, X., Bachmann, E.R.: Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking. IEEE Trans. Robot. **22** (6), 1216–1227 (2006)

8. Caputo, F., Greco, A., D'Amato, E., Notaro, I., Spada, S.: A preventive ergonomic approach based on virtual and immersive reality. In: Advances in Intelligent Systems and Computing, Proceedings of the AHFE 2017 International Conference on Ergonomics in Design, Los Angeles, CA, USA (2017)
9. Caputo, F., Greco, A., D'Amato, E., Notaro, I., Spada, S.: Human posture tracking system for industrial process design and assessment. In: Advances in Intelligent Systems and Computing, Proceedings of the IHSI 2018 International Conference on Intelligent Human Systems Interaction, Dubai, vol. 722 (2018)

# Laboratory Experiment on Visual Attention of Pedestrians While Using Twitter and LINE with a Smartphone on a Treadmill

Shigeru Haga(✉) and Taimon Matsuyama

Department of Psychology, Rikkyo University,
1-2-26 Kitano, Niiza, Saitama 352-8558, Japan
haga@rikkyo.ac.jp

**Abstract.** Effects of smartphone use for SNS's while walking were investigated in a laboratory setting. Participants walked on a treadmill for 3 min and performed a visual detection task at the same time while using (under the Twitter and LINE conditions) or not using (under the control condition) an iPhone SE. In front of the treadmill, there was a screen on which a video taken in a crowded underpass was projected. The detection task was to respond to a target (red circle) displayed on the screen 6 times at random intervals in the 3-min trial. Results showed that the number of missed targets was significantly greater and the reaction times to the visual targets were significantly longer under the Twitter and LINE conditions than under the control condition. The results indicated visual inattention of pedestrians using smartphones for Twitter and LINE while walking.

**Keywords:** Smart-phoning while walking · Pedestrian safety · Detection task Reaction time

## 1 Introduction

Pedestrians' inattention while on crowded sidewalks or underpasses sometimes causes injuries and other problems. Lately, many people have been using their smartphones for social networking systems (SNSs) such as Twitter, Facebook, and LINE. The purpose of this study was to collect data from a laboratory experiment on the phenomenon of inattention caused by using Twitter and LINE with a smartphone.

There have been many studies of distracted drivers using cell phones (e.g. [1–5]) while we found relatively few studies concerning cell phone use by pedestrians (e.g. [6–8]).

The first author of this paper has conducted several experiments on the inattention of pedestrians operating a cell phone [9–11]. In his recent study [12], participants read email messages (email-reading condition), exchanged messages through LINE (LINE-chat condition), or just held a phone (control condition) while walking on a treadmill. A movie made using a wearable video camera in an underground passage was projected onto a large screen in front of the treadmill. The participants pressed a hand-held button as quickly as possible when they saw a girl wearing a red cap. Reaction


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 226–232, 2019.
https://doi.org/10.1007/978-3-319-94619-1_22

times were longer and the number of missed targets greater under the email-reading and the LINE-chat conditions than under the control condition. Contrary to our expectation, however, the participants missed many more targets under the email-reading condition than under the LINE-chat condition. We had assumed that participants would be more distracted under the LINE-chat condition because texting messages should be more distracting than just reading messages.

This experiment had a shortcoming. Because the target (girl with a red cap) came into view by walking from a distance in a crowded underpass, timing of the stimulus onset was so ambiguous that the reaction time was not precisely measurable (Fig. 1). Therefore, in the present study, we superimposed targets generated and controlled by a computer program and considered that we could more precisely record responses of participants than with the previous procedure.

**Fig. 1.** Scene in the movie projected on the screen in the first author's previous study [12]. The girl wearing a red cap was the target to be detected.

## 2 Methods

### 2.1 Participants

Fifteen undergraduate students (7 males and 8 females, average age 20.73 years) participated in the experiment. However, due to noncompliance with the instructions, 3 students were excluded from the analysis. As a result, data from 12 participants (7 males and 5 females, average age 20.67 years) were analyzed. All had their own smartphones, with which they were familiar.

### 2.2 Visual Detection Task

A treadmill (Johnson Citta T82) was placed in front of a 120-in screen (2438 × 1829 mm). Distance from the center of the screen to the eyes of the participants was approximately 3 m (Fig. 2). A movie recorded in advance by one of the authors using a wearable video camera (SONY FDR-X3000R) while walking on a sidewalk on Rikkyo Street by the university campus was projected by a projector

(SONY VPL-CX6) hung on the ceiling while the experimental tasks were carried out. Figure 3 shows a scene from the movie.

The participants walked on the treadmill at a velocity of 3 km/h while they performed a visual detection task for 3 min.

The stimulus for reaction was a red circle 10 cm in diameter presented on the screen by a second projector (EPSON EB-535W) placed on the floor. The circle appeared 6 times during the 3-min trial with a duration of 4 s at an unexpected location (lower half of the screen) and randomly within a 30-s window. Participants were required to respond to the target as quickly as possible by pressing a button held in the hand that was not holding the phone. The button for reaction (Kokuyo ELA-FP1) was connected wirelessly to a laptop computer (Lenovo Thinkpad X1 carbon) that controlled the stimulus presentation. Every reaction time was recorded on the computer.

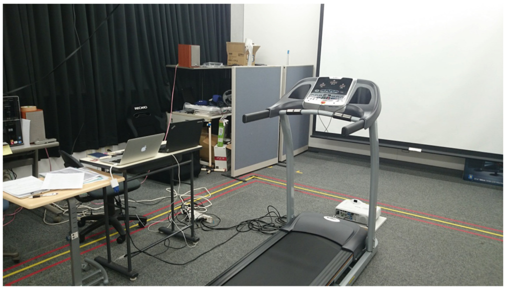


**Fig. 2.** The treadmill and screen used in the experiment. A projector hung on the ceiling projected the motion picture and a second projector placed on the floor in front of the treadmill displayed the visual targets.

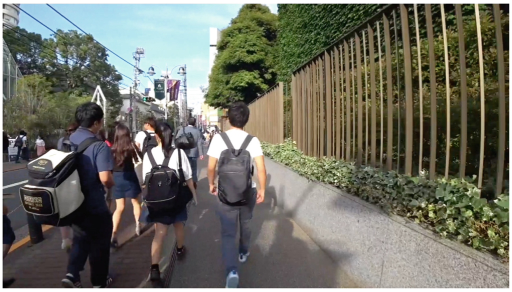


**Fig. 3.** Scene in the movie projected on the screen.

### 2.3 Smartphone Tasks

The participants performed the detection task while walking on the treadmill under the following three smartphone use/non-use conditions:

(1) control condition, participants only held the phone (iPhone SE) in one hand.

(2) Twitter condition, participants read a designated part of past timeline on an account that had been created by the authors. In order to make sure that participants read all the "tweets" in the timeline, they were asked about the contents of the tweets after the trial. In addition, participants were warned before the trial that they could be tested regarding the content of the tweets.
(3) LINE condition, participants were given a LINE account created by the authors and chatted with the experimenter, one of the authors, who sent the participants simple questions one after another. Approximately 15 Q&A's were exchanged throughout the experiment.

### 2.4 Workload Ratings

After the 3-min trial, the participants rated the subjective workload of the task with the Japanese version of NASA-TLX [13]. As with the original NASA-TLX [14], the rating scale consisted of 6 subscales: mental demand, physical demand, temporal demand, participant's own performance, effort, and frustration. Workload scores were calculated by averaging the ratings from 0 to 10 on a visual analogue scale for the total of the 6 subscales. It was reported that the average rating highly correlated with the formal weighted workload score using paired comparison of subscales according to the specific importance of the task [15].

### 2.5 Procedure

After giving informed consent for participation in the experiment and performing practice trials of the detection task and smartphone tasks, the participants performed three trials under each experimental condition: control, Twitter, and LINE conditions in a random order. The participants rated the workload after each trial, then rested for 3 min before the next trial. Figure 4 shows how the experiment has carried out.

**Fig. 4.** Using a smartphone while walking on the treadmill in the laboratory.

## 3 Results

A missed target was declared when the participant did not respond to the target within 4 s before it disappeared. No target was missed under the control condition. As shown in Fig. 5, mean number of missed target was significantly greater under the Twitter and LINE conditions than under the control condition. A statistical test using one-way analysis of variance (ANOVA) showed that the difference between the conditions was significant ($F(2, 24) = 17.24$, $p < .001$) and post-hoc analysis demonstrated that the differences between the Twitter condition and the control condition and between the LINE condition and the control condition were significant ($p < .05$ and $p < .001$ respectively).

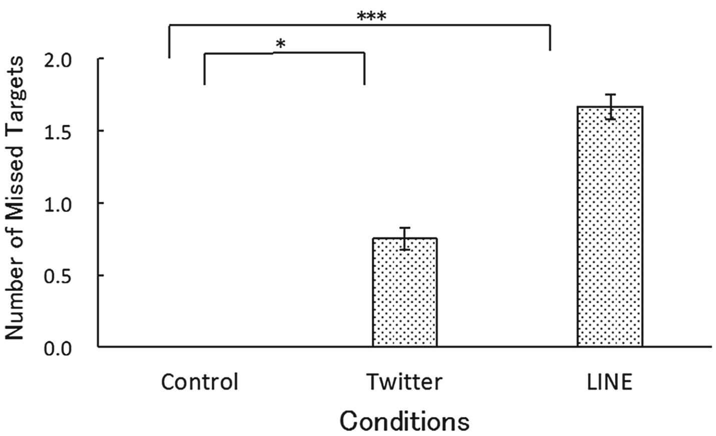



**Fig. 5.** Mean number of missed targets. No target was missed under the control condition. Error bars represent standard errors. $^{***}p < .001$, $^{*}p < .05$

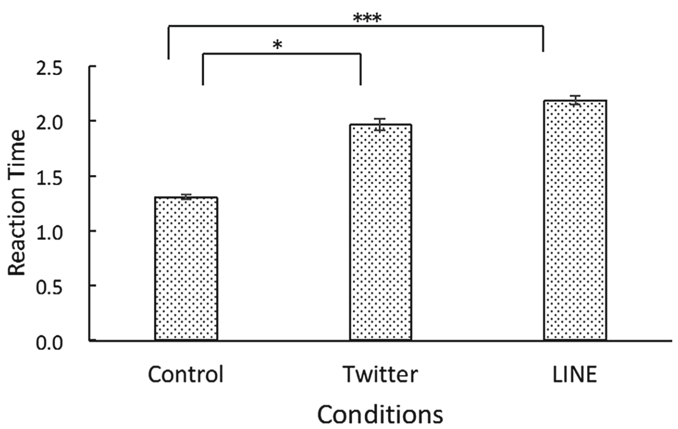



**Fig. 6.** Mean reaction time to the visual targets under three conditions. Error bars represent standard errors. $^{***}p < .001$, $^{*}p < .05$

Reaction times to targets detected within 4 s were averaged for each trial, then were compared between the conditions (Fig. 6). One-way ANOVA showed a significant difference ($F(2, 24) = 10.00$, $p < .01$). Post-hoc analysis demonstrated that the participants responded to the targets more quickly under the control condition than under the Twitter condition ($p < .05$) and LINE condition ($p < .001$). However, they reacted less quickly under the LINE condition than under the Twitter condition.

Workload ratings on the Japanese version of NASA-TLX were significantly different between the conditions ($F(2, 24) = 54.39$, $p < .001$) (Fig. 7). Post-hoc pairwise comparisons demonstrated that workload ratings were higher under the Twitter and LINE conditions than under the control condition ($p < .001$).

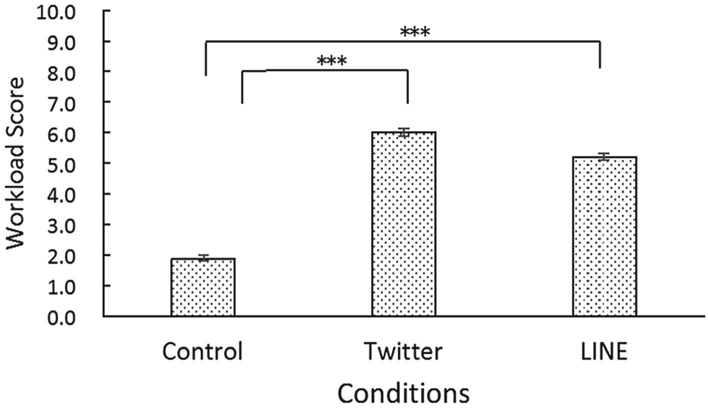



**Fig. 7.** Workload ratings for the experimental tasks under the three experimental conditions. Error bars represent standard errors. $^{***}p < .001$

## 4 Discussion

The results demonstrated that the participants detected fewer targets and reacted more slowly to the targets and that the workload of the experimental tasks was greater when they were operating a smartphone while walking on the treadmill. Performance was worse under the LINE condition than under the Twitter condition. This is understandable because the participants under the LINE condition had to text answers in response to questions given by the experimenter, while they only read texts under the Twitter condition although they rated the workload higher for the Twitter condition. These results suggest that visual attention of pedestrians is deteriorated through the use of a smartphone for SNSs while walking on the street.

Further research should be directed toward effective measures to regulate or discourage the use of smartphones while walking.

## References

1. McKnight, A., McKnight, A.: The effect of cellular phone use upon driver attention. Accid. Anal. Prev. **25**, 259–265 (1993)
2. Strayer, D.L., Johnston, W.A.: Driven to distraction: dual-task studies of simulated driving & conversing on cellular telephone. Psychol. Sci. **12**, 462–466 (2001)
3. McCarley, J.S., Vais, M., Pringle, H., Kramer, A.F., Irwin, D.E., Strayer, J.: Conversation disrupts change detection in complex driving scenes. Hum. Factors **46**, 424–436 (2004)
4. Beebe, K.E., Kass, S.J.: Engrossed in conversation: the impact of cell phones on simulated driving performance. Accid. Anal. Prev. **38**, 415–421 (2006)
5. Dula, C.S., Martin, B.A., Fox, R.T., Leonard, R.L.: Differing types of cellular phone conversations and dangerous driving. Accid. Anal. Prev. **43**, 187–193 (2011)
6. Hartfield, J., Murphy, S.: The effects of mobile phone use on pedestrian crossing behavior at signalized and unsignalised intersections. Accid. Anal. Prev. **39**, 197–205 (2007)
7. Nasar, J., Hecht, P., Wener, R.: Mobile telephones, distracted attention, and pedestrian safety. Accid. Anal. Prev. **40**, 69–75 (2008)
8. Hyman Jr., I.E., Boss, S.M., Wise, B.M., Mckenzie, K.E., Caggiano, J.M.: Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. Appl. Cogn. Psychol. **24**, 597–607 (2010)
9. Masuda, K., Sekine, Y., Sato, H., Haga, S.: Laboratory experiment on visual and auditory inattention of pedestrians using cell phones. In: The 28th International Congress of Applied Psychology, ICAP 2014, Paris, France (2014)
10. Haga, S., Sano, A., Sekine, Y., Sato, H., Yamaguchi, S., Masuda, K.: Effects of using a smart phone on pedestrians' attention and walking. Proc. Manuf. **3**, 2574–2580 (2015). The 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, Las Vegas, NV, USA. https://doi.org/10.1016/j.promfg.2015.07.564
11. Haga, S., Fukuzawa, K., Kido, E., Sudo, Y., Yoshida, A.: Effects on auditory attention and walking while texting with a smartphone and walking on stairs. In: Proceedings Part 1, 18th International Conference, HCI International 2016, Toronto, Canada, p. 186 (2016)
12. Haga, S.: Effects of smartphone use while walking on pedestrian's attention: a laboratory experiment using a treadmill. In: Proceedings for the 81st Conference of Japanese Psychological Association, Kurume, Japan (2017). (in Japanese)
13. Haga, S., Mizukami, N.: Japanese version of NASA task load index: sensitivity of its workload score to difficulty of three different laboratory tasks. Jpn. J. Ergon. **32**, 71–80 (1996). (Japanese with English abstract)
14. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index); results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183, North Holland (1988)
15. Miyake, S., Kumashiro, M.: Subjective mental workload assessment technique: an introduction to NASA-TLX and SWAT and a proposal of simple scoring methods. Jpn. J. Ergon. **29**, 399–408 (1993). (Japanese with English abstract)

# Biomechanical Load Evaluation by Means of Wearable Devices in Industrial Environments: An Inertial Motion Capture System and sEMG Based Protocol

Maria Grazia Lourdes Monaco[1(✉)], Agnese Marchesi[2], Alessandro Greco[3], Lorenzo Fiori[2], Alessio Silvetti[2], Francesco Caputo[3], Nadia Miraglia[4], and Francesco Draicchio[2]

[1] University Hospital of Verona, Verona, Italy
mariagrazialourdes.monaco@aovr.veneto.it
[2] INAIL, Monteporzio Catone, RM, Italy
[3] Department of Engineering, University of Campania L. Vanvitelli, Aversa, CE, Italy
[4] Experimental Medicine Department, University of Campania L. Vanvitelli, Naples, Italy

**Abstract.** Biomechanical overload is one of the main risk factors for musculoskeletal disorders among manufacturing workers and so far, it has been evaluated with observational methods. The aim of this research was to introduce a procedure for quantitative biomechanical overload risk assessment in which surface electromyography integrates with a motion capture system. The paper deals with actual test cases performed in an automotive company, using surface electromyography and a homemade inertial motion capture system. The quality of the data produced by it demonstrates that these devices can be integrated, worn in actual working conditions and are not influenced by electromagnetic interference.

**Keywords:** Wearable devices · Biomechanical load · Industrial ergonomics IMU · sEMG

## 1 Introduction

Biomechanical overload represents one of the main risks in the industrial environment and a possible source of musculoskeletal disorders and diseases (MSDs). Prevention of MSDs is a cornerstone of ergonomics. Even if the mechanical load on the human body in working life is not the exclusive causative factor, it is likely to constitute a major part of it. According to several studies, posture, range of motion, force and repetition must be considered in order to estimate risk exposure [1].

Specific attention has been given to working postures and movements: they are two important mechanical variables and load determinants.

Posture is the position in which someone holds his or her body when standing or sitting: it is influenced by the task, the workstation, the design of any work tools and


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 233–242, 2019.
https://doi.org/10.1007/978-3-319-94619-1_23

the anthropometric characteristics of the workers. Awkward posture refers to positions of the body (limbs, joints, back) that deviate significantly from the neutral position while job tasks are being performed. Posture is an important risk factor for the outbreak of low back pain [2–4]. Several laboratory studies have associated the possibility of the occurrence of low back pain with industrial work activities [5–8] while diverse epidemiologic research has shown the association between low back pain and awkward posture among automotive workers [9, 10].

Traditionally, there have been many subjective assessments on working postures and movements, using various observational protocols and checklists such as the Ovako Working Postures Assessment System (OWAS), the Rapid Upper Limb Assessment (RULA), and the Rapid Entire Body Assessment (REBA). Specific working methods, e.g. the European Assembly Worksheet (EAWS), have been developed in some work contexts, such as the automotive industry [11].

All of these assessment tools use on-the-job observation or video recordings to categorize the ranges within which each body segment falls, with obvious limitations in the characterization of physical exposure: subjectivity, rater bias, low precision, lengthy analysis periods and the requirement for highly trained observers. Their internal and external validity has also been questioned.

On the other hand, studies that use quantitative biomechanical measures could be more precise and reliable. An advantage of direct measurements is that they provide detailed and accurate values for jobs with varied work tasks.

Marras et al. first conducted a study in which they registered the posture taken from the trunk during work operations in an industrial environment, with the aid of real-time 3D monitoring systems [12]. The movements were analyzed using the Lumbar Motion Monitor (LMM), a sort of exoskeleton positioned along the vertebral column that accurately records and registers data on speed, acceleration, and position of the trunk. This study detected a relation between awkward trunk working posture and the onset of low back pain; a limit of this approach was the bulky tool used, which actually could hinder workers' movements. Another study of Norman et al. [13] highlighted the association between low back pain and trunk bending in the automotive industry, with a limit due to postural analysis that was based on 2D images from video.

Muscle activity performed during different occupational tasks is often evaluated by surface electromyography (sEMG) and presented as a percentage of maximal measured activity (% MVC). Surface electromyography is considered one of the most important tools for quantitative evaluation of biomechanical overload and provides possibilities of obtaining information, highly relevant from several ergonomic perspectives [14].

Recently, motion capture techniques to acquire human body movements have broadened their range of applications from sports training to rehabilitation, from animated graphics to ergonomic issues. For industrial applications, motion capture systems are largely used to acquire workers' movements, in order to assess ergonomic indexes and to improve working conditions, comfort and safety: motion capture systems composed of wearable inertial sensors represent the best solution for ergonomic application in manufacturing, being only slightly invasive and not an obstacle to workers' activities. Several researchers have introduced inertial measurement units (IMU) in order to evaluate body motion in actual working conditions [15].

The aim of the study is to introduce a new procedure for a quantitative risk assessment of biomechanical overload using an innovative wearable inertial motion-capture system together with surface electromyography.

# 2 Materials and Methods

## 2.1 Study Design, Setting, and Subject

The proposed protocol was tested during the "central tunnel cabinet assembly activity" performed in a real-work environment, in collaboration with Fiat Chrysler Automobiles Italy S.p.A. The electromyographic signals and trunk flexion-extension angles of two workers were registered during a single standard working task. The assembly task can be performed on both the left and right side of a car. Therefore, seven trials were collected from the left side and seven trials from the right side in order to evaluate the muscle activity and the posture angle trends under the two different task configurations. Several video shots were made at the same time to study the task.

**Working Activity Description.** The Fig. 1 shows the main work phases performed by a worker on the right side (RS) and left side (LS) of the automobile in the analyzed workstation where the central cabinet is assembled inside the car using screws, plugs, and cables. The cycle time lasts about one minute (58 s).

## 2.2 Procedure and Data Acquisition

### 2.2.1 Body Motion Study

A homemade inertial motion capture system has been used in this study. The system, developed by the research teams of Machine Design and Flight Control of the Department of Engineering of the University of Campania Luigi Vanvitelli, is composed of multiple micro inertial measurement units (IMU) [18]. A Kalman filter is used to compute the estimation of the attitude for each IMU, by combining a series of measurements affected by noise and other disturbance.

In order to include the investigated body segments (trunk, arms, forearm) together they will be referred to as "upper body". Considering the lower limbs in a steady state, the bones attitude estimation allows computing the whole upper-body posture. Each segment can be equipped with an IMU, composed of a tri-axial accelerometer, a tri-axial gyroscope, and a tri-axial magnetometer and used to estimate the orientation w.r.t. a fixed frame.

The system, in upper-body configuration, is composed of two independent modules. Each module is composed of 4 IMUs, positioned on the trunk, the arm, and the forearm. Data are recorded and pre-processed by a raspberry, powered by a battery. A camera, positioned on the goggles and synchronized with the raspberry, records the entire activity from the worker's field of view (Fig. 2).

An algorithm has been developed in order to provide attitude quaternions, attitude Euler angles per each IMU and, combining these data, the posture angle trends over the time considered: pelvis rigid rotation; flexion forward, lateral flexion and torsion of the trunk; elevation, lateral flexion and rotation of the arm; flexion and rotation of the

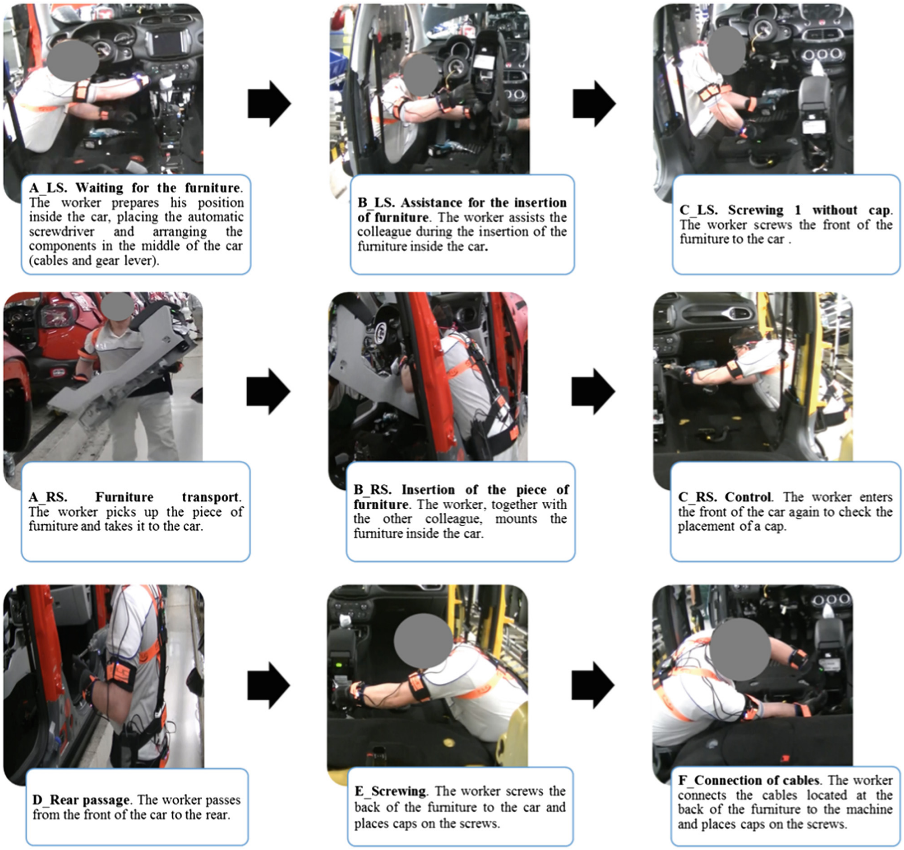



**Fig. 1.** Working activity

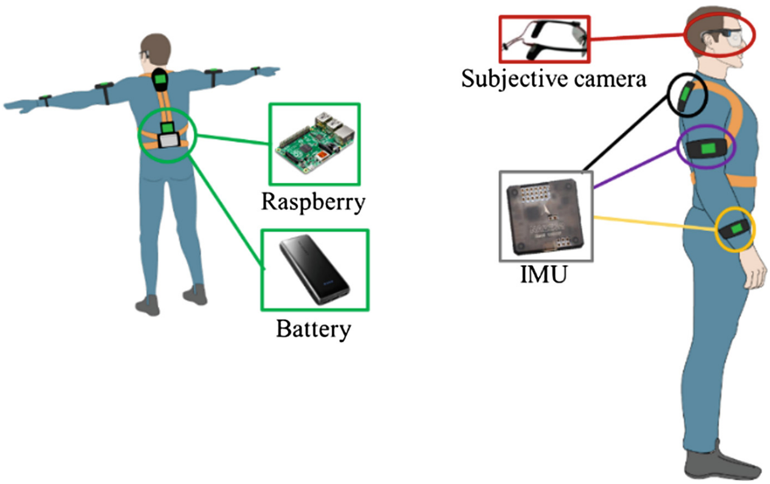



**Fig. 2.** Motion tracking system in upper-body configuration.

forearm. For this research, only the posture angles of the trunk, in particular, the forward flexion, have been investigated.

**Surface Electromyography Recording.** The enrolled subjects were studied using an 8-channel Wi-Fi transmission surface electromyograph (FreeEMG1000 System, BTS, Italy), with a sample rate of 1 kHz. The sEMG activity of trunk muscles was detected. After skin cleaning, bipolar surface electrodes Ag/AgCl (F9079, FIAB, Italy), prepared with electro-conductive gel (diameter 1 cm, distance between the electrodes 2 cm), were placed bilaterally along the direction of the muscle fibers of the Erector Spinae Thoracic region (EST), Erector Spinae Lumbar region (ESL) and Multifidus (M) (Fig. 3) according to the Atlas of muscle innervation zones of Barbero et al. [16].

Before starting the activity, each worker performed isometric maximal voluntary contractions (iMVC) of the muscles to determine the maximum level of muscle activation, which were used as reference during the signal processing phase. Each isometric contraction was repeated three times after a rest period of 3 min between trials [17]. To perform the iMVC, subjects were instructed to increase the force exerted from zero to their maximum and to hold it for about 5 s.

The acquired sEMG signals were processed using Analyzer software (Smart Analyzer, BTS, Milan, Italy) and MATLAB R2016b (The MathWorks, Inc., Natick, Massachusetts, United States). The sEMG signals were processed as follows: the raw iMVC and trials data were band-pass filtered (30–450 Hz), rectified and low-pass filtered (10 Hz). The processed signals of the working tasks were then normalized to a mean of the maximum values of the processed iMVCs. Amplitude parameters of these processed sEMG signals were then calculated in order to have a quantitative assessment of the muscle commitment during the working activity. Some amplitude parameters such as Average Rectified value (ARV), maximum value (MAX) and Root Mean Square (RMS) were evaluated within each registered trial.

The sEMG signals were obtained synchronously with the inertial sensor signals: the data, however, were separately processed. All the statistical analyses were performed using SPSS 17.0 software (SPSS Inc., Chicago, IL, USA).

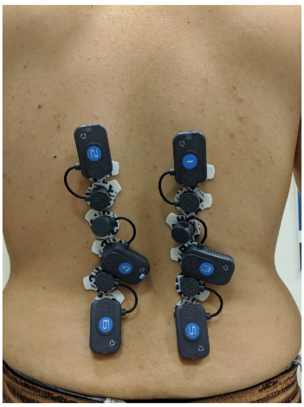


**Fig. 3.** Surface electromyography (sEMG) electrodes position.

## 3 Results

**Kinematic Recordings.** Figures 4 and 5 show four examples of kinematic measurements on one of the two subjects enrolled during the assembly activity analyzed. In particular, in each of the four graphs, the trend trunk flexion-extension angle is shown, normalized over time with respect to the duration of the cycle (Duty Cycle %), for each of work sub-phases (see Sect. 2.2.1). Measurements were conducted during the activities performed on the right side (RS) and on the left side of the vehicle (LS).

**Left Side Car Workstation (LS).** On the left side, the analysis of the first half of the work cycle showed a good reproducibility of the cycle, with some variations in the first phase (A_LS) due to the variability in the preparation of the workstation. As can be seen in Fig. 4, the maximum value of anterior flexion is approximately 60°, achieved in almost all the phases.

In B_LS the trunk flexion-extension has the same trend of increase, which leads, towards the final part of the cycle, to a trunk flexion of about 60°, as seen in both cycles shown in the figure, to then have a small and rapid decrease in the values of the angle up to about 50° (52° for 1_LS and 46° for 2_LS).

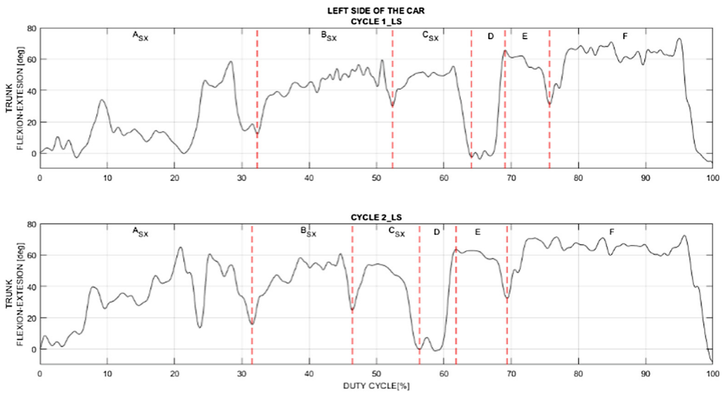



**Fig. 4.** Work subphases end their names are shown. The two graphs refer to two cycles of the activity on the left side of the car.

In C_LS, trunk flexion-extension shows an initial increase in the angle value, to then decrease sharply until it reaches 0°. Phases B and C are those in which a greater staticity is observed with average of flexion (m ± DS) of 53 ± 4°, 49 ± 4° and 69 ± 3°, respectively.

**Right Side Car Workstation (RS).** On the right side of the car, the work activities are associated, in the first phase (A_RS), with minimal postural variations of the trunk, having flexions that do not exceed 40°. The postural engagement instead increases in the next phase (B_RS) when the trunk is almost always bending around 45°, with small oscillations, and then returning towards 15–20° at the end of the phase. This is the most static phase. In the third phase, C_RS of the two cycles 1_RS and 2_RS, the flexion-extension of the trunk presents an initial increase in the angle value, passing from 14° to arrive at 59° for 1_RS and from 21° to 53° for 2_RS. This increase was followed by a decrease in the value of the trunk flexion angle, reaching 35° for 1_RS and 42° for 2_RS.

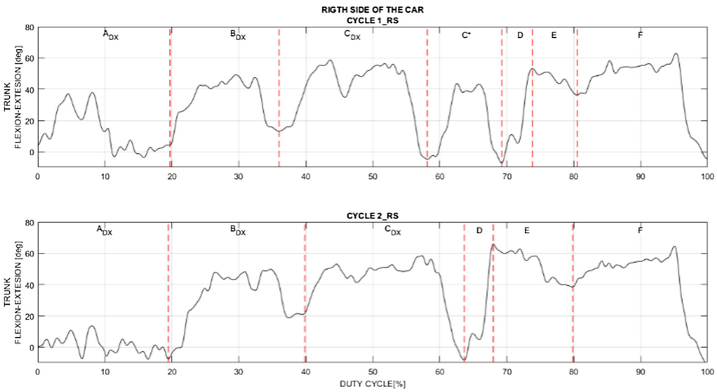



**Fig. 5.** Work subphases end their names are shown. The two graphs refer to two cycles of the activity on the right side of the car.

Finally, in the final part of the work phase, there is a further increase in the value of the angle followed by a stabilization and then a decrease in value, as in the B_DX cycle, up to −4° for 1_RS and −9° for 2_RS. In cycle 1_RS there is a further working phase, called C*, in which the flexion-extension of the worker's trunk shows a rapid increase of the angle from −4° to 44°, which is maintained for less than 4% of the cycle, and a subsequent decrease of the value from 43° to about −6°.

Finally, in phases D, E and F of all four cycles (1_LS, 2_LS, 1_RS and 2_RS) the flexion-extension angles of the worker's trunk has the same trend. In the fourth phase D, the flexion-extension angle of the trunk has a pronounced increase for the entire duration of the phase, with variable values that reach as much as 74° (in cycle 2_RS). Instead, in the fifth phase E the flexion-extension of the trunk has a trend opposite to phase D, that is, its value tends to decrease, but more gradually. Finally, in the sixth phase F the trunk flexion-extension trend shows an initial increase in the angle value, a more pronounced phenomenon in the cycles w.r.t the left of the car (from 31° to 67° for 1_LS and from 33° to 71° for 2_LS). After that, this value of the flexion-extension angle of the trunk is maintained for almost the entire subphase and ends with a small increment, followed by a rapid decrease until to 0°. The F phase is characterized by the static nature of the postures with angles that exceed 60°, working on both sides of the car.

**Surface Electromyography.** Figure 6 shows an example of the activity recorded from the Erector Spinae Thoracic region during the manufacturing phases carried out on the left side of the car assembly line. The graphs show the linear envelope of the sEMG signal normalized in amplitude (% MVC) and over time (Duty Cycle %), Work subphases end their names are shown in Fig. 6.

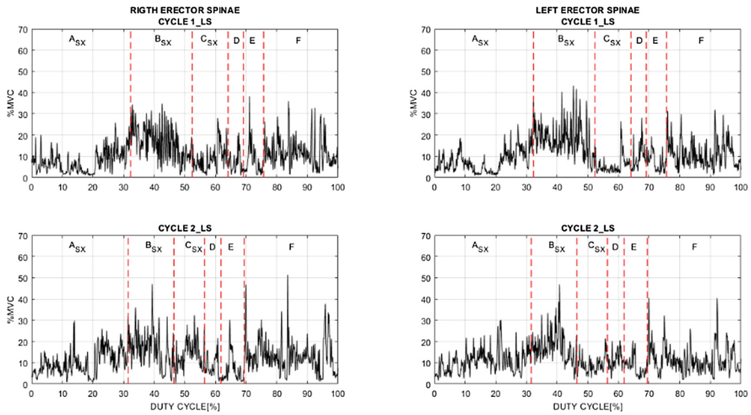



**Fig. 6.** sEMG signals of the Erector Spine thoracic region during the activity on the left side of the car, normalized in amplitude (% MCV): above the 1_LS cycle and below the 2_LS cycle. The red dashed lines divide the phases.

In the two work cycles 1_LS and 2_LS, in both the muscles, muscular activity in the A_SX phase shows an initial low value, then increasing at the end of the phase. The increase of the muscular activity is maintained even for half of the second work phase B_SX, in both cycles and muscles and the value decreases toward the end of the phase. The muscular activity in this second work phase, B_SX is greater in the left muscle.

In the third work subphase, C_SX, the trend of the muscular activity returns to increase in the right muscle, in both cycles 1_LS and 2_LS, then decreasing at the end of the phase. In the muscle of the left muscular activity remains low and constant for the entire duration of the subphase, with the exception of a pronounced peak of activity at the end of cycle 1_LS. In the subphase D, the muscular activity in both cycles of Fig. 6 has a slight decrease on the right, where the RMS values are of about 8% in 1_LS and 10% in 2_LS, and a slight increase on the left, where RMS values are 11% in 1_LS and 12% in 2_LS.

In the subphase E, muscle activity on the left has a slight decrease in both cycles, with RMS values of 9%. In the right muscles, the muscular activity remains low with a peak in the middle in both cycles showed in Fig. 6.

Finally, in subphase F, muscular activity has a growth in both cycles and muscles. In the right muscle, RMS values are of about 13% in 1_LS and 15% in 2_LS, while in left muscles are of about 12% in 1_LS and 13% in 2_LS.

## 4 Discussion and Conclusions

The aim of this study was the development of an integrated sEMG-IMU protocol for posture evaluation during work activities in an automotive environment. A few outcomes of the measurements made with electromyography and inertial sensors are reported. We observed a good correspondence between kinematic and electromyographic data. Both signals could be integrated into a wider evaluation of the biomechanical load due to awkward postures with or without manual load handling. Furthermore, sEMG together with kinematics can contribute to better analyze both static and dynamic postures.

Several research groups used these integrated approaches with the combination of sEMG and IMU sensors to assess biomechanical overload. The novelty of the proposed protocol is that it was tested in a real work environment rather than in laboratory settings [19].

The results of this study confirm the chance of using surface electromyography in conjunction with inertial sensors, overcoming the presence of potential electromagnetic interference that worsens the quality of the signals. Moreover, this pilot study demonstrates that the use of these technologies does not hinder the execution of the work activity in a production line, which requires speed and accuracy in performing movements in a rather short cycle (<1 min). The quality and reproducibility of the recorded signals allow a good comparison of the same subject at different times. The possibility of objectivizing and customizing the biomechanical risk assessment due to awkward postures leads to a more accurate management of occupational risks in assembly lines.

This study has some limitations such as technical characteristics of used inertial sensors that don't allow to investigate trunk torsion. The presence of this postural component was however evidenced by the electromyographic data analysis showing different activity patterns on the left and right side. The corrective measures, already made to the suite, will permit a more detailed analysis of motion patterns in the three space planes. Furthermore, the small sample size affects the full generalization of the results.

Further research is needed for a complete validation of the suite to create a predictive model of force and posture and to identify exposure limits. A modified version of the proposed technique could be implemented in order to perform muscle fatigue analysis [20–22], Exposure Variation Analysis [23] and Joint Analysis of Spectrum and Amplitude [24, 25].

## References


1. Bernard, P.B.: Musculoskeletal Disorders and Workplace Factors: A Critical Review of Epidemiologic Evidence for Work-Related Musculoskeletal Disorders of Neck, Upper Extremity, and Low Back. National Institute for Occupational Safety and Health, Cincinnati (1997)
2. Armstrong, T.J., Radwin, R.G., Hansen, D.J., Kennedy, K.W.: Repetitive trauma disorders: job evaluation and design. Hum. Factors **28**(3), 325–336 (1986)
3. Garg, A., Moore, J.S.: Epidemiology of low-back pain in industry. Occup. Med. **7**(4), 593–608 (1992)
4. Bazzini, G., Capodaglio, E., Panigazzi, M., Prestifilippo, E., Vercesi, C.: Rischi da posture incongrue. G. Ital. Med. Lav. Erg. **32**(3), 215–222 (2010)
5. Rowe, M.L.: Low back disability in industry: updated position. J. Occup. Med. **13**(10), 476–478 (1971)
6. Kelsey, J.L., White, A.A.: Epidemiology and impact of low-back pain. Spine (Phila. PA) **5** (2), 133–142 (1980)

7. Andersson, G.B.: Epidemiologic aspects on low-back pain in industry. Spine (Phila. PA) **6** (1), 53–60 (1981)
8. Pope, M.H.: Risk indicators in low back pain. Ann. Med. **21**(5), 387–392 (1989)
9. Punnett, L., Fine, L.J., Keyserling, W.M., Herrin, G.D., Chaffin, D.B.: Back disorders and non-neutral trunk postures of automobile assembly workers. Scand. J. Work Environ. Health **17**(5), 337–346 (1991)
10. Vandergrift, J.L., Gold, J.E., Hanlon, A., Punnett, L.: Physical and psychosocial ergonomic risk factors for low back pain in automobile manufacturing workers. Occup. Environ. Med. **69**(1), 29–34 (2012)
11. Schaub, K., Caragnano, G., Britzke, B., Bruder, R.: The European assembly worksheet. Theoret. Issues Ergon. Sci. **14**(6), 616–639 (2013)
12. Marras, W.S., Lavender, S.A., Leurgans, S.E., Rajulu, S.L., Allread, W.G., Fathallah, F.A., Ferguson, S.A.: The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. The effects of workplace factors, trunk position, and trunk motion characteristics on risk of injury. Spine (Phila. PA) **18**(5), 617–628 (1993)
13. Norman, R., Wells, R., Neumann, P., Frank, J., Shannon, H., Kerr, M.: A comparison of peak vs cumulative physical work exposure risk factors for the reporting of low back pain in the automotive industry. Clin. Biomech. (Bristol Avon) **13**(8), 561–573 (1998)
14. Merletti, R., Farina, D.: Surface Electromyography: Physiology, Engineering and Applications. IEEE Press/Wiley, New York (2016)
15. Filippeschi, A., Schmitz, N., Miezal, M., Bleser, G., Ruffaldi, E., Stricker, D.: Survey of motion tracking methods based on inertial sensors: a focus on upper limb human motion. Sens. (Basel) **17**(6), 1 (2017)
16. Barbero, M., Merletti, R., Rainoldi, A.: Atlas of Muscle Innervation Zones. Understanding Surface Electromyography and Its Applications. Springer, Milan, New York (2012)
17. Merletti, R., Botter, A., Troiano, A., Merlo, E., Minetto, M.A.: Technology and instrumentation for detection and conditioning of the surface electromyographic signal: state of the art. Clin. Biomech. (Bristol Avon) **24**, 122–134 (2009)
18. Caputo, F., Greco, A., D'Amato, E., Notaro, I., Spada, S.: A preventive ergonomic approach based on virtual and immersive reality. In: Proceedings of the International Conference on Ergonomics in Design. Advances in Intelligent Systems and Computing, AHFE 2017, Los Angeles, CA, USA (2017)
19. Peppoloni, L., Filippeschi, A., Ruffaldi, E., Avizzano, C.A.: (WMSDs issue) A novel wearable system for the online assessment of risk for biomechanical load in repetitive efforts. Int. J. Ind. Ergon. **37**(6), 563–571 (2015)
20. Lindstrom, L., Magnusson, R., Petersen, I.: Muscular fatigue and action potential conduction velocity changes studied with frequency analysis of EMG signals. Electromyography **10**, 341–356 (1970)
21. Lindstrom, L., Magnusson, R.: Interpretation of myoelectric power spectra: a model and its applications. Proc. IEEE **65**, 653–662 (1977)
22. Lowery, M., Vaughan, C., Nolan, P., O'Malley, M.: Spectral compression of the electromyographic signal due to the decreasing muscle fiber conduction velocity. IEEE Trans. Rehabil. Eng. **8**, 353–361 (2000)
23. Fjellman-Wiklund, A., Grip, H., Andersson, H., Karlsson, J., Sundelin, G.: EMG trapezius muscle activity pattern in string players: part II-influences of basic body awareness therapy on the violin technique. Int. J. Ind. Ergon. **33**, 357–364 (2004)
24. Luttmann, A., Jager, M., Laurig, W.: Electromyographical indication of muscular fatigue in occupational field studies. Int. J. Ind. Ergon. **25**, 645–660 (2000)
25. Luttmann, A., Jager, M., Sokeland, J., Laurig, W.: Electromyographical study on surgeons in urology. II. Determination of muscular fatigue. Ergonomics **39**, 298–313 (1996)

# Analysis of Physical Feature in the Course Turn While Walking

Ryota Sakashita[1(✉)] and Hisaya Tanaka[2]

[1] Infomatics Major, Kogakuin University Guraduate School, Tokyo, Japan
j214057@ns.kogakui.ac.jp
[2] Kogakuin University, 1-24-2 Nishi-Shinjuku, Tokyo 163-8677, Japan
hisaya@cc.kogakuin.ac.jp

**Abstract.** In this study, we analyzed the physical rotary quantity of the yaw axis while subjects walk on running-machine, because it was considered that this rotary quantity may be available for discrimination of the right and left walk intention on the running-machine. Four subjects were attached a gyro sensor of Enliven3D to a head, a shoulder, a waist and a foot. Then, they walked two kilometers per hour on running-machine. When subjects walked on the running-machine, we requested them to intend changing direction and going straight with looking straight. The shoulder, waist and foot in all subjects turned about 20° under the influence of turn-intention in the changing direction. There was individual difference of about 10° in the rotary quantity between every subject and each part. Each part without heads turned under the effect of turn intention, even if we requested subjects to look at the right and the left.

**Keywords:** Locomotion Interface · Walking running-machin · Gyro sensor

## 1 Introduction

Recently, Virtual Reality (VR) technology is studied actively due to improve Computer technology. VR technology is available to process the reality information using computer simulation. However, anyone can't always use comfortably VR technology. Some VR contents user occur VR sickness that is similar to Symptoms of car sickness [1]. It is important to improve presence for suppressing VR sickness.

Locomotion Interface (LI) is the technology to improve presence in VR contents. There was some study that using motion capture system, running-machine and Kinect to develop LI. However, motion capture system is too expensive to use normally. Running-machine is almost cheaper than motion capture system. However, running-machine has the problem that the walker can't walk except straight. Therefore, to use running-machine for LI, it is important to make it possible to change walking-direction. Previous study analyzed foot rotary quantity and acceleration for analyzing walkers root [2]. Therefore, we consider that walker's body change direction caused by turn intention.

In this study, we analyzed the physical rotary quantity of the yaw axis while subjects walk on running-machine. To use this quantity for LI, walker that using running-machine may be able to change direction in VR environment.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 243–250, 2019.
https://doi.org/10.1007/978-3-319-94619-1_24

## 2 Method

To develop Locomotion Interface (LI) that is available to change walking direction in Virtual Reality environment with running-machine, It is necessary to know any quantity that each part of body. Therefore, we focused on body rotation while subjects walking. This experiment was carried out in the following the order.

1. Walk on circle on the floor both clockwise and counter-clockwise in order to unify walking image.
2. Mount gyro-sensor on subject's head, shoulder, waist, and both tiptoes.
3. Stand on running-machine and hold handrail of it.
4. In one trial, first, stay over 3 s, next walk straight with looking straight, finally, intention walking direction.

To research relationship between body rotation, looking-direction and intention of change-walking-direction, we analyzed walking of 3 patterns intention and 3 patterns looking direction (Fig. 1). Subjects who mount gyro-sensors stand ready to walk running-machine. When we give them a sign to walk straight, subjects walk straight and look front on running-machine 2 km per hour for five seconds. After five seconds, we give subjects a sign to intent change-walking-direction.

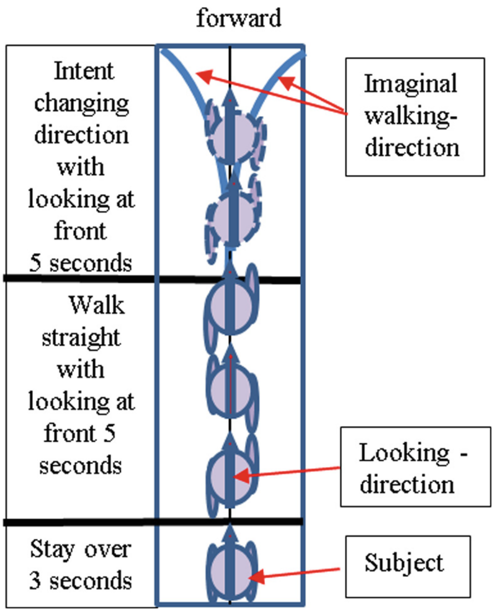



**Fig. 1.** Measurement flow

## 3 Measurement Environment

The measurement was done at Kogakuin University Shinjuku Campus Digital Contents Production Studio (Fig. 2). The gyro-sensor that used in measurement was Enliven-3D. Subjects wear the sensor on head, shoulder, waist, and both foot (Fig. 3). The running-machine that used in measurement was Running-Machin-R18 of IGUNIO Co. Subjects grab handle for safety while walking. Four subjects were twenties male.

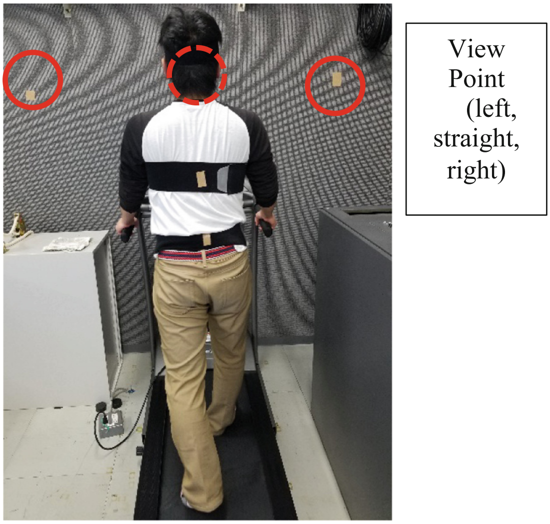



**Fig. 2.** Measurement environment

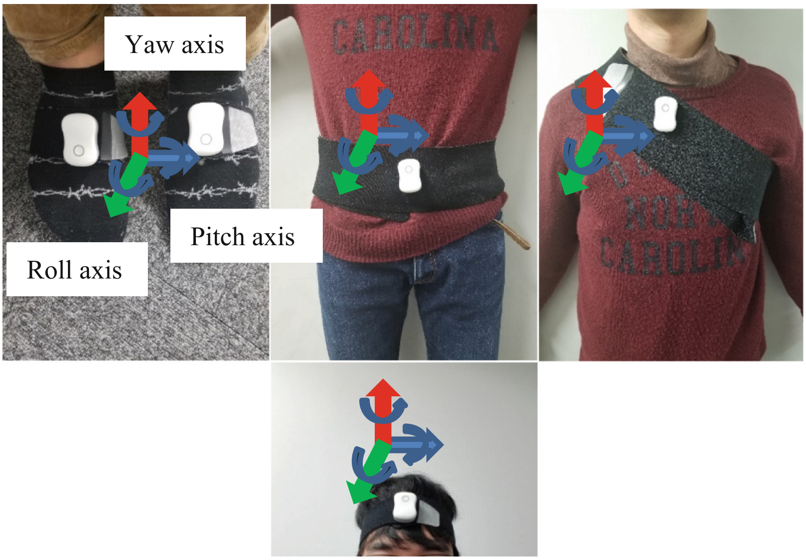



**Fig. 3.** Wearing sensor point. [Enliven 3D]

The sensor of Enliven-3D was produced by Enliven Japan Inc. [3]. It can measure 3 axis accelerations, 3 axis geomagnetic data, and 3 axis gyro data. The data measured by the sensor is sent to iPad by Bluetooth network. User can make simple 3D human model in iPad (Fig. 4). Connecting the sensor to iPad and adapting the senor that displayed in iPad and the part of human model, User can display the data that was sent by the sensor. In this measurement, sent the data of iPad to Computer, we analyzed only yaw axis gyro data.

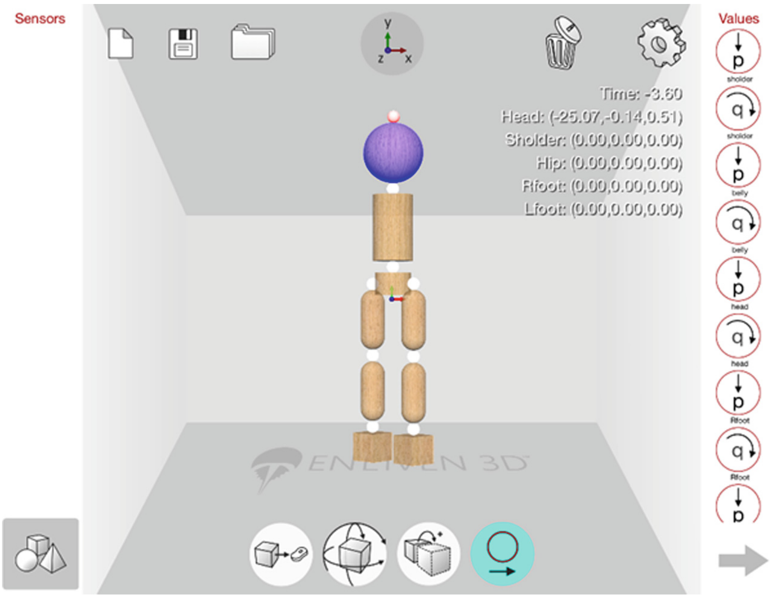



**Fig. 4.** Enliven 3D in iPad screen

## 4 Results

Figure 5 shows Result of head rotary quantity that one of subjects walking on running-machine (Fig. 5). Vertical axis shows measurement time. Horizontal axis shows head rotary quantity. If horizontal value is positive, subject's head rotate clockwise. If the value is negative, subject's head rotate counterclockwise. Blue line shows the rotary quantity when we gave subject a sign to intent change-walking-direction to left. Orange line shows the rotary quantity when we gave subject a sign to intent walking straight. Red line shows the rotary quantity when we gave subject a sign to intent change-walking-direction to right. Figure 6 shows result of shoulder rotary quantity. Figure shows result of waist rotary quantity (Fig. 6). Figure 7 shows result of waist rotary quantity (Fig. 7). Figure 8 shows result of left foot rotary quantity (Fig. 8) Fig. 9 shows result of right foot rotary quantity (Fig. 9). Each part rotated to left and right side in intent-time affected by a sign of intent change-walking-direction to left and right.

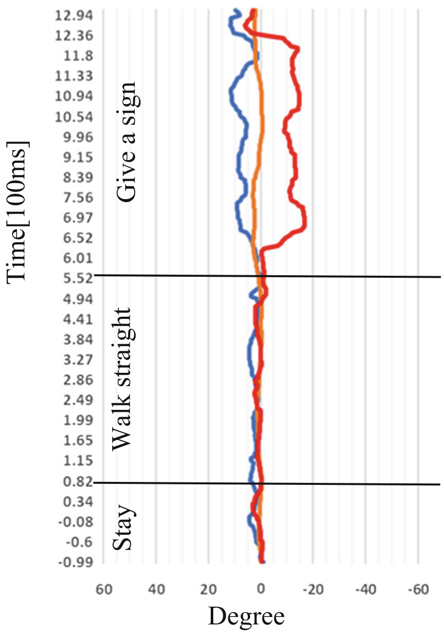



**Fig. 5.** Head rotary quantity

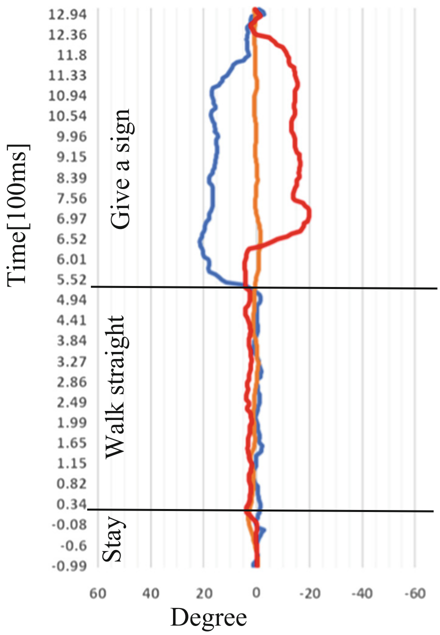



**Fig. 6.** Shoulder rotary quantity

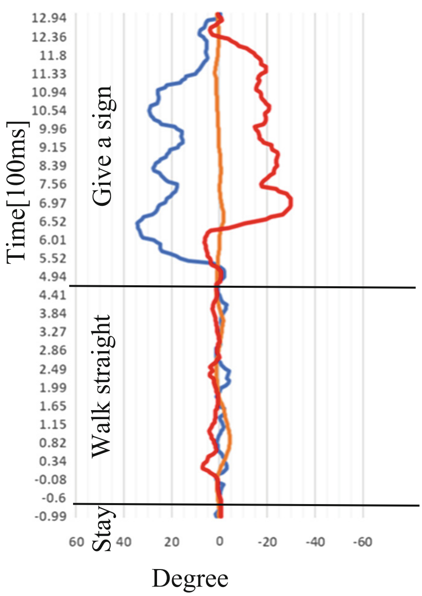



**Fig. 7.** Waist rotary quantity

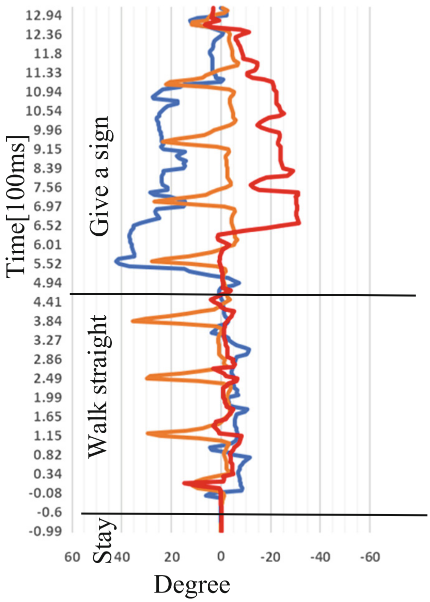



**Fig. 8.** Left foot rotary quantity

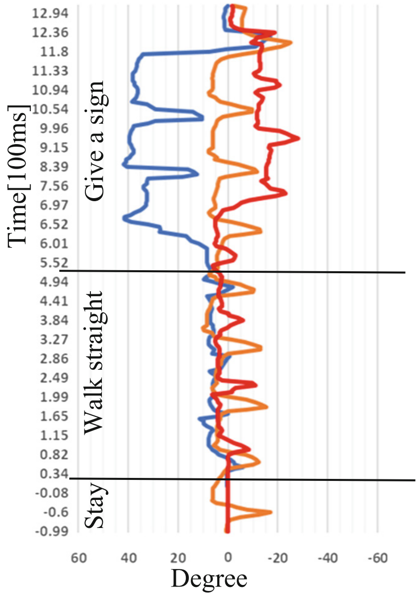



**Fig. 9.** Right foot rotary quantity

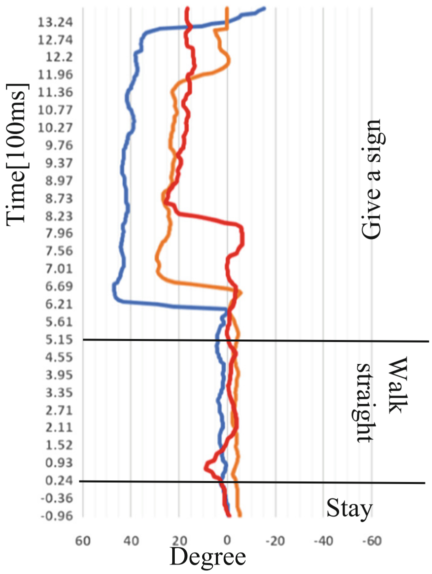



**Fig. 10.** Head rotary quantity (turn left)

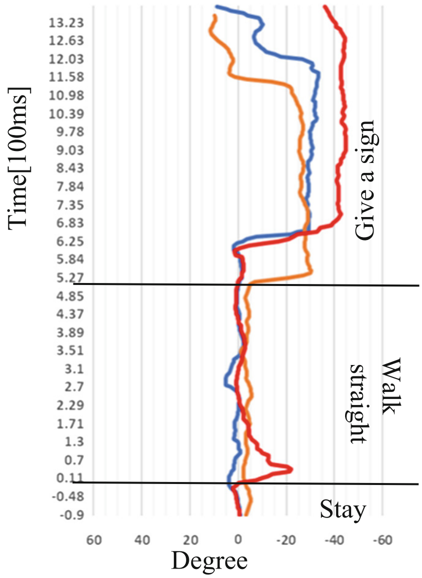



**Fig. 11.** Head rotary quantity (turn right)

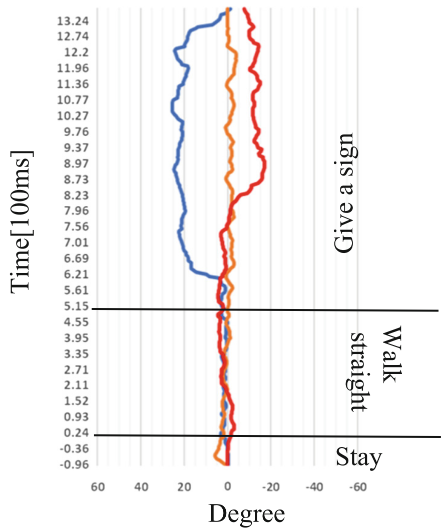



**Fig. 12.** Shoulder rotary quantity (turn left)

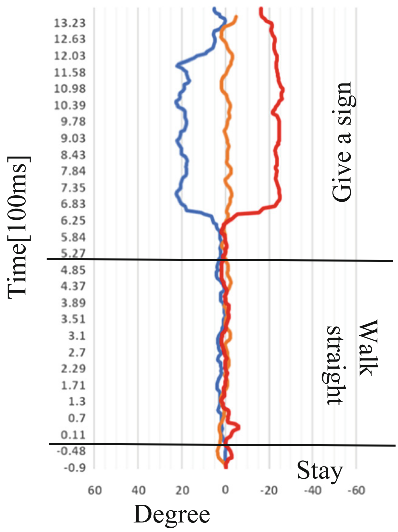



**Fig. 13.** Shoulder rotary quantity (turn right)

Figure 10 show subject's head rotate quantity when we give a sign to look left side (Fig. 10). Figure 11 show subject's head rotate quantity when we give a sign to look right side (Fig. 11). Subject's head rotated affected by a sign to look. The effect of a sign to intent change-walking-direction was little.

Figure 12 show subject's shoulder rotate quantity when we give a sign to look left side (Fig. 12). Figure 13 show subject's shoulder rotate quantity when we give a sign to look left side (Fig. 13). Shoulder rotary quantity was not affected by a sign to intent change-walking-direction.

## 5 Discussion

We analyze walker's rotary quantity while walking on running-machine. When subject walk on running-machine with looking forward, each body part was rotate effected by intention of changing direction. Subject's shoulder and waist was affected by a sign to intent change-walking-direction without being affected by looking direction. However, head rotary quantity was affected by only looking direction. Therefore, I consider that shoulder and waist rotary quantity are available to discriminate intention that walker intent to change direction. Because, their rotary quantity was not affected by looking direction. Head rotary quantity is available to discriminate looking direction. Therefore, if subject walk on running-machine that only walking straight, using shoulder rotary quantity, we can discriminate direction that subject intent for develop Locomotion Interface (LI). Previous study shows view information affect to walking direction [4]. When we develop LI using image-display, it is necessary to check rotary quantity.

## 6 Conclusion

In this study, we analyzed the physical rotary quantity of the yaw axis while subjects walk on running-machine. All subject's head rotary quantity was affected by only looking direction. Shoulder and waist rotary quantity was affected by a sign to intent change-walking-direction. These rotary quantities will be available to discriminate direction for Locomotion Interface that using running-machine.

## References

1. Tanaka, N.: A survey of countermeasure design for virtual reality sickness. Virtual Reality Soc. Jpn. **10**(1), 129–138 (2005)
2. Sagawa, K., et al.: Non-restricted measurement of three-dimensional walk trajectory by the integration of tiptoe acceleration. Soc. Instrum. Control Eng. **40**(6), 635–641 (2004)
3. Enliven Japan Inc.: Homepage. http://www.virtuix.com/. Accessed 09 Feb 2018
4. Yoshida, T., et al.: Guidance of human locomotion using vection induced optical flow information, IPSJ SIG Technical report (2006)

# Game Design for Learning and Training

# Experience from Indoor Fire Search and Rescue Game Design for Technology Testing

Jaziar Radianti(✉)

Centre for Integrated Emergency Management, Department of ICT,
University of Agder, Kristiansand, Norway
jaziar.radianti@uia.no

**Abstract.** No matter how good the fire evacuation plan for a building is, there is a risk of immovable victims being left behind due to smoke obscuration, inhalation of poisonous gasses, or other reasons. Firefighters usually perform the search and rescue for casualties in responding to fire hazards, besides the fire-fighting. Among of the challenges in the search and rescue operation is how to locate victims, and to keep monitoring the fire development so that it will not accidentally harm the fire personnel. This paper presents a game designed to test a smartphone app's feature that supports concurrent tracking of indoor victims and fire spread. 22 volunteers were assigned roles either as rescuers or victims. The game was organized in collaboration with the fire service personnel as observers together with the building security officers. The evaluation results are presented from both observers' and players' perspectives.

**Keywords:** Indoor fire · App testing · Search and rescue · Serious games

## 1 Introduction

Among the challenges faced by firefighters in the risk-taking mission of search and rescue (SAR) for victims in the indoor fire are unfamiliarity with the burning building and identification of victims' positions, coupled with keeping an eye on the fire development. Not to mention protecting their own lives. One among the technological supports for SAR in a burning building is infrared-based thermal imaging cameras that enable firefighters to see objects such as hidden fire and victims in low visibility, but indeed, there is no way to detect the existence of victims from a distance.

This paper reports the implementation of an Indoor Fire Search-and-Rescue Game (IFSG) created for testing an alternative smartphone app for locating victims and predicting fire development, which has been prototyped [1–3]. This app was created in the completed SmartRescue project, [2] that aimed at making use of advanced smartphone sensors to alert people in an indoor fire situation. The app can support firefighters to detect victims' locations both remotely or in the short distance as long as the victims have activated the app. Regardless of these app features, this paper rather focuses on how to design a game to demonstrate applicability and usefulness of the app


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 253–265, 2019.
https://doi.org/10.1007/978-3-319-94619-1_25

to a broader audience, and not at all to discuss its underlying technology. The indoor fire scenario was chosen to match the required app testing environment.

By all means, researchers have initiated and proposed to use Serious Game (SG) for crisis management as an approach for testing the increasing number of ICT (Information-Communication Technology) solutions for crisis management such as Meesters [4]. The author demonstrates the potential use of the SG approach for ICT-tool evaluation and a way to improve the practitioners' understanding on how these ICT-tool inventions can help responders, and ultimately will increase their future adoption.

At the core, IFSG was multidisciplinary-collaborative research. In fact, IFSG offered exciting, diverse opportunities for data collection and research standpoints—ranging from the organization, communication and collaboration, evacuation, workload, information exchange, usability, game design, user experience, user interface, and so on—making the planning phase of this game quite challenging and unique. Hence, inserting the multidisciplinary research, coupled with the experience for "packing" app prototype testing and evaluation into an SG design, can be regarded as the main contributions of this paper. Besides, this paper includes the brief evaluation of the game implementation and user experience to determine which part of the game was considered best, and what aspects need improvement.

This paper contains six sections. Section 2 presents literature review. Section 3 describes scenarios and tested technologies. Section 4 presents the game design and implementation. Section 5 comprises discussions and lessons learned from the case in this paper. Section 6 concludes this paper and reveals some limitations of this work.

## 2 Literature Review

The SG has, in fact, been used in different organizational contexts as a mean for training. The application areas are quite broad, ranging from the military training, organizational education, medical care emergency services, politics, business, and many other sectors of society. The "SGs" is used to denote the games used for non-entertainment purposes [5], or education [6]. Engaging and fun game(ful) experience is the heart of the game, but promoting learning and education, participation, behavioral improvement are among the characteristics of an SG [7]. As mentioned by Djaouti et al. [8], the first definition of SGs with a meaning that is closer with today's usage was introduced by Abt [9], although the authors have found an older definition of SGs dating back to 1950.

There is no agreement in the literature, how can we categorize an SG. This question has been a subject of discussion in the literature which mentions different terms referring to SG, definitions, and classifications [10–13]. However, there are two elements that most authors agree as parts of a SG's definition, i.e. the purposes, both for entertainment and learning or training. However, this does not entirely solve the disagreement as other authors also differentiate a "game for learning" from an SG [14], but the article does not explain further the differences between the two categories or even tends to treat "SG" as "game for learning". Also, there are tiny differences between the different games and game-like experiences, i.e. Game, SG, Simulation, Gamification and Game Inspired Design [15]. Lastly, it is not so trivial to classify a game, into different game genres simply because there is also no consensus how the game and the SG should be classified.

Susi et al. [10] try to sort it out and suggest that SGs can be grouped into Games (application focuses on simulation, learning, and fun) and *3D applications* that use 3D game technology to solve business problems. Still, Sousi et al. point out some overlapping in this category where there are 3D applications that are not games, or 3D applications that are games and games that are not 3D applications.

Marczewski [15] suggests a game design intent as a way to differentiate different game nuances, presented in Table 1. This table shows that a game category can be detected from various design intents, i.e. game thinking, game elements, virtual world, gameplay and non-purposeful. However, this framework intent-based framework only includes "virtual world" as a game environment, which is typically produced by 3D applications that are commonly found in video or digital games.

**Table 1.** Game and game-like experience

| | | Design intent | | | | |
|---|---|---|---|---|---|---|
| | | Game thinking | Game elements | Virtual/life-like world | Game play | Non-purposeful |
| Game and game-like experiences | Game inspired design | √ | | | | |
| | Gamification | √ | √ | | | |
| | Simulation | √ | √ | √ | | |
| | Serious game | √ | √ | √ | √ | |
| | Game | √ | √ | √ | √ | √ |

In sum, papers attempt to classify games, games for learning and SGs such as [11–13, 16] consistently link this term to video games or video toys[1] using various platforms, vary from video console, PC, online game, second life, mobile or alternate reality game. Thus, to be considered as a fully SG, the IFSG cannot fully satisfy all necessary elements of the design intent, as it includes a real instead of virtual environment.

Thus, while the division itself is useful, but it does not fully capture the SG design proposed in this article. If all gameplay requirements are fulfilled, but the design and application include a real building, real tools and devices, and real people – would it then be considered as a SG? Therefore, the life-like world is added to the virtual world as an alternative for defining the environment setting used for the SGs. It this paper, the life-like world is defined as the imitation of the operation of the real-world process, in non-virtual context representing the characteristics, behaviors, and functions of this selected environment.

[1] Exhaustive list of SGs can be found in http://serious.gameclassification.com/.

Djaouti et al. [11] propose the G/P/S or the Gameplay (G), Purpose (P) and Scope (S) model to classify the SG genres. The G/P/S model is very practical in that one can determine quickly whether a game falls into a specific type. The **gameplay (G)** has options *Game-based* or *Play-based*. The **purpose (P)** of the game can be message broadcasting (educative, informative, persuasive and subjective), training (mental and physical) or *data exchange*. The **scope (S)** contains market that consists of 13 categories, and the *public,* i.e. the general public, professionals and students. Thus, this model is useful for getting an overview of how each game is played and for what purpose it is designed. However, this framework is used for the practical purpose in education where a teacher can use it to quickly detect SGs that are relevant for teaching.

There is another known term related to the game approach, i.e. "gamification". Gamification refers to the use of game design elements in non-game contexts, to create an interactive system and environment that aim to motivate and engage end-users through the use of game elements and mechanics [7, 17, 18].

Conversely, referring to Table 1 framework on "Game" and "Game-Like Experience", the SG is closer to the approach has been adopted for IFSG. Most SG designs aim at improving the participants' skills [19] and behavior, for learning and educating [20] and informing. Many studies have linked the SGs with the digital games as we seen in the G/P/S model [11]. These types of SGs try to allocate information or education and offer a balanced possibility for authentic and playful learning. However, in this paper, the proposed SG application is a combination of role-playing tailored to the ICT technology testing environment for data collection. In fact, the SG ICT technology testing for disaster management is still rare [4]. This work will hopefully fill this gap and contribute towards further research in this area.

## 3 Scenario and Tested Technology

Recall that the indoor fire scenario was used in IFSG, and the university building was used as a game venue. A series of discussions about the local practice of SR in a case of fire with the fire emergency staff of the university and local firefighters were conducted before the scenario development. The IFSG focus was a hypothetical situation where the fire had grown, and several occupants were trapped inside.

The scenario was designed as closely as possible to the standard evacuation procedure in this university. In the scenario, we assumed that with 30% of maximum room occupancies in the game area, there are approximately 126 people daily during normal working hours. Hence, to assume that there would be 10–15 trapped victims were quite realistic.

The game was designed for testing an app that as a part of a project aims at studying the possible use of advanced sensors that have been embedded in most of the today's smartphones. Through the communication technologies, the data can be passed to others, especially among the rescue team members and responders for monitoring, tracking, and decision support. The main features of the app are:

- Fire assessment (to locate where the fire is and to see its development over time).
- Fire prediction (to predict the possible direction of the fire spread).
- Indoor localization (to find the exact location of the victims indoor).

The fire assessment and fire prediction features have been developed using Bayesian Network (BN) reasoning and K-nearest neighbor algorithm for indoor localization [2].

# 4 Game Design, Mechanics, and Implementation

## 4.1 Design and Mechanics

As IFSG was a game in nature, hence the fun and enjoyment elements, as well as goals, rules, challenge and interactivity, were undertaken as part of the design. The design would like to find a balance between "seriousness" and "learning" in one hand, and the "technology testing environment" on the other hand. It has game elements, i.e., components that can keep people engaged. There are many examples of and ways to define game elements [21, 22]. The following features: "challenge", "collaboration", "resources", "time", "reward or scoring" and "theme" were covered in IFSG. The game elements overlapped slightly with game mechanics, although game mechanics is more about rules and procedures that guide players to respond to the game challenges.

On the game learning and game mechanics, several models and theories have been proposed to show the interaction between game and learning patterns for a SG design. Game Object Model (GOM) [23] provides a theoretical basis for the design of educational games, while LM-GM [24] introduces the interplay between pedagogical patterns and game design patterns of the SGs.

**Table 2.** Learning and game mechanics in IFSG, using adapted LG-GM model

| ***Learning Mechanics*** | | | ***Game Mechanics*** | | |
|---|---|---|---|---|---|
| Instructional | Guidance | Action/Task | Cooperation | Role Play | Information |
| Explore | Participation | Feedback | Collecting | Collaboration | Story |
| Hypothesis | Observation | Incentive | Strategy/ Planning | Cascading Information | Feedback |
| Motivation | Identify | | Time Pressure | Resource Management | Simulate/ Response |
| | Plan | | Competition | Action Points | |
| | Experimentation | | | Movement | |
| | Reflect/ Discuss | | | | |
| | Responsibility | | | | |

For the description of the IFSG design, the author adopts the LM-GM model as it has several advantages. LM-GM considers the interactions between learning and entertainment as a core of SG mechanics. The model is flexible, allows users to "interpret" the linkage between learning and gaming mechanics to describe the SG situation. It can be done by drawing a map or filling the table. The original LM-GM model consists of two main components, i.e. the LM framework with 31 points of learning mechanics, and GM framework with 37 points of game mechanics, which are derived from various literature. In brief, it provides a concise overview of what should be taken into account when planning a SG mechanics. Table 2 depicts the relevant elements from LM-GM model that are applicable for IFSG.

Some elements of LM and GM are somehow related and interacted in IFSG as illustrated with colorful lines connect to various components. The implementation of the identified mechanics is explained below which is not fully organized consecutively with the linked LM-GM points in the table.

(1) **LM: Motivation-Participation-Incentive-Feedback.**

The *motivation* element was inserted in the game goal, i.e. to search and rescue as many victims as possible, to take the victims to the safe area, and to report the victim's condition to the Medical Care Unit (MCU). The learning point from participation was in terms of players' flexibility for taking different roles in two session games, and learning quickly and acting based on newly assigned roles. The incentive was mostly related to the game goals, i.e. how many people can be saved, and players who keep alive until the end of the game. The questionnaires and debriefings served both for the feedback mechanisms to encourage participants' learning from the game, and instruments for data collections to provide time to collect data from them (see point 3).

(2) **LM: Instructional-Guidance-Action/Task → GM: Strategy/Planning-Move-ment-Action Points-Story.**

The author acted as a game master, which is known in the game world as an organizer who set-up rules and moderates the game. The rules were organized as follow. In the pre-implementation stage, each player received written information about the game: the story, scenario, outline, roles and tasks, action points and expected game duration to accommodate these two session games, briefing, and debriefing, trying out the tools and apps, and the transition between games. The game organizer's team played themselves before the actual game day, ensuring every single plan would work. On the game day, all players attended a general briefing to ensure that the game rules, movement, and action points were understood. It was also to consolidate the groups; to test the equipment; and other necessary preparation before the game.

In the first session, the burning rooms were known from dynamic fire markers glued on a door. The fire markers were added from time to time by the game organizers, while in the second session, the fire spreading from room to room was visible in the app. In brief, we combined real smoke and "simulated" fire in the IFSG. In the session *without the app*, the players should check the room one by one and reported to the medical care unit if the room were clear—nobody inside. The communication mode was walkie-talkie software on the smartphone. In the session with the app, information about the victims was available in the smartphone app which displayed an indoor layout. Overview of the victims' locations was visible in the app.

The deployment of the app was conducted in two ways: by sending the app directly to the players to download in advance, and by preparing ten devices with the app installed. The app usage was explained in the briefing and repeated before the 2nd session was started. In fact, familiarizing the players with the app was crucial to the success/failure of the game goal.

(3) **LM Responsibility → GM: Role Play-Cooperation-Collecting-Collaboration-Competition-Cascading Information-Time Pressure-Movement-Information-Story-Feedback-Simulate/Response.**

The Role-Play element of IFSG is that the organization of the game imitated the real local practice of firefighting as seen in Fig. 1. There is a Crew Manager (CM), Smoke Diver Leader (SDL) and Smoke Diver (SD). *CM roles*: to monitor the evacuation process and tracking the victims if they are safe. His position was outside the fire zone and provided instruction. *SDL roles*: to coordinate/communicate the team members in the field, and with CM and to do rescue operation. *SD roles*: to carry out the rescue operation, check the victim's condition (based on the tag on the victim, save them based on the victims' situation), and to communicate with other rescuers. He/she may receive information from the leader, about the fire, or about the victim.

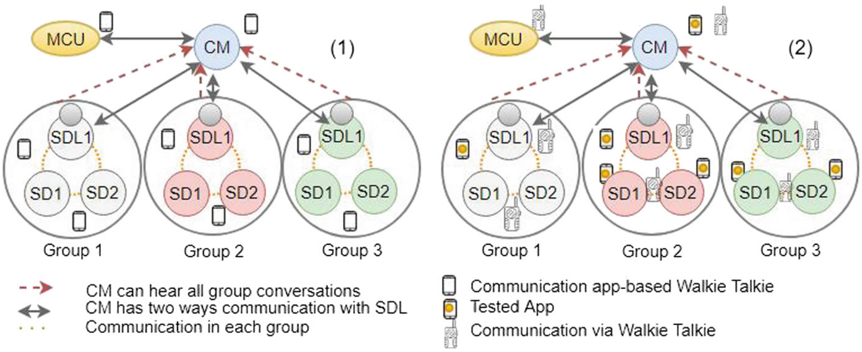



**Fig. 1.** Organization of the players in session 1 (left figure) and session 2 (right figure)

As communication modes, we used real Walkie Talkies and smartphone apps that functioned similarly to Walkie Talkie that could be used as information tracking tools. Careful preparation prior to the game was done to ensure that the groups interacted with each other through the correct channels based on the organizational communication hierarchy we had defined. In such way, cooperation, collaboration, information collection, information feedback, cascading information within each role and each group were granted. The competition occurred among groups in terms of their efficiency in managing the crew members and saving victims under time pressure.

A group of students was recruited to join the game, nine people were assigned as firefighters in three groups, and one person acted as CM, and one as MCU personnel. We had thirteen victims, of which two of them were dummies. Participants were asked, when they found a victim, read the tag, inform the SDL, and evacuate the victim to the North exit (which was predetermined and known to all). For adding the realism of this simulated environment, a couple of smoke machines were used and automatically triggered the building alarm as soon as the smoke was detected, which would not stop before the smoke had dissipated entirely.

(4) **LM: Explore-Hypothesis-Observation-Plan-Experimentation-Reflect/ Discuss-Feedback.**

As this game was created for learning and testing technologies, several research elements were embedded. At least five topics of research were integrated into IFSG, i.e. evacuation procedures/policy, communication, information flow and collaboration, usability, monitoring/tracking people, and hazard. Hence, we considered carefully of how not to give excessive burden to the players, and how not to reduce the enjoyment of the game despite these research elements added to the top of IFSG. The two-session game experiments permitted us to compare specific elements in each session (e.g. rescue performance, the number of victims saved, collaboration). The use of questionnaires and audio-video devices was informed to the players during the recruitment phase. Wearable cameras, video glasses for a usability test, and Walkie-talkie apps were considered as less obtrusive tools for data collection during the game and made the purposeful research less exposed. The players also had a chance to express their feeling, opinions, lessons-learned and reflections in the debriefing phase in each session.

### 4.2 Game Play

Two sessions of 30 min SR activities had been planned, i.e. one without and one with app support. We hypothesized that the rescue operation with app support (2nd session) would be faster than without app support (1st session). The IFSG goal was to search for victims trapped in the 3rd floor, and rescuing them by moving them all the way to the MCU located by the main entrance of the building.

All victims that were saved based on each victim's condition (safe/dead) had to be reported to the CM who monitored the overall progress of the rescue operation performed by all three rescuer groups. In these two scenarios, the players were permitted to swap the role, e.g. between leader and member of rescuers, or victims. No exact script was given to them as how to act, communicate and interact, except that they were informed on the outline of the roles, tasks, scenarios, prior to the game. Figure 2 summarizes the game timeline.

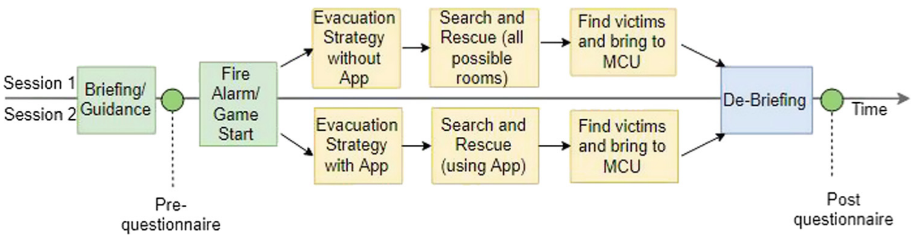



**Fig. 2.** The game timeline

## 5 Results

The results reported in this article highlight three points: the status of victims saved in the two experiments, evaluation of the app usage, and finally a comparison of the overall game implementation from observers' and players' point of view. Remember, we assumed that rescuers would work more efficiently *with app support* than *without app*, thus, more people would be saved. It was perhaps the most tangible indicator of the success or failure of our experiments.

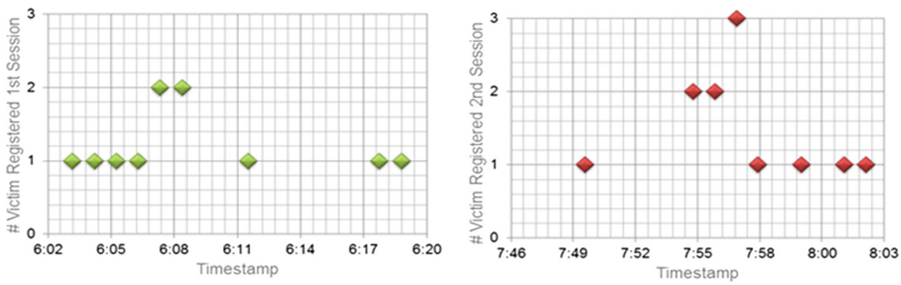



**Fig. 3.** Victim saved in session 1 and session 2

In Fig. 3 we can see a comparison of the victims that arrived and registered by the MCU in session 1 (top chart) and session 2 (bottom chart). The horizontal axis shows the timestamp while the vertical axis indicates the number of victims registered. As an illustration, from the chart in Fig. 3 we know there were two victims that arrived in MCU at 6:08 and 6:09 PM successively. In the session 1, it took 15 min from the first victim saved to last victim found, and two of the 13 victims were dead. In the session 2, it took 12 min (shorter time), but rescuers only managed to find 12 of 13, meaning 1 was missing, and 3 of them were dead. The time required to find the first victim in the second scenario was longer than in the first scenario (Fig. 3). All these results were unlooked-for.

It leads us to look more carefully at our questionnaire regarding the use of the app. Among the questions asked to the players concerning the app performance were: "Did the app crash during the game?", and "Did the markers of the victims' locations point to consistent locations?" The results are shown in Table 3:

**Table 3.** The app performance

| Role (N = 15) | Crash/not (in %) | | | Marker visible/not (in %) | | |
|---|---|---|---|---|---|---|
| | Yes | No | N/A | Yes | No | N/A |
| Victim | 57.2 | 42.8 | 0.0 | 14.2 | 28.6 | 57.1 |
| Rescuer | 75.0 | 12.5 | 12.5 | 50.0 | 25.0 | 25.0 |

Note that not all players used the app. Table 3 only includes those who had the app in session 2, i.e. eight rescuers and seven victims, or, 15 persons in total. 57.2% of victims and 75% of rescuers experienced app crashes. Some of them had to restart the app before searching could continue. Only 50% of rescuers saw the victims' markers while for the victims only 14.2%. We do not know if 57% who gave no answer were all experiencing a completely crashed app or only did not answer the question. It worth to mention the users with experience with app crash although only in the beginning is counted as a "crash" category. In other words, if the app crashed during the game, it does not mean the app does not fully work in the rest of the game. It is a rather technical disturbance at the beginning that prevented particular rescuers from acting promptly because they had to restart the app. On the whole, these can be either due to the app instability, or the smartphone's capability issue (e.g. if some players use older smartphones with lower memory or computing capacity).

The next question is of course: did it happen at all that the rescuers managed to help victims based on the app and did victims manage to inform about their position to the rescuers? Based on the debriefing discussion, after-game discussion with professional firefighters who watched the game, and interviews with a couple of players, there were, in fact, a few success stories where it was really the app that helped particular rescuers to find some victims. To put it differently, regardless of the revealed app stability problem, those who had the app running, could complete the task as expected.

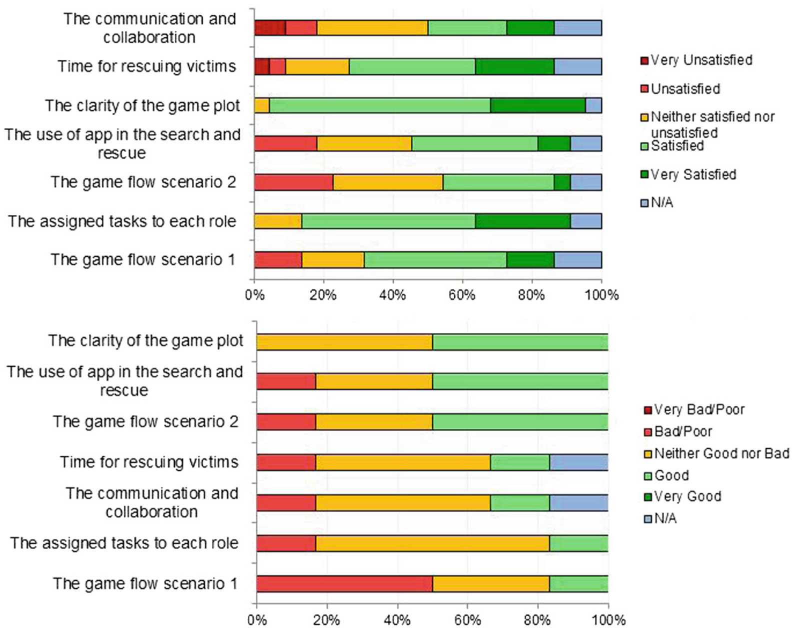



**Fig. 4.** Comparing evaluations on game implementation from observers' (bottom chart, N = 6) and players' perspectives (top chart, N = 22)

Likewise, one could see the great value of this app and its applicability, which was the main goal of the IFSG game implementation from both players' and observers' perspectives, they were asked to rate seven questions measured by a 5 step Likert scale. The observers judged based on "very good" to "very poor/bad" scale while the players' scale rated based on "very satisfied" to "very unsatisfied" scale. The results are presented in two stacked bar charts in Fig. 4, comparing the opinions of those two groups. The proposition of more positive respondents is marked with green or light green color in each bar while the most negative answers are colored red or light red. The yellow color indicates the proportion of neutral respondents while those who did not answer are marked with blue.

In general, most players were satisfied as indicated by longer green portions in each bar except "the game flow scenario 2". We look closer into the charts on the point that is rated poorly, which seems to come from the observers, i.e. the game flow 1 (see bottom chart, the longest red line). This is a bit counterintuitive as many players were more frustrated in scenario 2 shown from the graph where their satisfaction decreases. Moreover, we knew the reason was that the app did not fully work as expected for some users. Thus, a further discussion was conducted with the observers to find out what they saw that the player did not see so that they considered the scenario 2 to be better than the scenario 1.

One of the explanations is that the players played more systematically since they need to concentrate on the app, unlike in scenario 1 where the rescuers were rushed and tended to run, giving a chaotic impression. In reality, when working with fire, it is almost impossible to run, people have to walk with full concentration because they do not know where they are and the possible dangers ahead. If the rescuers run, they will also be exhausted too soon.

At the end of the game, the experience about the like and dislike side of each session of the game was collected through questionnaires. The players in general saw the positive sides of the session with app. However, the opinions vary more the second session such as "less chaotic", "new experience", "easy to report", "found fast" and other specific opinions about the functionality of the app. Based on the questionnaire results, the dislike of sessions with or without app came from several reasons. In the first session, the dislike part varies from personal feeling such as "noisy", "had no interesting role" or even unsatisfied that some players did not fulfill their role: "rescuer did not carry me all the way". Some opinions expressed the collaboration aspects such as lack of communication, or too little information. In the second session, the dislike part mostly directed to the technical difficulties found during the game such as "un-user-friendly", "crashed app". The non-technical comments are very likely to come from the victims that need to wait for the rescuers to come and help them. From these results, we know that there are many elements of the game that were appreciated, but there are weaknesses that we can learn from. Note that the reaction toward rescuing process is understandable since most of the players were not real firefighter and we could not expect them to act professionally as real firefighters.

## 6 Discussion and Lessons Learned

Designing SGs for testing an IT solution, coupled with extra research missions, is, in fact, an elusive task. There are several lessons to learn that we can discuss based on the results obtained in the previous section. First, the problem with the app used during the game may stem from unfamiliarity with the app, its user-friendliness, or compatibility issues with players' smartphones. The players had limited time to rehearse and to try out the app before the game.

Second, the slower reaction in response to the fire is highly recommended by firefighters, and more realistic. In reality, professional firefighters will walk slowly with full concentration on SR duty because they do not know where they are. Running will make the firefighters exhausted too early, which is undesirable. In the game, rushing and running created chaos. The undesirable effects of this behavior should be highlighted better to the players prior the game implementation.

Third, people make mistakes when exposed to the new technology, even if it is a mature one. The drill of thermal imaging camera usage introduced to the local fire service—for instance—was done multiple times before they could reap the benefits from it, and indeed, this will also be valid for our app testing. In other words, mistakes during technology testing are not entirely unexpected but repeated pre-game testing may minimize the errors in the actual game.

Indeed, testing the app in a SG context where the players are non- professional can be the primary limitation of this approach. Some players could not concentrate when exposed to the smoke, and were not trained to conduct multiple tasks. Hence, involving real users (in this context the firefighters) may be a better design for a SG intended for ICT-tool testing.

## 7 Conclusions and Future Works

In this study, we discuss a SG, designed for an app prototype testing. Despite some weaknesses in the stability of the app and the game implementation, the approach itself seems promising. Several improvements need to be considered, such as the importance of trial tests and better guide/technical information about the app. Future challenge if this technology becomes more mature, is how to increase the adoption of the tool, and how to use it within the existing SAR procedure.

## References

1. Lazreg, M.B., Radianti, J., Granmo, O.-C.: SmartRescue: architecture for fire crisis assessment and prediction. In: ISCRAM, Kristiansand, Norway (2015)
2. Lazreg, M.B., Radianti, J., Granmo, O.-C.: A Bayesian network model for fire assessment and prediction. In: International Workshop on Machine Learning, Optimization and Big Data. Springer (2015)
3. Radianti, J., Lazreg, M.B., Granmo, O.-C.: Fire simulation-based adaptation of SmartRescue app for serious game: design, setup and user experience. Eng. Appl. Artif. Intell. **46**, 312–325 (2015)

4. Meesters, K.: Towards using serious games for realistic evaluation of disaster management IT tools. In: Actes de la 2ème Journée AIM de recherche Serious Games et innovation (SG 2014). CEUR-WS, Paris (2014)
5. Metello, M.G., Casanova, M.A., de Carvalho, M.T.M.: Using serious game techniques to simulate emergency situations. In: GeoInfo (2008)
6. Linehan, C., et al.: There's no 'I' in 'emergency management team': designing and evaluating a serious game for training emergency managers in group decision making skills. In: Proceedings of the 39th Conference of the Society for the Advancement of Games and Simulations in Education and Training. Innovation North-Leeds Metropolitan University (2009)
7. Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. Int. J. Hum.Comput. Stud. **74**, 14–31 (2015)
8. Djaouti, D., et al.: Origins of serious games. In: Serious Games and Edutainment Applications, pp. 25–43. Springer (2011)
9. Abt, C.C.: Serious Games. University Press of America, Lanham (1987)
10. Susi, T., Johannesson, M., Backlund, P.: Serious games: an overview. Institutionen för kommunikation och information (2007)
11. Djaouti, D., Alvarez, J., Jessel, J.-P.: Classifying serious games: the G/P/S model. In: Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches, pp. 118–136. IGI Global (2011)
12. Yusoff, A., et al.: A conceptual framework for serious games. In: 2009 Ninth IEEE International Conference on Advanced Learning Technologies (2009)
13. Ratan, R., Ritterfeld, U.: Classifying serious games. In: Serious Games: Mechanisms and Effects, pp. 10–24 (2009)
14. Connolly, T.M., et al.: A systematic literature review of empirical evidence on computer games and serious games. Comput. Educ. **59**(2), 661–686 (2012)
15. Marczewski, A.: Game thinking. Even ninja monkeys like to play: gamification, game thinking and motivational design, p. 15 (2015)
16. Kirriemuir, J., Mcfarlane, A.: Literature Review in Games and Learning. A NESTA Futurelab Research report - report 8 (2004)
17. Werbach, K.: (Re)defining gamification: a process approach, in persuasive technology. In: Spagnolli, A., Chittaro, L., Gamberini, L. (eds.), p. 266–272. Springer International Publishing (2014)
18. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work? – a literature review of empirical studies on gamification. In: 2014 47th Hawaii International Conference on System Sciences (HICSS) (2014)
19. Di Loreto, I., Mora, S., Divitini, M.: Collaborative serious games for crisis management: an overview. In: 2012 IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE) (2012)
20. Tsai, M.-H., et al.: Game-based education for disaster prevention. AI Soc. **30**(4), 1–13 (2014)
21. Amory, A., et al.: The use of computer games as an educational tool: identification of appropriate game types and game elements. Br. J. Educ. Technol. **30**(4), 311–321 (1999)
22. Deterding, S., et al.: From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM, Tampere (2011)
23. Amory, A.: Game object model version II: a theoretical framework for educational game development. Educ. Technol. Res. Dev. **55**(1), 51–77 (2007)
24. Arnab, S., et al.: Mapping learning and game mechanics for serious games analysis. Br. J. Educ. Technol. **46**(2), 391–411 (2015)

# Serious Games in Virtual Environments: Cognitive Ergonomic Trainings for Workplaces in Intralogistics

Veronika Kretschmer[1(✉)] and André Terharen[2]

[1] Intralogistics and -IT Planning, Fraunhofer Institute for Material Flow and Logistics (IML), 44227 Dortmund, Germany
Veronika.Kretschmer@iml.fraunhofer.de
[2] Chair of Materials Handling and Warehousing, TU Dortmund University, 44227 Dortmund, Germany
Andre.Terharen@tu-dortmund.de

**Abstract.** Along with digitisation in the intralogistics sector, new methods of employee qualification arise. In this evaluation study a Virtual Reality-based serious game for workplaces in intralogistics was analysed due to usability, user experience, workload and motivational aspects. Results clarify that the virtual training is characterised by a good usability, a positively assessed user experience and a moderate mental workload. Intrinsic motivation of the study participants is positively associated with usability and user experience. Furthermore, mental workload is related with the usability of the VR game and technology-based personal information of the participants. The combination of serious gaming in a virtual training world is a new research field and offers a modern and efficient alternative to conventional training methods in intralogistics.

**Keywords:** Virtual Reality · Virtual training · Serious game · Human factors Cognitive ergonomics · Usability · User experience · Workload Intralogistics

## 1 Introduction

### 1.1 Virtual Trainings

Along with digitisation new methods for employee qualification arise. In various work contexts and economic sectors virtual training worlds can be found. Recent research shows that especially in the field of vocational trainings Virtual Reality (VR)-based simulations are suitable for exercising and assessing skills and furthermore reducing errors. The immersive effect of the VR technology enables the user to perceive situations realistically. Simultaneously, work situations can be simulated virtual that are cost-intensive, appear rather seldom or are difficult to train in the real working environment. Virtual learning environments and trainings are for example applied in aviation industry (Flight simulators), in the medical field (Training of medical procedures), in motoric rehabilitation, in in automobile construction, for exercising sports, or in construction safety management [1, 2].


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 266–274, 2019.
https://doi.org/10.1007/978-3-319-94619-1_26

### 1.2 Serious Games

A special type of digital knowledge transfer is the playful learning method of a serious game. The aim is to improve human motivation and performance regarding a special activity. Game mechanics are combined with suitable learning mechanisms to improve the learning success [3]. Digital games may enhance learning behaviour in comparison to non-playful alternatives [4]. Learning effects are associated with digital game design and complexity of interactions in the game environment. Recent research findings make clear that specific game design elements like badges, leaderboards, performance graphs, avatars or meaningful stories may have different psychological effects [5].

## 2 An Evaluation Study of a VR-Based Serious Game for Workplaces in Intralogistics

### 2.1 Research Objective

Going hand in hand with growing globalisation and flexibilisation of the working world, the logistics sector gains in importance. These days, due to digitisation, new information and communication technologies are increasingly used in companies. More and more processes are automated according to rising demands on entrepreneurial performance and costs in intralogistics. Nevertheless, for flexible and complex work steps, human beings are still an irreplaceable resource in the »Social Networked Industry« [6]. Therefore, a stress-optimised and thus human-centred design of human-machine systems is crucial to success of digitisation. To keep the human knowledge carriers in future, new variants of employee training have to be found. Nowadays, it is indispensable to provide cost-effective and time-saving alternative methods for occupational trainings. Thus, the research objective of our study was to evaluate a VR-based serious game of a packaging workplace. Our aim was to proof the operability and utility of a new technological solution for trainings in the intralogistics sector.

### 2.2 Description of the VR Technology

The VR-based serious game for workplaces in intralogistics was developed and designed at the Fraunhofer IML (see Fig. 1). Deployed hardware consisted of the HTC-Vive® system and a powerful desktop computer. To develop the VR-based serious game, the game engine Unity® in the Version 5.6.0f3 was used. Microsoft Visual Studio Professional 2015 was applied for creating und editing the scripts in C#.

### 2.3 Study Design

At first, for understanding and handling how to use the VR technology, all participants had to complete the Steam® HTC-Vive® Tutorial. Then, the participants played the VR-based serious game of a packaging process. For this purpose, a typical packaging process was chosen and designed according to a sufficient complexity level and different working steps. The logistical tasks were simulated close to reality and provided opportunities for task processing and identification for the player. The stationary

**Fig. 1.** Training environment of the Fraunhofer VR-based serious game of a packaging workplace (Source: Fraunhofer IML)

packing station (goods-to-person) had the advantage, in contrast for example to a person-to-goods picking process, of maintaining or reinforcing the immersion in the VR environment. Working steps were introduced via a headset.

The individual working steps of the packaging process were simulated realistically and consisted of the following steps:

- Requesting next order
- Selection of the specified box size and positioning on the packing table
- Opening the box
- Picking up the scanner and scanning the items to be packaged which are delivered by the automated guided vehicle (AGV)
- Putting the items into the box
- Selection of the specified packaging material and putting in the box
- Closing the box
- Sealing the package with the tape dispenser
- Placing the package on the target pallet next to the packing table

In the training environment the player received each next task which had to be performed via audio instructions with the help of a headset. As main game elements, visual aids, a running game time followed by a final displayed leaderboard, audio feedback during the scanning process and a final verbal feedback on the overall performance were used.

## 2.4 Survey Instruments

After the virtual packaging training participants were asked to fill out several validated questionnaires covering different areas: To evaluate the usability of the virtual game, the internationally known System Usability Scale (SUS) was deployed [7]. The SUS comprises ten items with a 5-point Likert scale from 0 to 4. The overall SUS score has a range from 0 to 100 by multiplying the total sum of the values of all items by 2.5 [8].

User experience was measured with an identically named questionnaire (UEQ) [9]. The subscales of the UEQ can be classified into attractiveness, pragmatic quality (Efficiency, perspicuity, dependability) and hedonic quality (Stimulation, novelty). Pragmatic quality describes task-based issues and the quality of use. Hedonic quality includes the quality of design. The UEQ consists of 26 unipolar items with a format of a seven stage semantic differential [9].

Subjective mental workload was recorded with the well-established and most widely used NASA-Task Load Index (TLX) [10]. The NASA TLX consists of six different subscales: Mental demand, physical demand, temporal demand, performance, effort and frustration. All six dimensions were rated on a bipolar scale from 0 (low) to 100 (high) with 5-point steps. To determine the overall mental workload, the raw TLX score was calculated by taking the average of the sum of the values of all six subscales [11].

To measure the intrinsic motivation during the packaging training, a German short scale (KIM) [12] was applied that is an adapted, time-economic version of the Intrinsic Motivation Inventory by Deci and Ryan [13]. The KIM consists of the four subscales interest/enjoyment, perceived competence, perceived choice and pressure/tension. While pressure/tension is a negative predictor, the remaining subscales load positively on intrinsic motivation. Each subscale consists of three items with a 5-point Likert scale from 0 to 4.

The Positive And Negative Affect Schedule (PANAS) [14] was deployed to assess the current affect situation of the participants. The PANAS consists of an adjective list with 20 items with a 5-point Likert scale from 1 to 5. The two dimension positive affect (PA) and negative affect (PA) contain ten positively and negatively connoted adjectives.

Dizziness, ocular and musculoskeletal strain were recorded by a validated questionnaire [15]. The scale dizziness consists of two items. Visual strain includes seven items and musculoskeletal strain comprise five items. Each item was rated on a 6-point scale from 1 to 6.

Furthermore, sociodemographic and personal data was captured. Technology-based personal information like the previous experience with technologies, the personal attitude towards technical progress [16] and the willingness to use new technologies [17] were asked for: Technology-based biography includes the two subscales avoidance tendency and innovation orientation [16]. The former is composed of seven items, the latter is formed of two items. Attitudes towards technology is composed of the two subscales potential threat (Two items) and the need for using technologies (Four items) [16]. The German short scale of technology commitment [17] consists of the subscales acceptance, control beliefs and agency beliefs regarding new technologies or modern technology development. Every subscale includes four items. Each item of the above mentioned technology-based subscales has a 5-point Likert scale from 1 to 5.

## 2.5 Sociodemographic and Technology-Based Personal Information

The study sample comprises 30 participants (50% male) with the age ranged from 19 to 29 years (M = 25.47; SD = 2.24). The knowledge in logistics varied between »none« (13.3%), »theoretical« (30.0%) and »theoretical and practical« experiences (56.7%). All participants were students working at the Fraunhofer IML.

Regarding the positively connoted subscales innovation orientation and need for using new technologies high mean values were achieved (see Table 1). According to the subscales with negative connotation, such as avoidance tendency and potential threat, rather low arithmetic means were reached (see Table 1).

**Table 1.** Descriptive statistics of the two subscales of technology-based biography.

| | *Min* | *Max* | *M* | *SD* |
|---|---|---|---|---|
| Technology-based biography | | | | |
| Avoidance tendency | 1.00 | 3.33 | 1.92 | 0.68 |
| Innovation orientation | 1.67 | 5.00 | 3.91 | 0.80 |
| Attitudes towards technology | | | | |
| Potential threat | 1.00 | 4.00 | 1.85 | 0.79 |
| Need for using technologies | 1.75 | 5.00 | 3.90 | 0.79 |

*Note. Min = Minimum, Max = Maximum,*
*M = Arithmetic mean, SD = Standard deviation*

The acceptance and agency beliefs regarding new technologies is rather high among participants (see Table 2). In contrast, control beliefs according to modern technology development is quite low on average.

**Table 2.** Descriptive statistics of the three subscales of technology commitment.

| Technology commitment | *Min* | *Max* | *M* | *SD* |
|---|---|---|---|---|
| Acceptance | 1.50 | 5.00 | 3.73 | 0.80 |
| Control beliefs | 1.00 | 2.75 | 1.69 | 0.64 |
| Agency beliefs | 1.75 | 5.00 | 3.68 | 0.79 |

*Note. Min = Minimum, Max = Maximum,*
*M = Arithmetic mean, SD = Standard deviation*

## 2.6 Usability and User Experience

The usability of the VR game reached an average SUS score of *M* = 71.81 (*SD* = 9.18) which can be interpreted as a »good« rating [8, 18]. The subjective assessments of usability varied from »bad« (SUS Score *Min* = 37.5) to »excellent« (SUS Score *Max* = 85.0) [8, 18].

Descriptive results of the user experience subscales attractiveness, stimulation, novelty, efficiency, perspicuity, and dependability are listed in Table 3:

**Table 3.** Descriptive statistics of the user experience variables.

| User experience | *Min* | *Max* | *M* | *SD* |
|---|---|---|---|---|
| Attractiveness | −2.00 | 3.00 | 1.71 | 1.15 |
| Stimulation | −0.50 | 3.00 | 1.84 | 0.93 |
| Novelty | −0.25 | 2.75 | 1.34 | 0.66 |
| Efficiency | −1.25 | 2.75 | 0.92 | 0.87 |
| Perspicuity | −0.50 | 2.50 | 1.68 | 0.58 |
| Dependability | −0.75 | 2.25 | 0.99 | 0.69 |

*Note. Min = Minimum, Max = Maximum, M = Arithmetic mean, SD = Standard deviation*

Subscales can be classified into the three dimension attractiveness, pragmatic quality (Perspicuity, efficiency, dependability) and hedonic quality (Stimulation, novelty). Attractiveness ($M$ = 1.71), pragmatic quality ($M$ = 1.20) and hedonic quality ($M$ = 1.59) of the VR game were assessed on average as positive. According to the user experience, the situation of affect during the VR training was also mostly positive (PA: $M$ = 3.60, $SD$ = 0.75; NA: $M$ = 1.35, $SD$ = 0.30).

## 2.7 Workload

The overall mental workload during the VR training was evaluated on average as moderate (Raw TLX: $M$ = 30.06, $SD$ = 12.12) [19]. The subjective strain regarding the subscales of the NASA TLX ranged from low to moderate (see Table 4).

**Table 4.** Descriptive statistics of the six subscales of mental workload.

| Mental workload | *Min* | *Max* | *M* | *SD* |
|---|---|---|---|---|
| Mental demand | 0 | 90 | 31.50 | 24.75 |
| Physical demand | 0 | 40 | 13.17 | 10.54 |
| Temporal demand | 5 | 70 | 35.17 | 19.80 |
| Performance | 0 | 90 | 36.67 | 28.72 |
| Effort | 0 | 70 | 35.00 | 17.81 |

*Note. Min = Minimum, Max = Maximum, M = Arithmetic mean, SD = Standard deviation*

The subjective perceived dizziness ($M$ = 1.73, $SD$ = 1.07), ocular ($M$ = 2.26, $SD$ = 1.21) and musculoskeletal strain ($M$ = 1.31, $SD$ = 0.32) caused by the VR glasses were valued as low.

### 2.8 Intrinsic Motivation

The subscales interest or enjoyment ($M = 3.41$, $SD = 0.81$), perceived competence ($M = 2.71$, $SD = 2.19$) and perceived choice ($M = 1.94$, $SD = 0.85$) are marked by a low to moderate average value. Furthermore, the average of subjectively perceived pressure or tension is quite low ($M = 1.12$, $SD = 0.79$).

### 2.9 Correlations Between All Variables

Bivariate correlation analyses indicate that the usability of the VR game was negatively related to mental workload (r = .39, p < .05). Furthermore, usability was positively associated with most user experience subscales (Attractiveness: $r = .67$, $p < .01$; Perspicuity: $r = .64$, $p < .01$; Efficiency: $r = .43$, $p < .05$; Stimulation: $r = .50$, $p < .01$; Novelty: $r = .55$; $p < .01$). Results make clear that the higher the usability of the virtual training was, the more positive was the current affect situation of the participants during interaction with the technology ($r = .55$, $p < .01$). Usability significantly increased with a higher degree of innovation orientation ($r = .38$, $p < .05$) and perceived need for using technologies ($r = .49$, $p < .01$). In addition, usability significantly decreased with rising potential threat.

Between mental workload and technology acceptance ($r = -.41$, $p < .05$) and the perceived need for using technologies ($r = -.44$, $p < .05$) a negative relationship could be found. Besides, results illustrate that the increase of subjective threat of technology was positively linked to overall workload ($r = .41$, $p < .05$).

Regarding intrinsic motivation, a positive correlation between usability and interest ($r = .42$, $p < .05$) and also perceived choice ($r = .37$, $p < .05$) could be determined. Equally, most user experience subscales were positively associated with interest (Attractiveness: $r = .78$, $p < .01$; Efficiency: $r = .61$, $p < .01$; Stimulation: $r = .78$, $p < .01$; Novelty: $r = .70$, $p < .01$) and perceived choice (Attractiveness: $r = .43$, $p < .05$; Efficiency: $r = .51$, $p < .01$; Stimulation: $r = .39$, $p < .05$; Novelty: $r = .45$, $p < .05$).

Between user experience and total subjective workload no significant correlations could be found. Furthermore, the scales dizziness, ocular as well as musculoskeletal strain were not related to mental workload or usability.

## 3 Summary and Outlook

All in all, the Fraunhofer VR-based serious game has a positive rated usability and user experience and is mentally and physically demanding only on a moderate level. Results of correlation analyses indicate that usability as well as user experience are positively associated with intrinsic motivation of the trainee. Furthermore, a user-friendly design of dialogue seems to be negatively related with the mental workload of the player. Equally, technology-related personality characteristics may influence the interaction between humans and technology.

With VR trainings process understanding and quality awareness can be provided. Furthermore, the immersion in a VR environment may lead to increased motivation and learning success of trained persons. Further advantages of digital workplace simulations are providing implicit knowledge, exercising rarely occurring work situations, comparing training results or promoting sensory-motor skills. In addition, with virtual trainings individual needs of each trainee can be addressed. The combination of serious games and virtual worlds is a new research field and offers a modern alternative to conventional training methods in the intralogistics sector. The evaluated VR game will continue to be developed at the Fraunhofer IML.

In future research, a wider age range of study participants should be investigated. As specific game design elements may have specific psychological effects [5], various configurations of game design elements should be tested and compared. To measure the degree of engagement in the virtual training world, further questions regarding immersion and presence should be used [20].

**Acknowledgments.** This study was founded by the »Centre of Excellence for Logistics and IT« and the »Innovationlab Hybrid Services in Logistics«. We thank Felix Moritz Bedarf, Alexandra Eichler and Vanessa Vogel for supporting us by the conduction of the study.

## References

1. Guo, H., Yu, Y., Skitmore, M.: Visualization technology-based construction safety management: a review. Autom. Constr. **73**, 135–144 (2017)
2. Miles, H.C., Pop, S.R., Watt, S.J., Lawrence, G.P., John, N.W.: A review of virtual environments for training in ball sports. Comput. Graph. **36**(6), 714–726 (2012)
3. Arnab, S., Lim, T., Carvalho, M.B., Bellotti, F., de Freitas, S., Louchart, S., Suttie, N., Berta, R., de Gloria, A.: Mapping learning and game mechanics for serious games analysis. Br. J. Educ. Technol. **46**(2), 391–411 (2015)
4. Clark, D.B., Tanner-Smith, E.E., Killingsworth, S.S.: Digital games, design, and learning: a systematic review and meta-analysis. Rev. Educ. Res. **86**(1), 79–122 (2016)
5. Sailer, M., Hense, J.U., Mayr, S.K., Mandl, H.: How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput. Hum. Behav. **69**, 371–380 (2017)
6. Tüllmann, C., Ten Hompel, M., Nettsträter, A., Prasse, C.: Social networked industry ganzheitlich gestalten. In: Future Challenges in Logistics and Supply Chain Management, vol. 6. Fraunhofer IML, Dortmund (2017)
7. Brooke, J.: SUS: A “quick and dirty” usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, A.L. (eds.) Usability Evaluation in Industry. Taylor and Francis, London (1996)
8. Brooke, J.: SUS: a retrospective. J. Usability Stud. **8**, 29–40 (2013)
9. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) HCI and Usability for Education and Work: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, 20–21 November, pp. 63–76. Springer, Heidelberg (2008)
10. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human mental workload, pp. 139–183. Elsevier, Amsterdam (1988)

11. Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 50 (2006)
12. Wilde, M., Bätz, K., Kovaleva, A., Urhahne, D.: Überprüfung einer Kurzskala intrinsischer Motivation (KIM) [Testing a short scale of intrinsic motivation]. Zeitschrift für Didaktik der Naturwissenschaften **15**, 31–45 (2009)
13. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Springer, New York (1985)
14. Krohne, H.W., Egloff, B., Kohlmann, C.-W., Tausch, A.: Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). Diagnostica **42**, 13–156 (1996)
15. Jaschinski, W., König, M., Mekontso, T.M., Ohlendorf, A., Welscher, M.: Computer vision syndrome in presbyopia and beginning presbyopia: effects of spectacle lens type. Clin. Exp. Optom. **98**(3), 228–233 (2015)
16. Claßen, K.: Zur Psychologie von Technikakzeptanz im höheren Lebensalter: Die Rolle von Technikgenerationen. Dissertation (2012)
17. Neyer, F.J.J., Felber, J., Gebhardt, C.: Kurzskala Technikbereitschaft (TB) [Technology commitment]. In: ZIS - Zusammenstellung sozialwissenschaftlicher Items und Skalen (eds.) (2016)
18. Bangor, A., Miller, J., Kortum, P.: Determining what individual SUS scores mean: adding an adjective rating scale. J. Usability Stud. **4**(3), 114–123 (2009)
19. Grier, R.A.: How high is high? A meta-analysis of NASA-TLX global workload scores. In: Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, pp. 1727–1731 (2015)
20. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence **7**(3), 225–240 (1998)

# Minecrafting Virtual Education

Stuart Lenig[1,2(✉)] and Nicholas Caporusso[1]

[1] Fort Hays State University, 600 Park Street, Hays, USA
{s_lenig,n_caporusso}@mail.fhsu.edu
[2] Columbia State Community College, 1665 Hampshire Pike, Columbia, USA

**Abstract.** Minecraft is an innovative video game that combines a simplistic, flexible, childlike interface with an open-ended, non-deterministic, playing environment that allows for a wide assortment of tasks and skills building while providing a platform for innovative game play, collaboration, and new formats for interaction and cultural creation. In this paper, we explore the dynamics of Minecraft and we discuss their implications in the context of virtual education.

**Keywords:** Minecraft · Education · E-learning

## 1 Introduction

In the recent years, Minecraft has exploded from an indie game to a best-seller platform with over 100 million copies distributed. Nowadays, the software invites creative participation from a total of 55 million monthly active players, programmers, and designers worldwide. Unlike point and shoot, sports, and specific strategy games (e.g., Doom, Call of Duty, The Sims, or Civilization), Minecraft comes with no official agenda, and once there, users can decide what end they wish to follow, similarly to real life. It offers an unarticulated schema when a player enters, and it is up to the user to transform the partial environment of the base game into something more personal. This leads to a culture of customization and innovation that has created massive communities that explore and create new aspects to the growing and continuously expanding virtual platform. Few proprietary gaming systems have allowed or encouraged this level of interactivity, expandability, and diversity. Most importantly, Minecraft demonstrated that unleashing individuals' creativity is even more important than providing users with adrenaline and compelling visual experience. On the contrary, pixel art succeeded as a democratization tool for user generated content.

The launch of Minecraft education edition represents a natural but important step in which programming, gaming, design, and educational resources have converged with the purpose of creating new technologies that seamlessly blend gaming and education. However, beyond this initiative, worldwide communities of educators independently created and keep generating new use cases for using this platform for teaching and engaging learners in exploring subjects in non-conventional ways. In this paper, we present a literature review of the most prominent applications of Minecraft for virtual education. To this end, we analyze scholarly articles and unpublished yet relevant best practices. In addition, we categorize currently available tools into education level and subject, with the objective of providing educators with both an entry point and a


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 275–282, 2019.
https://doi.org/10.1007/978-3-319-94619-1_27

reference to integrating virtual experiences in their teaching activities. Moreover, we analyze factors, such as, gamification, creative and spatial thinking, and integration between interaction between physical and virtual classroom, that might play a role in revitalizing education. Also, we identify directions for education research using gaming platforms and we present novel opportunities for designing learning environments.

By viewing Minecraft as a larger entity, and by analyzing how users are leveraging its potential to shape new products, to create communities, to engage with audiences, and to deliver experiences, other software houses could understand how to invest the resources spent in developing their game titles to generate an impact in other sectors, with a specific regard to education. Examples from Minecraft could lead to transforming games which are well-developed, mature, and adaptable to educational purposes, into cost-effective, collaborative, and expandable learning platforms.

## 2 Related Work

Minecraft differs from most games in that it is wildly anarchic in its construction and aims. Developer, Markus Persson wrote that the game is constantly under development [1], which, in turn, helped to keep the game a subject of interest. What is the subject of Minecraft, and indeed is there a single subject or objective in this game? It is more than a mere game and has morphed into an infotainment educational portal that has invaded classrooms across the globe. In [2] Minecraft was referred to as "an alluringly moving target to try to pin down, and so in order to assess how it is well-played, well-designed, and iteratively well-redesigned, we need to think more broadly about the approach Persson has taken toward the development of the game". That process has been circuitous and this has prompted some hesitation about the game and its implementation in the educational and scientific community.

The authors of [3] commented that introducing the game to a novice can be a difficult task due to the lack of instructions. A belief in rational processes presumes the game and its subsequent play needs to be directed towards a definitive goal. Minecraft's lack of destination can be worrying to a community that perceives a game design as requiring a requisite end goal.

Yet, it is precisely the open-ended nature of Minecraft that makes it a touchstone for innovative scientific and creative accomplishment. In structured gaming, goals and insight are limited. Minecraft is often associated with the burgeoning maker movement that is flourishing in communities and schools across the nation to teach innovation and entrepreneurship to a group of young creators that can empower them to be creatives rather than drones of an industrial oligarchy in the hands of a few corporations. In [4], the authors commented that Minecraft has participated in the building of makerspaces in libraries that provide students with hands-on skills. The perspective presented in [4] is that diverse forms of learning encouraged in an open environment provide the best educational opportunity. Authors of [4] write that, "whether the learning takes place through being a part of a team, building a robot, using maker kits, designing in 3D, or through individuals reading or group discussion". Minecraft offers this sort of generalized workspace where students can build without specified end goals.

Such open-ended constructions might seem less than optimal to produce significant results. Certainly, one subject area that has fascinated Minecraft scholars is the game's application to the fine arts and architecture. In [5], the authors postulate that the game's style of art creation, this open-ended free ranging design is precisely the format that Minecraft assails most capably: "rather than form a stable body, the immanent data-set moves as a flow of information through systems" [5]. Therefore, the authors imply that students are more challenged by such environments and this energizes the creative process, and they argue that this computer replication process merges the creative instinct of the creator with hard core data, describing the process as, "these operations include in general: generating a point cloud; building a mesh; creating a UV map, also called a skin or a texture". Thus, such operations foster more creativity than operations that do not have to reference real spaces in a real plane. What is built in Minecraft must be buildable and must conform to the rules, the geometry, and space of the Minecraft virtual world.

While many in the educational realm see production as proceeding through a series of concrete steps, in [5] authors argue that digital production need not follow this time-based progression sequential model derived from film and prior narrative structures, explained as "3D imaging is characterized by simultaneity – a function of parallel processing rather than cinematic succession" [5]. This form of creation violates some of the honored traditions of temporality, and therefore can be both embraced and questioned by academic institutions that are regulated through sequential processes and thus are wary of simultaneity, a factor of fast computer and machine processes.

City Building Games (CBGs) are described in [6], as changing the thinking models and methods of young scientists and artists and urges them to consider the whole topic of cityscapes earlier in their development thus providing a better entre for future city planners. The authors of [6] describe the games as, "highly successful cultural products that for millions of gamers have constituted their initial formative experience with urban planning and development". The authors of [6] are less concerned with the formal pedagogical models underpinning such games and prefers to view them as "social or cultural" devices. A central concept to the games' value is the ability of such games to urge players to think in terms of definite spatial dimensions [6]. They may not be working in a real world, but they must follow the rules of engineering in a real world even if their world is make-believe: "CBGs and other simulation/strategy games encourage the player to exercise adaptive critical reasoning while continually critiquing different theories and strategies" [6]. So, in this sense, Minecraft creators are dealing both with theoretical issues and practical problems. The only thing that is imaginary is the plane of creation that is a digital construct. In other regards their simulation contains elements of a real-world simulation.

## 3 Human Dimensions of Minecraft in Education

In the striving to make computers, gaming and education synonymous, much effort has been expended in finding games that are age appropriate. A delightful aspect of the Minecraft experience is that the game and its open-ended playing mechanism is that the game is age appropriate at almost any level, due to its innate capacity to be modified and the broad ranging community that engage in continually modifications of the game. Because of its complexity and online community, Minecraft is recommended for kids aged 8+. However, even younger children can enjoy the game with little adverse reactions, and therefore several suggest that the title is already suitable for kids at a very young level. Furthermore, a variety of games with similar game play and set of objectives that can occupy younger children and would not necessitate constant monitoring for students that desire to play Minecraft at a younger age. Essentially, the dynamics of Minecraft can be described as a sandbox game, which means that players are given a virtual construct with which they can build things, though survival mode and adventure mode offer more challenges and more dangers and even potential death for characters. As a result, the best approach is for parents who are concerned about younger children is to keep a watchful eye when they are involved in game play. If properly harnessed, Minecraft enables children to express their creativity in a 3D environment: they can create massive buildings, elaborate villages and cities, train systems, or even other creative objects and projects.

The wide disparity in age range for players and fans is one of the sources of inspiration for their players. As any product enjoyed and loved by children, parents need to feel comfortable and secure with the product. It is easy for parents to be concerned when a new product begins to draw their children's attention and consumes massive amounts of a child's time, as several researchers associated smart phone usage can degrade memory function [7]. However, researchers argue that reliance on a technology can stunt real world growth experiences. If this is true for people engaged in using smart phones is it true of people that spend long hours in video game play.

A big question concerns whether programs like Minecraft can teach conducive computer skills or must it ratchet our brains to new levels of stress, excitement and engagement that destroys constructive thought, concentration and the ability to get things done. Minecraft is a strongly collaborative learning tool for youth and parents and children need to approach it in that fashion. New improved version of the game, and specifically the cross platform, Minecraft Better together sadly evolves the game towards the industry that constantly depends on upgrades and new thrills to sustain audience involvement. While there is no rush to update to the new version, the existing Minecraft games will not stop working. However, parents need to know that it is the new version of Minecraft that children will want to play. However, while Minecraft can be obsessive as a hobby, it unlikely to be addictive and harmful the way shooting games are to mental development. The authors describe it as a "an amazingly creative experience", which has intrinsic educational value from problem-solving, storytelling and team building skills. According to its creator, the better together version allows kids to play together on console on tablet for the first time it also opens up services which of been unavailable on the console version.

Another aspect in Minecraft is the shelf-life of creative work as updates will happen as the game is developed in the coming years. While provisions for porting over previous constructions have been made for some versions of the game there is a potential that past creations, such as, old houses that are torn down in the redevelopment project of a town, may be swept away.

There are benefits to upgrades to new versions. Minecraft's transition to new gaming engines, such as, the Microsoft Bedrock engine, which happened after the acquisition of the platform, inspired a new look. Though the blocking sub-structure of the game remained, the visual textures and colors were enhanced by a gaming engine that rendered the colors and landscapes more vibrantly. Updates are also expected to improve game play. The Better Together update allows the Minecraft kernel to be loaded to the new Bedrock gaming engine which improves graphics and the look of Minecraft (famous for its blocking textures) and enhances cross platform play that allows users to play from one device to another device in seamless integration.

Recent articles explore the creation of a new novel that was produced by the Mojang company, the producing entity that has produced the Minecraft game and subsequent Minecraft forms of adventure. Minecraft: The Island is an adventure set in the Minecraft Universe. As reminded by the authors [8], the open-ended nature of the game is a big part of its appeal. Max Brooks, the author of the Minecraft novel is the son of Mel Brooks and Anne Bancroft and wrote the popular zombie novel World War Z. He combines a paranoid glee about end of the world apocalyptic visions with a Robinson Crusoe sense of adventure and innovation. In another sense, the expansion of Minecraft into the realm of fiction is appropriate for a game that has acquired a narrative sensibility and created its own adventure world.

Further change from conventional gaming is how Minecraft can steer students towards aesthetic adventures, and the creation of actual 3D objects. Many describe the potential of Minecraft to be more than a game, and to implement an aesthetic platform. Indeed, many game contributors are designing elegant landscapes and extreme architecture patterns out of the game's crude block construction. For instance, a company called Blockworks designs maps for the Minecraft platform and the company has even developed a coffee table book called Beautiful Minecraft which shows the most extreme game designs. The artistic community can thrive in the open gaming context of Minecraft, as are artistic opportunities flow from Minecraft's open structure, and by the fact that players, in addition to the classically video-game-like "Survival Mode", can also realize anything they want in the "Creative Mode", which removes any threats and turns Minecraft into a blank canvas.

Where Minecraft turns from a distraction to a device for learning is in the way the program is deployed and used at various schools across the nation. Gamification, the use of game strategies to increase learning has become a common practice in educational institutions. In a report [9] published by an educational firm discussing Minecraft's capability for teaching social and emotional learning through its platform. The company explained their research design, which involves meeting with teachers, school leaders, and students, and co-designing instructional resources. Testimonials and contributions from the viewpoints and experiences of dozens of educators from 11 countries across four continents enable shaping new educational tools.

While the findings of the report suggest that the game can help students with intellectual, creative, and emotional development, the schools seem to have focused on kids with strong backgrounds, good parenting, and strong resources that also could account for their strong personal development. One places a good student in a healthy educational environment and remarkably they tend to grow into healthy adults. One cannot ascertain from this scenario if Minecraft encouraged healthy development or if the research simply confirmed the impact of a good nurturing environment, with or without Minecraft. Despite this caveat, the researchers may have found impressive correlations between Minecraft and student development. As discussed in [10] "when educators intentionally work on the classroom culture using a game like Minecraft as a vehicle for that learning - when students are already intrinsically motivated to play - it can serve as a powerful learning tool. Studies point to learning as a cultural process, including complex aspects of development representing a larger human experience regardless of age, gender, class status, racial or ethnic group membership. It has global impact. It can be entirely non- verbal, and yet, it provides opportunities for rich verbal dialogue, story prompts, history, contextual understanding, science, fine arts and even world languages". So, the research suggests that the game can aid such developments in any children.

The education and therapy community have been interested in Minecraft's ability to engage students. So successful has the game been that therapists are now looking to the game design as a model for approaches to therapy. The authors of [11] explained that Minecraft's game design emphasizing an essential experience principle provided users with an emotive or cognitive connecting point that other games and other forms of therapy might not emphasize. The authors postulated that if therapeutic approaches could emulate the game's connectivity with audiences, the game and the therapy might be more effective. They quote [12] on video game design and explain that their notion of an essential experience principle of playing a game is comparable to the idea of "systems thinking" [12] or as the authors describe it, "the emergent, holistic, and immersive experience of the player that occurs when the tangible game elements are combined with the player's actual participation". [11] They believe that the experience comes before the playing and the winning and any other reward involved with the gaming mindset. The authors of [11] believe that one of the main concept in [12] is that the true purpose of playing a game is not to complete it or to win but rather to participate in its essential experience. Not all game authors adhere to this principle. Some work on building a world or making active game play, but Minecraft adheres to the heavy engagement model by providing a sandbox environment where concrete goals and specific problems are not dictated, raw materials must be manipulated and molded into tools and structures, and players can visit other worlds, build in them, allowing building and creating that can be spread and shared. Thus, Minecraft works on the psyche of the player in a much less fatalistic manner but at times in a more profound fashion the player's thought process is supported and aided by the interface, ambitions, and outcomes of the game, itself.

Still others have explored learning objectives and how gaming, and particularly Minecraft leverages gaming structures to participate in crafting learning environments. An article [13] mentioned a clear connection between the procedures and objectives of the technical communication field and the roles and missions of gaming structures, and they stated that the field already works at the intersection of the technical and the symbolic and games are both". The authors envision games as rule systems and frameworks for interaction and they see a very natural evolution from that process to the process of writing and thus interacting in that game world. Moreover, the authors speculate that creating manuals and communication materials to aid players in their gaming pursuits may draw on the writers' humanistic training to design systems that prioritize human/player experiences. The work of [13] suggests that not only can technical communication engage with traditional game play, it can also involve itself in game subversion by helping create technical writing hat can disrupt play. As an example, Minecraft griefing guides teach aspiring griefers how to enter social situations, gain trust, and then make people angry and disrupt game play for the spectacle and benefit of the person doing the griefing. Such insight into technical communication, the game Minecraft and gaming in general suggests a highly developed community of learners that study the game, discover the intents and concepts of the game, and find means to disrupt the game. This is a complicated series of interactions in a game world in which little is pre-destined or ordained as conventional game play or outcomes and suggests that even the practice of griefing demands a high level of originality to produce optimal disruption outcomes.

Owing to its expansive game play, Minecraft has evolved more complex audiences. People with disabilities and with special needs, and adult-non-gamers have engaged with the game to seek worlds that don't exist in a non-game environment but do exist very easily in the Minecraft environment. According to [14], limited budgets are driving individuals and small teams to experiment with minimalistic styles and explore themes and taboos that would normally be considered beyond the purview of mere games, such as mortality, morality, economics, and abuse. In their work, [14] explore how games like Minecraft cultivate an interest in exploration, creation, world-building, and the production of safe and attractive habitats: positive society building may be an outgrowth of non-confrontational non-aggressive game play. Society building on the scale of new civilizations could be one of the practices and outcomes of Minecraft's method of play as an act of creation.

But the alternative to deeply directed game play in videogames like Minecraft may be a form of human colonization of cyber worlds. Not horrific dystopias that offer a destiny of murder and dissolution in competitive violent worlds, such as, Grand Theft Auto, where situational ethics dominates our thinking, but perhaps the creation of utopian places where people congregate and set up an alternative existence. The authors of [14] introduce recent patterns of Minecraft consumption involving less of creating and having more of a focus on a minimalist existence a sense where more time is simply spent exploring and appreciating the landscape. Maybe the result of Minecraft's virtual world is a land of leisure and a sense of accomplishment, making a perfect world where none existed. However, what keeps audiences generating more worlds and more spaces in Minecraft is the incredible ability to simulate large chunks of the conventional world. According to the authors of [14], the simulation technology underlying the game is sophisticated enough to allow a recapitulation of the entire history of technological civilization, from primitive wood and stone tools, through agriculture, right up to factory-sized computers capable of running algorithms and displaying animated outputs on giant screens.

But apart from the ability of the game to synthesize and simulate our world, it has evolved, as we have, into a place where people can exist, something that is becoming far more difficult in the conventional non-cyber world: the game has become about finding, creating and fostering sacred spaces that evoke a sense of wonder, beauty and dramatic naturalism. Maybe Minecraft should be renamed MIND-craft, and maybe that world perfecting mindset is its greatest contribution to our desire to make better virtual worlds.

## References


1. Persson, M.: Minecraft. Game Dev. **18**(2), 24–30 (2011)
2. Duncan, S.: Minecraft, beyond construction and survival. Well Play. J. Video Games **1**, 1 (2011)
3. Short, D.: Teaching scientific concepts using a virtual world-minecraft. Teach. Sci. **58**(3), 55–58 (2012)
4. Sierra, K.: If you build it they will come. Knowl. Quest **46**(2), 42–48 (2017)
5. Minkin, L.: Out of our skins. J. Vis. Art Pract. **15**(2–3), 116–126 (2016)
6. Bereitschaft, B.: Gods of the city? Reflecting on city building games as an early introduction to urban systems. J. Geograph. **115**, 51–60
7. Wilmer, H.H., Sherman, L.E., Chein, J.M.: Smartphones and cognition: a review of research exploring the links between mobile technology habits and cognitive func-tioning. Front. psychol. **8**, 605 (2017)
8. Alter, A.: Minecraft the Island blurs the line between fiction and gaming. The New York Times, 28 July 2017
9. Getting Smart: How Minecraft Supports Social and Emotional Learning in k-12 Education. Getting Smart Research Report. Getting Smart Corporation, June 2017
10. dos Santos Petry, A.: Playing in Minecraft: an exploratory study 1. Rev. FAMECOS **25**(1), 1–18 (2018)
11. Folkins, J.W., Brackenbury, T., Krause, M., Haviland, A.: Enhancing the therapy experience using principles of video game design. Am. J. Speech Lang. Pathol. **25**, 111–121 (2016)
12. Schell, J.: The Art of Game Design: A book of lenses. CRC Press, Boca Raton (2014)
13. deWinter, J., Vie, S.: Games in technical communication. Tech. Commun. Q. **25**(3), 151–154 (2016)
14. Harron, N.: Fully destructible, exploring a personal relationship with nature through video games. Alternatives **40**(3), 16–22 (2014)

# Videogames, Motivation and History of Peru: Designing an Educational Game About Mariano Melgar

Ricardo Navarro(✉), Claudia Zapata, Vanessa Vega, and Enrique Chiroque

Grupo AVATAR-PUCP, Pontificia Universidad Católica del Perú, Lima, Peru
{ricardo.navarro, zapata.cmp, vanessa.vega, echiroq}@pucp.pe

**Abstract.** The use of video games as pedagogical tools that can be used within the classroom has been an initiative that has been in use for a few years with special emphasis since 2000. Therefore, using methodologies and tools that motivate students is an interest of the educational field. However, it is important to delve into the benefits of video games as educational materials that can motivate students. In addition, the motivation within a video game entails that the student can not only remember the facts and characters, but also that they have the initiative to ask or know more about it, which is extremely useful for active learning. This study aims to design a video game that improves students' motivation on the history of Peru. The game focuses on the historical figure of Mariano Melgar, hero forerunner of the independence.

**Keywords:** Human factors · Educational videogames

## 1 Introduction

Videogames are an inherently multidisciplinary field of study, since the very complexity of their development requires the joint work of computers professionals, educators, psychologists, sociologists, among others [1, 2]. Because of this, the construction of a consensual definition of "videogame" is a complex enterprise, since it can be approached from the perspective of the computer specialist (see [3, 4]) as from the perspective of the design of games or the very action of playing (see [5, 6]). Therefore, one possibility of integrated definition can be that which comprises the videogame as an electronic device that fulfills a mediating function of the playful action of the human being [7].

In line with the above, it is easy to see that videogames are the object of study of different disciplines, not only because of their popularity with children and adults, but also because of their proven benefits on cognitive, social and educational development [8–10]. These benefits are possible, among other reasons, because of the immersive characteristics of videogames, which generate a sustained interest of the users in the game, which in turn, allows the player to keep interacting with the device and develop


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 283–293, 2019.
https://doi.org/10.1007/978-3-319-94619-1_28

different abilities. Of course, these characteristics of immersion and sustained interest refer us to an important construct in the field of psychology: motivation.

Motivation is understood as the set of impulses that keep a person performing a certain behavior [11, 12]. These impulses can come from internal or external factors to the subject, so that in function of these the types of motivation can be classified as autonomous and controlled [11].

Autonomous motivation refers to the behaviors performed due to personal factors (I do it because I like it, or because I consider it important) [11]. Controlled motivation refers to behaviors performed influenced by factors other than the person, which can influence in an authoritarian way to force a behavior (I do it because they force me, or because if I do/do not do it, I will receive a reward/punishment) [11]. It is known that the autonomous motivation has the greatest potential to sustain a personal interest in people, which in turn generates positive consequences at the player's cognitive and affective level [13, 14]. Because of this, it is the kind of motivation that a videogame should try to generate.

A term associated to motivation, and studies about videogames, is immersion. The immersion with videogames can be understood as the cognitive, affective and emotional involvement that a person experience when interacts with this type of media [13]. Along with this concept, Deci and Ryan [15] report that an important aspect related to motivation is the feeling of competitiveness that the person experiences in the different activities that he/she performs. This also happens when playing a videogame [14].

There is evidence that aims that if we promote an autonomous motivation, which encourages the player's competition [14], will generate high levels of immersion and even generate Flow in the players. This is something desirable for any type of videogame, because Flow is a generalized sensation that the person experiences when acting with total immersion in an activity [16]. Its distinctive feature is the total involvement in the activity, which implies high levels of enjoyment [16, 17]. Therefore, it is desirable for any type of videogame to generate flow.

Although the differences characteristics of videogames have been studied to better understand their benefits [10], a field of study that has not been addressed –particularly in the educational context - is game design.

Game design is the structure from which a videogame starts to be developed, that is, the guidelines about the mechanics, rules, objectives, narrative, and characters of a video game [18]. The designs of video games, particularly educational ones, have been addressed in different ways [19], and it has been found that each one of these has a particular effect on the player [20]. Thus, mechanics can influence cognitive functions (First Person Shooter), narrative resources can influence knowledge (real-time strategy games, with history), visual effects in immersion (colors and graphics of any video game), etc. [19–21].

An important aspect in the development of educational videogames is to keep in mind the objective of learning in design. However, it is difficult if you do not have the north to guide the process. Therefore, an initiative that can be addressed is the use of thematic axes that serve as guidelines or guides in every decision or step that is made; thus, the use of "creative pillars" that fulfill this function is proposed. The pillars are "concept art" initiatives that can be useful for the design of videogames.

Within the field of study of the use of videogames for education, design must be oriented to the promotion of an intrinsic and immersive motivation that generates flow states in human beings. Thus, from the analysis made to the design created, the following model of steps to follow to design an educational video game is proposed and shown in Fig. 1.

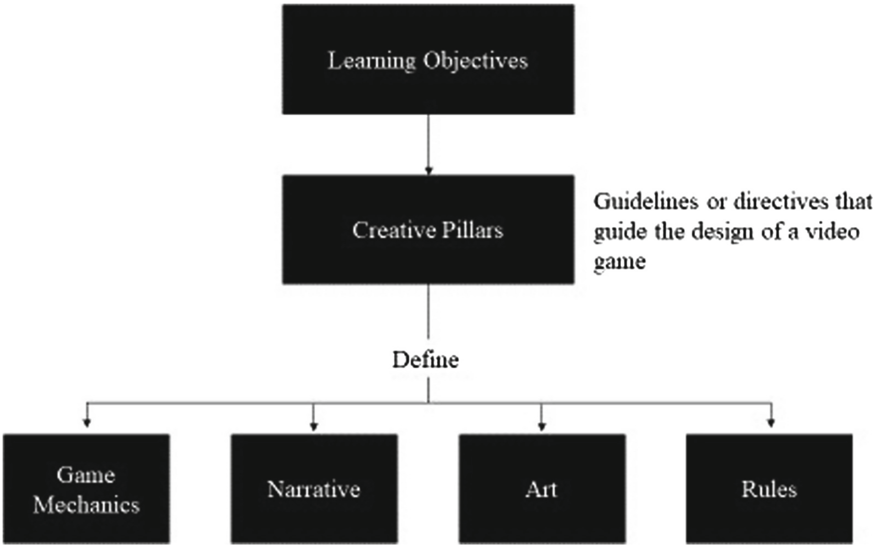



**Fig. 1.** Educational game design model

The learning objective should help to delimit the creative pillars of game design. Thus, they must generate guidelines that help filter game elements that help achieve the learning objective. This process must be considered part of an educational process; the complex part is to gamify that experience in a motivating and immersive videogame. In this sense, the present study seeks to design a videogame that can motivate students to learn history. With this purpose, it is proposed to approach and make a game design proposal based on the evidence collected. Thus, a videogame (in alpha stage) was developed to be used in a classroom. Likewise, class observations were made to identify if the game was being used properly and if the students showed interest or participated in the class. Finally, a focus group was held with five students who participated in the class with the video game.

## 2 Method

Based on the principles of motivation and immersion, a methodology was designed for game design and content development. In the first place, the use of thematic axes (creative pillars) that guided the content of the game using symbols associated with the historical content about the character of Mariano Melgar was established. also, it considered the cognitive and physical development of high school students who are directed the game. Another aspect was the use of colors and a background theme: the nostalgia and melancholy of the main character. Finally, the design based on the thematic axes was executed.

In addition, due to the needs of the project, the game design considered that the game can be incorporated naturally within a class. This was considered pedagogical requirements and learning objectives that helped the design of the video game itself.

After design and development, the video game was tested in a high school class of a public institution in Lima. We counted on the help of teachers to design a class based on the characteristics of the game and the requirements of game design. A focus group was conducted with five students (three men and two women), to identify the positive and negative aspects of the video game.

The participants pointed out that the video game liked and motivated them to continue playing and then to ask questions about the symbols. This proved that designing symbols of the historical contents, framing them as cross-cutting aspects of the videogame, allowed the attention of the students.

### 2.1 Participants

The participants were high school students, from two public schools in Lima. In the first school 23 students participated (9 women and 14 men) and in the second school 14 students participated (9 men and 5 women). For the focus group, there were 5 students from one of the schools.

The students who participated were between 14 and 15 years old and it was considered that during their school years none had received classes on the historical figure of Mariano Melgar, so the content to which they were exposed was new.

### 2.2 Materials

For this study, the following materials were designed:

- Interview guide: An interview guide was designed for a focus group, which focused on identifying if the video game attracted attention and interested the students. Likewise, it was sought to know the interest of this type of initiatives in comparison with the traditional classes and with traditional educational materials.
- Observation Sheet: Class observations were made, where we sought to identify if the teacher generated discussion spaces for the student to participate. In these, it was identified if the following was fulfilled:
  - Students participate in class, answering questions about the historical character.
  - Students make comments on personal aspects of the historical character.
  - Students use the videogame and ask questions about it.
  - Students talk to each other about how to move forward in the game.

### 2.3 Procedure

The GDD (Game Design Document) of an educational videogame about Mariano Melgar was made, following the model proposed from the review of the theory of motivation and immersion in videogames, as well as the principles of teaching and learning. The narrative, mechanics, rules and art were designed according to this model. Then, a videogame was developed in an alpha version that was introduced in a history class. The classes were observed by the researchers, who reported the behaviors of the teachers and students.

Finally, the opinion of the students was collected through a focus group, which lasted thirty minutes.

Details of the areas of the game design are reported are as follow.

**Learning Objective:** The learning objective is for the students to know the historical figure of Mariano Melgar, and his activities during the war for the independence of Peru. To achieve this, it aims to capture the attention and interest of students for the characters and content with which they interact in the game.

**Creative Pillars:** The creative pillars are a creative methodology that allows to filter the elements and contents that will be considered in the videogame. As the goal is that students are motivated to learn from a historical character (Mariano Melgar), it is necessary to rescue the most representative aspects of the character. Everything that is presented in the game must contribute to the learning objective. Based on the learning objective, the areas that should focus on the character that is intended to be taught are determined. Thus, from the information about Mariano Melgar, three important aspects of his life were recognized: (i) Curious Intellectual, where the academic dimension of the historical character is addressed: (ii) Romantic poet, where the romantic and idealistic dimension of the historical character is addressed; (iii) Patriotic hero, which addresses the political dimension and participation for the struggle for the independence of Peru.

**Narrative:** The narrative presents the figure of Melgar from the creative pillars of the videogame: Curious Intellectual, Romantic Poet and Patriotic Hero. The characters, relationships between characters, scenarios, enemies and other elements are inspired by the life, person and work of Melgar; but, in many cases the elements are symbolic of Melgar's life, person and work.

The narrative assumes that the players do not know anything about Melgar, thus, the story is told as the story progresses.

**Game Mechanics:** The mechanics developed for this game are those of a platform game. This allowed the player to better assimilate the symbols scattered in the game environment.

**Art:** The graphic proposal of the game responded to what was proposed by the creative pillars. It was required to have symbols that highlight the image of Melgar as an intellectual (see Fig. 2), poet (see Fig. 3) and hero (see Fig. 4).

**Rules:** The rules define the scenario where the game occurs. In that sense, many of the rules are linked to allowing the player to assimilate the symbols expressed in the art and characters.

**Fig. 2.** Melgar as intellectual


**Fig. 3.** Melgar as poet

**Fig. 4.** Melgar as hero

## 3 Results and Discussion

The objective of this study was to identify if the game design of an educational game, managed to capture the interest of the students, generating immersion in them.

From the results obtained, two (2) of the three (3) immersion levels indicated by Brown and Cairns [22] have been observed in the videogame Melgar: the last dream. First, it is possible to observe the level of Engagement immersion. Next, the appointments of the interview will be exposed where this level is evidenced:

Interviewer (I): "¿What did you like the most?"
Participant (P): "The videogame. The beginning of the videogame."
I: "¿What do you think about the duration of the class? ¿Very long…very short?"
P: "Too short, we wanted to play more. It should be longer."
I: "¿And what would you like to do?"
P: "That there are more levels of the game because it is very short (…)"
I: "¿And did you like the duration that was one hour of videogame and one hour of class?"
P: "Yes. I just wanted to play the videogame only."
(High school students, 14–15 years).

Brown and Cairns [22] point out that, in the first level of immersion, Engagement, the player needs to invest time, effort and attention to learn to play and use the controls. As it can be seen, Melgar: the last dream not only caught the attention of the students, but they invested the necessary time to understand the game, to successfully use the controls and finish it. Not only that, the students were willing to invest more time in the videogame, as they suggested more levels and mentioned that it was "very short".

In addition, it is possible to observe indications of the existence of the second level of immersion: Engrossment. However, this is not achieved because the students expected other elements in the video game.

I: ¿What things would you change in the game?
P: "More levels, more difficulty, multiplayer."
P: "More difficult, with more characters."
P: "The father of his beloved (Silvia) chasing Melgar."
(High school students, 14–15 years).

Likewise, the students compared the differences between this type of approach (with videogames) and the traditional method of teaching.

I: "Which differences did you notice between this session and a normal history class? (about the learning of Mariano Melgar)?"
P: "We would take longer to learn [in a regular class]. Because in the class they leave us books and we must read, instead here [with the videogame] they give us clues."
I: "¿And the graphics? ¿Everyone understood the instructions, what they have to do and were able to locate at the time?"
P: "Yes, it was not difficult."
P: "Yes."

P: “Yes, Yes.”
(High school students, 14–15 years)

Brown and Cairns [22] points out that, to enter this level, it is necessary that the characteristics of the videogame be combined in such a way that the player’s emotions are affected; they also point out that, in this state of immersion, the controls do not represent a difficulty and the player’s mind focuses on the story. As it can be seen, for the players of Melgar: the last dream, this did not happen due to the simplicity of the game, which generates a need for more difficulty in the player. This could be related to what Przybylski et al. [14] and Deci and Ryan [11] refer about the competition in this type of activities. If there is no feeling of competitiveness when performing the action (the game is difficult and challenges my skills to the limit) then no emotional responses will be generated in the player. Also, you cannot get the player to enter a state of Flow. This should be important for future studies and interventions in the design of educational videogames.

Although no evidence of the third level of immersion was found, Total Immersion, Jennett et al. [13] mention that this is rarely achieved and that it is rather fleeting experience. Also, it is possible that the short duration of the game and the context in which it was presented have diminished the chances of reaching this level. In addition to all that has been mentioned, it haven been observed that “Melgar: the last dream”, have positively influenced the academic motivation of its players. Next, the quotes of the interview where this influence is observed will be exposed

I: “¿How are your history classes here at school? ¿What do you think of them?”
P: “Well, interesting, pure theory. Pure work notebook,
copy the blackboard.”
I: “¿And what do you think of the way the game was integrated into class?”
P: “Well, because not only here we shared ideas, but also in the videogame they gave us clues to develop our own ideas. It was beneficial to the questions that they gave us.”
I: “¿Do you believe videogames are a useful tool for learning?”
P: “Yes, because we play and learn at the same time. Further, there is not like in those games that there is a space to be able to skip history, moreover we enjoy it.”
I: “¿Would you like the videogame to be in your history classes?”
P: “Yes, because you would learn a lot, it should be like that with each hero of Peru. More time to play, more characters.”
I: “¿As in a normal history class?”
P: “No, we would take longer to learn. Because in the class they leave us books and we must read, instead they give us clues here. When we are reading we get bored, words are new, and we cannot identify them.”
(High school students, 14–15 years).

González-Pienda et al. [23] mention that one of the dimensions of academic motivation is the expectation component, which includes “perceptions and individual beliefs about the ability to perform a task”. In this sense, we can observe that the students do not know and cannot identify the new words presented in the books or readings given in school, which could suggest that they do not feel capable of facing these tasks. However, the videogame allows students to feel more capable of performing specific tasks, because they mention that “here (in the videogame) they already give us clues”, that is, the game does not give explicit answers to the students, but rather gives them the pertinent help so that they can face these tasks and feel capable enough to solve them.

Likewise, [24] point out that another dimension of academic motivation is the affective component, in which Pardo and Alonso Tapia [25], cited in [24] suggest that people seek to understand and discover the causes of things that happen, the casual attribution that we give will trigger a series of affective. In this sense, “Melgar; the last dream” could positively influence the student’s attribution of success to the questions presented to himself, since the videogame “gave clues” that would allow him to develop his own ideas. Therefore, “Melgar; the last dream” could be a useful tool for students to solve specific tasks in a successful way, and to attribute that success to themselves.

This qualitative analysis was complemented by a class observation, to identify if there were interest in the contents presented in the videogame.

For this observation, there were two qualified professionals. In the observation, it was necessary to identify if the teacher allowed participation spaces for the students, as well as the student’s interactions during the use of the videogame. It was observed that the teachers allowed the active participation of the students, who made questions and observations to the contents of the game and the historical character (Mariano Melgar).

Also, it was found that, during the game, the students interacted with each other and asked questions about the events and symbology presented in the game, which was expected.

From the results, it can be identified that the game design had the expected effect on the students. This study allows to open the field of study about game design for videogames in educational contexts. It is hoped that, in this way, the playful content of an educational videogame will not be lost. On the contrary, it is expected that the model proposed for this study will be useful for the development of immersive educational videogames that motivate students.

## 4 Conclusions

The participants recognized that the videogame maintained their attention and interest for what happened in the class. They continued to play and learn more about the historical figure from the history of Peru (Mariano Melgar) and asked questions about what was happening. The teacher in charge of teaching a class with the videogame took advantage of this interest in the students to explain more about the historical figure.

The symbols and strategies used to design the videogame proved to have the desired objective by the designers and, in addition, fulfilled what was expected according to the theory on motivation and educational videogames.

## References


1. Henríquez, F.P., Zúñiga, T.A.: Hacia una conceptualización de los videojuegos como discursos multimodales electrónicos. Anagramas Rumbos Sentidos Comun. **15**, 51–64 (2017)
2. Salazar Oré, V.: Influencia de los videojuegos violentos y prosociales en la conducta de ayuda (2015). http://tesis.pucp.edu.pe/repositorio/handle/123456789/6271
3. Prensky, M.: Digital natives, digital immigrants part 1. Horizon **9**, 1–6 (2001)
4. Juul, J.: A dictionary of video game theory. Online Verfügbar Unter Httpwww Half-Real Netdictionary Zuletzt Geprüft Am. 1, 2010 (2005)
5. Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Grau, V., Lagos, F., López, X., López, V., Rodriguez, P., Salinas, M.: Beyond Nintendo: design and assessment of educational video games for first and second grade students. Comput. Educ. **40**, 71–94 (2003)
6. Mitchell, A., Savill-Smith, C.: The use of computer and video games for learning: a review of the literature (2004)
7. Bavelier, D., Green, C.S., Han, D.H., Renshaw, P.F., Merzenich, M.M., Gentile, D.A.: Brains on video games. Nat. Rev. Neurosci. **12**, 763–768 (2011)
8. Kato, P.M.: Video games in health care: Closing the gap. Rev. Gen. Psychol. **14**, 113–121 (2010)
9. Primack, B.A., et al.: Role of video games in improving health-related outcomes: a systematic review. Am. J. Prev. Med. **42**(6), 630–638 (2012)
10. Granic, I., Lobel, A., Engels, R.C.M.E.: The benefits of playing video games. Am. Psychol. **69**, 66–78 (2014)
11. Deci, E.L., Ryan, R.M.: The "What" and "Why" of goal pursuits: human needs and the self-determination of behavior. Psychol. Inq. **11**, 227–268 (2000)
12. Pintrich, P.R.: A conceptual framework for assessing motivation and self-regulated learning in college students. Educ. Psychol. Rev. **16**, 385–407 (2004)
13. Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. Int. J. Hum.-Comput. Stud. **66**, 641–661 (2008)
14. Przybylski, A.K., Rigby, C.S., Ryan, R.M.: A motivational model of video game engagement. Rev. Gen. Psychol. **14**, 154 (2010)
15. Ryan, R.M., Deci, E.L.: Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. **25**, 54–67 (2000)
16. Csikszentmihalyi, M.: The contribution of flow to positive psychology. In: The Science of Optimism and Hope: Research Essays in Honor of Martin E. P. Seligman, pp. 387–395. Templeton Foundation Press, West Conshohocken (2000)
17. Acha Abusada, M.P.: Necesidades psicológicas básicas, motivación y flow en estudiantes universitarios de arte (2014)
18. Salen, K., Zimmerman, E.: Rules of Play: Game Design Fundamentals. MIT Press, Cambridge (2004)
19. Dickey, M.D.: Game design narrative for learning: appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. Educ. Technol. Res. Dev. **54**, 245–263 (2006)
20. Clark, D.B., Tanner-Smith, E.E., Killingsworth, S.S.: Digital games, design, and learning: a systematic review and meta-analysis. Rev. Educ. Res. **86**, 79–122 (2016)

21. Young, M.F., Slota, S., Cutter, A.B., Jalette, G., Mullin, G., Lai, B., Simeoni, Z., Tran, M., Yukhymenko, M.: Our princess is in another castle: a review of trends in serious gaming for education. Rev. Educ. Res. **82**, 61–89 (2012)
22. Brown, E., Cairns, P.: A grounded investigation of game immersion. In: Extended Abstracts on Human Factors in Computing Systems, CHI 2004, pp. 1297–1300. ACM, New York (2004)
23. González-Pienda, J.A., González-Cabanach, R., Núñez, J.C., Valle, A.: Manual de Psicología de la Educación. Pirámide (2002)
24. González-Pienda, J.A., Núñez, J.C., González-Pumariega, S., Roces, C., González, P., Muñiz, R., Bernardo, A.: Inducción parental a la autorregulación, autoconcepto y rendimiento académico. Psicothema **14**, 853–860 (2002)
25. Pardo, A., Alonso-Tapia, J.: Motivar en el aula. Ediciones Univ. Autónoma Madr. Cantoblanco Madr. (1990)

# A Study to Improve Education Through Gamification Multimedia in Museum

Yidan Men[1(✉)], Robert Chen[2], Nick Higgett[2], and Xiaoping Hu[3]

[1] Art and Design, Guangdong University of Finance and Economics, Guangzhou, People's Republic of China
merdoor91@163.com
[2] Faculty of Art, Design and Humanities, De Montfort University, Leicester, UK
[3] Design School, South China University of Technology, Guangzhou, Guangdong, People's Republic of China
Huxp@scut.edu.cn

**Abstract.** The purpose of this paper is to study a game-based multimedia design promote museum education. The museum is an important contributor to the social education function, which should build a modernized public cultural service by using modern technology and new environment. The research methodology mainly adapted on three approach. (1) A literature review of the relationship between people and museums, digital media education and game-based learning. (2) An observation of the target groups behavioral habit in the in specific British Museums. (3) A summarize of the particularity and the popularity of the group. This paper focuses on three aspects to evaluate the practical significance of the various theoretical results and discussed these theories permeate the game-based multimedia design development in the museum.

**Keywords:** Gamification · Interactive multimedia · Usability
Museum education · Children and elderly · User experience

## 1 Introduction

### 1.1 Research Background and Motivation

Museums as an educational institution to transmit social cultures which are assumes a major responsibility as a public facility for the mass and can demonstrate the defects and problems of education at cultural level. The educational concept currently has limited energy and economic resources. Therefore, the role of the museum for education is more important than ever before. Museums are a form of theatre in which the varieties of human experiences and the complexities of the world are staged in ways they find intriguing and helpful [1]. Museums act as a high quality environment for public to both learn and relax in. Specially for children and elderly, which can absorb the knowledge within a museum through a variety of senses and in a natural environment to understanding the world. Digital media can enhance user relationships with the museum whilst also learning cognitive and social skills [2]. Some museums have


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 294–304, 2019.
https://doi.org/10.1007/978-3-319-94619-1_29

complex exhibition space with a mass of information about exhibits, which requires the audience to be receptive to understand. But for children, they only want to focus on more interesting things such as the digital screen and games and interactive items.

### 1.2 Aims and Objectives

The research aim is to effectively inform the audience of each museum exhibit. This would be particularly so for two specific groups, being the elderly and children. It is obligatory to encourage interaction between children and the elderly, to cultivate and stimulate the interest of children to explore the world they live in. Moreover, to enhance the museum directions and exhibits would help people to visit and view the museum in a more efficient manner.

The objectives to achieve these aims are:

- To explore the relationship between the museum and the user.
- To investigate the effect of multimedia design within the museum.
- To produce background research through literature research.
- To investigate current digital design within the museum.
- To identify the main problems of visitors and the museum.
- To research audience behaviour upon their visit to the museum.
- To identify the main sequential flow of the exhibition.
- To analyse effective information according to observation.
- To perform critical reasoning of the existing design.

### 1.3 Relationships Between Children and the Elderly Within the Museum

Since the last century, museums began to open their doors to children. Children are an important part of museum visitors and need a different way in which to view the museum compared to adults. Children see and understand the world much differently to adults. Children learn through touching, feeling, watching, listening and imitating. “In favourable circumstances, the museum is a place where the child can begin to understand the world in which he lives” [2]. The visual world is often intriguing to children because of its vivid colours and shapes. Visiting educational museums can help children to create a fundamental framework in the development of their minds. Many parents realise how important it is for their children to learn knowledge through museums. It has become an important part for th museum designer in considering the needs of children and the best ways for the museum to engage with children. However, most museums must start with the cooperation of schools and community organizations. Unfortunately there are still limitations for children to visit museums because of protocol, such as requiring an advanced reservation for the programme and being able to adapt to museum running schedules. There are convenient solutions in which to create more opportunities and multiple-choice within museums for children to visit by themselves. This would enable them to explore and discover the world without staff guidance, (for example manual guidance), through interactive games and animation displays.

The increasingly large group of elderly within society has brought huge challenges for the social services across the world. Museums, as an important cultural educational institution, should satisfy the needs of spiritual culture across different levels and in various forms for the elderly, for example, knowledge, understanding and aestheticism. Aged audiences are those of active participants and have a keen interest in museum exhibitions and other various activities. Whilst shouldering the responsibility of educating the next generation and scientific cultural knowledge instruction, the elderly could be exploited through the museum as having profound knowledge of reserves and rich life experiences. According to various activities, it could increasingly become a form of communication with the next generation and a sense of achievement for the elderly. In addition, it can be a form of continuous education for elderly people after they have retired.

There are a number of museum educational features that should be highlighted.

- The main roae of the museum is to attract the audience based on the exhibitionitself. The exhibits could composite visual images for visitors.
- Museum education is an intuitively reliable witness display of natural and human history, which is hard to replace by any other form of education.
- The characteristics of museum education are professional, sociable and popular. The museum environment is seen according to the themes which include the arrangement of the exhibition, the explanation of exhibits and interior decoration. Museum education should satisfy the requirement for service, cognition and the understanding of society.

### 1.4 Education Digital Media Within the Museum

Education is one of the most important functions of the museum. Traditional museums produce exhibition designs whereby the thinking mode is limited by technology, the concept and the exhibition site, which leads to more and more limited space and presentations. As a result, the consequences of the exhibition are less than satisfactory. Traditional museums only provide simple illustration labels of exhibition items which only show visitors in a visual way, making this a one-way style. It is hard for the audience to have a resonance when the information is shown in such a formal way without allowing relevant responses, which in turn gives the audience only a brief and rough visit without leaving any deep impression. More and more museums are starting to use media interaction design to transform the static visit guide, producing a multiple-sense interaction which mainly focuses on historic stories [1]. The way in which the exhibit information is usually conveyed in a single line, is transformed into information that will prompt the visitor whereby mutual consciences are exchanged. This method is committed to make the exhibition content and form unified and to enhance the communication effect, so as to keep the interaction and design in accordant. Present digital media museums most commonly exhibit ways in which to include virtual reality and thematic displays. The visitors not only enjoy a feeling of being personally on the scene, but can also have interaction during their visit through the emulated environment based on VR technology. The digital interaction equipment is featured with the capacity to get visitors easily immersed. The data regarding the museum itself, exhibition information and service contents can also be shown on the exhibition items. Various types of information can be broadcast to visitors in a timely and broad manner, which will encourage and enrich a deeper form of information across all parts of the exhibition hall. The museum visit can produce a scene display that is a process in which to lead and immerse visitors as being the principal part of the scene. The combination of scenes and the emotions related to it is an important factor in the design of multimedia interaction. Educational multimedia products require more thoughts on the learning environment for children [3]. On the part of design, humans are set to be a part of the environment, which allows visitors to choose the information they need under the multiple sensory interactions.

### 1.5 Gamification and Game-Based Learning

Development of the learning system of the future is not to use digital learning, but to change the ways in which people learn [4]. To combine education with games will change the conventional wisdom of learning drudgery. "Educational games" are defined as an instructive computer game program that cultivates the players' knowledge, skill, intelligence, emotion, attitude and values.

Interaction devices have a special educational meaning for children, such as touchable screens, video displays and games. Games can enhance the relationship between the user and the museum whilst at the same time learning cognitive and social skills [5]. Through the game mode and its nature, the museum exhibition is more normative and scientific, which can help visitors obtain good spiritual experiences such as interest, pleasure, motivation and successful learning. During the games, the players both cooperate and compete with each other. They win self-confidence and sense of accomplishment when finishing a task; they gain recognition and praise from companions in the interpersonal interaction; they will increase wisdom and improve skills in the process of finishing the tasks.

Game playing is not just to allow children to learn in a relaxed and pleasant environment through inspiring their interest of participation, but more importantly, the fun of the game can be used to develop a child's exploration spirit. The designer is required to have a full understanding of the exhibition items, the environment, historical and cultural background and to allow the visitors to explore things in their own way, to use their own thinking and creativity and to develop social contact and cooperation abilities. Therefore, the game design and planning is suggested to focus on user experience and perception so as to improve the engagement and participation of users and achieve the goal of gamification. Participation, relevance, interactivity and openness are highlighted from the two aspects of content and form. Therefore, the game design and planning is suggested to focus on user experience and perception so as to improve the engagement and participation of users and achieve the goal of gamification. Participation, relevance, interactivity and openness are highlighted from the two aspects of content and form.

### 1.6 Emotion Design Within in the Museum

Emotional design is a main point to be considered within interaction design, which is clearly and intuitively related to the view of the world and values of people. This is an essential role in which people can learn and understand new knowledge. Basically, emotion design is divided into three different levels, visceral, behavioral and reflective [6]. The visceral of emotion design is that of product appearance, which is related to the human sense of feeling and comprises of sight, smell, hearing and touch. In order to design a new product for the public, the appearance should conform to the aesthetic public demand. The meaning of behavioral design is that of function. Design is focused around life and can improve it. Most design is there to solve the inconvenient factor of daily life, and has become an existing necessity. It will require the designer to concentrate on the fluent and natural for their product. The last function of emotion design is reflective, which is a very important factor of personal feelings and thinking. The designer should pay attention to information, culture and effects of the product. For example, something prompted by personal memories, self-image and information can convey one to other. According to reflected experiences, knowledge and culture can establish one's personal status within society.

For the design of museum, the designer should immerse emotion into the product design, which by transition and transformation will develop the visceral, behavioral and reflective to be as perfect as possible. This is achieved through translation of the existing elements into the user demand. Colour is an important factor of the emotional transfer, as it can extract the expression of feelings by the audience. Colour has symbolic and national characters which can easily cause corresponding sympathetic chords and cultural association with the audience. Through the design of the museum, different colours can affect the emotions of people and have an impact on their first impression of the exhibits. The overall tone of the exhibition should have the correct colour choice which can create an appropriate atmosphere for the exhibition and to encourage the interest of the visitor. Most of the exhibition should portray a quiet and harmonious colour to avoid boredom for the visiting audience. However, the design for children should adopt big and bright colours to make it attractive to them. In addition to expressing emotional design through colour, it can also impact on the perspective of experience. It should clearly identify the requirements of the main user groups and fully inspire the instinct and intuition of the audience. Based upon User Center Design (UCD) methodology, priority should be given to the perspective of the user to deepen the impression of the product [7]. Therefore, research regarding the physiology and mentality of the user is a necessary process before the design is started. Confirmation of the target groups, analysis and the actual demand from target users is ultimate information required prior to the start of the project. As part of the help the designer create a suitable interaction design for users.

## 2 Method

The methodology is divided into five main aspects, namely secondary research, primary research and discussion.

1. Secondary Research: This is mainly focused upon in-depth literature research to gain an overview of the relationships between the elderly and the museum, children and the museum and digital principles. All gathered information is obtained from professional books, journal articles, newspapers and authoritative websites.
2. Target group: The age range of the target group is above 50 years old (elderly) and under 11 years old (children). This is the general retirement age for older people and the 'eager for knowledge' age for children.
3. Primary Research: To investigate the regularity of target group cognition, experience, behavior and environmental influence between museum interaction activities. To Define the interactive experience requirements of target groups in museum and find the main factor of Information dissemination according immersive interactive experience in museum.
4. Non-Interventional Observation: As a qualitative study, the observation will be recorded for Motion path, Activity behavior and interaction feedback from target group in public. This method will develop with standard operating procedures (SOP) to complete and collection date, to ensure the credibility, scientificity and standard of information collection.
5. Discussion: Analyze the data of the target group interactive process in the museum, and find out the general characteristics and pain points of the interaction experience between target group and museum environment.

## 3 Participants Selection

The target population focus group related to elderly people taking their children to the museum. Museums work as a public educational organization whereby children can learn a vast variety of knowledge from. According to participants the activities could inspire children's imagination and creativity, but at the same time appointments for activities are difficult for the elderly. The target population for this project is placed upon the elderly and children who visit museum exhibitions without staff to help or guide them. The age range of the target group is above 50 years old (elderly) and under 11 years old (children). This is the general retirement age for older people and the 'eager for knowledge' age for children. The intended customer profile would characterise two different target groups on the basic background and requirement of the museum facilities. The customer design preferences are an important factor to influence the final design outcome. According to the museum observation, the elderly are mostly more patient than younger people to discover and understand the detailed information of exhibition objects. They are much more willing to explore and share their views with others. The role of the elderly toward children is that of an interpreter. The entertainment element of this project is important in attracting the attention of children, through colourful displays and identification of shapes; therefore colour is an important requirement. Reading and safety risks should be avoided for both the elderly and children. Moreover, more attention is required by the elderly if their children are playing in a dangerous facility within the museum.

# 4 Primary Survey

The observation research will focus upon the sequential audience flow, behaviour of visitors and the design of exhibition space. In order to discover and analyse existing main problems within museums, the designer has carried out a detailed investigation of the Ancient Egyptians Exhibition in New Walk Museum. It includes the number of existing items, the audience visit line and the main products that receive the attention of visitors. The length of observation is four hours, respectively across the Wednesday, Friday and weekend, which is a very busy time for the exhibition. Visitors observed amounted to approximately 52, including a group of primary school students. Effective age research groups amounted to 11 of people over the age of 50 (elderly) and 27 under the age of 11 (children). Other numbers equated to family tourists without elderly people. In order to make convenient the observation and analysis of data, a SOP table will follow.

## 4.1 Observation Subject

The observation aims are to understand the sequential audience flow within the museum and the effectiveness of the existing exhibition space.

- To be aware of possible problems that may show up in the exhibition.
- To carry out a SOP before observation of the museum.
- To research audience behaviour upon their visit to the museum.
- To identify the main sequential flow of the exhibition.
- To analyse effective information according to observation.
- To summarize the observation research.

## 4.2 S.O.P

Standard operating procedure (S.O.P) is an important process in the preparation of observation which must be considered prior to observation taking place. The observation places the focus on target groups that include people over the age of 50 years (elderly) and under the age of 11 (children). The research starts upon entry to the museum and the time spent reading the introduction information about the exhibition of Ancient Egyptians. The next step is to observe the audience visiting process. Basically, this is whereby the audience plays focus upon the exhibit description, looks at the items and take photos. In particular, the process of reading the exhibit caption is a most important part of the observation. Whether visitors prefer to read digital media or traditional labels to understand the information will affect the design development process. This will allow more consideration to be applied to the way in which to attract the audience more toward the exhibition. Moreover, an interesting finding of the research related to the break time required by visitors. It discusses how many visitors would like to stay within the museum to relax and communicate ideas with their partner, or to leave the exhibition without any communication ideas. This ‘break time’ could show the way to audience exchange of views, which could influence the story of design required making it attractive enough for visitors. Finally, to research the possibilities of a gift shop should the exhibition be attractive enough for visitors to want to make purchases (Fig. 1 shows an example).

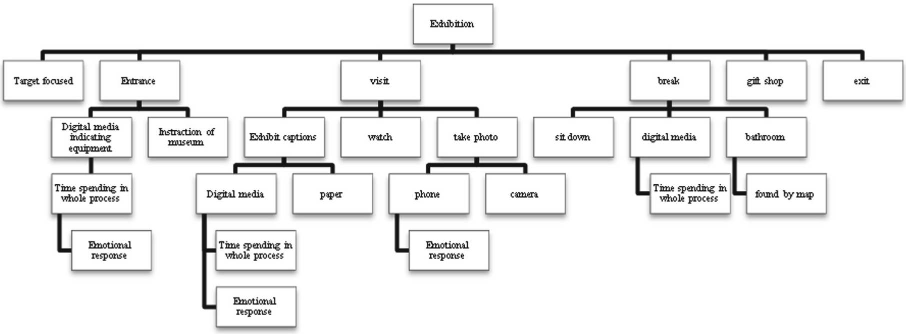



**Fig. 1.** S.O.P of observation.

## 5 Results Analysis and Discussion

According to observation carried out, there are more than 16 items on display in the exhibition (Fig. 2 shows an example). Because the exhibition space is narrow and dark, the label content is not easy to read and understand. The Ancient Egyptians Exhibition still uses traditional paper labels to display the information. However, the designer has created new content labels that are fun and interesting. The lighting problem has a huge impact on the ability to read any information. Many visitors prefer to change their idea to guess the meaning of the exhibit with their friends and family. The museum has created four special individual spaces for their audience, which can be seem in Fig. 2, numbers 2, 8, 9 and 13. In the centre of the museum are numbers 2 and 9, whereby there are a few chairs for visitors that are placed behind the touchable exhibit. However, this space is largely ignored by visitors because the display item cannot move nor shows any difference between other exhibits. The visitors are more interested in the model of a mummy, number 13, which has a lesser description in relation to the history of the mummy. Number 8 is an interactive space to demonstrate the life of ancient Egyptians. Interactive spaces could enhance the common exhibition items for the audience, but unfortunately the space and facility have fallen into disrepair because the overall outlook appears in a dilapidated condition. Most visitors are not interested in this area, the exception being children who will play in this space.

The concept of sequential flow is originated in the architectural design, which refers to the behaviour of the mobile trajectory within the space, based on personal behaviours. It is important to ink activities of people to satisfy visitor demands in the pattern of space. This can be expressed as the embodiment of function requirement. The sequential flow of the museum design is placed on the behaviour of people and combined with scientific theory to combine and produce the outcome.

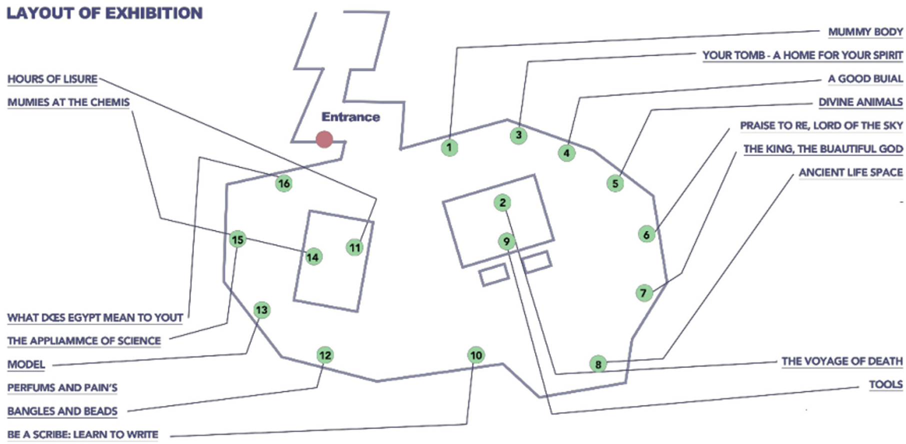



**Fig. 2.** Layout of exhibiton.

Based upon the observation, there are large numbers of exhibits that are displayed in a small space. The museum is passive in its direction of people viewing the exhibition. The role of the museum should be one to provide guidance to visitors to make their visit comfortable and easy. The situation in the New Walk Museum shows the main route for visitors is made clear, but some visitors may be attracted by separate stages (Fig. 3 shows an example), but there is no signage or instructions to lead visitors correctly through the exhibition of Ancient Egyptians. Museums should develop public service functions to lead visitors to their museum.

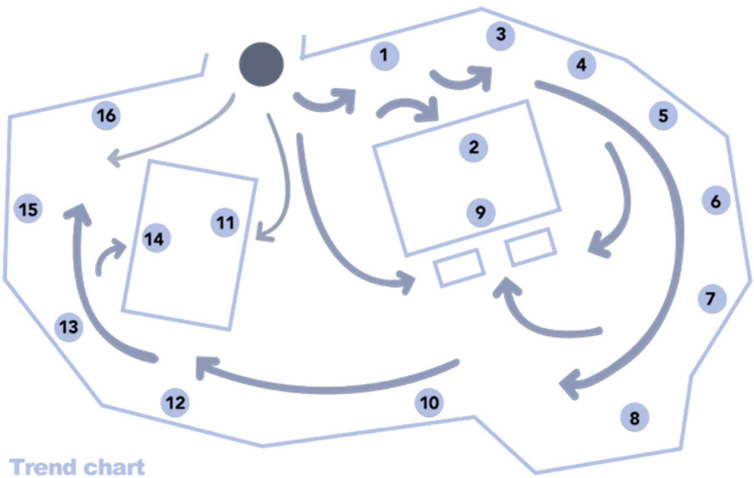



**Fig. 3.** The sequential flow of exhibition.

Results from the observation suggest that the data may be useful for future design development. The results are show below:

1. Correct education concept and sense of worth

For the interactive exhibition design based on learning-through-play ideas, games come as just thread to link scientific and technological knowledge. It is knowledge within the exhibition hall that should become the focus of attention. While designing games, we should make the best of connections among different items displayed in the hall and give full play to the potential that scientific and technological knowledge have to games. To some extent, gamification products are guiding and persuasive. Therefore, the design of games should be based on exhibition information of museums themselves and guide visitors with correct educational views and values. This way, unconsciously visitors will be properly educated while playing games.

2. Realised the initiative of visitors

During the interaction, a participant of the game needs to completely immerse himself in the museum scene, give free rein to his imagination and apply his mind and hand before his independent creativity is unleashed. In this process, the exhibition staff should restrain themselves from prompting to allow a natural learning and a full exploration of the flexibility, curiosity and spontaneous impulse of the visitors.

3. Create a passive visit within the museum

Visitors should be able to focus on objects and contents without being concerned about the route of visit. The designer should present the exhibition information and route directions easily and conveniently during the visit.

4. Provide logical information

The exhibition introduction should require clear logic to avoid repetition or missing exhibition objects.

5. Reasonable exhibition space

Ensure sequential flow is unimpeded to avoid convection and repetition, based on the exhibition information. In order to avoid traffic congestion and collision with different audient flows, there should be only one guide flow throughout the exhibition space.

## 6 Conclusion

Through deeper research for this major project design, the role of education within the museum is to display diverse and various multimedia and facilities to show the exhibits. It could enhance the impression and knowledge for the audience, particularly to improve the relationship between the museum and visitors. An interactive design can offer visitors the opportunity to share experiences and ideas. For the elderly and children, gamification could help to increase communication and encourage children to learn about the exhibition history and culture.

To conclude, the interactive multimedia design for a museum is a new concept which will transform the way traditional information is conveyed. It is a necessary development for museum education design in the future, and will give people a deeper consideration toward the exhibition environment and user emotions. Although the project was successfully completed on time, according to the research and design results the project design can still be improved and extended.

**Acknowledgements.** This work was supported by a grant from the Innovative Talents Training Program of South China University of Technology (Grant No: yjjg2016004).

## References

1. Roppola, T.: Designing for the Museum Visitor Experience. Graphics, USA (2012)
2. Winstanley, B.R.: Children and Museums. A.R. Mowbray & Co., Oxford (1967)
3. Christian, S.: Exhibitions and Displays. Birkhäuser GmbH, Swiss (2009)
4. Robertson, J.: Usability and children's software: a user-centered design. Comput. Child. Educ. **5**(3–4), 257–271 (1994)
5. Prensky , M.: From digital game-based learning. Comput. Entertain. **1**(1), 21 (2003)
6. Norman, A.D.: Emotional Design. Basic Books, New York (2005)
7. Abras, C., Maloney-Krichmar, D., Preece, J.: User-centered design. In: Bainbridge, W. Encyclopedia of Human-Computer Interaction. Sage Publications, Thousand Oaks (2004)

# Game Design Methodology, Usability and Applications

# A Usability Review of the Learning Master Serious Game in Support of the US Army Jumpmaster's Course

Tami Griffith[1(✉)], Crystal Maraj[2], Jeremy Flynn[2], and Jennie Ablanedo[2]

[1] Army Research Laboratory, Simulation and Training Technology Center, Orlando, FL, USA
tamara.s.griffith.civ@mail.mil
[2] Institute for Simulation and Training (IST), University of Central Florida (UCF), Orlando, FL, USA
{cmaraj, jflynn, jabaned}@ist.ucf.edu

**Abstract.** The use of serious games for training can promote engagement and encourage curiosity. While there is controversy over how to assess the effectiveness of using serious games for training, research efforts have linked instructional effectiveness to instructional objectives and game characteristics. This paper describes these components in relations to a serious game for students in the United States Army Jumpmaster course. The authors collected data on usability through user feedback from surveys after the use of a serious game called Learning Master. From the data collected, users perceived the serious game to be simple to use, function appropriately, and have an intuitive design. A majority of the users expressed confidence in using the system. The user responses called for improvements in the visual display and simplified access to the application. The design team plans to use this information to enhance and finalize the serious game for future testing.

**Keywords:** Serious games · Jumpmaster · Game for training
Instructional effectiveness · Game characteristics for learning · Learning master

## 1 Introduction

The past decade or so has seen a major growth in the use of games for training, also termed *serious games*. Serious games provide training in healthcare, in corporate training, within academic institutions, and within the military (e.g., America's Army, DARWARS Ambush and Full Spectrum Warrior [1]. A serious game is a game that primarily serves to educate or instruct [2, 3]. Unlike traditional training methods such as classroom-based instruction, which rely solely on the instructor, or simulation-based training, which is a replication of a real-world scenario, serious games act as an instructional approach with a defined learning objective using game elements to promote active learning [4, 5]. The authors of this paper believe that serious games can complement instruction by implementing meaningful serious game characteristics, such


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 307–317, 2019.
https://doi.org/10.1007/978-3-319-94619-1_30

as, competition, goals, rules, challenging activities, and choices [6]. Different combinations of these characteristics can innovate game design to promote learning.

A sample of literature on the training effectiveness of serious games [7] reveals a lack of consensus. Hayes [8] found that research in the use of serious games for training is often fragmented and enigmatic. One attempt to explain the training effectiveness of serious games focused on curriculum content to illustrate its benefits for promoting engagement and curiosity [9]. Other research focuses on empirical studies, such as usability studies, playability aspects, and performance outcomes to measure the effectiveness of serious games [10]. Despite the inconsistent viewpoints, there is a consensus that serious games are entertaining-yet-educational, with the purpose of improving learning. A technical report on instructional games suggests that effective learning requires congruency between instructional objectives and game characteristics (e.g., competition, goals, and rules) to produce instructional effectiveness [8] (see Fig. 1).

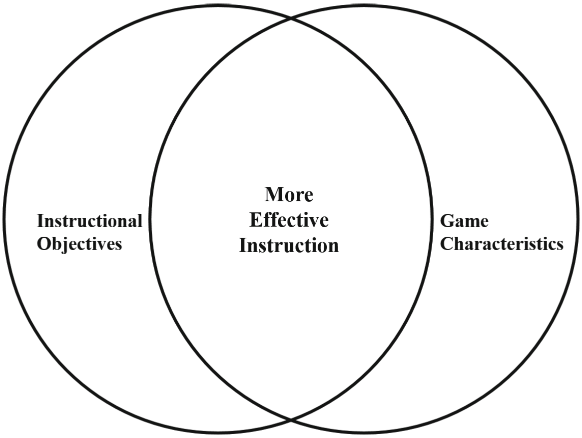



**Fig. 1.** A Venn diagram illustrating instructional effectiveness as an overlap of instructional objectives and game characteristics. Re-created from Technical Report by Hays [8].

## 2 Problem Statement

The U.S. Army Research Laboratory Simulation & Training Technology Center (ARL STTC) designed and implemented a serious game, dubbed Learning Master (LM). The Learning master functions in much the same way that the commercial games Trivia Crack and Quiz Up work, with the key difference being that the instructor has access to the server-side data. This access allow the instructor to add, edit or delete questions on the fly and to know at any time what questions are challenging or ambiguous to the students so that the information can be discussed in face-to-face meetings. The application (also referred to as a serious game), uses competition to encourage exposure to questions. Students can play on their own or challenge other students. Responses are timed and the best performers show up on a leaderboard. The instructor can recognize the leaders in the class. This provides additional reinforcement in using the application. The more students are exposed to the learning material, the better their performance in an assessment [11].

The development of the serious game answered a challenge from the Leadership of the Maneuver Center of Excellence, who sought to enhance student performance in the Jumpmaster Course. A Jumpmaster has the responsibility of teaching students the tactics, techniques and procedures associated with jumping from an airborne vehicle. Specific responsibilities of the Jumpmaster include “rigging individual equipment containers and door bundles and inspection of all equipment” [12]. The Jumpmaster course trains personnel “in the skills necessary to jumpmaster a combat-equipped jump and the proper attaching, jumping, and releasing of combat and individual equipment while participating in an actual jump”. [13]). Students enrolled in this course undergo a three-week, intense training curriculum. During the first week, they take a written exam, the current pass rate of which is 70%. The LM application was applied to promote learning effectiveness, and improve the pass rate without lowering learning standards. The LM application works on the student’s own Android and iOS device. The student is able to download the application and use it to learn the content for the nomenclature exam. The nomenclature exam tests familiarity on the required equipment to perform a jump. The LM application leverages characteristics from serious games to create an engaging, entertaining, and competitive tool that supports classroom-based instruction. The training application went through an iterative design process incorporating serious game elements. The focus of this paper assesses the initial user assessments of the prototype’s usability, and offers design considerations for subsequent updates.

Past research methodologies concentrated on summative assessment (e.g., pre-post-test) to assess user performance [14] and by association, instructional effectiveness. This proves advantageous for measuring learning directly through serious game applications. The Systems Usability Scale (SUS) [15] was used to evaluate the user’s responses. This paper examines the feedback from user’s responses on this scale to the Learning Master (LM) application.

## 3 Design Recommendations

The design requirements established by the team of engineers, scientists, and artists, with feedback from the training community, are illustrated in Table 1. The table presents the design categories leveraging characteristics of serious games and instructional objectives for inclusion into the LM application to improve instructional effectiveness.

**Table 1.** The LM application - design recommendations

| Design category | Recommendations |
|---|---|
| Challenge | *Goals*<br>• Game goals must be consistent with the instructional objective<br>• Performance feedback should be provided soon after the student achieves the goal<br>*Competition*<br>• The game should provide complexity and promote learning through competition<br>• Competition would be achieved using live opponent or using a criterion score<br>•Game scores must be related to the goal of the game |
| Curiosity | *Media*<br>• The LM application should include audio and visual effects to promote instructional content and reinforce learning<br>• Elements of informational complexity should stimulate learning and reduce boredom |
| Instructional quality | *Instructional structure*<br>• Game elements must provide foundational knowledge structure to reinforce pattern recognition learning<br>• The students should be provided with feedback to understand inconsistences and adjust learning based on corrections<br>• Debriefing and feedback should be provided by the classroom instructor as an additional layer of instruction and to support the instructional objectives |
| Instructional presentation | *Instructional strategy*<br>• A leaderboard serves as instructional strategy to present the competitive criterion scores<br>• Achieving a criterion score should be a factor for advancing to a higher difficulty level or starting a new game |
| Logging capabilities | *Data capture*<br>• Logging capabilities include capturing the number of correct and incorrect responses, questions frequency missed, response time, and the duration of time in the application from start to finish and allow the instructor to provide in-class feedback |
| Training platform | *Commercial mobile platform*<br>• A platform that enables rapid development across Android and iPhone mobile devices. The platform should be able to leverage platform specific hardware acceleration and allow for compiling native performance |

# 4 Method

The participants, instruments, surveys, material, and procedure are included below.

## 4.1 Participants

The study was conducted at Fort Benning Columbus, Georgia. A sample of 63 participants were recruited from an existing Jump Master class. There were 60 males and 3 females, whose ages ranged from 23 to 49 ($M$ = 30.8, $SD$ = 5.8). The participants received no monetary compensation for participation in this study.

### 4.2 Instruments

To understand the feasibility of the LM application, perceptual data was collected in the form of surveys

**Surveys**

*Demographics Questionnaire:* This questionnaire captures biographical information, such as age, gender, education level, and experiences.

*Systems Usability Scale (SUS):* The SUS evaluates the usability of the LM applications. Participants were asked to rate 10 items, using a scale ranging from strongly agree (1) to strongly disagree (5) [15]. The SUS serves as an initial data point to examine the systems utility.

### 4.3 Materials

***Learning Master Application:*** Participants downloaded the LM application on their own Android and Apple smart phones. The application presented nomenclature test content, which allowed the student an opportunity to practice prior to the course exam. Figure 2a. Captures the image shown when downloading the LM application. Figure 2b. Illustrates the leaderboard exhibited in the LM application.

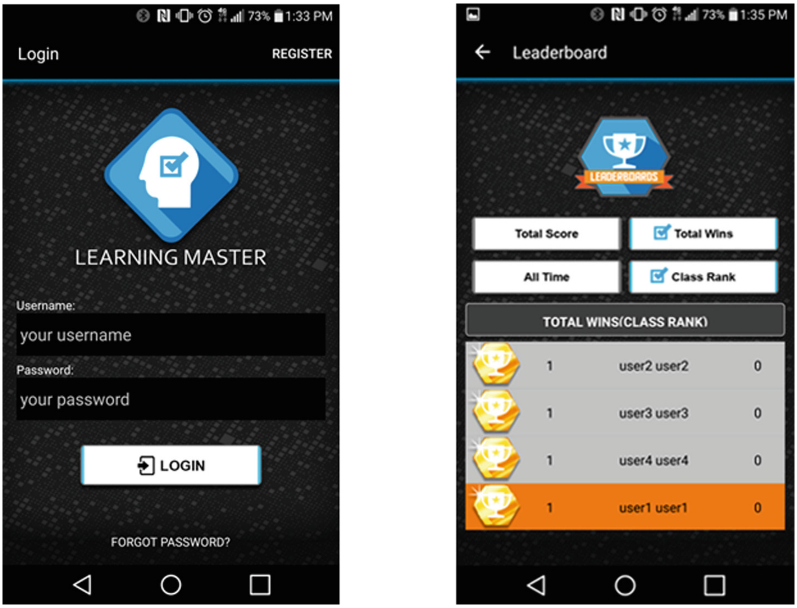



**Fig. 2.** (a) Login page after downloading LM game application; (b) leaderboard screen displayed in LM game application

### 4.4 Procedure

At a classroom in Fort Benning, the experimenter greeted the students and described the research. Students then read, dated, and signed the informed consent. Next, the experimenter asked the participants to complete the demographics questionnaire. Following completion of the demographics questionnaire, a member of the research team provided the participants with guidance on how to download the LM application to their smart phones. Participants without a smart phone had one provided to them with the application preloaded. The participants were asked to use the application on their device to practice for the nomenclature test. The following day, the nomenclature test was presented in a classroom-based setting with test facilitators. The test was presented to the student using a paper-based format. Next, the experiment team returned to the classroom to collect participant's ratings through the SUS.

## 5 Results

Descriptive statistics were calculated using the Statistical Package for Social Sciences (SPSS) to interpret the data findings. The SUS scores reported by users for the LM application found a ($M = 57.95$, $SD = 15.04$). When correlated with a seven-point adjective-anchored Likert scale by Bangor et al. [16], the outcome rating was "Ok" in usability. Next, the SUS statements were tallied to display the frequency scores (Table 2).

**Table 2.** Frequency response to SUS statements about LM.

| Survey statements | Frequency rating | | | | |
|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Neither | Agree | Strongly agree |
| 1. I think that I would like to use this system frequently | 6 | 9 | 20 | 12 | 3 |
| 2. I found the system unnecessarily complex | 7 | 9 | 24 | 2 | 3 |
| 3. I thought the system was easy to use | 3 | 4 | 20 | 12 | 6 |
| 4. I think that I would need the support of a technical person to be able to use this system | 15 | 13 | 13 | 4 | 1 |
| 5. I found the various functions in this system were well integrated | 3 | 6 | 31 | 5 | 0 |
| 6. I thought there was too much inconsistency in this system | 5 | 10 | 22 | 6 | 2 |
| 7. I would imagine that most people would learn to use this system very quickly | 1 | 2 | 19 | 16 | 9 |
| 8. I found the system very cumbersome to use | 6 | 10 | 24 | 3 | 1 |
| 9. I felt very confident using the system | 1 | 7 | 23 | 7 | 7 |
| 10. I needed to learn a lot of things before I could get going with this system | 10 | 11 | 16 | 7 | 1 |
| 11. The application was aesthetically pleasing | 2 | 5 | 30 | 8 | 0 |

From the results of Table 2, the responses in the SUS showed a majority rating for "neither" when assessing the LM. In order to dissect the responses, the neutral statements were barred, resulting in the extrapolation of the four ratings (i.e., strongly disagree, disagree, agree and strongly agree; see Table 3.)

**Table 3.** Post-barred ratings for SUS

| Survey statements | Data extrapolation |
|---|---|
| 1. I think that I would like to use this system frequently | Barring the neutral statements of neither and collapsing the poles into a binary, this comes out with equal amount, 30% for and 30% against the statement |
| 2. I found the system unnecessarily complex | Barring the neutral statements of neither and collapsing the poles into a binary, this comes out with 11% for and 36% against the statement |
| 3. I thought the system was easy to use | Barring the neutral statements of neither and collapsing the poles into a binary, this comes out with 40% for and 16% against the statement |
| 4. I think that I would need the support of a technical person to be able to use this system | Barring the neutral statements of neither and collapsing the poles into one, this comes out with 11% for and 61% against the statement. |
| 5. I found the various functions in this system were well integrated | Barring the neutral statements of neither and collapsing the poles into one, this comes out 11% for and 20% against the statement |
| 6. I thought there was too much inconsistency in this system | Barring the neutral statements of neither and collapsing the poles into one, this comes out with 18% for and 33% against the statement |
| 7. I would imagine that most people would learn to use this system very quickly | Barring the neutral statements of neither and collapsing the poles into one, this comes out with 53% for and 6% against the statement |
| 8. I found the system very cumbersome to use | Barring the neutral statements of neither and collapsing the poles into one, this comes out with 9% for and 37% against the statement |
| 9. I felt very confident using the system | Barring the neutral statements of neither and collapsing the poles into one, this comes out with equal 32% for and 17% against the statement |
| 10. I needed to learn many things before I could get going with this system | Barring the neutral statements of neither and collapsing the poles into one, this comes out with 18% for and 46% against the statement |

## 6 Discussion

Based on the results from the SUS survey, the following paragraphs present the usability assessment of the LM application in the prototype stage. Examining scores for statement 1 (i.e., "I think that I would like to use this system frequently") and statement 3 (i.e., "I thought the system was easy to use"), there was a general conclusion in using the LM application. Although the user interface design had a straightforward layout, the users did provide ways to improve the usability of the LM application. These include:

- Changes to the visual display text could be enlarged in size to increase readability.
- Images could be enlarged to maximize use of the display screen.
- The addition of more color in the images to drive the user's attention towards important features (and help identification on the nomenclature test).

The survey responses to question two (i.e., "I found the system unnecessarily complex") and question four (i.e., "I think that I would need the support of a technical person to be able to use this system") revealed that the users perceived the LM application as easily discernable, which reduces the need for technical support to use the system. This is likely attributed to the pre-training provided to the class on how to use the application. The pre-training quality may have promoted curiosity and challenge, while reducing task complexity, frustration, and failure. Further, the pre-training presentation matched the task and provided explicit knowledge, which reduced the need for additional support in using the LM application.

The user responses to question five (i.e., "I found the various functions in this system were well integrated") showed that there was room for improvement in the application's functionality. At times, the presentation of the icons were difficult to locate on the application (for example, arrow location). To improve functionality, aligning the arrows to indicate the correct response would reduce the number of misunderstood commands and promote positive training transfer. Another way to improve functionality is to have expandable screen content, through finger-pinching the screen to adjust the level of screen magnification. This may make the application more intuitive and allow the content to be easily viewed.

Concerning question six (i.e., "I thought there was too much inconsistency in this system"), the majority of the users reported an inconsistency with the LM application. Perhaps those who felt that there were inconsistencies within the application were referencing training content and not the application. During the assessment of the applications, there were some inconsistencies where incorrect answers were displayed as correct and correct answers as incorrect. The questions and answers were provided by the schoolhouse. The incorrect responses have since then been corrected and updated in the application. One system improvement that could mitigate this issue in the future is to have a "report error" button on each screen that will capture the issue then send it to the instructor for review and repair.

The survey responses to question seven (i.e., "I would imagine that most people would learn to use this system very quickly") suggest that users perceived the LM application as user-friendly. This lends support to instructional presentation of the material received, which reduces the need for additional technical assistance. Future research may consider investigating the instructional presentation using an alternative method to determine the effectiveness on user performance outcomes.

The survey responses to question eight (i.e., “I found the system very cumbersome to use”) and question number ten (i.e., “I needed to learn a lot of things before I could get going with this system”) suggests the users were satisfied using the LM application, as it was straightforward to use. Perhaps the LM application allowed the user to build on their knowledge base. The LM application incorporates game characteristics, such as rewards for correct answers and improvement in leaderboard rankings for better performance. These characteristics are expected to enhance training delivery and encourage future learning.

The survey response to question nine (i.e., “I felt very confident using the system”) indicated that the majority of users felt some level of trust using the LM application. User’s evaluation for trustworthiness of the technology may be based on a variety of factors related to the user, technology and task [17]. Such support was found in early stages of the prototype, as highlighted by the user’s survey responses to the LM application. User’s confidence in using the LM application may foster trust in preparation for the nomenclature exam. The next iteration for testing the usability of the LM application should consider comparing a course with the application and one with only the traditional lecture to assess user performance, and to evaluate trust in technology and levels of engagement.

## 7 Limitations

During data collection efforts, the following challenges were identified. One challenge encountered was that the roster for the class is not final until the morning the class begins. This makes it impossible to pre-populate a registration database for the students prior to their arrival. The students are very limited in available time during the course, so all activities had to be squeezed into very tight lunch times, dinner times or other downtimes. Students wrote their email address and their student identification number. This was used to establish a class-specific database and to allow their performance on the application to be linked to their test scores. After the database was created, they were able to download the software during their lunch hour. Even the student’s lunch hour is busy, so some students did not bother to download and set up the software.

Another challenge the project team faced was that the application required downloading directly onto each student’s personal android or apple smart phones. The project team noticed that in the course of transcribing written emails to the database, typos were made, to some extent this was worsened by the fact that the students needed to provide their personal email rather than their government-issued email addresses. To mitigate this issue, the project team asked the students to self-register using their personal email address. In spite of the challenges that arose, the project team was able to obtain data that served as a baseline for future research investigation.

## 8 Conclusion

This paper described the initial steps for advancing LM as a tool for training in the Jumpmaster course. A primarily analysis of the data provided observational trends and usability assessment of the LM application. From the data, the following conclusions are summarized and recommendations are offered. According to the user's response to the initial prototype, the system offers potentially usable application for training students in the Jumpmaster course. The LM application has a simple, functional, and intuitive ergonomic design. The pre-training support appears reduce the need for technical support. However, there are design recommendations for the next phase; these include improvements to the visual display and streamlining the administration of the LM application. From these recommendations and improvements, the next iteration of LM application may refine the efficacy of the application and return-on-investment to the military.

**Acknowledgments.** This team would like to express our deepest appreciation for the leadership and instructors at the Ft. Benning Maneuver Center of Excellence for their support in the development and evaluation of this training application. Their professionalism and support has always been inspiring as we modernize our fighting force.

## References

1. McLeroy, C.: History of military gaming. In: Soldiers Magazine, pp. 4–6, September 2008
2. Michael, D., Chen, S.: Serious Games: Games that Educate, Train and Inform. Thomson, Boston (2006)
3. Zyda, M.: From visual simulation to virtual reality to games. Computer **38**, 25–32 (2005)
4. Bruner, J.S.: Some elements of discovery. In: Learning by Discovery: A Critical Appraisal, pp. 10–113. Chicago, Rand McNally (1966)
5. Johnson, D.W., Johnson, R.T., Smith, K.A.: Active Learning: Cooperation in the College Classroom. Interaction Book Company, Edina (1991)
6. Charsky, D.: From edutainment to serious games: a change in the use of game characteristics. Games Cult. **5**(2), 177–198 (2010)
7. Guillen-Nieto, V., Aleson-Carbonell, M.: Serious games and learning effectiveness: the case of It's a Deal! Comput. Educ. **58**, 435–448 (2012)
8. Hays, R.T.: The effectiveness of instructional games: a literature review and discussion. Naval Air Warfare Center, Training Systems Division (2005)
9. Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., Berta, R.: Assessment in and of serious games: an overview. In: Advances in Human-Computer Interaction (2013)
10. Bachvarova, Y., Bocconi, S., van der Pols, B., Popescu, M., Roceanu, I.: Measuring the effectiveness of learning with serious games in corporate training. Proc. Comput. Sci. **15**, 221–232 (2012)
11. Weibell, C.J.: Principles of learning: 7 principles to guide personalized, student-centered learning in the technology-enhanced, blended learning environment. WordPress, Dissertation (2011)
12. U.S. Army Maneuver Center of Excellence, "Jumpmaster," 27 November 2017. http://www.benning.army.mil/infantry/rtb/1-507th/jumpmaster/. Accessed 21 Feb 2018

13. Scanlan, P.: Is Parachute capability still relevant to modern expeditionary operations? Aust. Army J. 37–55 (2012)
14. Allen, L., Seeney, M., Boyle, L., Hancock, F.: The implementation of team-based assessment in serious games. In: Proceedings of the 1st Conference in Games and Virtual Worlds for Serious Applications, Coventry, UK (2009)
15. Brooke, J.: "A quick and dirty" usability scale. In: Usability Evaluation in Industry, pp. 189–194. Taylor & Francis, London (1996)
16. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. Int. J. Hum.-Comput. Interact. **24**(6), 574–594 (2008)
17. Xu, J., Le, K., Deitermann, A., Montague, E.: How different types of users develop trust in technology: a qualitative analysis of the antecedents of active and passive user trust in a shared technology. Appl. Ergon. **45**(6), 1495–1503 (2014)

# Games and Business: Human Factors in Gamified Applications

Luís Filipe Rodrigues[1(✉)], Abilio Oliveira[1], Carlos J. Costa[2], and Helena Rodrigues[3]

[1] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisbon, Portugal
lfrodrigues0502@hotmail.com,
{Abilio.oliveira,HMCNC}@iscte-iul.pt
[2] ISEG (Lisbon School of Economics and Management), University of Lisbon, Lisbon, Portugal
cjcosta@iseg.ulisboa.pt
[3] Instituto Universitário de Lisboa (ISCTE-IUL), Business Research Unit (BRU-IUL), Lisbon, Portugal

**Abstract.** The user participation in the software development process is a key to improve the efficiency of business process, and more than likely, improve the user experience. This research aims to examine the users' perceptions and expectations in the development of gamified applications. We outlined different categories of banking software to represent two types of software game analogies (soccer versus blackjack, respectively). Through user discussion, groups and a survey questionnaire with open questions. Content and thematic analyses, with Leximancer were used to analyse the text and to identify the key drivers of the gamified software. The findings from the textual analysis revealed themes (game, design, innovation, information, engagement, purchase, and analogy) that represents the gamification key drivers for software development in business. This is significant to understand the user's perception about gamification to attract, retain, and user's engagement.

**Keywords:** Gamification · Software development · Banking · Leximancer
Content analysis · Human factors

## 1 Introduction

The purpose of this research is to examine the customers' perceptions and expectations in the development of gamified applications. E-business software development is a human centric and socio-technical activity [4]. In software development, communication between stakeholders is at least as important as programming, yet project development activities often emphasize the programming over the communication aspect because customers usually do not have the time or patience to participate in requirement gathering and discussion activities [25].

Human-computer interaction (HCI) community developed a large variety of user-centered design (UCD) techniques, according to typical human behavior and systems specifications. However, these methods are still underused and difficult to understand by software development teams and organizations [22]. The task of


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 318–324, 2019.
https://doi.org/10.1007/978-3-319-94619-1_31

managing a software project can be extremely complex, drawing on many personal, team, and organizational resources [20, 21].

The business requirements always evolve, because of changes in technology, customer needs, business domains, or even thinking on the engagement of new customers [24]. The success of software development process is founded on three basic pillars: processes, technologies and people [8]. Given the importance of people, significant efforts to support and analyze decision making in these projects have been made.

One important aspect is the diverse decision-making methods of professionals and management teams driving the projects. The study we will present here may help to develop and manage teams in order to understand the key drivers for gamification, and adequacy of the developments in terms of design content and buying process. The findings will provide business inside with feedback on customers' attitudes towards gamification design in finance. The results of the study may be guidance for developers and marketers in order to improve the quality of the gamification software and finally the customer acceptance and usage.

The importance of involving end-users and customers in the software development process is widely recognized and was the primary motivation in the early 1980s for work on "end-user" software development [2]. Much of this early work focused on developing mechanisms that would allow end-users to directly contribute to the development of requirements and designs, and understand the artifacts created by software designers. The work at that time focused on developing requirements and design notations that are "customer-friendly", that is, that can be used to create requirements specifications and system designs that provide customers with significant and clear insight into those same characteristics.

## 2 Literature Review

What is missing so far in the recent trend to develop gamified business applications, is more research to understand customers' preferences in terms of expectations – to know what they really would like to see implemented or have available –, trends and software design. By trends, we mean the key drivers that lead to the engagement of customers with a gamified application. Software design has to do with the concept of applying game mechanics and game design techniques in serious business applications, involving, and motivating customers to use these applications, and buy products - for instance, in the case of business applications.

Researches that only focus usage and technical aspects of an application, frequently, not just fail to engage participants with it, but they can indeed negatively affect participants interest for the service or product offered [16], ruining possible business – and even the public image of a company. Therefore, a good application, or website, cannot just show, some items available for customers, with a nice look. It is also necessary, though, is not enough, to be easy to use. However, again, that is not sufficient. In fact, attract the interest of customers, as potential buyers, and stimulate their intention to use an application and buy online, is usually a more elaborated (or even complex) process.

To a great extent, these are good reasons why research on gamification is often focused on the conception and development of well-designed games, with its elementary elements and stimulating features for users, such as badges, rewards and rankings (e.g. [5]). It is not that gamified business applications cannot work well if only centered in design components (such as games characteristics and elements), but to be successful, they must take into account the customers' preferences on games design and processes, from the moment of conception, and along all the development and implementation process. A successful application does not just have a nice design with games components [19].

Gamification is not a simple substitute for a friendly and interactive game design application. More than this, it may be a very valuable alternative to frame a business process. It is essential that business companies progressively understand the crucial importance of customers, and the relevant role they may have on business success, considering their implicit (or explicit) participation even in the adoption of some information technologies, features and other options. How can, this may be achieved? Asking the users and possible customers about their needs, expectations and feelings (about a product and how they percept it may be presented in a more desirable way). Some surveys may help in this sense.

We notice that we need to understand the motivations of customers in order to fulfill desires goals. It is also essential that we enhance what they value the most, also observing the type of information they search for, their buying decisions, always considering their opinions, suggestions and needs, as well as applications usability and attractiveness. In addition, if a gamified application can contribute to the literacy of users the better it is [6]. In sum, we find out that we need to listen and interact with customers to better identify key drivers with valued skills. They have an important place in research, in fact, they are the top reason for research.

We verified that most examples of gamification do a remarkably poor job of interpreting where the potential rewards come from, and none of those applications adequately account the ways in which people and business contexts differ [5].

To date, research in gamification has mainly concentrated on how making gamification systems fun and engaging for end-users [1, 10, 14]. An increasingly number of gamification systems illustrates this fun, or enjoyable, design approach. Theories are emerging to support subjects that involve business context [14], and some researchers even feel that the purpose of gamification is to advance instrumental business goals [12].

Nevertheless, focusing on business is still seen as a symptom of meaningless gamification [14]. Therefore, and in that respect, we observe a large opportunity for research into what is important for customers, but poorly understood, which is the area of business gamification [15].

In the present paper, we aim to fill a research gap in the field of gamified systems, examining the development key drivers for customers' experiences in a gamified business application. Thus, we may help to fulfill this gap by describing on a broad contextual investigation underlying the customers' key drivers for the design and processes of gamified business, in the financial sector. We report the events, highlighted in our study that significantly threatened the success of a gamified business system. Our study may help developers and managers to evaluate the quality and adequacy of gamified business applications, and their development, in terms of design, contents, and business targets.

## 3 Method and Sample

The research is centered on the analysis of customer reviews after they experienced, and tested, a gamified software demo, developed for a digital bank, in Portugal, and offers an explanation about what is more appreciated by them, and what lead customers to use a gamified business application – or, in general, what may drive customers to gamification usage.

In this study, we outlined different categories of financial software (mutual funds and warrants) to represent two types of game analogies (soccer versus blackjack, respectively). Therefore, we created two customer groups (soccer – composed by 8 bank customers, and blackjack – 7 bank customers; total N = 15). So, fifteen customers (85% male and 15% female) were represented in this study, having more than 25 years old (33% with age more than 40).

Each type of software was presented, and tested, by the participants in each of the two groups: soccer and blackjack. After the participants' testing experience, data were collected through a paper questionnaire with six open questions:

1- What did you felt? 2- What did you most like? 3- What did you least like? 4- What changes do you propose? 5- What is missing or what you would you remove from the game? 6- Other suggestions for the game.

After, all answers were compiled in a Microsoft Excel spreadsheet, representing a row for each customer/participant. We highlight that 2,036 words were obtained, considering all participants written feedback.

To analyze all the data collected from customers' answers to the questionnaire, we used content and thematic analyses, with Leximancer, a statistical package adequate for qualitative processing of data [18] - it has sometimes been applied in software development contexts.

Leximancer is a software that conducts an unsupervised analysis of natural language texts, given in an electronic format. Leximancer provides an inductive identification of topics (or themes) with minimal manual intervention by researchers [23] in two types of steps, or procedures: semantic, and relational. The text analysis capabilities provide an overall conceptual structure for qualitative research, and detailed information on specific topics, such as, software requirements and qualitative data analysis [13].

The Leximancer software can identify clusters of words – according to how frequently they occur in sentences – that generally go together, close to a main text – a possible concept –, which is then grouped together in a net with all processed concepts. The concepts are so clustered into higher-level themes, shown as colored circles, which are heat-mapped - in a concept map – to show the relationships between concepts, and the proximity of the themes with each concept – the closer is a theme to a concept, the stronger it is linked with that concept [11].

## 4 Results and Discussion

The analysis of customer reviews, according to the data collected produced a concept map with 7 themes, as follows: game, design, innovation, information, engagement, purchase, and analogy (see Fig. 1). Moreover, themes were associated with important concepts, like game, client, rules (game theme), colors, graphic, application (design theme), idea, process (innovation theme), appealing, simple (information theme), interest, investment (engagement theme), 'game' theme, analogy (analogy theme), and purchase (purchase theme).

Therefore, the narrative for each concept, within each theme, confirms that the key drivers that characterize the gamification usage (derived from the customer reviews after their experience using the gamified application) are quite homogeneous and require the attention of software developers and digital managers to keep a user's attention.

We must note that customer feedback about software can help to measure software quality [3], and improve customers' acceptance [17]. Nowadays, digital company managers have to manage their businesses in a very a competitive environment. Thus, to know customer feedback is an essential component in any modern software development toolkit, in in order to a manager or company remain as a strong

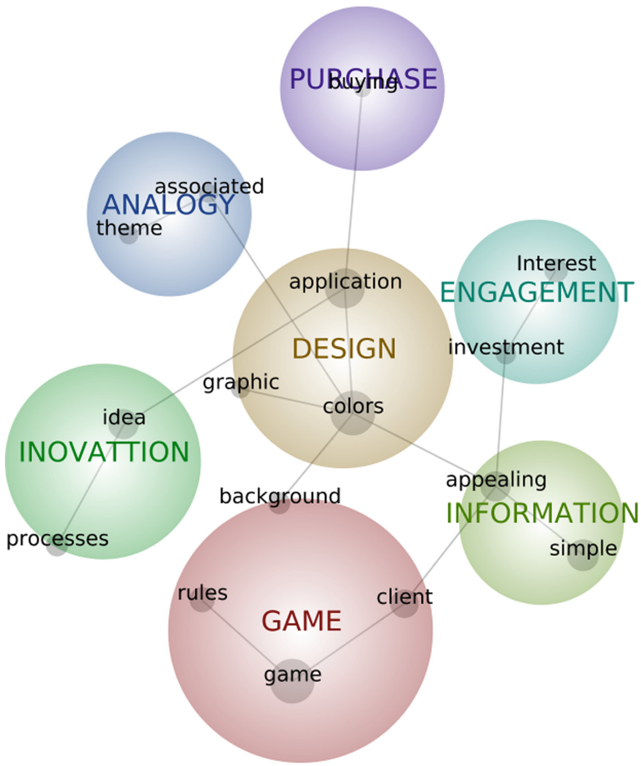



**Fig. 1.** Concept map of gamification: testers perceptions and expectation of e-business gamification test development process (Source: Own elaboration).

competitor, and that implies keeping customer's happy [7]. The perceptions of customers, and their feedback about a software application (for instance, after experiencing a demo), can help managers, and software developers, to measure and improve customers' satisfaction, identify unhappy customers, reduce churn, and get valuable insight about how a product or service can be improved and optimized [9].

This research enriches the existing literature by emphasizing the importance customers' perceptions and opinions in gamification software development. So, we suggest that a gamified application, rather than just supported by data stimulated by the researcher, should be tested by potential users/customers, to obtain their spontaneous feedback (according to their reviews) that may be compiled through their answers to a given questionnaire (with open-questions). Considering their perceptions towards gamification in finance, an application may be improved by software developers, web programmers and line teams, and finally made available, being more attractive and useful to customers, and stimulating its use, and the intention to buy a product or a service.

The data text obtained from users may be processed with Leximancer, which is a statistical software that processes content analysis of qualitative data, and produces quantitative results, that may be shown in graphic modes, to highlight the main themes and concepts in which perceptions of customers are categorized.

This study also contributes to digital business research by presenting a novel approach for extracting, analyzing, and understanding customers' perceptions in a gamification context. So, customers' feedback is essential to improve the quality of business management software.

## 5 Conclusion

We conclude that in reference to customers' perceptions and expectations on (the development of) gamified applications, we may consider 7 main themes (game, design, innovation, information, engagement, purchase and analogy) and 15 concepts (game, client, rules, colors, graphic, application, idea, process, appealing, simple, interest, investment, game theme, analogy, and purchase), that should be considered during gamified software development, since they reveal a strong importance among customers preferences. Hence, gamified applications may be optimized, to be more attractive and effective, offer ease of use and enjoyable experience, and stimulate customers to use them more often - and thus improving business.

Future research should examine the customers' perceptions of other types of gamified software in digital finance industry.

## References

1. Amir, B., Ralph, P.: Proposing a theory of gamification effectiveness. In: Companion Proceedings of the 36th International Conference on Software Engineering, pp. 626–627. ACM (2014)
2. Briefs, U., Siborra, C., Schneider, L. (eds.): Systems Design For, With, and By the Users. North-Holland, Amsterdam (1983)
3. Cao, L., Ramesh, B.: Agile requirements engineering practices: an empirical study. IEEE Softw. **25**(1), 60–67 (2008)

4. Casado-Lumbreras, C., Colomo-Palacios, R., Soto-Acosta, P., Misra, S.: Culture dimensions in software development industry: the effects of mentoring (2011)
5. Deterding, S.: Gamification: designing for motivation. Interactions **19**(4), 14–17 (2012)
6. DomíNguez, A., Saenz-De-Navarrete, J., De-Marcos, L., FernáNdez-Sanz, L., PagéS, C., MartíNez-HerráIz, J.J.: Gamifying learning experiences: practical implications and outcomes. Comput. Educ. **63**, 380–392 (2013)
7. Fabijan, A., Olsson, H.H., Bosch, J.: Customer feedback and data collection techniques in software R&D: a literature review. In: Fernandes, J., Machado, R., Wnuk, K. (eds.) Software Business, ICSOB 2015. LNBIP, vol. 210, pp. 139–153. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19593-3_12
8. Hernández-López, A., Colomo-Palacios, R., García-Crespo, Á., Soto-Acosta, P.: Team software process in GSD teams: a study of new work practices and models. In: Professional Advancements and Management Trends in the IT Sector, p. 154 (2012)
9. Kannadasan, S., Aravazhi, D.: Customer satisfaction in marketing. Clear Int. J. Res. Commer. Manag. **6**(8) (2015)
10. Kappen, D.L., Nacke, L.E.: The kaleidoscope of effective gamification: deconstructing gamification in business applications. In: Proceedings of the First International Conference on Gameful Design, Research, and Applications, pp. 119–122. ACM (2013)
11. Leximancer Pty Ltd.: Leximancer User Guide, Release 4.5 (2016)
12. Mollick, E.R., Rothbard, N.: Mandatory fun: gamification and the impact of games at work. The Wharton School Research Paper Series (2013)
13. Myers, M.D.: Qualitative research in information systems. Manag. Inf. Syst. Q. **21**(2), 241–242 (1997)
14. Nicholson, S.: A user-centered theoretical framework for meaningful gamification. Games + Learning + Society, **8** (2012)
15. Richards, C., Thompson, C.W., Graham, N.: Beyond designing for motivation: the importance of context in gamification. In: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play, pp. 217–226. ACM (2014)
16. Rigby, S., Ryan, R.: Glued to Games: How Video Games Draw Us In and Hold Us Spellbound. Praeger, Santa Barbara (2011)
17. Rising, L., Janoff, N.S.: The Scrum software development process for small teams. IEEE Softw. **17**(4), 26–32 (2000)
18. Rodrigues, H., Brochado, A., Troilo, M., Mohsin, A.: Mirror, mirror on the wall, who's the fairest of them all? a critical content analysis on medical tourism. Tourism Manag. Perspect. **24**(2017), 16–25 (2017)
19. Rodrigues, L.F., Costa, C.J., Oliveira, A.: Gamification: a framework for designing software in e-banking. Comput. Hum. Behav. **62**, 620–634 (2016)
20. Rose, J., Pedersen, K., Hosbond, J.H., Kræmmergaard, P.: Management competences, not tools and techniques: a grounded examination of software project management at WM-data. Inf. Softw. Technol. **49**(6), 605–624 (2007)
21. Seffah, A., Metzker, E.: The obstacles and myths of usability and software engineering. Commun. ACM **47**(12), 71–76 (2004)
22. Soto-Acosta, P., Casado-Lumbreras, C., Cabezas-Isla, F.: Shaping human capital in software development teams: the case of mentoring enabled by semantics. IET Softw. **4**(6), 445–452 (2004)
23. Tkaczynski, A., Rundle-Thiele, S.R., Cretchley, J.: A vacationer-driven approach to understand destination image: a Leximancer study. J. Vacation Mark. **21**(2), 151–162 (2015)
24. Turk, D., France, R., Rumpe, B.: Assumptions underlying agile software development processes. arXiv preprint arXiv:1409.6610 (2014)
25. Wiegers, K., Beatty, J.: Software Requirements. Pearson Education, London (2013)

# Investigating the Human Factors in eSports Performance

Daniel Railsback(✉) and Nicholas Caporusso

Fort Hays State University, 600 Park Street, Hays, USA
drrailsback@mail.fhsu.edu, n_caporusso@fhsu.edu

**Abstract.** Although its history dates to the beginning of the videogame era, in the recent years eSports has rapidly been maturing into a professional scene that has captured a large and growing audience: the 2017 world finals of one of the most successful pro gaming platforms reached a peak of 75.5 million unique viewers, compared to the NBA finals, which averaged 20 million spectators. Several studies already supported eSports as a legitimate form of sports. In this paper, we focus on identifying the human aspects of eSports and we discuss their relevance in competitive games. To this end, we discuss interviews with sports and eSports professionals, and we analyze the results of a survey distributed to athletes who are pursuing their career in traditional sports and in pro gaming.

**Keywords:** Human factors · eSports · Pro gaming

## 1 Introduction

Electronic sports (eSports) is a novel type of competition and spectator entertainment that pits individuals or teams playing video games in front of a large crowd attending the show in presence or remotely. The birth of eSports can date back to many competitions but the one that stands out is the First National Space Invaders competition in 1980. Taking place in New York city with 10,000 participants, the concept of eSports spectatorship was born by being broadcasted on live television [1].

In the recent years, eSports has continued to gain popularity and grow into an unprecedented case in the history of computer and entertainment. Specifically, 2017 has been a successful year for eSports. League of Legends, the title of one of the fastest-growing pro gaming scenes, held the knockout stage for the world finals, from October 19th to November 4th. Statistics from this 20-day event include over 4 billion hours of total time watched, a peak of over 106 million viewers during the semi-finals, over 75 million unique viewers during the finals, and over 33 million concurrent viewers throughout the tournament. To put these numbers into perspective, the 2017 NBA finals averaged 20 million viewers, only [2]. Indeed, the large numbers reported by statistics help understand the magnitude of eSports in terms of popularity. Nevertheless, their growth is even more significant if compared to numbers from 2015. Riot Games, the publisher of League of Legends, reported 36 million unique viewers for the world finals [3]. These statistics indicate two-fold increase in popularity of spectating live and online for eSports, in just two years.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 325–334, 2019.
https://doi.org/10.1007/978-3-319-94619-1_32

As other traditional sports, eSports has turned into an extremely popular spectator sport. Many fans from all around the world flock to live eSports events to watch their favorite players compete in person. In October 2013, one of the competitions sold out the 10,000 available tickets in less than one hour [4]. Fans and live events are a driver for the industry, and this is no different than traditional sports, such as, soccer or basketball. As other competitions, eSports tournaments are broadcasted with play-by-play commentary, slick graphics, player and team statistics, human interest pieces, post-game interviews, and even instant replay analysis [5]. Over the history of eSports, the popularity of team eSports has taken the spotlight leaving most single player competitions behind [1]. With the rise of eSports players are turning competitive video games into a career. Moreover, popular eSports, such as, Starcraft and League of Legends, started to create a culture where best players become super stars.

Not only are eSports events drawing the focus of fans and spectators, but it is also attracting a large amount of revenue. As the history of eSports has evolved, so has its prize pools. In 1999 its average was under $100,000, while it increased to over $600,000 per eSports event, in 2015 [1]. Moreover, eSports are converging towards a structured and lucrative business model in which players are organized into professional leagues and tournaments. Starting in the spring of 2018 the North American League of Legends Challenger Series (NALCS) will be moving to a traditional format for their league: franchises will cost a flat $10 million for existing NALCS organizations [6].

Consequently, it becomes clear that it is not a question of whether eSports will change the way we view and look at sports but how it will change the sporting entertainment industry altogether.

## 2 Related Work

eSports has been maturing into a very professional scene that has captured a huge audience that continues to grow at a rapid pace. The popularity and growth of the industry plays a very important role in defining what exactly an eSports is for those unfamiliar with the industry. Because the eSports industry is relatively new, it is hard to create a definition for it, though achieving a common understanding for eSports is very important for its recognition as a sport. However, one of the official definitions of eSports (i.e., “Competitive tournaments of video games, especially among professional gamers”) does not incorporate several crucial aspects of pro gaming. To this end, authors created a check list on defining a sport [7]. Moreover, in the last years, several studies addressed the debate about whether professional computer gaming can legitimately be defined sports and compared to traditional athletic challenges. Indeed, the evolution of eSports itself is contributing to the discussion [8]. Following some of the major trends of the industry, eSports will probably continue to grow at a healthy rate, and the line between eSports and traditional sports will continue to fade [7, 9, 10].

Indeed, eSports and traditional sports share several differences and similarities. Nevertheless, only a limited number of studies focused on understanding the characteristics that distinguish the two disciplines and the ones which make them similar. In [11] the authors analyzed the importance of physical abilities and concluded that motor skills are a defining element of eSports.

Therefore, in this paper, we focus on the human aspects of eSports, and we use both qualitative and quantitative methods to investigate the nature of the similarities of electronic and traditional sports. Also, we detail the key psychological and physical factors for agonistic success in pro gaming, and how they relate to other sports.

# 3 Study

In our study, we investigate the human factors that are relevant for athletes' success in pro gaming and in traditional sports, presented as similar categories. Our analysis does not focus on contributing to the debate about whether eSports is a legitimate type of sport. Conversely, we try to identify the common psychological and physical traits and challenges of athletes in the two types of disciplines.

## 3.1 Materials and Methods

To this end, we realized interviews with four professionals working as managers or coaches in sports and eSports. We selected respondents from Kansas Wesleyan University which, in addition to traditional athletics, developed competence in progaming, in the last years. Specifically, two individuals from conventional sports and two from the eSports scene were selected for the interviews: C1 is the head coach for the football team, they have extensive experience in coaching student athletes, and they have a history of competing as a NCAA division 1 athlete; C2 is the head coach of the woman volleyball team, and thanks to their track record in terms of team victories, they took the position of the most successful volleyball coach in the history of the University, in late 2017; C3 is a former member of a high ranking eSports League of Legends team, and they currently hold the position of head eSports coach at the University; C4 holds the position of eSports manager for the National Association of Collegiate eSports (NACE), and they are the core founder of the eSports program at the University, and during their undergraduate studies supported the hiring of a coach for their program, which resulted to be C3.

Although the interviews were unstructured, we asked the similar questions to both groups and we encouraged respondents to lead the conversation. Conversations with coaches and industry professionals provided us with the opportunity of analyzing common patterns emerging from individual experiences and perspectives. This, in turn, helped us acquire a unique level of understanding of the key human factors for comparing eSports and traditional sports. Moreover, it supported identifying the fundamental aspects for developing a questionnaire, which was utilized to investigate the key factors among athletes in traditional sports and in pro gaming.

Specifically, given their involvement with a very physical sport, C1 provided great insight regarding the physical requirements for a top-level sport athlete. During the interview, C2 focused on coaching success and demonstrated how sports operate and the crucial aspects of how to build successful athlete and program. The conversation with C3 resulted in a deep understanding of what it takes to build an eSports culture, given their experience in creating a new program completely dedicated to pro gaming. Also, by discussing about aspects which are relevant for coaching, C3 highlighted the individual human factors in becoming an eSports athlete, with a bottom-up approach. Finally, the managerial experience of C4 integrated the study with a top-down perspective on what defines an eSports and how eSports organizations function across the United States and worldwide.

Subsequently, results from qualitative interviews were utilized to design a questionnaire specifically aimed at exploring the human factors eSports by comparing traditional sports athletes and pro gaming participants, with the hypothesis that both groups share similar motivations, attributes, and traits in approaching their activity and, thus, human aspects play the most crucial role in significantly characterizing both types of sports as a single group. Specifically, the survey contained a total of five attributes that were utilized to describe the key human factors of sports and eSports, and items that aimed at getting more insights into the dynamics of players. The survey included questions focused on the topics that emerged in conversations with coaches and managers, and answers were collected using a Likert scale to achieve quantitative results and to standardize the methodology across the two types of sports. Also, subjects were asked to rank factors in terms of importance and impact on performance. The questionnaire was submitted to participants selected among students. They were divided in two groups: G1 included 48 individuals between the age of 18 and 25 who are involved in traditional sports, whereas G2 featured 48 eSports participants within the age range of 18–30.

## 4 The Human Factors of eSports

Interestingly, both C1 and C2 were aware of the growth and popularity of eSports. Specifically, when asked about the difference between traditional sports and pro gaming in terms of human factors, C1 commented “I cannot wait for the day that two people sit down to play a competitive game of football through eSports; one of these people has played football their entire life, been on a team through college and maybe even in the professional leagues, whereas the other person has never physically played football but competes in football via eSports and beats the player that competed in the traditional sport; that is one of the reason that I think eSports is so interesting”.

When asked about a definition of eSports, C3 defined them as “a popular video game that is competitive in the fact that there is a clear winner”. Therefore, the main points that C3 bases this definition on is the concept of a winner and a loser. In this regard, our respondents’ perception of pro gaming with the common definition of traditional sports, which is an “athletic activity requiring skill or physical prowess and often of a competitive nature”. Although this sentence supports the hypothesis that both types of sports being are in the same group, more accurate analysis is required to expand on the meaning of sports beyond the moment of the competition and identify whether other factors are relevant in approaching sports. For instance, C2 commented that “the first thing that pops into my head when I think of the word is technology; this type of competition seems to evolve around visual equities, fast reflexes along with a high level of hand eye coordination”.

However, as several other aspects are involved in eSports, our investigation of the physical and psychological components might provide insightful information to reshape the definition of pro gaming to take into consideration additional features that specifically characterize eSports. Also, our analysis includes results from interviews and surveys to highlight the similarities and differences between traditional sports and pro gaming, and to identify the most important human factors that play a role in considering eSports as a legitimate sport (Fig. 1).

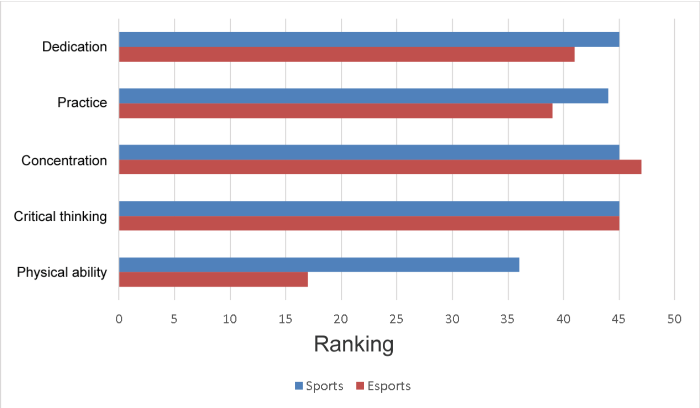



**Fig. 1.** Ranking of important factors for eSports and sports.

### 4.1 Competition

As mentioned in the previous section, competition is considered at the very core of the definition of any sport activity. The presence of a goal that players have to achieve by identifying strategies and the fact that the game will ultimately result in a winner and a loser, seems to be the main motivation for playing. Sport, both in its traditional format and in the eSports version, creates a scenario in which players deeply experience the dynamics of competition, which are among the most important psychological drivers for any animal species. In this regard, League of Legends is a very clear example of an environment where teams contend for resources. Moreover, C2 considered competition as the most important, must-have feature of sports as they revealed: "throughout my entire life I have been playing sports; if I had to pick one word to describe sports, it would have to be competitiveness; I believe that competitiveness alone drives many people into the world of sports". In this regard, C1 agreed with C3, as both stated: "when deciding whether something is a sport or not, we need to look at how it ends; if the outcome is strictly through competition resulting in a winner and a loser, then it is a sport, while if the outcome is decided by opinion, then it is art; I believe that eSports are more of a sport than gymnastics because the outcome results in decisions made by facts". Moreover, C1 reported that the most important trait determining sport is competitive integrity: "a winner and loser will be determined by only internal sources following the rules of facts".

### 4.2 Inclusivity

During the interviews, respondents referred several times to items, which can be associated with inclusivity, a very unique and distal factor of eSports. Surprisingly, inclusivity was not taken into consideration by other studies, and it receives less attention. However, its ability to be unbiased helped eSports grow to the level of popularity that it has today. Surprisingly, the topic was introduced by a traditional sports professional: C1 revealed that “eSports is one of the most inclusive sports that exists. It is completely unbiased toward gender, size, weight, race, or even religion: anyone can sit down and be included”. As eSports players start their career autonomously and games develop in a collaborative environment that augments a single-person physical setting, almost anyone can start playing eSports. Then, when they reach competitive or professional levels, their skills and ranking are the only factors that matter. This feature, which eSports inherently has, is increasingly being introduced in traditional sports, to make them diverse and inclusive in terms of levels of physicality, ethnicity, origin, gender, and status. C1 continued supporting the analogy between the two types of sports by mentioning their efforts in improving the inclusivity of their football program: “this is one of the things that I love most about sports; in my football program, I have almost every type of demographic other than females. I have players that line up together, go to meetings together, even laugh and bond together. I get the privilege to bring all these people together in harmony to accomplish one thing; eSports is very similar to this but on a much larger scale”. However, our findings contrast with other studies about diversity in eSports [12], as our participants might not be aware of statistics about gender and ethnicity, and therefore, their beliefs might be biased.

### 4.3 Concentration

Concentration was another factor evaluated during the survey, following suggestions given by the four respondents. Concentration resulted in the highest-ranking factors in terms of importance and impact on performance for both traditional sports and pro gaming. The results of the survey showed that concentration is more critical in eSports than in traditional sports. However, this is not a statistically significant difference. Moreover, the differences in duration and dynamics of eSports matches might explain this difference: traditional sports have breaks and time-outs that help players rest and relax to a certain extent, whereas pro gaming tournaments consist of uninterrupted sets. This is supported by C4, who discussed an example in League of Legends: “a match can last anywhere from 20 to 60 min; in this time there are no time outs, no breaks or rest, and all the players are experiencing very intense situations; a lack of focus or concentration for a single moment can lose the whole match for them; concentration is something that a professional eSports player cannot lack”.

### 4.4 Critical Thinking

Critical thinking ranked second as a key success factor for both traditional sports and eSports. Both groups were aligned in attributing the same value to this item. Indeed, as sport implements a type of game, teams' strategies significantly contribute to victory or loss, and individuals' tactics are important in improving the outcome of a coalition. Therefore, both sports and eSports have an intellectual component which is more important than physical aspects, as confirmed from our results (see Physical preparation). Among respondents, C2 had a great perspective on the contribution of critical thinking to a competition: "when you look at the best college players in the nation and at elite level players, those individuals have exceptional critical thinking attributes". Moreover, this factor is directly associated with competition, as C2 referred "what drives most players is more passion and heart, but when a player has great critical thinking abilities that is what brings them to be a whole new level as a competitor". This is in line with findings from previous studies, which addressed the dynamics of specific games, such as, League of Legends [12–14].

## 4.5 Dedication

Participants ranked dedication as the third factor in terms of importance and impact on performance, with a slight difference between traditional sports and eSports, with the latter scoring lower. During the interview, C4 commented that "almost all of the most successful teams pour their soul into eSports; if there is anything that these players do have it is dedication". However, other participants had the same perception in regard to sports. As shown by our findings, the difference in dedication between eSports and traditional sports is extremely limited. As a result, we can conclude that high level of dedication is extremely important for both categories.

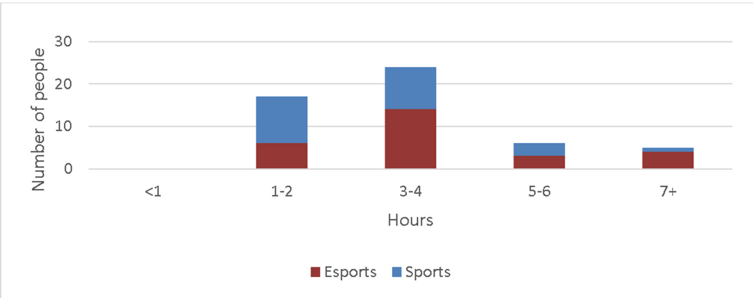



**Fig. 2.** Hours dedicated to training.

## 4.6 Practice

The survey explored practice, as during the interviews this emerged several times in respondents' sentences. Specifically, C3 commented: "we practice at least five times a week inside and out of the game; a lot of the time that we put into practice consist of game plan, strategy, and improving overall individual performances within the game; often, when people think of competitive video games, they see it to be nothing but fun, unfortunately this is not the case; the main focus for our team is to get better and produce good results in competition; once this is accomplished, we do try and have a little fun”. However, similar comments were made by professionals from traditional sports. Data from the survey (see Fig. 2) show that pro gaming players perceived training as less important than subjects in the other group.

### 4.7 Physical Ability

Physical ability was discussed extensively during the interviews, and it was included in the survey. C1 supported the hypothesis of a great similarity of the two types of disciplines by stating that “eSports competitors are athletes; they are using their brains and their body to compete”. This is an especially interesting perspective on the need for physical ability in eSports and sports, as C1 is a professional in the traditional sports space”. The importance of physical ability in eSports was reiterated by C4, who made several comments on the physical factors that play a role in competitions. They stated that “many people underestimate the physical requirements of eSports; without fast reaction time, and hand eye coordination an athlete won’t find much success in eSports”. Although this category had the largest disparity between the two groups of responders (with eSports players attributing less than half the score of traditional athletes), both ranked physical ability as the least important factor.

### 4.8 Audience Engagement

During the interviews, respondents reported several elements related to players’ engagement in the game, such as, competition, strategic thinking, and technical preparation. Moreover, they spontaneously introduced distal items that expand outside the individuals involved in the game, which, in turn, can be associated with the concept of audience. Although very marginal in some conversations, this factor can be considered one of the defining aspects of sports, and it deserves some consideration in the definition of the word. Furthermore, it is one of the fundamental components of the success of a specific type of sport, or game. Although this factor was not included in the questionnaire, the interviews revealed that when investigated further, the presence of an audience sets the boundary between what can be considered sport (e.g., basketball or League of Legends) and what is just an activity, such as, playing with cards. Specifically, C2 stated: “it is not a sport if there is not any level of entertainment for spectators; the first thing that comes to my mind when thinking about sport, is the word spectator: will anyone want to sit down and watch this? What type of value does this bring outside of the competition”. Also, the presence of an audience of viewers creates a generative loop which is beneficial to the development of a sport: higher levels of engagement drive commercial interests which, by injecting resources, promote dissemination to a larger audience.

### 4.9 Continuous Training

The survey asked athletes from pro gaming and traditional sports about their training habits and practices in the off-season, to investigate how they prepare to get ready for competition. We offered four different options, and we gave respondents the opportunity of adding their own opinion. Results are shown in Fig. 3, and they demonstrate a common perception of the level of training required by their discipline.

Also, we focused on time spent on training. Figure 2 illustrates how many hours the two types of competitors spend training each day, and it highlights the level of dedication and commitment both groups of players must have.

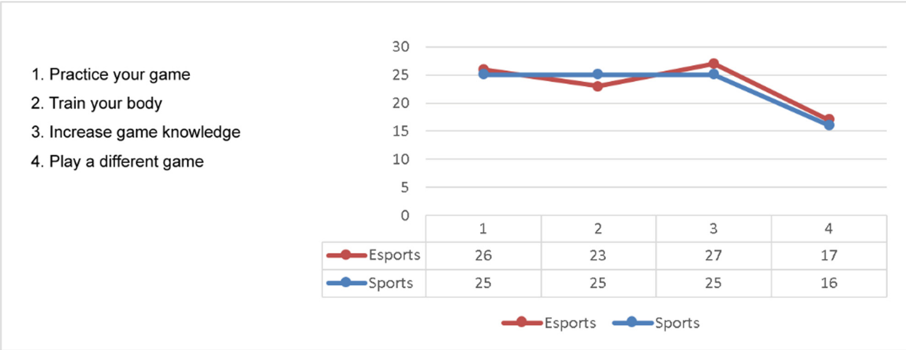



**Fig. 3.** Type of training and strategy implemented in the off-season period.

## 5 Conclusion

eSports is a form of competitive sports that heavily relies on the use of technology and video games. In the recent years, its level of popularity has started to spread outside of the eSports culture and into more traditional sports settings. As a result, several studies identified the similarities between sports and eSports, and made a legitimate case for the latter being considered as a sport.

Our study investigated the human factors involved in eSports, and we addressed the topic from the perspectives of coaches and athletes of traditional and electronic sports. To this end, we conducted interviews with professional coaches and players of student teams and eSports leagues. Also, we realized a questionnaire in which we addressed 5 aspects related to individuals' agonistic preparation to tournaments, that is, dedication, practice, concentration, critical thinking, and physical ability.

From our findings, we can conclude that professional coaches share a similar perspective on the challenges and dynamics which traditional and electronic sports have in common, though the intensity of physical activity might be perceived as different. Moreover, specific advantages of eSports, such as inclusivity and the ability of being unbiased, were praised by both groups of coaches. Furthermore, data from our survey shows that 73% of sport players (both groups) recognize pro gamers as athletes, whereas 27% of respondents do not consider eSports as an agonistic activity. In contrast, results from our quantitative analysis reveal that there is no statistical difference between traditional and electronic sports in 4 out of the 5 dimensions of agonistic preparation considered by our study.

As a conclusion, achieving a better understanding of the common human factors, dynamics, and challenges of sports and professional gaming might help define the interpretation eSports and leverage the opportunities that arise from the similarities between traditional and electronic sports.

## References

1. Olsen, A.H.: The evolution of eSports: an analysis of its origin and a look at its prospective future growth as enhanced by Information Technology Management tools. arXiv preprint arXiv:1509.08795 (2015)
2. Holloway, D.: TV Ratings: NBA Finals Continue to Rise From 2016, Game 3, 13 June 2017. http://variety.com/2017/tv/news/tv-ratings-nba-finals-1202464230/. Accessed 27 Feb 2018
3. Heaven, D.: Rise and rise of esports, 16 August 2014. https://doi.org/10.1016/S0262-4079(14)61574-8. Accessed 26 Feb 2018
4. Schmidt, S., Shreffler, M.: Motivations for eSport consumption: a road map for traditional sports online spectating. In: 2015 Sport Marketing Association Conference (2015)
5. Lee, D., Schoenstedt, L.J.: Comparison of eSports and traditional sports consumption motives. ICHPER-SD J. Res. **6**(2), 39–44 (2011)
6. Khan, I.: Riot releases details on NALCS franchising with $10 M flat-fee buy-in, 01 June 2017. http://www.espn.com/esports/story/_/id/19511222/riot-releases-details-na-lcs-franchising-10m-flat-fee-buy-in. Accessed 27 Feb 2018
7. Jenny, S.E., Manning, R.D., Keiper, M.C., Olrich, T.W.: Virtual (ly) athletes: where eSports fit within the definition of "sport". Quest **69**(1), 1–18 (2017)
8. Wagner, M.G.: On the scientific relevance of eSport. In: International Internet Computing and Computer Game Development Conference Proceedings, Las Vegas (2006)
9. Hewitt, E.: Will eSports ever become widely accepted as official sports and how will they affect the way we entertain ourselves if they do? In: Sharpe, J., Self, R. (eds.) Computers for everyone, pp. 81–83 (2014)
10. Hollis, K.: Time to be grown-ups about video gaming: the rising eSports industry and the need for regulation. Ariz. Law Rev. **57**, 823–847 (2015)
11. Hilvoorde, I.V., Pot, N.: Embodiment and fundamental motor skills in eSports. Sport Ethics Philos. **10**(1), 14–27 (2016)
12. Kim, S.J.: Active 21st century, 01 August 2017. Gender inequality in eSports participation: examining league of legends. http://hdl.handle.net/2152/62914. Accessed 26 Feb 2018
13. Schubert, M., Drachen, A., Mahlmann, T.: eSports Analytics Through Encounter Detection Other Sports (2016)
14. Ferrari, S.: From generative to conventional play: MOBA and league of legends. In: DiGRA Conference (2013)
15. Winn, C.: The well-played MOBA: how DotA 2 and league of legends use dramatic dynamics. In: DiGRA Conference, May 2015

# Impressions and Congruency of Pictures and Voices of Characters in "The Idolmaster"

Ryo Takaoka(✉), Naoto Hayash, Yosuke Nakagawa, and Masashi Yamada

Graduate School of Engineering, Kanazawa Institute of Technology, 7-1 Ohgigaoka, Kanazawa, Ishikawa 921-8812, Japan
{b1402904, b1223334, b1205879}@planet.kanazawa-it.ac.jp, m-yamada@neptune.kanazawa-it.ac.jp

**Abstract.** The impressions of pictures and voices of the characters in "The Idolmaster" were investigated. In Experiment 1, pictures of 61 characters who appear in the mobile game, "The Idolmaster, Cinderella Girls Starlight Stage" were used as stimuli. Moreover, voice recordings of the title calls spoken by voice actress were used as stimuli in Experiment 1. The participants were requested to rate their impressions of the stimuli. The results of factor analysis showed that the impression space was constructed by three factors: "activity", "coolness" and "evaluation". In Experiment 2, the congruency between the pictures and the voices was investigated. The results showed that coincidence in both the activity and the coolness factors was important for the congruency between the pictures and voices.

**Keywords:** Congruency · Character design · Semantic differential method Voice actress

## 1 Introduction

Animated movies (anime), video games and cartoons have developed within the sub-culture on Japan. However, in recent years, these contents have been called "Cool Japan" and Cool Japan contents are recognized by the government as an export-oriented manufacture of Japan. In fact, the Japanese Ministry of Economy, Trade and Industry supported overseas operations of Cool Japan contents under a supplemental budget of 34.4 billion yen in 2012 (Japanese Ministry of Economy, Trade and Industry, 2012) [1].

In the present study, impressions of pictures and voice recordings of the characters in "The Idolmaster" were investigated. Moreover, the correlations of the congruency between character designs and voices were also investigated. The Idolmaster is one of the successful series of video games in Japan [2]. It was first released as an arcade game in 2005, and new contents have been developed for consumer and mobile games over ten years. In recent years, anime content was also created for TV series and movies. The gameplay and story follows the career of a producer in charge of training prospective pop idols on their way to stardom.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 335–342, 2019.
https://doi.org/10.1007/978-3-319-94619-1_33

In this series, various types of young girls appear as prospective idols. Each girl was designed with a different face, style and fashion. Moreover, a different voice actress voiced each one. In the present study, perceptual experiments were conducted using the pictures and voices of the characters in the mobile game, “The Idolmaster, Cinderella Girls Starlight Stage” [3].

# 2 Experiment 1

## 2.1 Methods

Experiment 1 consisted of two sections; Sects. 1 and 2. In Sect. 1, pictures of 61 characters who appear in the mobile game, “The Idolmaster, Cinderella Girls Starlight Stage” were used as stimuli. In this game, 61 different voice actresses individually call the title “The Idolmaster, Cinderella Girls Starlight Stage” as the characters. In Sect. 2, these 61 voice recordings were used as stimuli. Twenty students from Kanazawa Institute of Technology participated in the experiment. One half of the participants were familiar and the other half were unfamiliar with the game. They rated their impressions for the 61 pictures of the characters in Sect. 1, and for the 61 voice recordings in Sect. 2, using sematic differential method. Twenty-four seven-step bipolar scales shown in Table 1 were used in the rating. The pictures were presented on a computer display EIZO FlexScan SX2462W-PX, in Sect. 1, and the voice recordings were presented through headphones STAX SR-404 in the level of LAeq = 75.5 − 83.3 dB, in Sect. 2. The participants sat on a chair in a dark sound-proof room and were requested to view or listen to the stimuli.

## 2.2 Results and Discussion

The rated values for each scale and each stimulus were averaged over participants. Factor analysis with principal factor method and varimax rotation was performed using these averaged data. The results showed that the three-factor solution accounted for 78% of the data variance. The three factors were labeled “activity”, “evaluation” and “coolness” respectively, after the scales which showed high absolute values in the factor loadings. The pictures with high values in activity tended to show active poses wearing clothes that were easy to move in, while those with high values in evaluation tended to wear proper and feminine clothes. The pictures with high values in coolness tended to show fine-looking faces with up-angled eyes and gentle poses. The voices with high values in activity tended to have clear accents and large intonations, while the voices with high values in evaluation tended to have clean timbre. The voices with high values in coolness tended to show high intelligibility.

## 3 Experiment 2

### 3.1 Methods

A picture was selected from each quadrant of the three-dimensional impression space in Experiment 1. The corresponding voice recordings were picked up. They were called original voices in the present study. Moreover, eight spoken voices which placed in the shortest distances from the eight original voices were picked up, and called similar voices in the present study. Each original voice and similar voice were placed in the

**Table 1.** Semantic differential scales and their factor loadings.

| SD Scale | | | Factor | | |
|---|---|---|---|---|---|
| | | | **Activity** | **Coolness** | **Evaluation** |
| Powerless | - | Powerful | .892 | .330 | .014 |
| Weak | - | Strong | .717 | .648 | -.008 |
| Unfit | - | Healthy | .830 | .434 | .209 |
| Dull | - | Vivid | .957 | .035 | .118 |
| Dark | - | Bright | .911 | -.056 | .310 |
| Boring | - | Joyful | .911 | -.159 | .311 |
| Calm | - | Aggressive | .957 | -.101 | -.044 |
| Gloomy | - | Cheerful | .928 | -.011 | .269 |
| Gentle | - | Showy | .950 | -.057 | -.064 |
| Sober | - | Flamboyant | .898 | .007 | .203 |
| Unimpressive | - | Impressive | .712 | -.016 | -.247 |
| Lethargic | - | Ambitious | .924 | .079 | .221 |
| Childlike | - | Mature | -.159 | .889 | .133 |
| Loose | - | Tight | .221 | .906 | .198 |
| Girlish | - | Boyish | .521 | .803 | -.135 |
| Round | - | Sharp | -.021 | .945 | .008 |
| Hard | - | Soft | .139 | -.909 | .259 |
| Unpleasant | - | Pleasant | .311 | .025 | .884 |
| Dirty | - | Clean | .180 | .213 | .899 |
| Clouded | - | Unclouded | .270 | .033 | .895 |
| Ugly | - | Pretty | .347 | -.382 | .789 |
| Rough | - | Smooth | -.380 | -.204 | .821 |
| Recusant | - | Dutiful | .156 | -.659 | .620 |
| Banal | - | Unique | .460 | -.285 | -.550 |
| **Cumulative Contribution Rate** | | | .462 | .702 | .874 |

same quadrant of each corresponding picture. Table 2 shows the list of the characters and voice actresses. The eight pictures were synthesized with 16 voices (eight original and eight similar voices), resulting 128 audio-visual stimuli were constructed. The 20 participants who participated in Experiment 1 rated their impressions for each of the 128 stimuli using semantic differential method. The 24 scales shown in Table 1 were used in the rating, in the present experiment again. Moreover, the participants were requested to rate the degree of the congruency between the picture and voice for each stimulus using a seven-step unipolar scale. Each stimulus was presented through headphones STAX SR-404 in the level of LAeq = 75.5 − 83.3 dB and a computer display EIZO FlexScan SX2462W-PX. The participants sat on a chair in a dark sound-proof room and were requested to view and listen to the stimuli.

**Table 2.** List of selected characters and voice actresses. The voices with odd numbers are the original voices and those with even numbers are the similar voices.

| | Character name | Voice actress | ID |
|---|---|---|---|
| Picture | AYAME HAMAGUCHI | | P-1 |
| | UZUKI SHIMAMURA | | P-2 |
| | AKI YAMATO | | P-3 |
| | KIRARI MOROBOSHI | | P-4 |
| | SYUKO SHIOMI | | P-5 |
| | CHIERI OGATA | | P-6 |
| | ASUKA NINOMIYA | | P-7 |
| | SYOKO HOSHI | | P-8 |
| Voice | AYAME HAMAGUCHI | Masumi Tazawa | V-1 |
| | MIO HONDA | Sayuri Hara | V-2 |
| | UZUKI SHIMAMURA | Ayaka Ōhashi | V-3 |
| | YUMI AIBA | Juri Kimura | V-4 |
| | AKI YAMATO | Tomo Muranaka | V-5 |
| | RANKO KANZAKI | Maaya Uchida | V-6 |
| | KIRARI MOROBOSHI | Rei Matsuzaki | V-7 |
| | MIRIA AKAGI | Tomoyo Kurosawa | V-8 |
| | SYUKO SHIOMI | Lu Ting | V-9 |
| | RIN SHIBUYA | Ayaka Fukuhara | V-10 |
| | CHIERI OGATA | Naomi Ōzora | V-11 |
| | ARISU TACHIBANA | Amina Sato | V-12 |
| | ASUKA NINOMIYA | Shiki Aoki | V-13 |
| | KANADE HAYAMI | Yuko Iida | V-14 |
| | SYOKO HOSHI | Satsumi Matsuda | V-15 |
| | KOUME SHIRASAKA | Chiyo Ousaki | V-16 |

### 3.2 Results and Discussion

The rated values for each scale and each stimulus were averaged over participants. Factor analysis with principal factor method and varimax rotation was performed using these averaged data. The results showed that the three-factor solution accounted for 87% of the data variance. Table 1 shows the factor loadings. The three factors were labeled "activity", "evaluation" and "coolness" respectively, after the scales which showed high absolute values in the factor loadings. These three factors were identical to the results in Experiment 1.

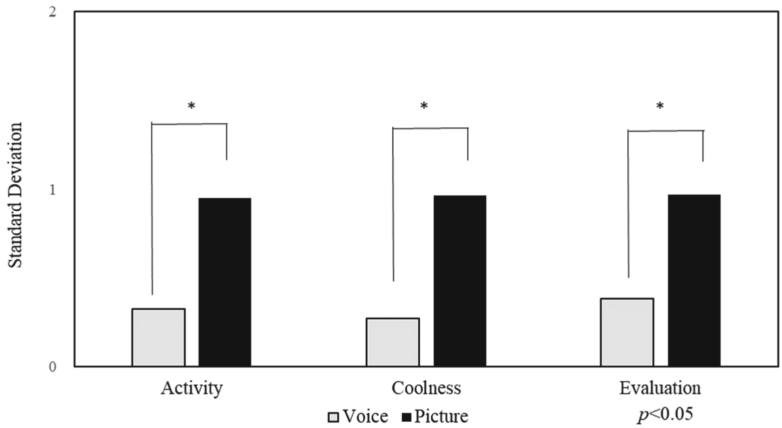
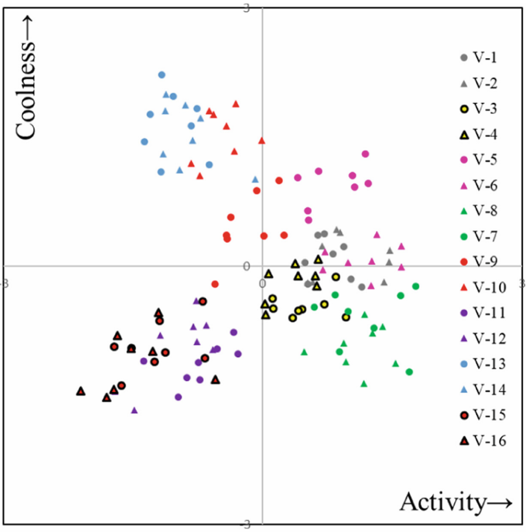


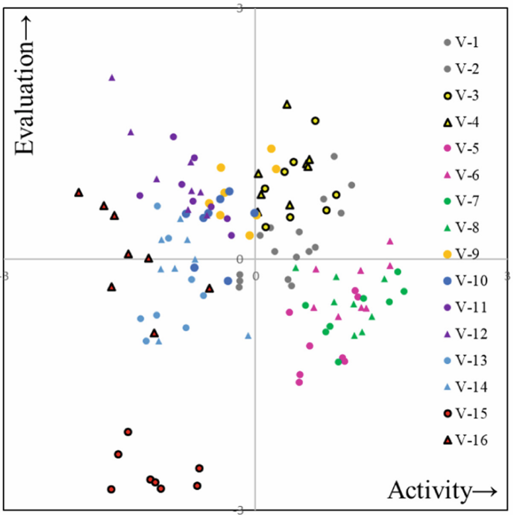

**Fig. 1.** Standard deviations were compared between the voices and the pictures for each factor of the activity, coolness and evaluation.

The audio-visual stimuli were plotted in the three-dimensional impression space. The plots of the stimuli tended to be grouped by the voices rather than the pictures in the space. To confirm this, the centroid and standard deviations were calculated for each voice and each picture. Figure 1 compares the standard deviations between the voices and the pictures for each factor. The results of *t*-tests showed that the value of the standard deviation of the voices was significantly smaller than that of the pictures for each factor ($p < 0.05$). This implies that the effects of the voices on the impressions were larger than the pictures. Figure 2 shows the plots of the audio-visual stimuli in the three dimensional space, spanned by activity, coolness and evaluation factors. The stimuli are identified by the voices in Fig. 2.

Figure 3 shows the congruency for eight cases of coincidence between the voices and pictures on the three factors, activity, coolness and evaluation. The congruency was compared among the cases using one-way ANOVA and multiple comparison tents with Tukey's HSD. As shown in Fig. 3, there was no significant difference between the case where the voice and picture coincided in all factors and the case where the voice and picture coincided in the activity and coolness factors, in the significance level of $p < 0.05$. Moreover, the congruency in the case where the voice and picture coincided in the activity and coolness factors shows a significantly higher value than the cases where the voice and picture coincided in activity and evaluation factors and in coolness and evaluation factors. These results imply that the coincidence in both the activity and coolness factors are important for the congruency between the voices and the pictures. The coincidence in the evaluation factor is less important for the congruency.




(a) Activity – Coolness plane




(b) Activity – Evaluation plane

**Fig. 2.** The audio-visual stimuli were grouped by the voices and plotted on the three-dimensional space, which is spanned by the activity, coolness and evaluation factors

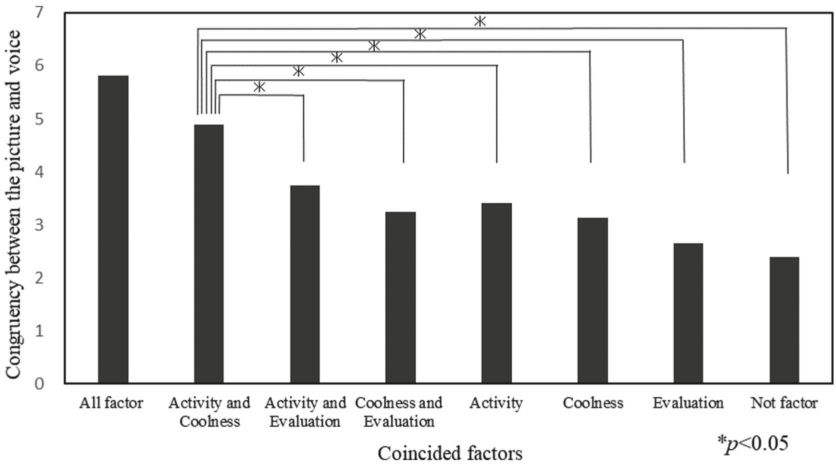



**Fig. 3.** Averaged degree of congruency between the pictures and voices were plotted as a function of cases of coincided factors.

Finally, the participants were divided into familiar group and unfamiliar group with the game. Moreover, the stimuli were divided into combinations where the original voices were paired with the pictures and combinations where the similar voices were paired with the pictures. Then, the degrees of congruency were compared among the four cases by two-way ANOVA as shown in Fig. 4. The results showed that there were no significant differences between the familiar and unfamiliar groups and between the original and similar voices, in the significance level of $p < 0.05$. These results imply that the familiarity with the game did not affect the congruency and the voices which possessed similar impressions could be replaced to the original voices.

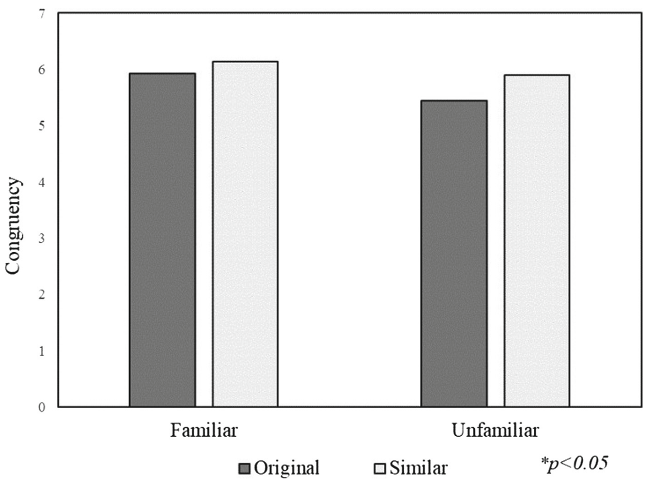



**Fig. 4.** Congruency between the pictures and voices for the participants familiar and unfamiliar with the game, and the original voices paired with the pictures and similar voices paired with the pictures.

## 4 Conclusion

In the present study, the impressions and congruency of the pictures and voices of characters were investigated in the context of the Idolmaster. The results showed that the impressions of the pictures and voices were illustrated by the three-dimensional space which was spanned by the activity, evaluation and coolness factors. It was also shown that the effects of the voices on the impressions were significantly larger than the pictures. Moreover, coincidence in both the activity and the coolness factors was important for the congruency between the pictures and voices.

## References

1. Japanese Ministry of Economy, Trade and Industry, Support for overseas operations of Cool Japan contents under supplemental budgets in 2012. http://www.cas.go.jp/jp/seisaku/cool_japan/dai1/siryou5.pdf. Accessed 28 Jan 2014
2. The Idolmaster official web. http://idolmaster.jp/. Accessed 5 June 2017
3. The Idolmaster Cinderella Girls: Starlight stage official site. http://cinderella.idolmaster.jp/. Accessed 5 June 2017

# Case Study: Game Character Creation Process

Cassiano Canheti(✉), Flávio Andalo, and Milton Luiz Horn Vieira

DesignLab, Universidade Federal de Santa Catarina, Campus Universitário, CCE, sala 101, Florianópolis, SC 88040-535, Brazil
canheti@gmail.com, {flavio.a,milton.vieira}@ufsc.br

**Abstract.** Although there are many types of game development methodologies, there are few academic studies about actual processes used by game creation studios. To fill this gap, this paper presents a study around a character development during the production of the game Heavy Metal Machines, a multiplayer competitive game about battle of machines, developed by Hoplon Infotainment company, localized in Florianópolis, Santa Catarina, Brazil. This case study used a qualitative research method, by watching and evaluating for ten weeks one of the character's development. This research was executed with the studio's team of developers. The author followed the team during the character development and finalization, watching the development meetings, and followed the creation and delivery of each artifact related to the character. All this process will be presented in this paper, showing in full depth the steps listed above, helping game designers and students improve their workflow, using industry proven methods.

**Keywords:** Multiplayer game · Character · Develop process

## 1 Introduction

The objective of this paper is to describe the character development process of the Heavy Metal Machines free to play game, developed by Hoplon Infotainment, a Brazilian game developer company based in Florianopolis, at Santa Catarina province. The Heavy Metal Machines game is a computer game of a gender known as MOBA, that means Multiplayer Online Battle Arena. The MOBA is a gender focused in competitive play and digital championships.

The digital game industry has grown to global scale in the last decades, since the first videogame, Magnavox Odyssey [1]. This industry makes more than U$ 90 billion each year [2]. Brazil entered and has been growing in the game development business, with companies across the country, like Hoplon Infotainment. Founded in 2000, Hoplon is a game developer company with more than 70 employees and has developed international games like Taikodom and Taikodom: Living Universe. It's focused in develop computer, multiplayer, free to play games.

The Heavy Metal Machines game has multiple characters, each of them with unique styles and different roles in the game. In order to reach the established objective, that is the game character development process, it was required verify the team


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 343–354, 2019.
https://doi.org/10.1007/978-3-319-94619-1_34

involved, the creation and development steps and each deliverable at each step, resulting in a fully, completed character.

The methodology adopted was the company's development process observation, crossing with some process founded in industry's most used bibliography; and the analysis of the deliverables on each development step.

## 2 Methodology

The objective of this article, from the above text, is to present the game character development process in a practical case, and to watch the applied methodologies, the involved professionals and the steps deadlines. In order to make this qualitative research, it has used the field research method, point by Gil [3].

It was chosen the Heavy Metal Machines game because of the physical proximity made the dada collecting easier, and thanks to the good willing of the company's team who has agreed in offered their space to analysis and data collect. From this, Duarte and Barros [4] said the intentional choice is the proper research type for this case, because the availability and proximity, and causing no loss or damage on the study.

First, it has built a bibliography study regarding the game development process and methods. Second, it has defined the company and game as the study object, and after it has analyzed how the company uses the methodologies and develops its characters in a practical matter. Therefore, the case study is the proper choice because it is efficacious to provide a detailed analysis of the researchable object through real events qualitative data. And it describes this phenomenon within the context [5].

## 3 Characters and Games Development Process

The game develop is a complex task that requires technical knowledge from multiple, different fields. Although each company and each project have its own unique traits, like the game's scope, technology, company's size and business model, in any game develop company you can find a similarity related to the team's role.

Chandler [6] said that the professional roles you find in a game develop company are: engineering, art, design, audio, quality assurance and production. Each role has its own, unique technical features and if they are used together it is possible to develop a game. But depending of the project size, a single person may assume multiple roles, thus cumulating game develop functions. In other cases, the professionals work together in the same role or field, forming sub teams. And each team has its own leadership – a team leader who guides the team through the company's and project's goals. But for the final goal, that means the game development, each role is relevant and important at the same.

The engineering team is composed of programmers, computer engineers, computer scientists and similar. They are experts in construct the programming code that will make the game works. Therefore, they should work with every single aspect of the game – game mechanics, animations, graphic elements, audio elements, content creation tools, client-server technology and so on. They establish the technical art requisites for the games, with the art team. They also need to optimize the game's technical performance. So, they need to implement all rules, features and tools of a game.

The art team is responsible to create all graphic artifacts of a game: characters, scenarios, visual interface, animations, visual effects and any graphic material used both in or out the game, but related to it. Besides, Chandler details:

> *"The art works with the designers together to define the worlds and cinematic that will be used in the game, and also work with the engineering to determine how the technology can be used easily in the art develop pipeline. If it is necessary to create several graphic assets, it is likely that the artists number would be bigger than other team members in a proportion of two to one"* [6].

There are some different artistic work specializations, such as 2D concept and animation, 3D modeling, texturing and animation, concept art, technical art, visual effects and others. Each game and each team has its own artists division according with the necessities.

The game design team is responsible to create the game mechanics, rules, usability, content, storyboard, and all the aspects of the game involved with the user, or the gamer, will experiment when playing the game. Usually the game design has its own divisions like system design, dialog and texts, level design and game balance. Fullerton [7] points that the game designer is the one who advocates for the player, that is who sees the game with the gamer's eyes. So, the game designer must focus on the gameplay – the act of playing a game. Most of his work is to maintain the focus in the gamer experience during the gameplay and not be distracted with other development issues. His objective is assuring that the game will have the best possible gameplay.

Most of the game design process to achieve a good gameplay experience lays in the iterative process. Fullerton [7] explains that these iterations are cyclic steps starting with the ideas generation, following by implementation and tests. After the tests the results are evaluated and the process starts over, making adjustments on the original idea based on the problems found during the tests, and so on until the experience reaches level desired by the game designers. Only through this iterative cycle it is possible to polish the gameplay and improve its quality. The game designer is responsible for each step of this process.

The audio, usually, is outsourced in a game project. Even so it's a very important area and it has a major role in building the atmosphere, narrative tone and environment of the game as much as the art and animations does. The audio team is responsible to construct and apply the sound effects, the soundtrack, the songs used in the game, the in-game scenario songs that compose an atmosphere, and the voices for the characters and narrators. The amount of audio assets used depends on the game's size.

The production is the field dedicated to manage and follow-up the game development. Chandler [6] points the main objective of a producer is holding the team united and productive. The producer has the responsibility to coordinate the game developing team effectively. Usually the producer is the team's leader and he is involved with the other leaders for both resolving creation problems and keeping the development schedule on time.

One of the main qualities of a game producer is the communication. The producer must keep the team motivated with the work. And he also needs to deal with multiple tasks, work conditions, pre-requisites and deadlines. Discipline and a good sense of organization are good qualities for a game producer be able to handle everything.

The last, but not least, game development area is the Q.A. team, or the quality assurance. The Q.A. team is composed of game testers who conduct several game tests, including exploratory tests and user case tests to search game problems and flaws. Wilson [8] said that both tests are important, and both achieve different goals. Besides, the test team tests the game back from the beginning until the end of the project. It is the last line of defense to find and resolve problems before the game is launched to the public.

Another important aspect of the game development is the development process itself. Moore [9] points that the game development process is divided in three major cycles: the pre-production, the production and the postproduction.

The pre-production is the starting step. It is the step when the following are defined and created: game concept and gender, the game platform, the basic gameplay, the project documents overview, budget, risks analysis and prototypes. The construction of up to three gameplay prototypes during the pre-production is fundamental to evaluate and approve the project as a whole, said Sztajer [10].

When the project was approved, the production step gets started. The team starts to construct the game code and its artifacts. Thus, it is during the production that the game is built. Similar to the prototyping, the game construction is an iterative process, with cycling creation, tests, analysis and evaluating steps.

When the game code is finished and the game content is completed, the next step, postproduction, is the game polishing and packaging. Rouse [11] alleges the post-production it's the moment when the game is transformed from good to excellent. Usually the team is focused in make final touch in the game experience, in gameplay, animations, graphics, usability, audio, wherever it is. The iteration process happens all the time during postproduction, to validating the polishing adjustments. And the packaging is the process to prepare the game to its launching, and archiving this code, assets and files to latterly use in other projects.

## 4 The Heavy Metal Machines Game

"Heavy Metal Machines" (Fig. 1), according to its official homepage, it is an online computer game, that a team with four players plays the game against other four players. Each player chooses an unique character and pilots its vehicle with exclusive abilities, and the game objective is to take a bomb located at the center of game area and deliver it to the enemy base. The team that delivers the bomb three times is the winner. The game is based on fictional universe of its own, of a post-apocalyptic scenario when the humanity lives in insane battles with modified vehicles, with heavy metal as their soundtrack.

Since Heavy Metal Machines is a free-to-play game, there is no fee requirement to play the game. Its business model relies in its in-app purchases, or IAPS, said Luton [12]. The Heavy Metal Machines IAPs are mostly different visual model variations of


**Fig. 1.** The "Heavy Metal Machines" game cover picture.

the character's vehicles, with only aesthetic purposes in the game. Other IPAs Heavy Metal Machines have are speed booster for experience points and virtual currency, with helps players to leveling up and buying other rewards quickly.

When a player starts a match, he or she is directed to a game screen to choose on which character he or she wants to play with. This character selection works for just one match. On each new match the players must select their characters. It's impossible to choose a character that has already been chosen in the same team, but it's possible to repeat characters chosen in previous matches.

The game characters are divided in three different categories. Those categories were created according to their main function in the Heavy Metal Machines' gameplay. The categories are: *transporters*, *interceptors* and *supports*. The transporters are the best characters to deliver the bomb to the enemy's base. The interceptors are specialized to attack the enemies and intercept the enemy bomb carrier, and the supports are those who helps their allies to survive, repair them or giving them some temporary buff during the match.

Each character has 3 weapons: the main weapon, used more frequently; the secondary weapon, stronger than the first, but which is used once at 10 s average; and the special weapon, that cause greater effects in the game, but it has to be charged if a player wants to use it. To charge the special weapon, the player has to fight the enemies, repair allies and carry the bomb. Besides the weapons, each character has an unique advantage to make it even more different from the others. This advantage isn't an active command; instead it works automatically, like for example an ability that increases the vehicle's speed, and so on.

These character weapons and advantage are complementary each other. They were built to be used together in a way to maximize the character's function during the gameplay. But even when different characters share the same function, each character has its own gameplay strategy, including use of some tactics depending on the enemies and allies of a specific match.

A player can buy weapon's and vehicle's upgrades during the match. It can be done by gathering a specific resource called metal, which is used as currency to buy these upgrades. The player can gather metal by destroying others player's vehicles and by delivering the bomb. These upgrades increase the weapons and vehicle's capabilities, like an upgrade that allows the weapon not only cause damage but also apply a slow effect on the targeted vehicle for a few seconds. At the end of each match, all purchased upgrades are discarded, with the excess metal. These resources don't persist to other matches.

# 5 Heavy Metal Machines Character Construction Process

The researcher has used the research method of reviewing the develop team during the construction of one of the Heavy Metal Machines character, from the beginning up to the character final steps, when it has been ready to public launch. The object of this case study is the character called Stingray, an Asian teenager.

Each character is a modular object of Heavy Metal Machines game. It can be developed separately from other game features and content. Furthermore, it is possible to include new characters in the game without major gameplay changes, with, if applicable, minor impact in usability and other, secondary, game mechanics. The biggest impact of a new character in the game is on its game balance, especially regarding the old characters interaction with the new one.

Therefore, a character can be fully developed in parallel with any other activity of the Heavy Metal Machines team. So, even the professionals involved with develop a specific character works exclusively at it, though some members don't have to work in each single step of the character development process. Each studio member joins with the character team just in specific moments, according to the development process. But firstly, it is necessary to understand what a character in the Heavy Metal Machines game is. Comprehend what are the artifacts that should be built and delivered on each step.

## 5.1 What Is a Character?

The Heavy Metal Machines character is a result of many different features built for the same purpose and theme, according to a careful review of the character development process. That means, every single character has: its own background history, part of the Heavy Metal Machines fictional universe; its own vehicle and character's name; its own avatar and its own vehicle – the core gameplay object (see Fig. 2). This vehicle has its own weapons and abilities: three weapons, one passive advantage and nine exclusive upgrades. The characteristics like vehicle speed, weapons damage and vehicle hit points are unique per character as well. Lastly, each character also has: its own scenario – a visual concept of the character's lair; its own in game dialogs with exclusive voice; and a theme soundtrack, inspired on heavy metal music style.

Each one of these features requires a group of artifacts that should be built. And it's important to keep all features within the same character's vision; the team leaders work with their teams to ensure the character's coherence and unity.

The game designers should write all the character related texts used in the game: history, descriptions of weapons and upgrades, narrative used in model descriptions and tips how to play with the new character. All of these are documented in a text editor. Also, the game designers should determine all the character game statistics numeric values. The result is compiled in a math calculator sheet.

The artists draw the character's scenario, avatar and vehicle, all of these in many different pose. Each drawing should respect the technical specifications established by the engineers to integrate it with the game software. The artists also should model the vehicle's 3D model. The character's vehicle has some different visual models, and the artists should create them. At last, they should create the vehicle's visual effects.

**Fig. 2.** The avatar of "Stingray", one of the Heavy Metal Machines character (left) and its vehicle (right).

The audio team should compose the character's soundtrack and sound effects and save both in audio files. They also record the character's dialogs, using dubbing actors in a recording studio.

The engineers should deliver the character's game code working properly: its weapons, vehicle, upgrades and any new type of visual special effect. Also, the new character should be included in the content edition tool that is integrated within the game software, so the game designers can register the character's numerical values and texts in the game. Table 1 presents each artifact, divided for each team involved.

**Table 1.** Character's artifacts divided for each team.

| Team | Artifact |
|---|---|
| Game design | Text of characters, weapons, upgrades, tips and models descriptions<br>The characters' vehicles, weapons and upgrades numerical values |
| Art | The character's drawings of its avatar, vehicle, weapons and scenario<br>The vehicle 3D model, texture and animations<br>The vehicle and weapons visual effects<br>The vehicle additional visual model styles' drawings and 3D texture |
| Audio | Character theme soundtrack<br>Weapons and vehicle audio effects<br>The character's in game voice dialogs |
| Programming and engineering | Special effects, weapons and vehicle mechanic programing code<br>Content tool integrated with the game software |

## 5.2 Character Development

The character development structure is similar to the game itself: it is divided in preproduction, production and postproduction, and it process is iteration-guided, with the analysis and approval cycles.

The preproduction starts with the new character's gameplay mechanics definition. During two to three days the game designers have defined the new character's gameplay style, first from its role – support, interceptor or transport – and, then establishing its weapons, upgrades and advantage. These definitions are not random: the weapons, upgrades and advantage should be compatible each other, thus the character's gameplay has some synergy with its abilities and making its game experience unique, even with the same role of other characters.

In order to guide the creative process, the game designers establish keywords for the character. The keywords work for a high description of its gameplay. From the keywords, the game designers discuss and propose the character's abilities and weapons (Fig. 3). Because the team records all weapons and properties of all characters used in the game, they avoid unnecessary repetitions to assure each character has an uniqueness gameplay.


**Fig. 3.** Stingray using one of his weapons during a game match.

However, since the initial steps are mental and imaginative task, it's possible to multiple interpretations from each team member. As the game designers descript each new character, there were always multiple interpretations; each member had their own character view. There was no scientific basis on the character concept and style, with a completely empirical process. To solve this, the game designers applied Jung archetypes, according to Krawczyk and Novak description [13].

**Defining Archetypes.** Since then, the game designers have been adopting the Carl Jung archetypes [14] found in both mythology and literature, to create their characters. The use of Jungian archetypes is important because helps the team to share the same vision about the character and dictates its history behavior and visual style. The archetype is constructed by images of collective nature, occurring through all the planet as myth components and as individual product of unconscious origin. The archetypes are mostly used within products, brands and even histories. Mark and Pearson [15] points that even the Star Wars movies series have used the archetype model, inspired in the Joseph Campbell's The Hero with a Thousand Faces book, to create heroic figures and mythic plots.

Since the game designers has begun to use archetypes to create their characters, the conflicts of understanding the characters diminished expressively. And, with different archetypes, they could change the character styles and create difference between them. There are twelve different archetypes, established by Mark and Pearson [15], used in nowadays brands, products, plots and fictional characters. Table 2 presents these archetypes, each one with its slogan and goal.

**Table 2.** The twelve archetypes and their characteristics.

| Archetype | Slogan | Goal |
|---|---|---|
| Innocent | Free to be you and me | Be happy |
| Regular Guy | All men and women are equal | To be part of |
| Hero | Where is there will, there is a way | Make the world better |
| Caregiver | Love the others like yourself | Help others |
| Explorer | Don't seize me | Experiment a free world |
| Outlaw | Rules are made to be broken | Undertake what is not working |
| Lover | You are unique | Find and give love |
| Creator | If you can imagine it, it can be done | Accomplish a vision |
| Jester | You just live once | Have fun |
| Sage | The truth will set us free | Understand the world |
| Magician | I'll make the things happen | Transform |
| Ruler | Power it's the only thing | Control |

Since Heavy Metal Machines is a futuristic, dystopic, post-apocalyptic game, and with some dose of humor, this scenario made it easier to introduce some characters using the archetypes of Outlaw, Jester and Hero. That doesn't mean the other archetypes could not be used, but only they would be used with more exotic characters. So, of the thirteen playable characters, there are three Outlaws, three Jesters, two Heroes and one of each Creator, Magician, Explorer, Innocent and Ruler.

### 5.3 Artifact Construction

At the end of the character concept step, the game designers document the mechanics and present the new character for the rest of the team. Then, the art and engineering joined the process and their work has begun. The art should create concept arts for avatar, vehicle and even the weapons of the new character. And the programmers should evaluate the mechanics and build a playable prototype. At this point, usually one programmer and two artists join the character's team.

Meanwhile, the game director must decide which musical style will be used for this new character, choosing the bands, songs and heavy metal styles related to it. It's a very important step, since Heavy Metal Machines is a game strongly related to heavy metal culture, and each character has its own heavy metal influence. This choice will guide both the avatar's and the vehicle's visual concept. The game director chooses the heavy metal references from the archetype chosen for this new character.

During the preproduction, the programmers built the gameplay prototype. Then, the game designers test the prototype and evaluate the character's gameplay experience: its weapons, upgrades, and basic gameplay. During this step occurs many cycles of iteration, adjusting and improving the gameplay through tests and validation. At the same time, the art team creates visual effects to use in these gameplay tests. The preproduction takes three weeks to be done. The objective after three weeks is to finish the gameplay and the approved avatar, vehicle's and weapons' concepts.

Gradually the team begins the production step, as each member finish its preproduction artifacts. During the production the team focused in built every single artifact that compose the character: the vehicle 3D model and its visual variations, the avatar and scenario images, the mechanics programming code, the description texts of the character's weapons, abilities and history. If necessary, the graphics computer team can be included during production step to deliver solutions for new visual effects.

The audio team also joins with the character's team during the production step. The Hoplon studio outsources all the audio production, so the game director describes the character concept, the archetype chosen and the music references of this new character. The audio team starts to compose the character's song theme and the gameplay audio effects. Also, they search actors and actress for the character's in game voice.

The game designers should write the description texts and keep testing the character. These tests data are registered and the game designers analyzed them, with the intention of balancing the new character within the rest of the game. The game balance is a highly iterative process, with constant balance adjustments, because each new character affects the game dynamics and because the players community has always been trying to found new gameplay tactics and ways of play thus interfering the balance between the characters.

The art team should work with the vehicle's 3D model, animations and texturing. They should draw the avatar's many poses, build the vehicle visual variants and produce the weapons special effects.

The Q.A. team should help the game designers testing the new character and reporting any problems and bugs they found during the tests.

The production step lasts in 5 weeks. The team gradually enters to the postproduction step as the team members were concluding their artifacts. Then, they should solve the problems found and make the final adjustments on the gameplay, animations and effects. The audio artifacts are integrated with the character. Finally, all the character texts are translated to other languages: English, Spanish, French, German and Russian. The postproduction step lasts in 3 weeks. So, the character takes 10 weeks to be constructed (Table 3). At the end of the tenth week, if there are no delays or problems, the character is finished and ready to be published in a game update.

Every time they complete a character, the studio head gathers all the team leaders (game designer, art, programmer and Q.A.) to a *post morten* meeting. They've made a

**Table 3.** Character development weekly work schedule.

| Activity | Week |
|---|---|
| Preproduction: character mechanics and archetype definition | 1 |
| Preproduction: character prototype construction and concept art definition | 2 to 3 |
| Preproduction: prototype tests and validation | 2 to 3 |
| Production: artifact building and code implementation | 4 to 8 |
| Production: gameplay tests | 4 to 8 |
| Post production: visual polishing, mechanics final adjusts and code optimization | 9 to 10 |
| Post production: quality assurance tests | 9 to 10 |

total review of the character's process, discussing the positive and negative aspects, what was right and what was wrong, and what can be done to improve the process for new characters. That way, they are always improving the development process.

## 6 Conclusion

The objective of this paper was to evaluate how a theoretical game development process is used in practical work, by careful reviewing a specific part of the game.

The data collecting process was the witnessing of a full character development, since its beginning up to its end, when the work has finished and making the character available to public launching. After ten weeks of reviewing the character construction, we could verify Hoplon uses a bibliography-driven, well known, game develop process with preproduction, production and postproduction. They divided the teams according to each expertise area, one leader each team. This method helps to assure a better consistency with the development, and at the end of each character it's possible to compare it with other character's development, correcting problems and making adjustments to increasing the process even more.

However, it's an adaptable process. The three-step process assume each step is clearly separated from each other, but this is not the case. The transition step is gradual, according to each task and each team member expertise. And also, the internal, technical, process from each task itself are divided in the three steps of preproduction, production and postproduction. The macro step division works to guide the team regarding the current development status and helps them to keep on schedule.

The Heavy Metal Machines is a game that depends on many characters frequently, each character unique and different from each other. So, its fundamental a fast, efficient develop method. Therefore, the adopted process it's very important, and also important is the necessity of revise it. That's the reason when each time a character develop is finished, the team leaders evaluates and adjusts its process in one or more meetings. Then they can optimize the schedule, diminish the cost of each character and even improve the delivered quality.

For futures discuss and researches we can review other companies' character development process from games with similar mechanics, and compare both process, artifact delivery, team involved and work schedule.

**Acknowledgments.** We would like to thank the CNPQ institution for its financial support. We also want to thank the Universidade Federal of Santa Catarina and its employees, who provide us the structure and support to write this paper. And also, thanks to our DesignLab colleagues with their patience and helpfulness to providing so many tips and adjusts for this. At last, thanks for Hoplon company and its team of passionate developers, our friends, who kindly authorized the information we used here.

## References


1. Kent, S.L.: The Ultimate History of Video Games – The Story Behind the Craze that Touched Our Lives and Changed the World. Three Rivers Press, New York (2001)
2. Adrenaline: Mercado de games gera US$ 99,6 bilhões mundialmente e PC é a plataforma mais lucrativa, indica pesquisa. Disponível em: http://adrenaline.uol.com.br/2016/04/30/42730/mercado-de-games-gera-us-99-6-bilhoes-mundialmente-e-pc-e-a-plataforma-mais-lucrativa-indica-pesquisa/. Acesso em 30 de junho de 2017
3. Gil, A.C.: Como elaborar projetos de pesquisa. Atlas, São Paulo (2002)
4. Duarte, J., Barros, A.: Métodos e técnicas de pesquisa em comunicação. Atlas, São Paulo (2005)
5. Yin, R.K.: Estudo de caso: planejamento e métodos, 4th edn. Editora Bookman, São Paulo (2010)
6. Chandler, H.M.: Manual de Produção de Jogos Digitais, 2a ed. Bookman, Porto Alegre (2012)
7. Fullerton, T.: Game Design Workshop – A Playcentric Approach to Creating Innovative Games. Morgan Kaufmann Publishers, Burlington (2008)
8. Wilson, D.: Quality Assurance: A Methodology for Wide-Spectrum Game Testing. http://www.gamasutra.com/view/feature/132398/quality_quality_assurance_a_.php. Acesso em 30 de junho de 2017
9. Moore, M.E.: Basics of Game Design. CRC Press, Boca Raton (2011)
10. Sztajer, P.: The 15 Steps of (Particular) Pre-Production. http://www.gamasutra.com/blogs/PaulSztajer/20130324/189171/The_15_Steps_of_Particulars_Preproduction.php. Accessed 30 June 2017
11. Rouse III, R.: Game Design: Theory & Practice, 2nd edn. Wordware Publishing, Texas (2005)
12. Luton, W.: Free 2 Play - Making Money from Games you Give Away. New Riders (2013)
13. Krawczyk, M., Novak, J.: Game Development Essentials: Game Story & Character Development. Cengage Learning (2007)
14. Jacobi, J.: Complexo, arquétipo e símbolo na psicologia de C.G. Jung. Editora Vozes, Petrópolis (2016)
15. Mark, M., Pearson, C.S.: O Herói e o Fora-da-Lei: Como construir marcas extraordinárias usando o poder dos arquétipos. Editora Cultrix, São Paulo (2001)

# Disruptive Games: Power and Control or Fantasy and Entertainment

Ekaterina Emmanuil Inglesis Barcellos[1(✉)], Galdenoro Botura Jr.[1], Eric Inglesis Barcellos[2], Milton Koji Nakata[1], and Lívia Inglesis Barcellos[1]

[1] Universidade Estadual Paulista - UNESP/FAAC - Campus de Bauru, Av. Eng. Luiz Edmundo Carrijo Coube, 14-01, Bauru, SP 17033-360, Brazil
kettyinglesis@bol.com.br, galdenoro@gmail.com, livia.barcellos@gmail.com, milton@faac.unesp.br

[2] Fundação Armando Álvares Penteado - FAAP, Rua Alagoas, 903, Higienópolis 01242-902, SP, Brazil
eric.barcellos@gmail.com

**Abstract.** Video games are a complex form of culture, creative entertainment and representative art, which surpasses, in technology and relevance, other forms of audiovisual expression. Since the conception of the first games, through computer technology and military knowledge, games have become an entertainment, translated and molded as interactive graphic images. The initial aim, which was to create a kind of distraction, outgrew itself and made possible to simulate and provide users with the psychological sense of control and power over fear, decisions and over the imminent danger of new challenges (real or imagined), all at same time. From the pioneer arcade to the modern consoles and multiple platforms seen today, the games evolved to meet users' expectations in search of simulation, interaction and immersion experiences. In this trajectory, developers, programmers and designers applied all the technological advances possible in attempt to achieve complete man-machine interaction. This article seeks to identify if the games in the current stage, loaded with hyper-realism, in relation with their 'cartoonized' counterpart, aimed at a new generation of users eager for disruptive technologies and for interactions that assure the sensation of influence and control over fear.

**Keywords:** Games · Graphics · Hyper-realistic design
Human-systems interaction

## 1 Introduction

Video games originated in the Cold War era, when they were developed as a form of strategy stimulation, and later adapted to relieve the tension derived from the post-war period. In addition to creating a new type of distraction, games simulated and provided users with the psychological sense of control and power over fear and the imminent danger of new confrontations [1, 2].

The first interactive digital electronic game called the "Tennis for Two" was created by Higinbotham in 1958, with very simple graphics. It is considered a precursor


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 355–365, 2019.
https://doi.org/10.1007/978-3-319-94619-1_35

(although there were two other records, “Bertie the Brain” (1950), and “OXO” (1952), since its movements and graphics’ frame rate refreshed in real time [1]. In this period, video games were developed in gigantic computers, restricted to high-cost scientific or military research, inaccessible to ordinary users. Outside of the military and academic environment, Baer and Russell, familiar with new technologies and military designs, were pioneers in popularizing electronic games. Baer created the home console (Brown Box, 1968/Magnavox Odyssey, 1972), adapted to TV sets, while Russell, at MIT, created games with comic book narratives such as “SpaceWar” (1968), simulating the destruction of enemies [3]. Until this period, the graphics were restricted to presenting simple themes of war, army and fictions on the conquest of space (begun in the 60s). From the 1970s onwards, with the production of the electronic chip and other technologies that emerged in Silicon Valley, a great change began in the creative profile of games and the offer of systems and consoles started to grow.

Since the release of the first home video game console, platforms started to differ from themselves and the pioneer arcade, and systems such as Atari, Nintendo DS/Wii/Switch, PlayStation, XBOX, PC, among others, for almost 50 years. Games evolved from technological hobby to interactive visual media of great relevance, maintaining themselves as systems in which rules interact. Players need to model in their minds how codes act, so that they can use those interactions and achieve goals. In this sense, it is inductive that games provide the user with a sense of control and power, dealing with strategies and mastering tactics, or even interacting with a fanciful entertainment [4].

The maximum and best user interaction and immersion experience within the game is the target of the game [5]. In order to retain players and create a title or series of success, it is essential to meet the expectations of the user to be offered by the proposal of the game. In addition to well-known characters and designs such as Marvel franchises or DC Comics franchises, we can cite countless games that have standardized the players’ preferences, with FPS proposals, which, for a part of the players, helps and increases the realism and the immersion [5]. Another important aspect that does not affect this research is the sound quality. For some players, the sound more than potentiates and interferes with the simulation and interactivity that the game achieves, defining the level of credibility and immersion, and being also a factor of withdrawal, when it does not reach reciprocity. In this aspect, poorly drawn and poorly defined graphics may justify disinterest in certain games [6, 7].

To a greater degree, this research seeks to define the importance of graphic styles, and how much they determine the choices of games, and their performance [8]. Subtly realistic graphics mixed with animation effects, as seen in games such as GTA, Counter Strike, or those with a more realistic touch, such as Last of Us, are being bombarded with scripts and hyper-real design; or the fanciful hyper-realism of Tomb Raider, Final Fantasy, Assassins Creed, God of War or Resident Evil. On the other hand, ‘cartoonized’ or pixelated, timeless and nostalgic titles such as Donkey Kong, Super Mario and its franchises, Zelda, Tetris, Pokemon, continue to be as appreciated, as the new favored games like League Of Legends and Overwatch rise. Despite the evolution of gaming graphics, a nostalgic resumption adds new players, loyalty to younger audiences in consecrated fields. In addition to these, upcoming indies titles, from companies outside the mainstream sphere (dominated by gaming giants like SONY, Microsoft and Nintendo), are following this style.

In this sense, the optimization, the visual graphic appeal, the agility of actions and levels of the stages, the inserted technique and the gameplay are determining factors to obtain the desired result, providing a satisfactory gaming experience, where the user acquires and extrapolates limits and capabilities through qualified entertainment [7, 9].

The games extrapolated greatly its purpose and the advances obtained with the insertion of games and technologies of gaming like VR or Hologram [10] with disruptive proposals have expanded to diverse areas, from entertainment to health and education resulting in anthropological and social studies built upon its functionality, besides scientific and academic studies [10]. We analyze all the benefits and strategic qualities already recognized in games, but also the disadvantages of its excessive use by many users. It is likely that the identification with hyper-reality, present in the graphic realism of current games is so close, that the user becomes over influenced by the actions chosen and stories in the game environment, generating distortions and transferring them to the environment, where the constant and continued exposure to simulations and immersions are amplified [11]. Meaning, there are consequences acquired from gaming, such as optimized reasoning and reaction, the insertion of new social and ethical codes, created from the interpretation of a simulated reality, or even affecting the gamers in their personal and professional life, more prepared, more aggressive or acquiring skills of strategy and reaction [11, 12].

With this objective and facing multiple questions, the article proposes to analyze and question how the design and graphics aspects of the games affect, contribute or change users' behavior to the proposing a visual, emotional and physical evaluation, based on habits and preferences of gamers. It seeks to understand, through the outline of the answers, and its respondents, to what extent the images, the design, and the graphics, accompanied by the narrative, evolves the feeling of immersion in a simulated and parallel game reality, and provide or improve user expectations. In a conclusive opinion, he suggests proposing more assertive paths in design proposals and in the evolution of graphics and stories from user-oriented design and a better Human-machine interaction between them and games.

## 2 Presentation and Analysis of the Results

The results presented from graphs, performed through the crossing of the data provided, are extracted from the analysis performed on a sample of 102 responses, obtained through a questionnaire sent through WEB social networks, and answered online by players of electronic games between 18 and 25 years in Brazil. Initially, the stratification was done for hours dedicated to the games and by the type of platform chosen, showing that the vast majority, about 48%, use their personal computers (PCs) to play. In this case, those who play between 5 and 10 h a week predominate, with 33% of the total, followed by those who practice more than 35 h per week with 23% of the total. The second platform most used in the country, within the age group analyzed, is the PlayStation console, with 28% preferably among all players, especially among those who play up to 10 h/week, with 52% of those who have opted for this console. Practitioners who do not prefer any platform, that is, play in any of them, correspond to 24%, especially among those who practice 5 and 10 h/week, which corresponds to 24% of these. Other gaming platforms have not achieved a significant percentage of responses. Figure 1 shows more information.

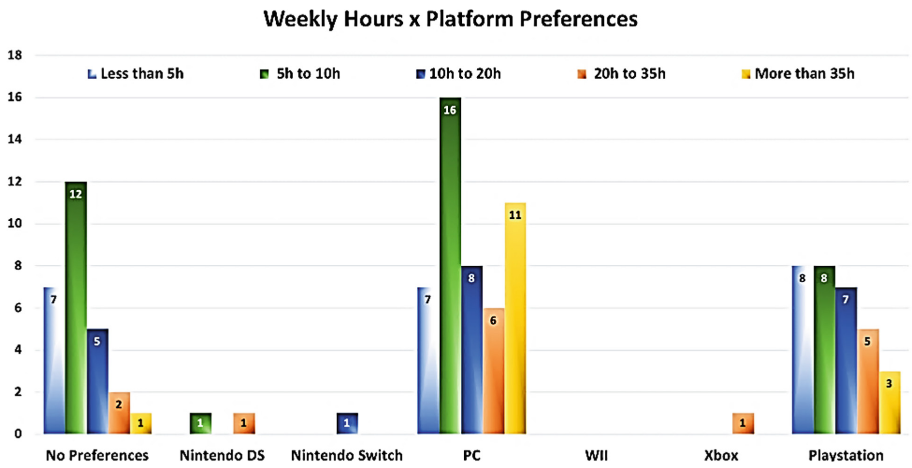



**Fig. 1.** Relation between "platform preference" and "number of hours played weekly". Author's chart, 2018.

When asked about their preferred platform as a function of their playing time, the choice for playing using personal computers (PCs) also prevailed, with 43% preference, especially among those with more than 15 years of gaming experience; 23% of those who chose by this platform, showing that the acquired habit (in a time when there were no other ways to play) prevailed over the possible facilities offered by more modern consoles. Nintendo has chosen the platform for few players being, however, practicing for over 15 years, with the DS for 5 to 20 h per week and the switch to a player with more than 35 h a week. The Wii wasn't chosen as a preference.

The PlayStation platform ranks second among the platforms preferred for respondents, with 28%, especially among those with experience over 15 years of games, showing that for these the preferred platforms correspond to the personal computer (PC) and the console PlayStation, with 72% of the total. The Nintendo DS is a choice for gamers who play more than Those who do not care about the type of platform correspond to 24% of respondents, as shown in Fig. 2.

Research has shown that interest in a game, for most, regardless of experience they have. What generates greater interest is the characteristic that the game presents in a global way, without highlighting any of them separately. No special features (among the suggested ones), such as technological, sound or visual effects, alone, lead them to take an interest in a given game. For those who preferred to highlight a particular feature, 13% preferred scenarios and characters close to reality. This aspect reinforces the interest of the research that seeks to observe the effect of the graphics and the characters on the players (Fig. 3). Nowadays, developers linked to the big film

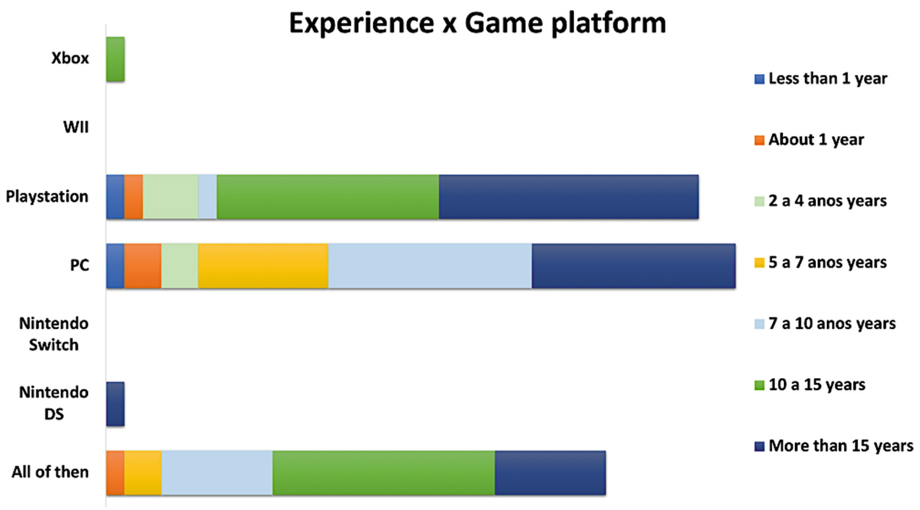



**Fig. 2.** Choice of platform based on number of years played. Author's chart, 2018.

companies have used techniques and effects of this industry applied to the games, exploring graphics that tend more to the visual realism and the visual hyper-realism, in the scenarios and personages, as in the effects.

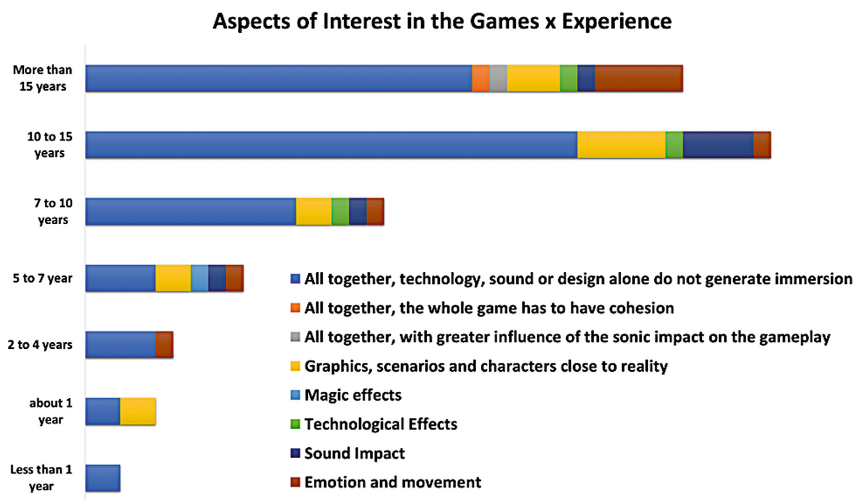



**Fig. 3.** Aspects that generate interest in a game based on previous experience. Author's chart, 2018.

The choice of the respondents confirms a good acceptance of this type of visual proposal. When asked why they play electronic games, 82% of them revealed that the practice of playing games is a form of distraction and entertainment; 62% do so in search of relaxation. Another 41% seek challenges and overcoming, while for 46% it is a form of abstraction. The sense of escape from reality was the choice of 55% of them through their practice. Another 30% say they try to isolate themselves through games. The answer to this question was not exclusive, and the respondents were given the option to tick more than one option in the proposed form. Figure 4 presents the data analyzed.

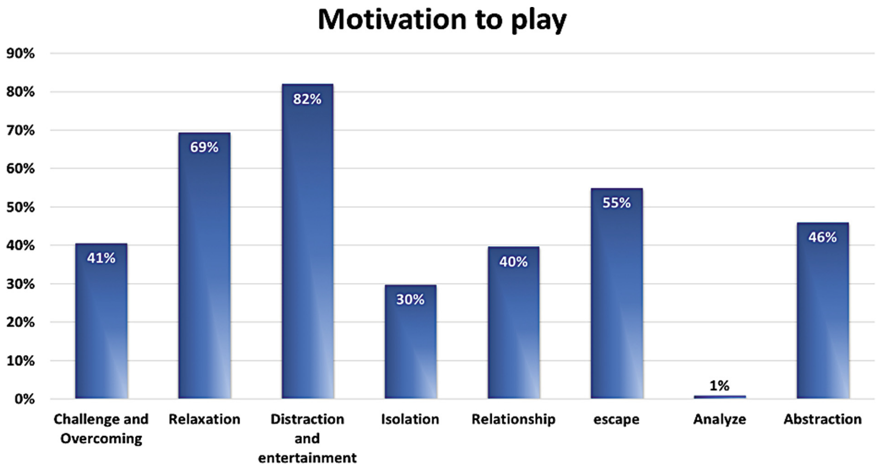



**Fig. 4.** Motivation for the video game practice. Author's choice, 2018.

The capacity for personal change through practice gained from gambling was one of the questions asked in one of the questions. What happens when analyzing the results from the data obtained is that there is the awareness that the continuous practice of the games generated an improvement of reasoning, learning and the development of strategic thinking for those who play, such as more identified skills. The improvement in reflexes (22%), decrease in dispersion (26%), improvement in relationships with other people (32%) and focus (41%) were also observed by them, as shown in Fig. 5.

Behavioral changes were identified by respondents, with 44% reporting that they became more attentive after they started playing. Relaxation and a greater sense of calm, they scored 12% and 14%, respectively. Only 4% of the data show that the aggressiveness factor was increased by those practicing this type of activity. The report that the electronic game does not affect the personal and/or professional aspects was present in 20% of the answers, as shown in Fig. 6.

The immersion in games, generating a greater human-system interaction is the factor that most calls attention in the evaluation performed, being present for 91% of the players. When the immersion comes accompanied by the evolution of the stories, they motivate 43% of the players; when it comes along with the evolution of gameplay, another 48% are interested, as Fig. 7 shows.

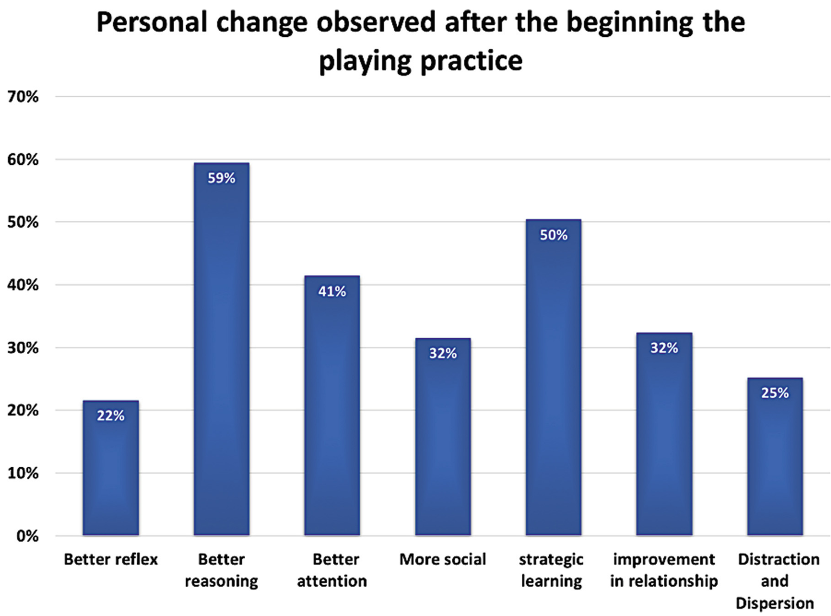



**Fig. 5.** Personal changes observed by players. Author's chart, 2018.

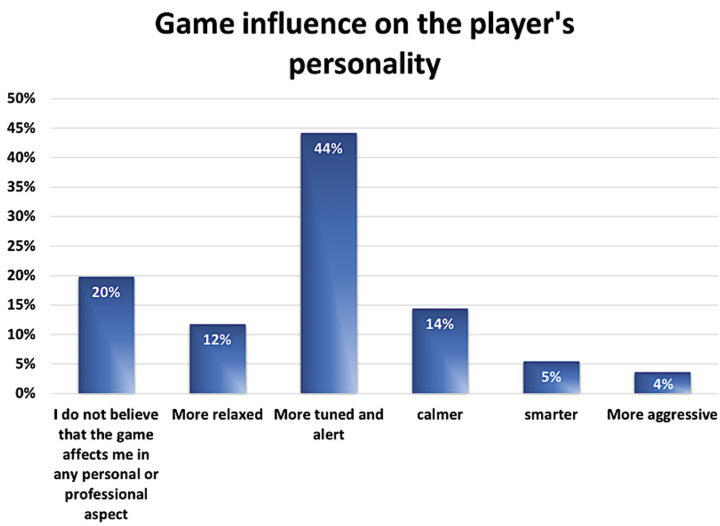



**Fig. 6.** Game influence on the player's personality. Author's chart, 2018.

Interactivity with other players is another important factor as motivation for 35% of respondents. New technologies that allow entering other scenarios or experiencing new realities, either through virtual reality or holograms, still do not motivate the game to happen, perhaps because they are very incipient so far.

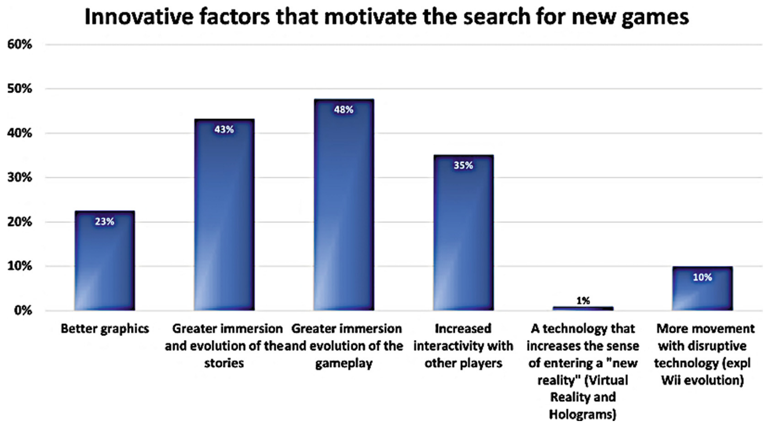



**Fig. 7.** Innovative factors that motivate the search for new games. Author's chart, 2018.

Surprisingly, improving graphics is a predominant factor for only 35% of players, as shown in Fig. 7. In other words, better graphics and shocking design coupled with evolution of weak stories and insufficient gameplay do not motivate players.

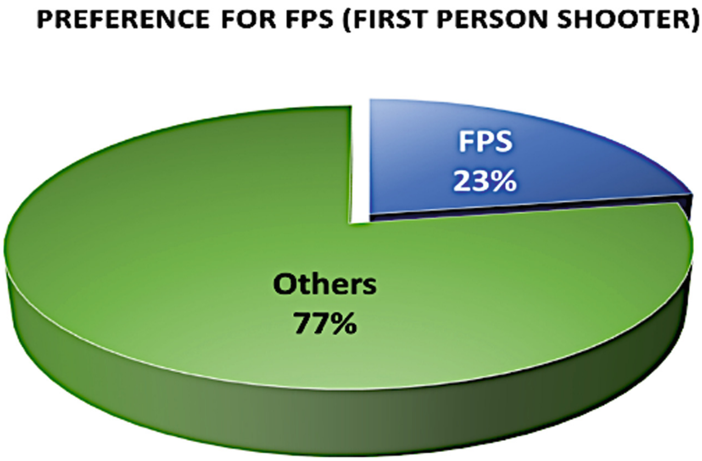



**Fig. 8.** Preference for FPS style (First Person Shooter) games compared to others. Author's chart, 2018.

The FPS game style, Fig. 8, is very common between titles and players, being a style of play and a mode of interaction and visualization of scenes and characters as first-person shooter. Research in Brazil, has shown that among 18–25-year old, only about 1/4 of them choose this type as their preferred game mode.

Another essential aspect of the research approach, presented in Fig. 9, is the preference in terms of style and design, when comparing graphs and scenarios.

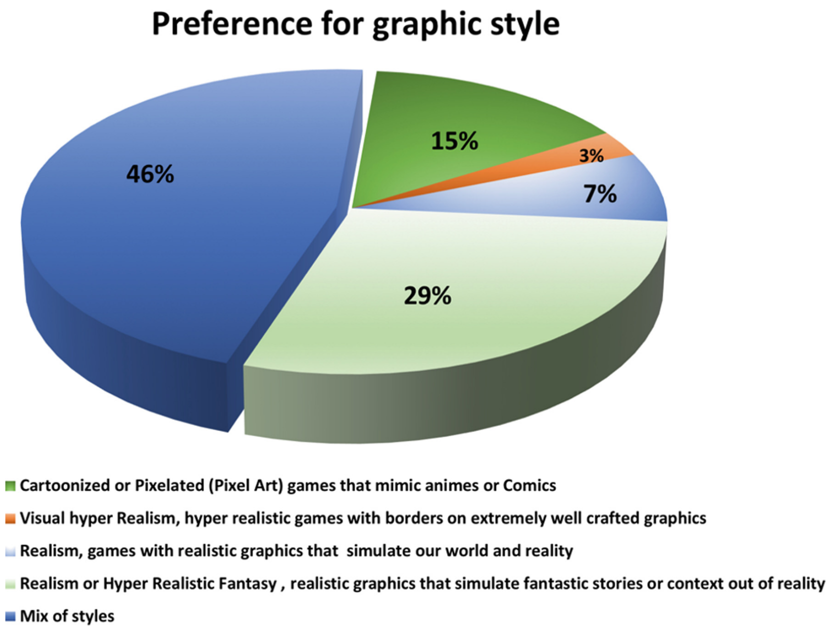



**Fig. 9.** Graphic Styles Preferences. Author's chart, 2018.

Between five different options and categories, the 1st are the 'cartoonized' or pixelated games (pixel-art): where design and the graphics, characters and scenarios use caricature, playful or pixelated drawing techniques. This item includes titles that refer to children's or comic characters. This style represented 15% of the overall choice. The 2nd style was graphic visual hyper-realism with stunning scenes and characters that surpass reality with graphically hyper-realistic images, shocking and provoking a certain impact, because they are confused with cinematographic effects of films, with about 3% preference. The 3rd was the graphic visual realism focusing on scenarios and characters simulate real life and approach reality, but without excesses, that reaches 7% of preferences. The fourth was visual realism or fantastic visual hyperrealism and fantasy, focused on fantastical and unrealistic scenarios and characters, approaching fables and mystical and powerful or supernatural characters, being futuristic or linked to the past and the medieval times, with aspects far removed from reality. This style reached almost 30% of the preferences. This style was almost 30% of the preferences.

It was surpassed only by the fifth alternative with 46% of the total choices, which suggested the choice of a mixture of all styles, that is, suggests that the player likes to vary several types of styles between games, and not stick to one them only.

## 3 Final Considerations

Changes in the behavior of people between the ages of 18 and 25, practitioners of electronic games in Brazil were identified through a survey applied to those who have access to the web, predominantly within the State of São Paulo. Even most of the time the changes that can be considered beneficial to the individuals, such as the claim of greater concentration and reasoning ability and more refined reflexes, identifies a worrying part of the players, seeking through them the escape from reality and the isolation in proportion that should not be disregarded. On the other hand, for a large number of practitioners, the habit of playing has improved the way they relate to other people, making them more social and capable of overcoming challenges.

In the analyzed age group, the vast majority use the personal computer or a platform to play, preferring games that have integrated technology, sound and image, that allow them immersion, evolution of the stories and gameplay. This search for greater immersion is what leads the vast majority to look forward to meeting other games and those developers who provide a new set of games/platform, using technologies that offer greater interactivity between man-system and between systems and multiplayer in interaction, and bring them closer to experiencing characters and scenarios closer to reality, adding new technologies to the wishes of their audience, will make this innovation disruptive, modifying the scenario. Even disruptive proposals like Wii or switch and complements like VR or hologram wasn't chosen as preferences and will run out if there are no stories and evolutions of gameplay that sustains them for a long time. It is also important to point that players are more interested in realism or hyper-realism in a fantastical or out of this world setting than in a hyper-realistic simulation by itself, meaning that simulating a completely realistic environment do not captivate the players attention as much. Further studies and research could define exactly the details that justify and characterize the meaning of this profile. But it is conclusive that, given these results obtained in the sampling of about 100 respondent gamers between the ages of 18 and 25, the disruptive aspects are less important than the sense of control and power, for interactions that assure the sensation of influence and control over fear and it seems that they haven't exhaust themselves in the interactive possibilities presented, coupled with fancy touches guaranteeing entertainment.

**Acknowledgments.** The authors thank the Foundation for the Support of the Research of the State of São Paulo – FAPESP, through Process Nº 11169-4 FAPESP, for the support provided to carry out this research.

## References

1. Halter, E.: Sun Tzu ao Xbox: War and Video Games. Thunder's Mouth Press, New York (2006)
2. Amoroso, D.: Tecmundo - Cultura Geek: A História dos Videogames do Osciloscópio aos gráficos [Vídeo] (2009). https://www.tecmundo.com.br/xbox-360/3236-a-historia-dos-video-games-do-osciloscopio-aos-graficos-3d.htm
3. Jordão, F.: Tecmundo - Cultura Geek, A evolução dos gráficos dos jogos [Vídeo] (2011). https://www.tecmundo.com.br/video-game-e-jogos/14987-a-evolucao-dos-graficos-de-jogos-video-.htm
4. Wolf, M.: Encyclopedia of Video Games: The Culture, Technology, and Art of Gaming, vol. 2. Ed. by Mark, J.P. Wolf. Santa Barbara, CA, Greenwood (2012)
5. Souza, C.A.P.: Imersão e presença nos jogos FPS: uma aproximação qualitativa. Tese de Doutorado. Pontifícia Universidade Católica de São Paulo – PUC/SP (2012)

6. Lenoir, T., Lowood, H.: Stanford.edu. History of computer game design (2015)
7. Kennedy, J.: The Evolution of Video Game Graphics from the 1950s to today [Video] (2016). https://www.siliconrepublic.com/play/video-game-graphics-evolution
8. Slater, M., Khana, P., Mortensen, J., Yu, I.: Visual realism enhances realistic response in an immersive environment. IEEE Trans. Visual Comput. Graph. **20**(4), 606–615 (2009)
9. Rodrigues, A.: As complexas narrativas da nova geração de jogos eletrônicos. Caderno Ilustríssima - Folha de São Paulo (2015). http://www1.folha.uol.com.br/ilustríssima/2015/02/1586204-as-complexas-narrativas-da-nova-geracao-de-jogos-eletronicos.shtml
10. Barcellos, E.E.I., Botura Jr., G.: The interactive holography as metaphor and innovation in optical representation in design. Procedia Manuf. **3**, 754–761 (2015)
11. Krcmar, M.: PHYS.org. Games and Realism, Wake Forest News. Wake Forest University, 22 December 2010 (2010). http://news.wfu.edu/2010/12/22/video-games-and-realism/
12. Jaguaribe, B.: O choque do real: estética, mídia e cultura. Rocco, São Paulo (2007)

Check for updates

# Semantic Congruency Between Music and Video in Game Contents

Natsuhiro Marumo(✉), Yuji Tsutsui, and Masashi Yamada

Graduate School of Engineering, Kanazawa Institute of Technology, Ishikawa, Japan
b1339194@planet.kanazawa-it.ac.jp, m-yamada@neputune.kanazawa-it.ac.jp

**Abstract.** Emotional features have been illustrated by a two-dimensional model, which was spanned by valence and arousal axes, in the simplest way. In the present study, the correlations between semantic congruency and the emotional coincidences on the valence and arousal factors between music and videos were clarified, in the context of game contents. Participants rated the degree of congruency between music and videos. The semantic congruency was very high when the emotions of the music and video were coincided in both factors. In the cases where emotional feature of a musical piece did not coincide with a video in the valence or arousal factor, the congruency significantly decreased. When the emotions coincided neither factor, the congruency showed the lowest values. The results implied that both the valence and arousal factors in the emotional features were equally important for the semantic congruency between musical pieces and videos.

**Keywords:** Emotion · Semantic congruency · Audio-visual
Semantic differential method

## 1 Introduction

Music accompanies video in various audio-visual contents, e.g., films, TV programs, and video games. Music emphasizes the emotional features of a video, when the music is congruent with the video. There exist two types of congruency between music and video, i.e., structural congruency and semantic congruency. Semantic congruency refers to the congruency between emotional features of music and video. In most of the studies on semantic congruency, musical and video materials showed opposite emotions, e.g., happiness and sadness. Combining the musical and video materials, audio-visual stimuli were constructed to reveal the correlations between semantic congruency and the emotional coincidence between music and videos. For example, Boliver et al. prepared friendly and aggressive materials of musical excerpts and videos to construct audio-visual stimuli [1]. They showed that the musical and visual materials were perceived as congruent, if the emotional features in the musical materials coincided with the visual materials. Russell illustrated emotions by a two-dimensional plane spanned by valence and arousal [2]. However, the correlations between semantic


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 366–372, 2019.
https://doi.org/10.1007/978-3-319-94619-1_36

congruency and the emotional coincidences have not been clarified on the two-dimensional factors yet.

In the present study, the correlations between semantic congruency and the emotional coincidences on the valence and arousal factors between music and videos were clarified, in the context of game contents.

## 2 Experiment

### 2.1 Methods

Tsukamoto et al. prepared 100 pieces of game music [3]. Then, listeners rated the emotional features using the semantic differential method. The results of factor analysis showed that the emotional features of the game music were illustrated by a two-dimensional space, which was spanned by valence (pleasant-unpleasant) and arousal (arousing-calm) factors. In the present study, one musical piece was chosen from each quadrant (A1–A4). Then, a video of game-play, which was accompanied by each musical piece, was recorded for 60–70 s (V1–V4). The video materials included sound effects but no music. The four videos were combined with four musical pieces, then 16 audio-visual stimuli were constructed (A1V1, A2V1,…, A4V4). Moreover, four more musical pieces which showed similar emotional features to A1-A4 were chosen and labeled A1’–A4’ (Fig. 1). Then, four more audio-visual stimuli A1’V1, A2’V2, A3’V3 and A4’V4 were constructed. The four video, four musical and 20 audio-visual stimuli are listed in Table 1.

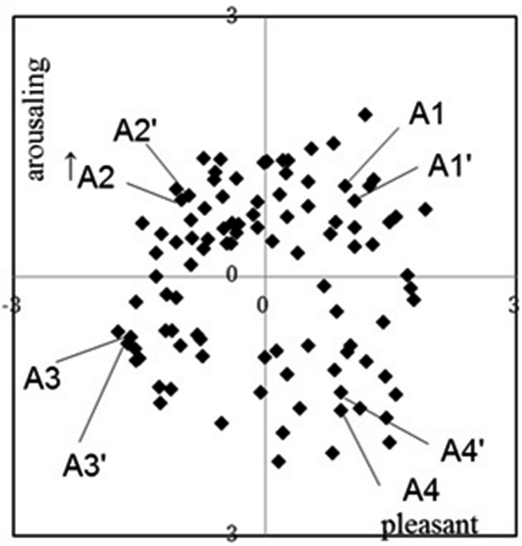



**Fig. 1.** Emotional futures of the game music

**Table 1** List of the stimuli

| Game Title | | Abbreviation | | |
|---|---|---|---|---|
| Video | Music | Video | Music | Stimulus |
| Samurai Warriors 2 | - | V1 | - | V1 |
| Biohazard0 (train) | - | V2 | - | V2 |
| Biohazard0 (Chapel) | - | V3 | - | V3 |
| Monster Hunter 3 | - | V4 | - | V4 |
| - | Honnoji | - | A1 | A1 |
| - | ED-ZOMBIE | - | A2 | A2 |
| - | ChapleⅢ | - | A3 | A3 |
| - | Village on the Sea, Moga | - | A4 | A4 |
| - | For Achieve | - | A1' | A1' |
| - | Ganado Ⅱ | - | A2' | A2' |
| - | The Basement Of Police Station | - | A3' | A3' |
| - | Old Man Orange's Appearance | - | A4' | A4' |
| Samurai Warriors 2 | Honnoji | V1 | A1 | V1A1 |
| Samurai Warriors 2 | ED-ZOMBIE | V1 | A2 | V1A2 |
| Samurai Warriors 2 | ChapleⅢ | V1 | A3 | V1A3 |
| Samurai Warriors 2 | Village on the Sea, Moga | V1 | A4 | V1A4 |
| Samurai Warriors 2 | For Achieve | V1 | A1' | V1A1' |
| Biohazard0 (train) | Honnoji | V2 | A1 | V2A1 |
| Biohazard0 (train) | ED-ZOMBIE | V2 | A2 | V2A2 |
| Biohazard0 (train) | ChapleⅢ | V2 | A3 | V2A3 |
| Biohazard0 (train) | Village on the Sea, Moga | V2 | A4 | V2A4 |
| Biohazard0 (train) | Ganado Ⅱ | V2 | A2' | V2A2' |
| Biohazard0 (Chapel) | Honnoji | V3 | A1 | V3A1 |
| Biohazard0 (Chapel) | ED-ZOMBIE | V3 | A2 | V3A2 |
| Biohazard0 (Chapel) | ChapleⅢ | V3 | A3 | V3A3 |
| Biohazard0 (Chapel) | Village on the Sea, Moga | V3 | A4 | V3A4 |
| Biohazard0 (Chapel) | The Basement Of Police Station | V3 | A3' | V3A3' |
| Monster Hunter 3 | Honnoji | V4 | A1 | V4A1 |
| Monster Hunter 3 | ED-ZOMBIE | V4 | A2 | V4A2 |
| Monster Hunter 3 | ChapleⅢ | V4 | A3 | V4A3 |
| Monster Hunter 3 | Village on the Sea, Moga | V4 | A4 | V4A4 |
| Monster Hunter 3 | Old Man Orange's Appearance | V4 | A4' | V4A4' |

The experiment consisted of three sections. In the first section, the emotional features of video stimuli were rated. In the second section, the emotional features of musical stimuli were rated. In the third section, the emotional features of the audio-visual stimuli were rated. For the audio-visual stimuli, the congruency between the video and music was also rated. Ten students from the Kanazawa Institute of Technology participated in the experiment. The participants were requested to rate their emotional features of the audio-visual stimuli using 23 semantic differential (SD) scales listed in Table 2. The scales were bipolar seven-step scales, e.g., "very weak", "fairly weak", "slightly weak",…, "very strong". The scales appeared on a response sheet in a random order for each combination of participants and stimuli. The participants rated the congruency between the video and the music, using a unipolar seven-step scale, “completely incongruent”, “quite incongruent”,…, “perfectly congruent”.

Each stimulus was presented through headphones, STAX SR-407 in the level of LAeq = 74–84 dB and a computer display, EIZO FlexScan SX2462W-PX. The participants sat on a chair in a dark sound-proof room and were requested to view and listen to the stimuli.

**Table 2** Semantic differential scales and their factor loadings

| SD scales | | | Factor | |
|---|---|---|---|---|
| | | | Valence | Arousal |
| Cheerful | - | Gloomy | -0.970 | -0.060 |
| Hard | - | Soft | 0.847 | -0.470 |
| Delightfil | - | Dull | -0.854 | -0.410 |
| Tense | - | Relaxed | 0.828 | -0.523 |
| Warm | - | Cold | -0.986 | 0.095 |
| Heavy | - | Light | 0.962 | -0.008 |
| Clean | - | Dirty | -0.953 | 0.192 |
| Brilliant | - | Fuzzy | -0.967 | -0.051 |
| Dark | - | Bright | 0.980 | 0.089 |
| Cute | - | Uncute | -0.942 | 0.259 |
| Pleasant | - | Unpleasant | -0.979 | 0.044 |
| Sordid | - | Fresh | -0.953 | 0.242 |
| Agitated | - | Calm | 0.100 | -0.964 |
| Majestic | - | Feeble | -0.119 | -0.740 |
| Powerful | - | Powerless | 0.221 | -0.949 |
| Tranquil | - | Restless | -0.325 | 0.914 |
| Loose | - | Tight | -0.506 | 0.804 |
| Excited | | Unexcited | -0.031 | -0.988 |
| Speedy | - | Slowly | -0.026 | -0.948 |
| Impressive | - | Unimpressive | 0.092 | -0.825 |
| Weak | - | Strong | -0.190 | 0.977 |
| Vivid | - | Simple | -0.287 | -0.920 |
| Neat | - | Mixed | 0.620 | -0.609 |
| Contrivution Rate | | | 0.499 | 0.410 |

## 2.2 Results and Discussion

Numbers −3 to 3 were given for each of the seven categories on the SD scales. The rated scores were averaged over the participants and the obtained mean values were used for factor analysis with the principal factor method and varimax rotation. Table 2 shows the resulting factor loadings. The results of the analysis showed that the two-factor solution accounted for 91% of the data variance. The two factors are labeled "valence" and "arousal" respectively, after the scales which show large absolute values in the loadings for these factors. In Fig. 2, each stimulus is plotted on the two-dimensional space spanned by valence and arousal.

Figure 2 shows that each video (V1–V4) is plotted in the same quadrant as each original musical stimulus (A1–A4), respectively. It is also shown that each similar music A1'–A4' is plotted in the same quadrant as each original music A1–A4, respectively.

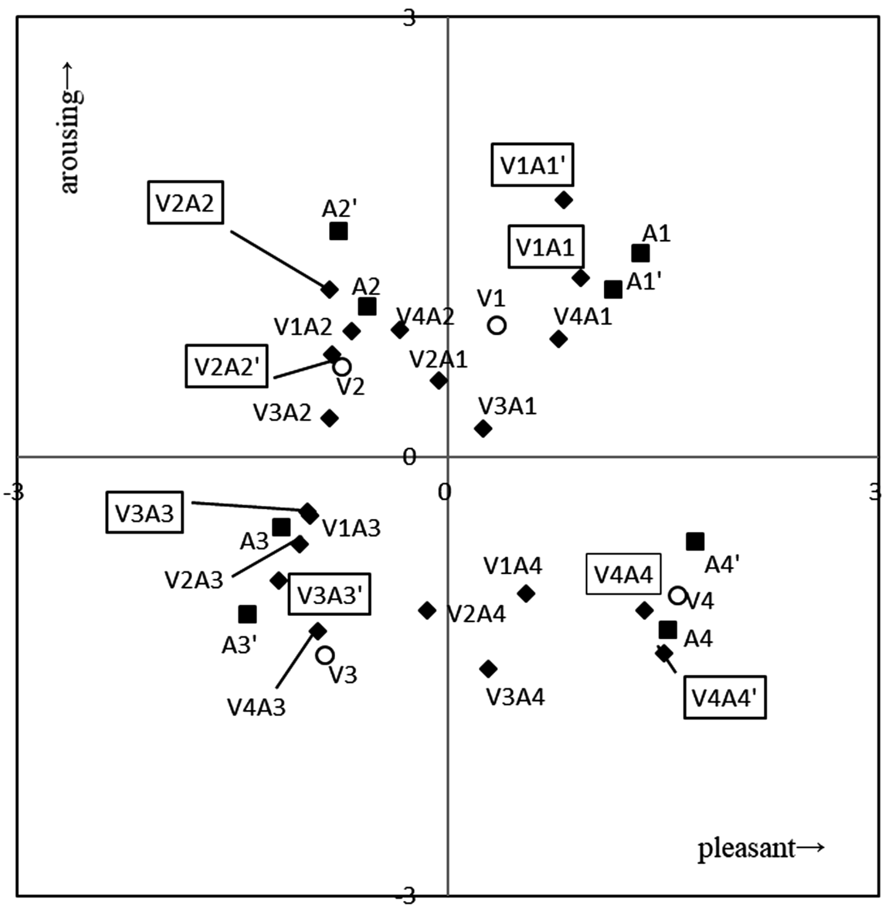



**Fig. 2.** Emotional features of the stimuli

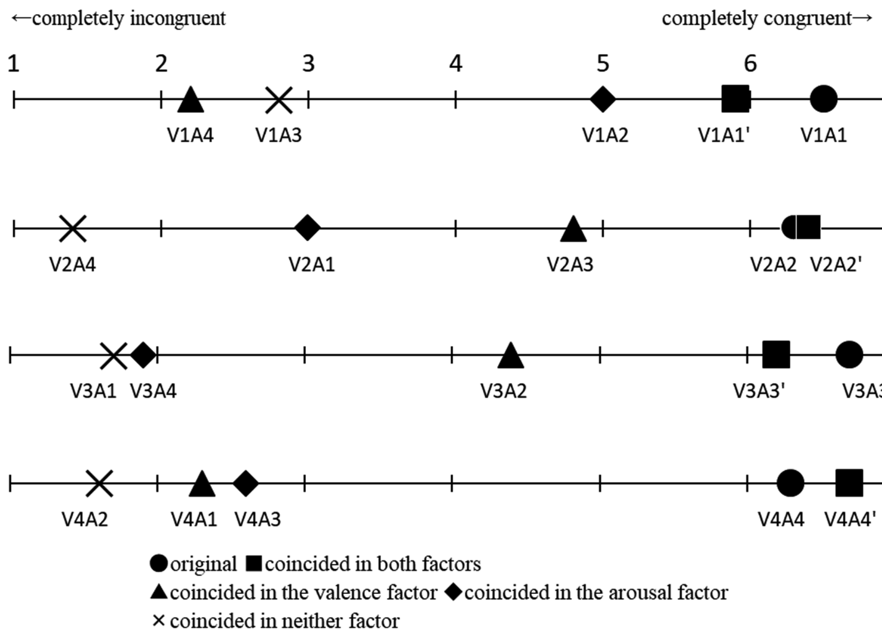



**Fig. 3.** Averaged values of the congruency

Figure 2 shows that the distances between the audio-visual stimuli and video stimuli are larger than the distances between the audio-visual and musical stimuli, on both the valence and arousal axes. This implies that the emotional features of the audio-visual stimuli are strongly affected by music, but not by video. The results in the present study are consistent with the results of Cohen and Iwamiya [4, 5]. Figure 2 also shows that the positions of the audio-visual stimuli, where the original musical pieces are replaced by the pieces from the same quadrants, are placed near the original audio-visual stimuli. This implies that the emotional features of audio-visual stimuli do not change significantly even if the original musical piece is replaced by a musical piece which shows similar emotional features.

Figure 3 shows the degrees of the congruency for each audio-visual stimulus. The congruency for the audio-visual stimuli where original musical pieces are combined with the videos indicates very high values as well as the stimuli where the similar musical pieces replaced the original pieces. This suggests that the videos are perceived as congruent with musical pieces which possess similar emotional features both in the valence and arousal factors. The degree of the congruency largely decreases in the cases where the emotional features coincide in one factor, but not in the other factor. The stimuli are perceived as incongruent when the video and music coincide neither in the valence factor nor in arousal factor. The results of ANOVA and multiple-comparison tests showed that the stimuli where the emotional features of the video and music coincided both in the two factors were significantly more congruent than the stimuli where they did not coincide either in the factor ($p < .05$). There was no significant difference between the cases where the video and music coincided only in the valence factor and the cases where they coincided only in the arousal factor. These results imply both the valence and arousal factors are equally important for the congruency between videos and music.

## 3 Conclusion

In the present study, the videos of game-play and game music were used to construct audio-visual stimuli. The two-dimensional plane, spanned by valence and arousal, illustrated the emotional features of the music, videos and the audio-visual stimuli. The results also showed that the emotional features of the audio-visual stimuli were strongly affected by music but not by videos, both in the valence and arousal factors. Moreover, the results showed that both the coincidence in valence and arousal factors are important for the congruency between the video and music, at least in the context of video games.

## References

1. Bolivar, V.J., et al.: Semantic and formal congruency in music and motion pictures. Psychomusicology **13**, 28–59 (1994)
2. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**, 1161–1178 (1980)
3. Tsukamoto, M., Yamada, M., Yoneda, R.: A dimensional study on the emotion of musical pieces composed for video games. In: Proceedings 20th International Congress on Acoustics, Sydney, Australia (2010)
4. Cohen, A.J.: Associationism and musical soundtrack phenomena. Contemp. Music Rev. **9**, 163–178 (1993)
5. Iwamiya, S.: Multimodal Communication Revised Version of Music and Video. Kyushu University Press, Fukuoka (2011)

# Proposal for Video Game Using the Concept of Becoming the Dotted Main Character

Namgyu Kang[1(✉)], Sadayoshi Takaki[2], Nobuyasu Kaito[2], and Aki Yamauchi[2]

[1] Future University Hakodate, 116-2 Kamedanakano, Hakodate, Hokkaido, Japan
kang@fun.ac.jp
[2] Graduate School, Future University Hakodate, 116-2 Kamedanakano, Hakodate, Hokkaido, Japan

**Abstract.** In recent years, with development of VR technology, many video games have attracted attention that offer gamers the experience of being in the world of the game. In this study, we focused on a method other than VR technology to offer gamers the experience of directly entering the game world. To this end, we created "DOT.GATE," in which gamers can feel a different type of enjoyment, and with the concept of making real people appear in the virtual game world. "DOT.GATE" was designed to enable real gamers to become a character in the virtual game world, so they can feel enjoyment unlike that of previous games. Intended for everyone to enjoy without needing special skills, "DOT.GATE" was designed as a retro game like those of the NES (Nintendo Entertainment System) console. When a user stands in front of the "DOT. GATE" display, the user's whole body is dotted in using a computer program and shown in the "DOT.GATE" game world. An evaluation experiment was conducted with some students at Future University Hakodate to clarify users' impressions of the concept of "DOT.GATE." From the experiment, we found that many gamers preferred the concept and felt it was very fun and offered a different type of pleasure than previous games. In future, it will be necessary to shorten the time needed for a gamer to be dotted in by the program, and to add more rich game content.

**Keywords:** Game design · Becoming a game character · Kansei evaluation

## 1 Introduction

In recent years, the number of products and services available for enjoying games and having fun has been increasing. In particular, there is a growing demand to immerse users in a virtual world using VR technology. One example is PlayStation VR by Sony, which is different from conventional games that are played via images output on displays such as TVs or PC monitors. Rather, gamers can feel they are playing in 3D space, surrounded by the game in 360° due to the VR system [1]. Moreover, gamers can feel they are truly entering the game world with 3D audio technology. In this study, we reversed the context and focused on a method other than VR to provide gamers with


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 373–377, 2019.
https://doi.org/10.1007/978-3-319-94619-1_37

the experience of directly entering the game world. A good example of a game that allows users to have fun in a way unlike most existing games is "Photoshop Live." While waiting at bus station, users can take a photo and use Photoshop to edit it, displaying their art work on a monitor at the bus station as a live show. Users who are waiting for a bus at the bus station feel fun in being immersed in the virtual world of artwork.

Given this background, the purpose of this research is to clarify whether gamers can feel a different type of enjoyment through the concept of making real people appear in the virtual game world. Moreover, we clarify whether appearing as a main character in the game gives gamers the experience of "really being in the game world (Fig. 1)."


**Fig. 1.** Photoshop live street retouch prank

## 2 Proposal for "DOT.GATE"

Based on background described above, we proposed the prototype of a new game called "DOT.GATE," in which gamers can feel a different type of enjoyment through the concept of making real people appear in the virtual game world. In other words, the proposal's concept was to make the gamer appear as a main character in famous games. "DOT.GATE" was designed to allow real gamers to become main characters in the virtual game world, giving them a different feeling of fun than existing games. "DOT. GATE" was designed as a retro game like those of the NES (Nintendo Entertainment System) console for everyone to enjoy without needing special skills. For the "DOT. GATE" game, a display, a camera, two PCs, and a background cloth for chroma keying were used Fig. 2. The "DOT.GATE" character is generated based on a photograph taken against the chroma keying background cloth by two administrative members, with the gamer in a specified pose Fig. 3.

When a user stands at chroma keying background cloth to take a photo in front of the "DOT.GATE" display, the user's whole body is dotted in using a computer program and shown in the "DOT.GATE" game world. A new game is then started with the new main character. Figure 4 illustrates the flow of "DOT.GATE."When the game is over, the result of the game score is displayed on the screen with the dotted game character of each gamer oneself (Fig. 5).

**Fig. 2.** The chroma keying background cloth to take a photo

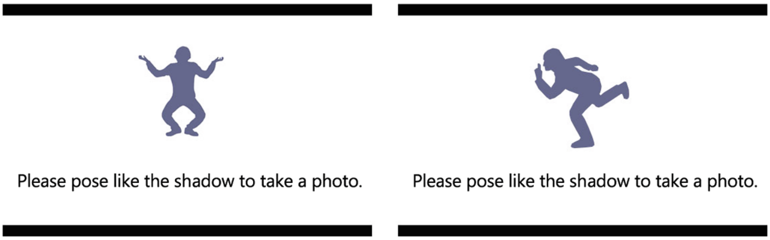



**Fig. 3.** Example poses for taking photos


**Fig. 4.** Example flow of "DOT.GATE"

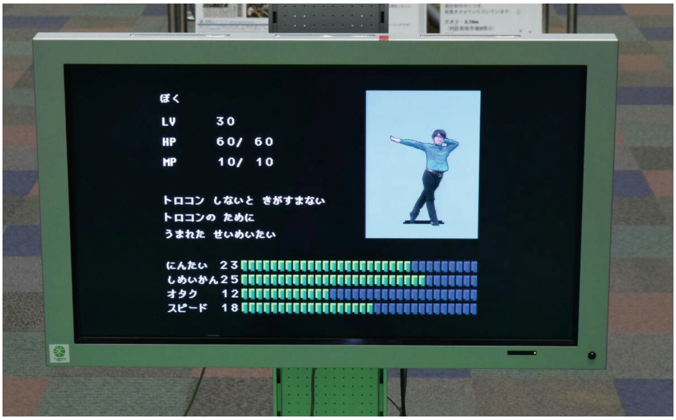



**Fig. 5.** Example of game score with the dotted game character

## 3 Evaluation of "DOT.GATE"

Evaluation was conducted to confirm whether gamers could enjoy a feeling of being immersed in the game. The experiment to evaluate "DOT.GATE" was conducted with 21 participants at Future University Hakodate on January 18, 2018. Each participant posed for a photo against the chroma keying background cloth. The photos were processed into the dotted main character of the existing famous games such as 'Street Fighter'. Figure 4 shows an example of the proceed characters using the taken photos. Each participant enjoyed a very short and simple game in which they themselves became the main character of the game. Subsequently, the participant evaluated

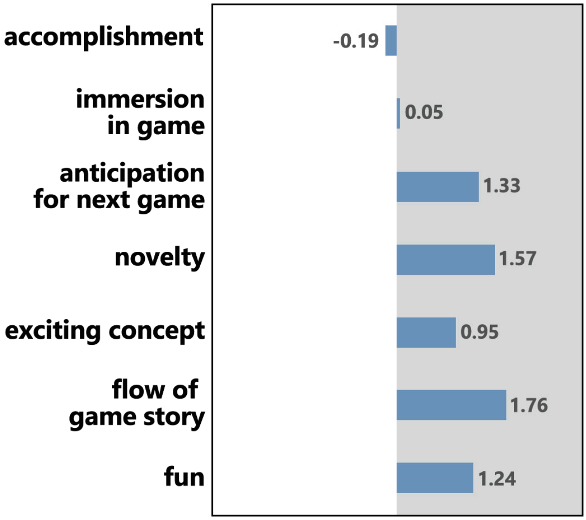



**Fig. 6.** Results of evaluation

"DOT.GATE" using the SD method with 6 levels for 7 evaluation items: (1) feeling of accomplishment, (2) feeling of being immersed in the game, (3) anticipation for the next game, (4) novelty, (5) exciting concept, (6) flow of game story, and (7) fun. These 7 items were selected based on the textbook for experiments in the *Kansei* engineering field [3].

## 4 Conclusion

Results of the experiment (Fig. 6) showed that "DOT.GATE" was evaluated especially highly in "flow of game story" and "novelty." The items "anticipation for the next game," "fun," and "exciting concept" were also highly evaluated. However, "feeling of being immersed in the game" received only weak positive evaluation. Moreover, "feeling of accomplishment" was evaluated negatively.

From the evaluation experiment, "DOT.GATE" using the concept of becoming the dotted main Character of game was clarified as fun and exciting concept. Although it includes three different game stories, each game is completed in 3–5 min. Additionally, each game has only a few sequences directly controlled by the gamer, and many cut scenes in which the gamer just watches the flow of the game. In other words, "DOT. GATE" was evaluated as a passive game. However, the flow of the game was considered very good, and concept novel, fun, and exciting.

The results of evaluation show that "DOT.GATE" is a novel game with a good story flow, and that the concept is effective for inciting users' enjoyment. However, there are some points that must be developed with respect to immersion in the game and gamers' feeling of accomplishment. It will be necessary to clarify the impression made and influence of the different various game characters on enjoyment. In the future, the "DOT.GATE" will be developed with various game contents and game characters to increase the gamers' sense of immersion, so gamers can enjoy more aggressively with positive experiences.

**Acknowledgments.** This research was conducted in the class of 'Introduction to Information Design 2' at graduated school of Future University Hakodate. We thank the students of Future University Hakodate who cooperated for this research with the experiment.

## References

1. PlayStation VR. http://www.jp.playstation.com/psvr/. Accessed 19 Mar 2018
2. Photoshop Live-Street Retouch Prank. https://www.youtube.com/watch?v=BRAM8MpqIeA. Accessed 19 Mar 2018
3. Tadahiro Hukuda (Human Performance Laboratory): Ergonomics Guide, How to Science Kansei. Scientist Press, Japan (2009)

# Designing Educational Games Based on Intangible Cultural Heritage for Rural Children: A Case Study on "Logic Huayao"

Yuanyuan Yang(✉), Duoduo Zhang, Tie Ji, Lerong Li, and Yuwei He

School of Design, Hunan University, Changsha, China
41924207@qq.com

**Abstract.** Intangible Cultural Heritage (ICH) is one of the application fields of educational games. However, most of these applications are provided for the general public of cultural experience or information access, rather than designing for rural children's cultural learning. We propose that rural children are critical learners and inheritors of ICH. ICH education has to be future-oriented and integrated with rural children's learning needs and learning styles. This research is conducted in rural China with local children. We observe the learning behaviors of rural children through an experiment, and classify learners into activists, reflectors, negatives and entertainers. With the design case of "Logic Huayao", we present a design model for educational games with cultural learning. In designing process, the players, culture, games, and learning contents should be considered as an inclusive model and integrated into the final interaction design process. Finding the relationship between them will enable us to develop educational games that include cultural meaning and problem-solving skills.

**Keywords:** Intangible Cultural Heritage · Educational games · Learning style Cultural learning · Rural children

## 1 Introduction

The idea of using digital games for supporting and enhancing Intangible Cultural Heritage (ICH) education is not new. An overview of recent serious games in cultural heritage can be classified into three types of applications: reconstruction of history, virtual museums, and cultural demonstrations [1]. However, most of these applications are offering the opportunity to remote cultural experience or information access for the general public, rather than designing for locals. Like virtual museums, although they are helpful, still lack a powerful mechanism to engage the large public into an active state of learning where spectators are motivated to create their own knowledge rather than to receive information passively [2]. ICH is not only a fascinating subject need to safeguard, but also a vivid and real experience inherited by locals. According to UNESCO, "Intangible cultural heritage, transmitted from generation to generation, is constantly recreated by communities and groups in response to their environment, … … promoting respect for cultural diversity and human creativity [3]. ICH education for local children can take the advantage of learning and practicing in cultural space,


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 378–389, 2019.
https://doi.org/10.1007/978-3-319-94619-1_38

promoting children's cultural identity and creativity. And above all, it empowers locals to sustain their culture's regeneration ability.

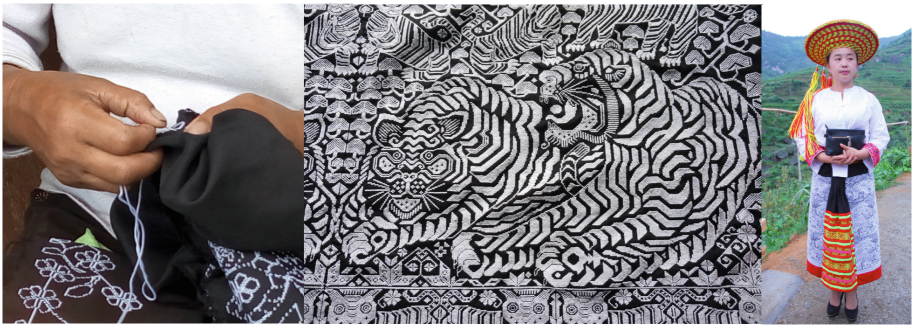
In recent years, some traditional handicrafts and music in China have begun to enter the rural primary schools in the form of "Bringing the Intangible Cultural Heritage into School", hoping to establish the connection between children's education and the unique regional culture. However, for many reasons, it is often difficult to achieve the desired results in practice. As we are learning, living and working in a changing world, facing the challenge: how to help local communities and groups to "constantly recreate" ICH culture? Pöllänen and Urdziņa-Deruma put forward with "Future-Oriented Reform of Craft Education", which should focus on collaborative learning, active participation and the use of different tools and technologies to create new knowledge [4]. Future-oriented learning expands our understanding of ICH's meaning and value. Therefore, this paper probes into the possibilities of new type educational game with future-oriented cultural learning. We propose that educational games designed on the basis of ICH education are culturally meaningful to rural children and conforming to children's cognitive schema construction.

The study has been started with Huayao cross-stitch (Fig. 1), a national Intangible Cultural Heritage in a remote village of Longhui County, Hunan Province in China. Huayao, literally Flowery Yao, is a branch of the Yao minority. The reason they got the name is because local girl's long straight skirts are usually decorated with intricately detailed cross-stitch patterns. Using white or colored thread and dark color homespun cloth as materials, Huayao girls are skillful on creating patterns. Because Yao has no written characters, they pass down their ethnic history and daily life by patterns with cross-stitch. There are more than 1000 patterns, which can be divided into four categories: history, tradition, animals and plants. Huayao woman needs to spend about half a year to finish a cross-stitch skirt, and embroiders two or three hundred thousand needles on average. One thing is very impressive: during the long process of embroidery, following the warp and weft grid on homespun cloth, Huayao women can cross-stitch beautifully and complicated patterns without models and drafts but only by memory. It is the creativity and logical thinking that conveyed by the "Thinking Hand". Can we extract the elements of Huayao cross-stitch and integrate them into children's mathematical logic training and creativity, and finally transform it into educational games through appropriate game mechanics?

Thus, the research group, have designed an educational game called Logic Huayao. We look forward to helping rural children to develop the sense of mathematics and logical thinking. At the same time, guiding them get familiar with patterns and structures of Huayao cross-stitch under the local cultural contexts. With the design case of "Logic Huayao", we propose a design model for educational games with cultural learning.

## 2 Related Work and Theories

At present, researches about the application of ICH in children's educational games are still very less. However, with the development of constructivist theories of learning in recent years, socio-cultural factors become increasingly important in educational game



(a) (b) (c)

**Fig. 1.** (a) Huayao cross-stitch; (b) Details of the "Tigers downhill" cross-stitch skirt; (c) The picture of Huayao ethnic dress skirts.

design. Nielsen proposes the third generation of educational games, which implies an inclusive triangle of player, culture, and computer games to understand the full spectrum of educational use, and it stresses the key role of providing a social context that facilitates asking the right questions and going the right places [5]. Early experiments by Kam et al. [6] in rural India have shown that compares with foreign western games, English games designed according to the local traditional game mechanisms avoid cultural problems and are more accessible to rural children. Some researchers try to get inspiration from the cultural elements to promote a better learning experience. Tian et al. [7] present the design and implementation of two culturally inspired mobile group learning games—Multimedia Word and Drumming Strokes to help children learn Chinese characters. They found that in group collaborative games, children learn better through observation and reflection. Inspired by Chinese intangible cultural heritage Shadow Play, Lu et al. [8] design the ShadowStory game for promoting creativity, collaboration, and intimacy with traditional culture among children. These present studies report the educational game design inspired from the traditional culture, and discuss the cultural influences in learning experience, but has not mentioned how to promote cultural learning and recreation, as well as how to design educational game based on rural children's learning situation and learning styles.

To develop a successful educational game, learner's differences, social cultures, game genres, learning activities and content, and learning styles are all important issues. According to Killi's game experiencing flow model [9], the task itself, the use of artifacts or individual differences should be taken into account while designing educational games. It is very important to provide suitable challenges for learners to improve learning experience. Honey and Mumford [10] learning style is one of the well known experiential learning, they classify learners into activists, reflectors, theorists, and pragmatists. Prensky [11] proposes the relationship of learning content, learning activities and possible game types, and analyzes how to design corresponding learning activities and game genres according to different learning contents. Chong et al. [12] believe that each person has his or her own way of converting, processing, storing, and retrieving information. So, it is very important to understand individual learning styles to predict learners' responses and feelings in diverse learning environments. After

Prensky and Chong et al., Rapeepisarn et al. [13] further analyze the relationship of learning style, behavior when playing game, and showed the comparison and matching of learning techniques, learning activities, learning styles to possible game genres. The depth and scope of research in learning styles are developing rapidly, however, as a very big learner group, children in developing areas are still less to be reported.

Inspired by the third-generation educational game design theory by Nielsen and the game experience design model by Killi, we propose a design model for educational games with cultural learning (Fig. 2). While designing educational games, the four elements of the players, culture, game and learning contents should be considered inclusively under real socio-cultural situation and learning situation, and integrated them into the final interactive design flow. Rural children are very special learner owing to diverse culture and environment in which they live. Next, through the "Logic Huayao" design practice, this paper shows how to extract the cultural elements from Huayao cross-stitch for children's education, and analyzes the playing and learning styles of rural children, establish relationships between them with proper game genre.

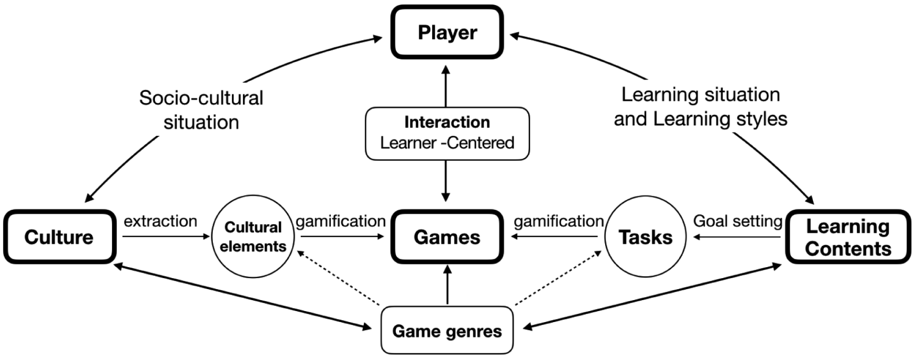



**Fig. 2.** Design model for educational games with cultural learning.

## 3 Methods

According to Nielsen, children' s interaction with technology is different depending on their age and cognitive ability [14]. We chose the children aged 9–12 in Baishuidong Primary School at Huayao as the research subjects. The whole research consists of three stages: pilot study, collaborative innovation and design. (1) Pilot Study Stage: through questionnaires and interviews, we investigated the digital environment, game preferences, current situation and difficulties of mathematics learning of rural children, and further pursued understanding of the pattern memorization and creative schema of Huayao cross-stitch; (2) Collaborative Innovation Stage: we invited the Intangible Cultural Inheritor to teach a Huayao cross-stitch lesson, and then carry out collaborative innovation experiments with multiple participants to test the target game and give suggestions; (3) Design phase: we designed the game Logic Huayao based on

Huayao cross-stitch, confirmed the level of the challenge through the analysis of children's game learning behavior, instilled the storyline into the game and so on.

The pilot study's questionnaire survey was undertaken in a campus environment with the help of teachers. Thus, the researchers became acquainted with these children very quickly and were able to observe their behaviors closely in the next stage. On collaborative innovation stage, children were divided into several groups, each group include 5 to 6 children and an observer sitting together with them. For better observation, we used Multiple-scan with tabular records and video graphics approach (Fig. 3). Multiple-scan is a sampling method for social-behavioral research on a group of children [15]. In this research, each child has a corresponding record sheet, and the observer observes each child in his/her group by turn, recording children's behavior and the time of occurrence. The observer records the learning behavior of children from four aspects: knowledge acquisition and comprehension, interaction with classmates and teachers, concentration and motivation, motor skills and attitude. In addition, we used mobile phone's cameras to capture gestures, facial expressions and other non-verbal interactions, while a fixed position camera served as an aid to recording the behavior that may be overlooked.

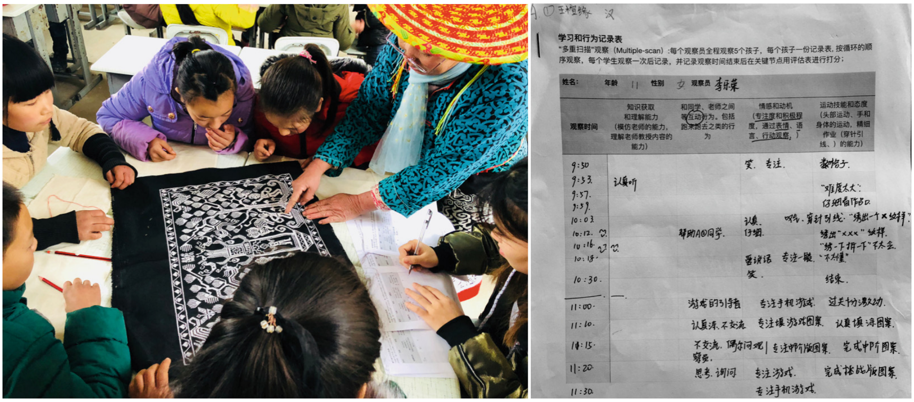



**Fig. 3.** Multiple-scan method in the stage of collaborative innovation

We take advantage of game design patterns, that are usually found in earlier successful games, as basic building blocks for informing new game designs [16]. The game genre is positioned as a logical puzzle game. Nonograms is a classic picture logic puzzle from Japan. In a grid, each row and column have a number corresponding to the location of the black block, the player according to the number of groups to fill or leave the grid, finished to reveal hidden pictures. Compared with Nonograms, counting yarn skill of Huayao cross-stitch has the same rules in some way (Fig. 4). Nonograms requires some elementary knowledge of integer calculations [17], as the difficulty increases, the player also needs certain contradictory reasoning ability. At the same time, Nonograms is a kind of graphical logic exercises, which suitable for children who prefer concrete imaginal thinking.

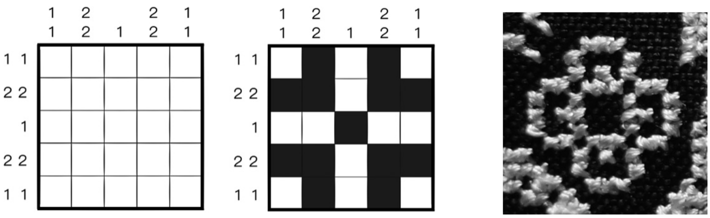



**Fig. 4.** Japanese Nonograms and Huayao Cross-stitch

## 4 Research and Results

### 4.1 Pilot Study

To investigate the current situation and issues in rural children, we conducted informal interviews with Baishuidong primary school principal and five teachers at first, the interview questions focused on the children's learning situation and problems. Also with permission, we checked the latest math test scores and home visit records of students, and conducted a questionnaire survey on 109 pupils (aged 9–12, 61 boys and 48 girls). The team also conducted a scene interview with ten Huayao cross-stitch women to understand the pattern of memorization and creative thinking of the cross-stitch. And find a game mechanism which can combine the Huayao cross-stitch and meet the need of children's learning.

**The Current Situation of Rural Children's Math Learning and Digital Skills**
Under the policy of integrating urban and rural education in China, Baishuidong Primary School is taught and tested with standard teaching materials. According to the latest math test scored out of 100 points, the average score of grade 5 was 35.37, only 8 (19.51%) students passed; and the average score of grade 6 was 23.82, only 6 (17.6%) students passed. Combine the practical teaching experience from math teachers, we summarized three main reasons for rural children's math learning difficulties: Firstly, most rural children cannot keep up with the pace of teaching, and have low interest in the curriculum. Secondly, most children are poor at abstract logical thinking, but prefer concrete imaginal thinking and work better with graphic teaching content. Thirdly, most of them are left-behind children, lack of family supervision and have been unable to develop good study habits. We need to guide children participate actively.

According to Kam's experience, designing such games that target low-income learners in the developing world is considerably more challenging due to their lower levels of familiarity with technology and video games [18]. However, our survey found the informatization of China's rural areas has developed rapidly in the past two years. 84.4% of the 109 rural children surveyed had at least a smartphone or computer in their family, of which 25.7% had a home network. They often play games together after school. And the top three favorite digital games are the King of Glory, Minecraft and Miracle Nikki. The survey showed that there is no obvious difference between urban children and rural children in digital skills. Math teachers also observed that compare to children's class performance, they appeared more capable of playing digital games, which convinced us in using educational games to motivate the learning potential of rural children.

**The Feature of Huayao Cross-Stitch's Learning Process**
To find the cultural elements suitable for enhancing children's abstract logical thinking and creativity from Huayao cross-stitch, we studied the thinking pattern in cross-stitch creation to find the corresponding game mechanism. We interviewed ten Yao women (6 aged 35–45 and 4 aged 50–65) in their home, and asked them questions as the following: *How did you learn the cross-stitch? Do you need a reference pattern before embroidering? How do these patterns come from? How do you remember these patterns? Can you explain how you created this new pattern? etc.* We found that the learning and creating process is continuous and progressive. First learn counting yarn and master the basic stitches, then get familiar with traditional patterns, creating new patterns. And counting yarn is a very crucial skill in pattern memory and creation. Over the longtime embroidery work, Yao women keep continuous reflective observation, they often pause to observe the shape of the pattern through counting yarn. About new pattern creativity, some Yao women said: *"There were many animals in mountains, such as tigers, birds, rabbits and so on, we embroider what we see", "Now children go to school, I like to embroider new pattern according to the pictures in their textbooks".* This shows they have been inspired by their environment or daily life, and constantly recreation.

## 4.2 Logic Huayao Collaborative Innovation Experiment

In order to further observe the learning behavior of rural children, and explore game innovation, we arranged a collaborative innovation experiment in a classroom with multiple participants, which included a math teacher, 27 children volunteers (aged 11–12, 8 boys, 19 girls), 5 observers, 2 organizers, and Ms. Feng (aged 58) who is the intangible cultural inheritor of Huayao Cross-stitch. 27 children were divided into 5 groups by random, and each group have an observer sitting with them.

**Procedure.** This experiment is divided into three steps as Huayao Cross-stitch learning experience (45 min), game experience (45 min) and pattern design (20 min) (Fig. 5). First, Feng gave a cross-stitch lesson for children. She provided some cross-stitch pieces for children to observe the pattern closely, and taught children counting yarn and basic stitches skills step by step. Game experience part was started after a break. To help children understand the game rules quickly, we provided a download nonograms game–Logic Pic [19] for children play on mobile phones and ensure they understand the rules basically. Then every child asked to play a demo game on A4 paper. There are 4 tasks from easy to hard, which are redesigned according to the Huayao cross-stitch patterns. And a blank grid paper is presented to children to draw patterns what they want in the last session. Finally, each child's game outcomes and self-created patterns on paper are photographed, combined with learning behavior records, video recordings, home visit records and other data for the final assessment.

**Results.** This experiment helped researchers to further understand different learning styles of rural children, and we found some inspiring results. Firstly, according to children's learning behaviors, we classified learners into activists, reflectors, negatives


**Fig. 5.** (a) Huayao cross-stitch learning experience; (b) game experience; (c) pattern design

and entertainers (Table 1). It's a little different from Honey and Mumford's classification, because our research focused on rural children and tested in one game genre.

**Table 1.** Rural children's learning style

| Learning style | Learning behavior |
|---|---|
| Activists | Active, acting before thinking, prefer teamwork, seeking help or help others frequently, but sometimes lack independent thinking |
| Reflectors | More cautious, acting after thinking, learning through observation and independent thinking, sometimes get stuck due to self-limitation |
| Negatives | More passive, acting with lack of thinking, prefer trial and error method, easy to complain and give up when face difficulties |
| Entertainers | Active, just enjoy playing part, little care about learning outcomes, prefer play in their own way when face difficulties |

We also found that children's math ability and game learning outcomes are basically proportional (Fig. 6). There are some noteworthy exceptions, for example, an active child who turned from "poor" to "excellent" during the game-based learning, he only got 10 points at math test, but finished 3 tasks and mastered the game strategies.

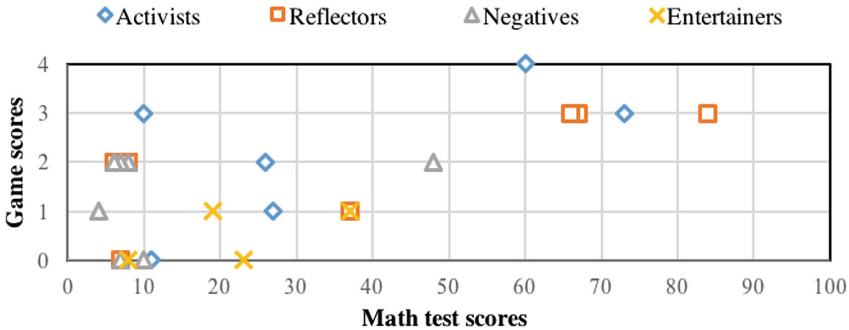



**Fig. 6.** The distribution of Math ability, game performance and learning styles.

Learning from mistakes and learning from peer are very common behaviors in all types of learners. However, trial-and-error behavior can't gain the positive knowledge without reflective thinking. Many children encountered difficulty when they play the demo game on paper because they can't pass through just keep clicking. Therefore, to increase the reflective thinking part to avoid too much trial-and-error behavior is very important when designing educational games for children, as well as to provide more tips and feedbacks to help children learn game strategies.

In the pattern design part, we observed that every child enjoyed drawing freely, just like Huayao cross-stitch, they create what they like, such as their names, love, and transformers etc. (Fig. 7). And they felt relax and began to give various suggestions: *"I think the game should to be more colorful", "I hope to have a background music in there", "I want to use this pattern to decorate my gun", etc.* So, free creation part is a good way for children to stay concentration, increase participation, and bring in new ideas.

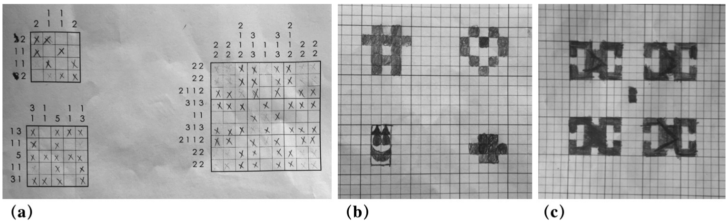


**Fig. 7.** (a) Game demo on paper; (b) "Love" drew by a girl; (c) "Transformers" drew by a boy

As the intangible cultural inheritor, Ms. Feng is a specialist of pattern memorization and creation. She associates the grid on paper with warp and weft on fabric very naturally, and followed the embroidery sequence by habit, she drew different patterns on paper. When she played the demo game, she guessed the meaning of the pattern at first: "Is this the Octagon Flower?", and then she finished the task through memorization, even though she didn't know the game rules exactly. It's a very meaningful discovery which confirmed that this game mechanism can be associated with cross-stitch skills. Thus, we hypothesis that following some visual and meanings clues, game experience and cross-stitch experience could be associate together, mutual combination, mutual promotion.

### 4.3 "Logic Huayao" Game Design

Based on above analysis, we designed the educational game Logic Huayao for rural children (Fig. 8). Learning goal is set based on the analysis of learning contents, and then presented to the learner with the form of game task. Huayao cross-stitch elements are also contained in the learning content as part of cultural learning, which is transformed into the game story, meanings, visual elements and creative parts. We designed a cute Huayao girls who wear folk costumes. She constantly encourages children to keep going, provides strategies and tips more accessible, leads children to avoid too many trial-and-error behaviors lack of reflective thinking. The creative function enables children to create new Huayao cross-stitch patterns even they didn't know how to embroidery in practice. In addition, after finishing pattern creation, children can upload it online to generate new game tasks, share and interact with friends. We think this is a way to stimulate their enthusiasm to learning and creation. Player's learning style, cultural elements, game mechanism and learning contents are fully considered in this game and integrated into the final interactive design flow.


**Fig. 8.** Logic Huayao game design. (a) Game task based on Huayao pattern and logical exercise; (b) game reward with cultural meaning; (c) and (d) cross-stitch pattern creation and upload

## 5 Discussion

Playing and learning under socio-cultural situation, it is a way for guiding learners learn their own culture in naturally, stimulate their enthusiasm and enhance their creative ability. From future-oriented perspective, ICH can become rural children's educational resources as well as a part of socio-cultural contexts in educational games. This study explored a new form of ICH educational game, which associated cultural elements in ICH and learning needs for rural children both in cultural learning and diversified skills exercise. When design educational games, learning style, culture, game genres and learning contents should be considered inclusively and connectedly under real socio-cultural situation and learning situation. Finding the relationship between them will enable us to develop educational games that include cultural meaning and problem-solving skills.

Rural children are very special learners and players. Contrary to the poor learning situation on school curriculum, they have impressive enthusiasm and potential on digital game-based learning. Their learning styles can be classified into activists, reflectors, negatives and entertainers, and each style has different learning issues and need to be considered in design process. In a word, designing more free creation spaces, collaboration opportunities, appropriate challenges, accessible game strategies and reflective thinking time are crucial methods for improving their engagement and learning outcomes.

## 6 Conclusion

The diversity of education and culture could be mutual promotion each other. Educational games based on ICH may leading the ‘constantly recreation” of the culture. Just like Huayao cross-stitch, patterns are constantly recreating by local women. In Logic Huayao, the game experiences and the Huayao cultural experiences are likely to be further transformed through practice and promote learning transfer. We look forward to generating and prompting more vitalities to Huayao cross-stitch through the game by rural children. A further discussion and research is necessary, the design model need to be verified in different culture and conditions, that will be the content of our next study.

**Acknowledgments.** The project is funded by the Department of Development Planning, Ministry of Education P.R.C, under project no. (2017)304; Qipai Intangible Cultural Heritage Protection and Research Fund, Academy of Arts & Design, Tsinghua University, under project no. (2015)05. We would like to thank Zhang Bo, Hu Yuzhe, Li Yongbin and their students in Baishuidong Primary School for their participation on the study.

## References

1. Laamarti, F., Eid, M., El Saddik, A.: An overview of serious games. Int. J. Comput. Games Technol. **2014**, 1–15 (2014)
2. Mortara, M., Catalano, C.E., Bellotti, F., Fiucci, G., Houry-Panchetti, M., Petridis, P.: Learning cultural heritage by serious games. J. Cult. Herit. **15**, 318–325 (2014)
3. Convention for the Safeguarding of the Intangible Cultural Heritage (2003). http://portal.unesco.org
4. Pöllänen, S., Urdziņa-Deruma, M.: Future-oriented reform of craft education. In: Reforming Teaching and Teacher Education, pp. 117–144. Springer (2017)
5. Egenfeldt-Nielsen, S.: Third generation educational use of computer games. J. Educ. Multimed. Hypermedia **16**, 263 (2007)
6. Kam, M., Mathur, A., Kumar, A., Canny, J.: Designing digital games for rural children: a study of traditional village games in India. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 31–40. ACM (2009)
7. Tian, F., Lv, F., Wang, J., Wang, H., Luo, W., Kam, M., Setlur, V., Dai, G., Canny, J.: Let’s play Chinese characters: mobile learning approaches via culturally inspired group games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1603–1612. ACM (2010)
8. Lu, F., Tian, F., Jiang, Y., Cao, X., Luo, W., Li, G., Zhang, X., Dai, G., Wang, H.: ShadowStory: creative and collaborative digital storytelling inspired by cultural heritage. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1919–1928. ACM (2011)
9. Kiili, K.: Digital game-based learning: towards an experiential gaming model. Internet High. Educ. **8**, 13–24 (2005)

10. Honey, P., Mumford, A.: The Manual of Learning Styles. Peter Honey, Maidenhead (1992)
11. Prensky, M.: Computer games and learning: digital game-based learning. In: Raessens, J., Goldstein, J. (eds.) Handbook of Computer Game Studies, pp. 97–122. The MIT Press, Cambridge (2005)
12. Chong, Y., Wong, M., Thomson Fredrik, E.: The impact of learning styles on the effectiveness of digital games in education. In: Proceedings of the Symposium on Information Technology in Education, KDU College, Patailing Java, Malaysia (2005)
13. Rapeepisarn, K., Wong, K.W., Fung, C.C., Khine, M.S.: The relationship between game genres, learning techniques and learning styles in educational computer games. In: Pan, Z., Zhang, X., El Rhalibi, A., Woo, W., Li, Y. (eds.) Technologies for E-Learning and Digital Entertainment, pp. 497–508. Springer, Heidelberg (2008)
14. Nielsen, J.: Children's Websites: Usability Issues in Designing for Kids (2010)
15. Li, Y.: Play and Child Development. Zhejiang Education Publishing House, Hangzhou, eBook (2008)
16. Kreimeier, B.: The case for game design patterns (2002). http://www.gamasutra.com/features/20020313/kreimeier_01.htm
17. Batenburg, K.J., Kosters, W.A.: Solving nonograms by combining relaxations. Pattern Recogn. **42**, 1672–1683 (2009)
18. Kam, M., Agarwal, A., Kumar, A., Lal, S., Mathur, A., Tewari, A., Canny, J.: Designing e-learning games for rural children in India: a format for balancing learning with fun. In: Proceedings of the 7th ACM Conference on Designing Interactive Systems, pp. 58–67. ACM (2008)
19. Logic Pic. http://tappsgames.com/app/logic-pic/

# Game Design Methodology Considering User Experience in Comprehensive Contexts (Trial on Inducing Player to Terminate Game Contentedly by Motivation Control)

Eisuke Hironaka(✉) and Tamotsu Murakami

Department of Mechanical Engineering, The University of Tokyo, Tokyo 113-8656, Japan
hironaka1132@mail.design.t.u-tokyo.ac.jp

**Abstract.** In conventional game designs, only "positive" user experience (UX) such as how to make the game fun and attractive has been drawing attention. The play of a game, however, belongs to not only a context of a game but also comprehensive contexts including life and society. Problems such as being got a game machine away by parents because of too much game play and causing troubles and accidents by playing a game while walking or driving might be considered as "negative" UX in such comprehensive contexts. The objective of this research is to propose a game design methodology of preventing negative UX of the game in comprehensive contexts without reducing positive UX of the game. This paper reports a trial on controlling the user's motivation by changing the game parameters of "Breakout" such as ball speed and paddle size and inducing the user to terminate the game contentedly.

**Keywords:** Game · User experience · Motivation · Context · Flow theory

## 1 Introduction

Since a game such as a video game is a product for entertainment, most conventional studies on games focus on how to attract people into playing games and how to keep people playing games as much as possible. Recently, the concept of user experience (UX), which is defined as all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur before, during and after use of product or service [1], is becoming important in the field of product design. How people are attracted play a game (before the game play) and how people enjoy and continue playing the game (during the game play) should be considered as "positive" UX of games in the context of game play. The play of a game, however, belongs to not only a context of a game but also broader contexts of life and society (Fig. 1), and there could be "negative" UX in comprehensive contexts of not only game but also life and society, such as,


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 390–402, 2019.
https://doi.org/10.1007/978-3-319-94619-1_39

- being got a game machine away by parents because of too much game play and resulting poor academic performance,
- causing troubles and accidents by playing a game while walking or driving.

Since these negative UX of a game might be not only personal but also social problems, there is a movement of defining a pattern of gaming behavior (digital-gaming or video-gaming) characterized by impaired control over gaming as "gaming disorder" [2].

With those backgrounds, the objective of this research is to propose a game design methodology of preventing negative UX of the game in comprehensive contexts including life and society without reducing positive UX of the game in a context of game. As the first stage of the research, we examine possibility of inducing a player to terminate a game appropriately. As a method of controlling the continuation/interruption of game play, an application of parental control has been developed which allows parents to monitor the usage of a game by children and restrict browsing websites or playing games from a smartphone [3]. Although such a method of forcibly terminating the game might be able to avoid negative UX in a comprehensive context, it is not an ideal goal for our purpose because it may make the user dissatisfied and deteriorate a positive UX of the game. Moreover, parental control approach cannot be applicable to adult players of games. Therefore, our methodology should enable game designers to design games in which the player can terminate playing the game voluntarily without becoming uncomfortable as they intend.

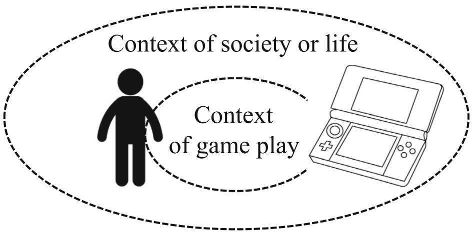



**Fig. 1.** Comprehensive contexts.

## 2 Motivation of Game Play

### 2.1 Related Studies on Motivation

To achieve the goal of this study, we focus on motivation that explains behavioral principle of people, examine game design factors that motivate users to play game, and propose methods of game design that can control users' motivation in game play. Motivation is generally understood in terms of willingness. "Motivation may be defined as the energization (i.e., instigation) and direction of behavior" [4]. Motivation can be classified based on the different reasons or goals. We can classify motivation into two types, i.e., intrinsic motivation and extrinsic motivation. "Intrinsic motivation is defined as the doing of an activity for its inherent satisfactions rather than for some separable consequence” and “extrinsic motivation is a construct that pertains whenever an activity is done in order to attain some separable outcome” [5].

Not only motivation classification based on general theory but also motivational factors in the game have been studied. A study [6] focused on the extent to which individual behavior is self-motivated and self-determined, and proposed self-determination-theory (SDT) explaining that motivation is increased when three psychological needs “competence”, “autonomy”, “relatedness” are satisfied. They applied SDT to video games and explained the factors that satisfy psychological needs in game play [6]. Factors of games that meet the needs of “competence” are important as a field of level design even in the current game. In SDT, the flow theory [7] is related to “competence”. In the level design of the game, the flow theory explain the reason why it is desirable that the degree of difficulty of the game should be balanced with the user skill.

The following three conditions are necessary for the occurrence of the flow experience. “First, flow tends to occur when the activity one engage in contains a clear set of goals”, “second precondition for flow is a balance between perceived challenges and perceived skills”, and “finally, flow is dependent on the presence of clear and immediate feedback” [7]. Flow experience is explained as “flow is an experience so gratifying that people are willing to do it for its own sake, with little concern for what they will get out of it, even when it is difficult or dangerous” [8]. It is desirable for the user to experience flow during completely game play.

Motivational factors caused by the structure of the game have also been studied. Game provides goals according the rules [9]. The goals provided by the game are the states or the result that game finally reaches. We think the achievement of goals and the process of achieving goals are accompanied by intrinsic motivation and extrinsic motivation. For example, when the users achieve the goal, they get a sense of accomplishment, clear up the disappointing failure, and get items, prize money and so on attached to the goals, even in the process of achieving the goal, reward and punitive stimulation may occur. For example, users enjoy work in the process of reaching goals like flow experiences, or the work themselves are bored and the user stop playing the game [10].

Juul said that motivation factors for the game are in failure experiences during game play. The failure depends on the success path to the goals. Juul classified the paths to success into three types as follows [11],

– Skill: “When we fail in a game of skill, we are therefore marked as deficient in a straightforward way: as lacking the skills required to play the game”,
– Chance: “Failing in a game of chance therefore marks us as being on poor terms with the gods (or Fortuna), or as simply unlucky, which is still a personal trait that we would rather not have”, and
– Labor: “A game of pure labor (such as Stat builder) may not even allow a player to fail. Lack of success in a game of labor therefore does not mark us as lacking in skill or luck, but at worst as someone lazy (or too busy) For those who are afraid of failure, this is close to an ideal state”.

In many cases, the game consists of a combination of these three successful routes. Besides these success routes, recently there is also a route to purchase items of the virtual world in game and to advance the game. It is seen in free-to-play games.

Now that we sort out motivational factors from previous research, next we examine how to control the degree of motivation by manipulating motivational factors. Temporal Motivational Theory (TMT) model explains the relationship between motivational factors and the degree of motivation. TMT model expresses the degree of utility or motivation by integrating researches on motivation and behavioral economics such as Prospect Theory, features of Pico Economics, Expected Utility Theory, and Needs Theory [12].

$$\mathrm{Motivation}(\mathrm{Utility}) = \frac{Expectancy \times Value}{1 + Impusiveness \times Delay}. \tag{1}$$

The numerator of Eq. (1) multiplies the success probability or the sense of self-efficacy (*Expectancy*) and the reward obtained from the result (*Value*) [13]. The denominator multiplies the sensitive to the time to get reward (*Impulsiveness*) and its time (*Delay*). The flow experience is considered as internal reward in the process of achieving the goal of games of skill and *Value* of TMT. The degree of labor in the process of achieving the goal of games of labor is corresponding to the *Delay* of TMT. Therefore, here we hypothesize that motivation could be weakened by breaking the balance of difficulty or by enlarging the labor. If motivation can be controlled by independently manipulating difficulty and labor without deteriorating the user experience, it is possible to interrupt the user's game play for various games.

### 2.2 "Breakout" as a Game Example

At the first stage of our research, a simple game with rather small number of game parameters to describe motivation factors in our hypothesis is suitable. As a game to satisfy that condition, we adopt "Breakout" as a game example in this study. The degree of difficulty can be manipulated by the size of the paddle. In Breakout, the time required to achieve the goal is connected with labor, but there are no parameters that can manipulate labor alone. By simultaneously manipulating the speed of the ball and the size of the paddle such as faster ball speed and larger paddle, we manipulated the degree of the time while keeping the difficulty constant (Table 1). When the two parameters are manipulated, it was necessary to quantify the degree of difficulty of the game in order to keep the degree of difficulty constant. In Sect. 3, we conducted simulations and experiments to quantify the difficulty of Breakout when manipulating

**Table 1.** Breakout parameters and influence on motivational factors.

| Manipulated parameters | Influenced factors | |
|---|---|---|
| | Difficulty | Labor |
| Speed of the ball | X | X |
| Size of the paddle | X | |
| Speed of the ball × Size of the paddle | | X |

the ball speed and paddle size by using the concept of configuration space. In Sect. 4, we conducted experiments to control the motivation by manipulating the difficulty and labor of Breakout based on quantified the degree of the difficulty as a result in Sect. 3.

# 3 Quantifying Difficulty of Breakout by Configuration Space

The configuration space (here in after C-space) is a parameter space showing the range of movement restricted by obstacles and possible attitudes and positions and typically used for robot route search and collision avoidance [14]. From the values of each parameter such as the rebound angle, speed of the ball, the position and speed of the paddle, the area where the player can return the ball in the C-space is described, and the area is used as the quantified difficulty of Breakout. We associate the result of the user test with the degree of the quantified difficulty of Breakout so that we can evaluate how difficult the Breakout with specific parameter values in reality by the simulated area size in the C-space.

## 3.1 Simulation

Figure 2 is Breakout model to simplify calculation of its difficulty. We suppose the ball bounces off the center of the upper frame, and ignore the bouncing of the ball on the left and right frames. The kinetic parameters of C-space use the rebound angle $\theta$ of the ball and the position $x$ of the paddle. The $x = 0$ point is the center of the lower frame in Fig. 2. In the model, the ball bounces at angle $\theta$ and velocity $u$ and the paddle of size $L$ can move with speed $v$. Using these parameters, we calculate the area size depicting conditions for the paddle to return the ball in C-space as follows.

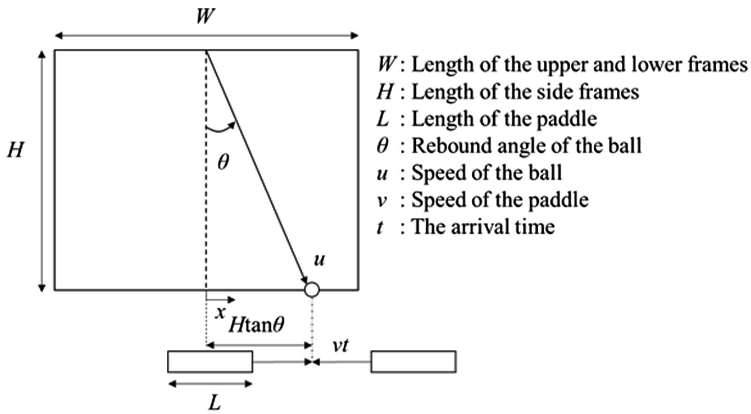



**Fig. 2.** Breakout model to make C-space.

Since the reflection on the left and right frames of the ball is ignored, the domain of $\theta$ is within the angle range of A and is described by Eq. (2).

$$-\tan^{-1}\frac{W}{2H} \leq \theta \leq \tan^{-1}\frac{W}{2H}. \tag{2}$$

Because of the range of the frame and the size of the paddle itself, the domain of the position $x$ of the paddle is represented by Eq. (3).

$$-\frac{W}{2}+\frac{L}{2} \leq x \leq \frac{W}{2}-\frac{L}{2} \tag{3}$$

The arrival position of the ball is $H\tan\theta$. The arrival time $t$ of the ball is $H/u\cos\theta$. The distance that the paddle moves during $t$ is $vt$. Since the position $x$ of the paddle is the center of the paddle, user can return the ball even if the there is a distance to the ball by the size of the paddle. The range of the paddle position $x$ (free area) at which the ball can be returned be expressed by Eq. (4).

$$H\tan\theta - \left(\frac{Hv}{u\cos\theta}+\frac{L}{2}\right) \leq x \leq H\tan\theta + \left(\frac{Hv}{u\cos\theta}+\frac{L}{2}\right). \tag{4}$$

We assume that the extent of this free area represents the ease of returning the ball and the domain represents the difficulty of returning the ball. The area $S$ of the free area and the area $D$ of the domain are expressed by Eqs. (5) and (6). *S/D* was used as the theoretical value of Breakout difficulty.

$$S = \frac{2vH}{u}\ln\frac{1+\sin\left(\tan^{-1}\frac{W}{2H}\right)}{1-\sin\left(\tan^{-1}\frac{W}{2H}\right)} + 2L\tan^{-1}\left(\frac{W}{2H}\right). \tag{5}$$

$$D = 2\tan^{-1}\left(\frac{W}{2H}\right) \times (W-L). \tag{6}$$

## 3.2 Comparison of Simulation and Experiment

We selected the evaluation index of the difficulty of the game measured by the experiment. We set the difficulty level of the game to be “easy”, “ordinary” and “difficult”. The difficulty was based on the number of stages cleared. The threshold as to whether the user can clear the stages can be calculated by the probability that the user misses the ball when trying to return the ball.

$N$ is total of the number of balls shot (including both firing and rebounding). *Efficiency* is the number of times the ball hit the blocks per shot. The number of broken blocks is represented by Eq. (7), and $N$ is represented by Eq. (8).

$$Blocks = N \times Efficiency \tag{7}$$

$$N = hits + misses + 1 + stageclears \tag{8}$$

$N$ is the sum of the first one ball (1), the first 1 ball when you cleared the stage (*stage clears*), the misses that the user restart after mistaking (*misses*) and the rebounding balls (*hits*). The probability of missing is *Miss Ratio* (*misses/N = Miss Ratio*). The stage clear condition is that the number of *misses* is less than the number of life (*life*).

$$N \times Miss\ Ratio = misses \tag{9}$$

$$\frac{Blocks}{Efficiency} \times Miss\ Ratio = misses \leq life \tag{10}$$

$$\therefore \frac{Miss\ Ratio}{Efficiency} \leq \frac{life}{Blocks} \tag{11}$$

In Eq. (11), the condition of the user's performance (*Miss Ratio* and *Efficiency*), which is required to clear, is determined by the value depending on the structure of the stage (*life and Blocks*) is determined. So in the user test, *Miss Ratio/Efficiency* was used as an evaluation index.

Four students from the University of Tokyo (3 males, 1 female) participated in the user test. As shown in Table 2, *u* and *L* were changed under 5*6 conditions. Participants played 30 stages in order, and continued to play until they had returned or made mistakes a fixed number of times.

**Table 2.** Parameters of breakout and values for user test

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| *H* | 24.5 | *L* | 5, 10, 15, 20, 25, 30 |
| *W* | 39.0 | *Blocks* | 21 |
| *v* | 40 | *Life* | 4 (0, 1, 2, 3) |
| *u* | 20, 30, 40, 50, 60 | A fixed number of times | 35 |

### 3.3 Result

Figure 3 shows the correspondence between the *S/D* value and the measured value by the user test. There is region where the variation is somewhat large. As the value of *S/D* increases, the measured value decreases in order. Although *S/D* is not a strict difficulty index, we calculated the degree of the difficulty using S/*D* for convenience and decided the experimental level in Sect. 4.

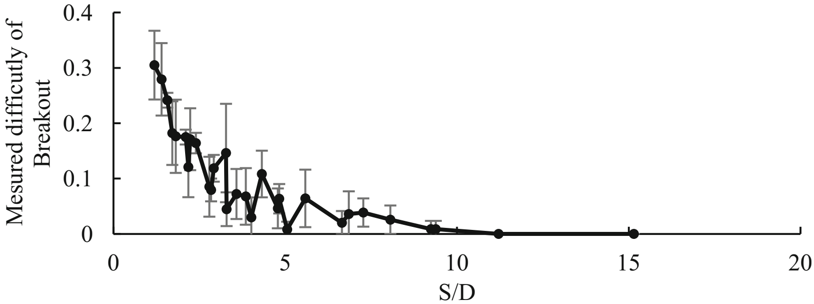



**Fig. 3.** Correspondence between *S/D* and the results of user test

# 4 Control of Motivation by Manipulating Difficulty and Labor

## 4.1 Method

In this section, we verify whether we can control motivation controlled by manipulating difficulty and labor independently, and behavior of stopping playing the game could be induced. We also aim to check whether there is a motivational control method that does not worsen the user experience. We implemented Breakout of 5 Stages and 4 Life using Unity (Fig. 4). If the experiment participants complete all the stages, they may lose their goal and their motivation may be weakened. The game presented the score ranking where the participants' scores are reflected when participants are over the game or have completed 5 stages. In addition, each stage was made so that the stage became difficult little by little. The threshold of performance for completing each stage is calculated as shown in Table 3.

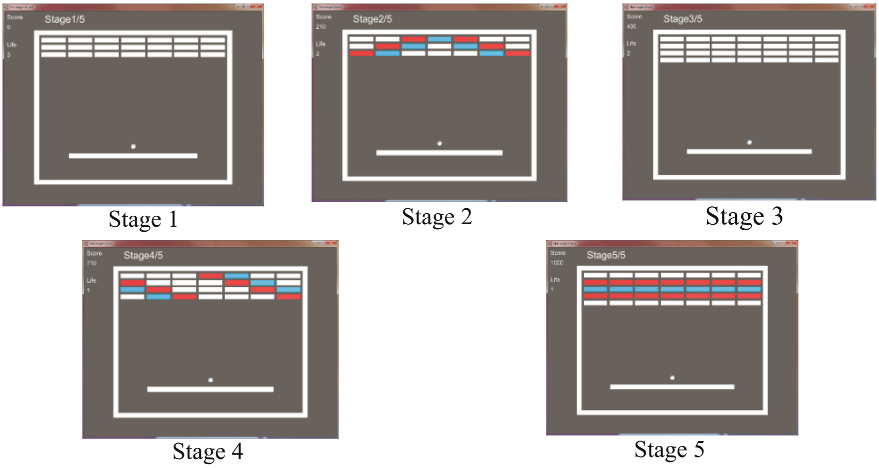



**Fig. 4.** Stages of breakout.

**Table 3.** Required number of collisions per stage and clear threshold.

| Stages | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Cumulative number of blocks | 21 | 48 | 76 | 111 | 160 |
| Threshold | 0.190 | 0.083 | 0.052 | 0.036 | 0.025 |

**Table 4.** Experimental level

| Levels | Difficult | Easy | Medium | Large labor | Small labor |
|---|---|---|---|---|---|
| *u* | 45.0 | 45.0 | 45.0 | 18 | 60 |
| L | 4.49 | 26.68 | 19.83 | 4.075 | 22.46 |
| S/D | 1.5 | 6 | 3.5 | 3.5 | 3.5 |
| Return time [sec] | 1.08 | 1.08 | 1.08 | 2.72 | 0.82 |

Based on this clear threshold, we decided levels of the experiment. *S/D* was set so that *S/D* of difficult game is correspond with the threshold of stage1, *S/D* of medium game is correspond with the threshold of stage 3 and *S/D* of easy game is correspond with the threshold of stage 5. Used level values are shown in Table 4.

When the difficulty of the game was manipulated alone, the speed of the ball was constant and the difficulty was adjusted only by the size of the paddle. When the labor of the game was manipulated alone, *S/D* was constant while the speed of the ball and the size of paddle being manipulated. The degree of labor was set to three levels, Large, Medium and Small based on time for reciprocating when the ball bounces vertically.

We ask participants to keep distance between the display and their eyes about 50 cm so that the visual angle is constant. In Unity, camera setting is Orthographic and Size is 20. The resolution of the screen was set to $1024 \times 768$. Under this condition, the resolution per one unit is 25.6 [pixel/unit]. The paddle can be moved with the left and right keys on the keyboard. For Input setting, Gravity is 10, Dead is 0.001 and Sensitivity is 10 so that the speed of the paddle become 40 unit immediately after pressing left and right keys. Twenty students from the University of Tokyo (19 males, 1 female, mean average 22.95) participated in the experiment. The experiment participants played five kinds of games (medium game + 2 types of difficulty control + 2 types of labor control) in random order. Participants was asked as "It is okay to stop the game when you want to stop the game for some reason, such as when you cannot enjoy the game". The game automatically terminated when 15 min passed. We set 3 min break during each game. Participants answered the questionnaire described below after finishing the game. Motivation and enjoyment are in 5-point Likert type scale, and difficulty and progress speed are by 5-point Semantic Differential method.

- Motivation: 1: I wanted to stop already/2: I could have stopped a bit more quickly/3: I could quietly stop/4: I was good at continuing/5: I wanted to continue more.
- Enjoyment: 1: It was not fun at all/2: It was not fun somewhat/3: I cannot say either/4: It was a little fun/5: It was a very fun
- Difficulty: 1: very easy/2: somewhat easy/3: neither/4: somewhat difficult/5: very difficult.
- Progress speed (labor): 1: very slow/2: slow/3: neither/4: fast/5: very fast.

In addition, participants wrote their impression about game other than above items. Their play statuses were recorded each operation, score and time.

### 4.2 Result

The averages of questionnaire answers are shown in Fig. 5. The results of ANOVA and Multiple Comparison by Bonferroni are also shown in Tables 5 and 6. There was no significant difference in labor evaluation at the time of the difficulty operation. A significant difference was observed in the difficulty evaluation at the time of the labor operation. Motivation of the difficult and large labor games was significantly reduced compared to the medium game. Evaluation value of Enjoyment was reduced with high difficult and large labor games.

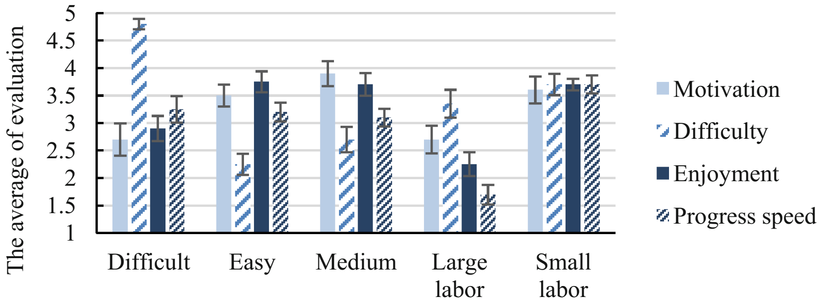



**Fig. 5.** Answers of questionnaire.

**Table 5.** Result of ANOVA (**: $p < .01$, *: $p < .05$)

| | | Df | SS | MS | F | Prob > F |
|---|---|---|---|---|---|---|
| Motivation | Difficulty | 2 | 14.93 | 7.467 | 6.488 | 0.00377** |
| | Samples | 19 | 23.27 | 1.225 | 1.064 | 0.42088 |
| | Residual | 38 | 43.73 | 1.151 | | |
| | Labor | 2 | 15.6 | 7.8 | 6.209 | 0.00464** |
| | Samples | 19 | 19.07 | 1.004 | 0.799 | 0.69432 |
| | Residuals | 38 | 47.73 | 1.256 | | |
| Difficulty | Difficulty | 2 | 74.1 | 37.05 | 70.749 | 1.54E−13** |
| | Samples | 19 | 17.25 | 0.91 | 1.734 | 0.0734 |
| | Residuals | 38 | 19.9 | 0.52 | | |
| | Labor | 2 | 10.3 | 5.15 | 6.741 | 0.00312** |
| | Samples | 19 | 29.92 | 1.575 | 2.061 | 0.02852* |
| | Residuals | 38 | 29.03 | 0.764 | | |
| Enjoyment | Difficulty | 2 | 9.1 | 4.55 | 4.16 | 0.0233* |
| | Samples | 19 | 8.18 | 0.431 | 0.394 | 0.9837 |
| | Residuals | 38 | 41.57 | 1.094 | | |
| | Labor | 2 | 28.03 | 14.017 | 21.622 | 5.37E−07** |
| | Samples | 19 | 13.52 | 0.711 | 1.097 | 0.391 |
| | Residuals | 38 | 24.63 | 0.648 | | |
| Progress speed | Difficulty | 2 | 0.233 | 0.1167 | 0.16 | 0.853 |
| | Samples | 19 | 14.983 | 0.7886 | 1.079 | 0.407 |
| | Residuals | 38 | 27.767 | 0.7307 | | |
| | Labor | 2 | 42.13 | 21.067 | 36.761 | 9.32E−10** |
| | Samples | 19 | 11 | 0.579 | 1.038 | 0.446 |
| | Residuals | 38 | 21.2 | 0.558 | | |

Figure 6 shows the number of interruption behaviors at each level. Only the maximum of nine players interrupted games at each level. As a result of examining the ratio of the number of interruption, a significant difference was observed between the

**Table 6.** Multiple comparison (Bonferroni)

| | Diff-med | Med-easy | Easy-diff | L-med | Med-S | S-L |
|---|---|---|---|---|---|---|
| Motivation | 0.023 | 0.642 | 0.042 | 0.0026 | 1 | 0.0973 |
| Difficulty | 4.50E−07 | 0.17 | 2.40E−10 | 0.0853 | 0.0016 | 0.8194 |
| Enjoyment | 0.137 | 1 | 0.068 | 8.90E−05 | 1 | 3.30E−05 |
| Progress | 1 | 1 | 1 | 6.80E−06 | 0.043 | 1.70E−06 |

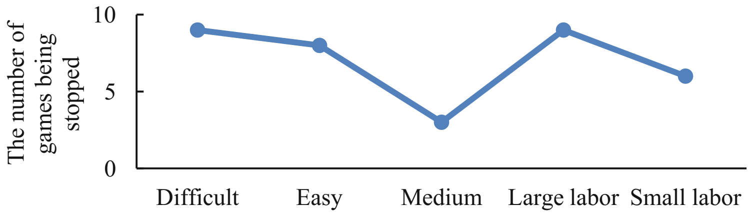



**Fig. 6.** Number of games being stopped.

difficult, easy, and large labor games, and the medium game ($p < .05$, $z$ = N(0,1), difficult: $z = 2.45$, easy: $z = 2.24$, large labor: $z = 2.83$).

### 4.3 Discussions

According to *S/D* of the C-space, the difficulty of the game was set to three levels, difficult, medium and easy, but in the difficulty evaluation about easy game, there was a tendency but there was no significant difference. Because participants played a difficult game in the same experiment, they did not feel the difference of the difficulty between medium and easy game. A significant difference was observed in *Miss Ratio/Efficiency*. There is no significant difference in labor (Progress speed) between the levels, and the preconditions for constant labor are satisfied. The labor was set to large, medium, and small so that the *S/D* of the C-space was constant, significant differences were recognized in the subjective evaluation, so the main effect was confirmed. Experiment participants felt more difficult to play large and small labor games than medium game. From free form questionnaire, for example, the participant pointed the difficulty of player had to predict the trajectory of the ball because the ball is too fast. Difficulties that cannot be expressed in the current Breakout model occurred. Necessary are improvements incorporating models such as human reaction time and adjustment of paddle position.

In difficult games and large labor games, motivation was reduced and interruption behaviors increased compared to the medium games. This result supports the theory of previous researches (Flow, TMT). They were unable to maintain the participants' enjoyment. In difficult game and large labor game, negative emotions such as giving up, getting tired and boredom were caused.

In small labor game, participants' motivation did not decrease much. This also supports previous research. However, it should be considered that the consistency of the difficulty was not kept. As a result, the small labor game was relatively fun and variation in the evaluation was small. The tempo of gameplay is presumed to be an important factor for enjoyment.

In the easy game, there was no significant difference in the subjective evaluation value of the motivation with medium game, which was different from the previous study. In addition, the easy games maintain enjoyment of the game. In the free answer, there was an opinion that "It is fun to be able to clear even if it is easy". It is presumed that enjoyment was maintained because a sense of achievement was surely obtained in the easy game. In the easy games, however, the number of interruption increased more than the medium games. There were two types of people, who were satisfied with completing the five stages and stopped playing the game and who continued the game aiming for score ranking. If the relationship between satisfaction and motivation reduction/end of behavior is known, there is a possibility that the game may be terminated well at the end.

At all levels more than half of the total participants did not finish the game within 15 min. In situations where it is not necessary for the experiment participants themselves to consider the comprehensive contexts, the action of stopping the game is difficult to be induced. In addition, this experiment does not consider interaction between difficulty and labor. In the future, it is necessary to consider combining control of difficulty and labor, making users aware of context, and ways to change goal itself.

## 5 Conclusion

The results of this paper are summarized as follows.

1. We proposed a quantification method of the difficulty of Breakout using configuration space.
2. Large labor and difficult games make users stop playing the game, but the user experience gets worse.
3. Easy games cannot make user's motivation weaken, but there is a possibility that users can stop playing the game at the appropriate timing because they are satisfied with complete of game goals.

## References


1. ISO 9241-210:2010: Ergonomics of human-system interaction—Part 210: Human-centered design for interactive systems (2010)
2. World Health Organization: Gaming disorder. http://www.who.int/features/qa/gaming-disorder/en/. Accessed 28 Feb 2018
3. Nintendo: Make Nintendo Switch part of the family. https://www.nintendo.com/switch/family-fun/. Accessed 28 Feb 2018
4. Elliot, A.J., Covington, M.V.: Approach and avoidance'. Educ. Psychol. Rev. **13**(2), 73–92 (2001)
5. Ryan, R.M., Deci, E.L.: Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. **25**(1), 54–67 (2000)

6. Ryan, R.M., Rigby, C.S., Przybylski, A.: The motivational pull of video games: a self-determination theory approach. Motiv. Emot. **30**(4), 344–360 (2006)
7. Csikszentmihályi, M., Abuhamdeh, S., Nakamura, J.: Flow. In: Elliot, A. (ed.) Handbook of Competence and Motivation, pp. 598–698. The Guilford Press, New York (2005)
8. Sweetser, P., Wyeth, P.: Game flow: a model for evaluating player enjoyment in games. ACM Comput. Entertain. **3**(3), Article 3A (2005)
9. Juul, J.: Half-Real: Video Games Between Real Rules and Fictional Worlds. MIT Press, Cambridge (2011)
10. Endoh, M., Mikami, K., Kondo, K.: Why do players drop out of the game? - Reason due to game design. Digital Game Research Association Japan, pp. 15–18 (2014)
11. Juul, J.: The Art of Failure: An Essay on the Pain of Playing Video Games. The MIT Press, Cambridge (2013)
12. Steel, P., Konig, C.J.: Integrating theories of motivation. Acad. Manag. Rev. **31**(4), 889–913 (2006)
13. Steel, P.: The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. Psychol. Bull. **133**(1), 65–94 (2007)
14. Lozano-Pérez, T.: Spatial planning: a configuration space approach. IEEE Trans. Comput. **C32**(2) (1983)

# Enhancing Usability and User Experience of Children Learning by Playing Games

Fatima Masood(✉), Mahnoor Dar, Muhammad Sohaib Shakir, Muhammad Ahmed, Imran Kabir, Muhammad Hassan Shafiq, Zaheer Mehmood Dar, and Hamna Zakriya

University of Lahore, Lahore, Punjab, Pakistan
Faati67@gmail.com, Mahnoordar367@gmail.com, Hafizsohaib01@gmail.com, m.ahmed.3705@gmail.com, ims.manil993@gmail.com, hassan.shafiq@hotmail.com, zaheerdar@icloud.com, hamna.khawaja@yahoo.com

**Abstract.** As children is an important part of the society. Nowadays, children can easily use any interactive system. Its need of time to develop interactive systems that can help children in their learning, enhance their knowledge. CCI is vast area of research that focuses on the affiliation among kids and their gadgets and their correspondence advantages. Here, we use 5–12 years aged children to evaluate the UX and usability of games by using different techniques. There are number of games developed on daily basis having various characters. Children are more likely to play that game, where they can behave, as they want. Because there are some children who can act and behave normally but in games, they want some action. Our focus is to enhance the usability and user experience of children to play games that educate them. The idea of this paper is "how children can learn by playing games".

**Keywords:** Usability · Child Computer Interaction · User experience
Playing games · Learning

## 1 Introduction

The Human Computer Interaction (HCI) association is never-endingly developing. At first ascending out of the ergonomics and socio-specific study inspiration likewise, providing supervision and examination into actions based arrangements, the HCI society has reliably been occupied with deviations in the all-inclusive community and methodology that are under its transmit. Deviations in people pass on, learn, work and play have all influenced on the HCI design. For instance, investigate is when in doubt now based on customer involvement relatively on convenience, on carefree nature relatively on effectiveness and on correspondence relatively governor.

Child Computer Interaction (CCI) is an examination territory inside HCI that worries the phenomena about the relationship among kids and computational and correspondence progresses. CCI centers on Play, Learning and Communicating regarding PLC. Empowering advances are accessible for learning through play. In 2011, Read and Bekker [1] acquainted with CCI as the "examination of the Activities,


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 403–409, 2019.
https://doi.org/10.1007/978-3-319-94619-1_40

Behaviors, Concerns and Abilities of Children as they participate with PC headways, much of the time with the intervention of others (fundamentally adults) in conditions that they for the most part (in any case, all around don't totally) control and oversee". It generated from takes of effort, generally decided from goods in advance practice inside preparing and schools, in its particular underlying ages, beforehand framing into an recognizable gathering advantaged the HCI which establish a strategy of forming into its own prepare with its specific related procedures and courses of action.

## 2 Literature Review

It is hard to pinpoint when CCI started; it could impartially battle that the foremost huge workings by Papert and later Kafai, Resnick and Ackermann. In the era of late-80s to mid-90s there was amazing work done in the wide-ranging variety of children's capabilities to practice PCs - a lot of this was being distributed, in HCI, and in addition in diaries about PCs and preparing; one case is the work on youths' data limits by Revelle and Strommen. [1] In a comparative time traverse, there were numerous examinations investigating, how adolescents made do with scrutinizing from screens and the possible results for affirmation, especially talk affirmation were largely seriously investigated.

There are six standards in building up the CCI as far as typified association gadget, material collaboration gadget, and multimodal connection gadget. These standards are: 1. Poking up a component upon a fingertip into a show, 2. organizing customer's want, 3. making reaction self-evident, 4. getting ready for mistakes, 5. being unsurprising to deliver association, and 6. being open-completed courses of action identifying with system ask. In light of these standards, we built up the CCI regarding exemplified connection, material communication, and multimodal association. These standards were connected during the time spent this examination. The creators additionally mirrored that innovation ought to be intended for the tyke's needs and capacities, and can be utilized as a part of the learning condition given empty space and constrained assets. As indicated by these outline standards, this investigation welcomed youngsters to take an interest in the plan procedure by request systems with open-ended inquiries.

## 3 Problem Statement

### 3.1 Too Much Innovation

Innovation is not as important for kids as for developers. Because children is only concerned about games, that is easy to play. Rather than focusing on innovation, developers consider their end-users.

### 3.2 No Feeling of Control

It is specifically identify with dodging tension. As indicated by a Psychologist, while taking an interest in ideal encounters, individuals can control even in troublesome circumstances. As some games have no proper difficulty level e.g. Easy, Classic, Hard etc. Children are helpless to play that game by themselves.

### 3.3 Mindless Games

There are games that have no connection with reality and real world. Children just sit around day and night and play games. Instead of learning new and productive thing, children lost his engagement in any physical activity. Physical activity makes children healthier and it activates brain cells.

### 3.4 Not Properly Defined Objectives and Feedback

Mostly games have no proper objectives for their players. Feed is also an important part. As there is no proper section for feedback, players leave the game and un-installed it from their gadgets.

### 3.5 Transformation of Time

At the point when individuals take an interest in ideal encounters, they encounter a change of time, which appears to go significantly speedier. This also make problem for players. As Children does not maintain their attention on studies.

### 3.6 Built Too Much Emotional Connection

When children attached emotionally with any game, they want to play it all time.

They do not know that game's significance on their behavior and cognition. If kid lost his game or stuck at a game level, he worries and forgets other things.

## 4 Methodology

The proposed methodology is first select a group of children that aged 5–12 years. Take a brief look on their social development skills. Using scale Low- Medium- High, we select appropriate choice for their motivation, ability and trigger level. Evaluation of identified behavior of children, help us to select the right option (Table 1).

**Table 1.** Evaluation of identified behaviors of children at aged 5–12 years

| Social development skill (Target behavior) | Motivation | Ability | Trigger |
|---|---|---|---|
| Shyness | L | L | L |
| Time management | M | L | L |
| Spatial arrangements of personal resources | M | L | L |
| Social abilities | L | L | L |
| Concentration and attention | H | L | M |
| Sibling responsibility | M | L | L |
| Controlled T.V. exposure | L | M | M |

In the above table L represent for Low, M for Medium and H for High. Using the mentioned table, we have to propose a game that can influence individual behavior including shyness, time and resource management, their social abilities, concentration etc. (Fig. 1).

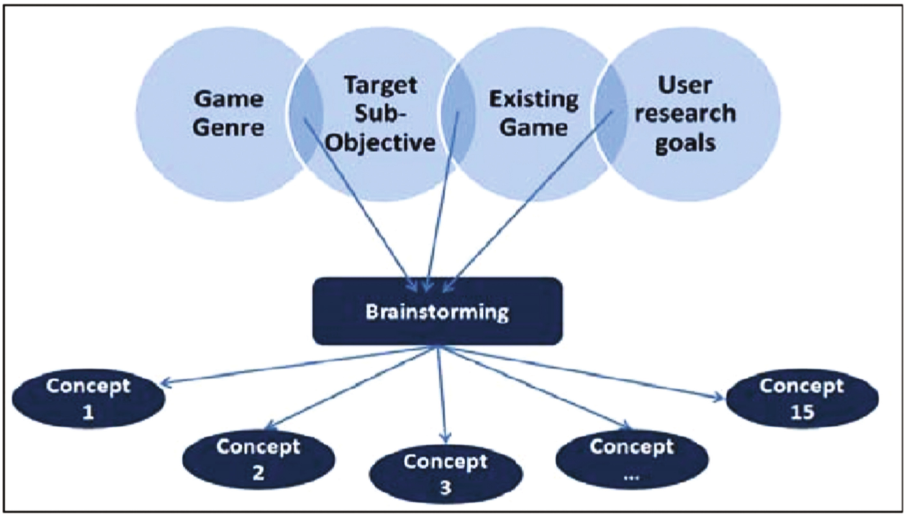



**Fig. 1.** The technique used to generate instant abstract game concepts [7]

Using above concept, brainstorm the group of children to understand what type of game they want to play, how the interface should be. There must be number of concepts generated by this brainstorming. Brainstorming regarding game genre, target objective, goals etc.

**Game Genre:** Classification of the game they want to play I-e, adventure game, real time strategy, puzzle, action etc.

**Target Sub-objective:** It is purely base on behavior of the child. It also includes how much time an individual require to perform a specific target.

**Existing Game:** This question is enough to brainstorm their likes and dislikes about particular game. As I believe children are best evaluators of any interactive product.

**User Research Goal:** Every one of these builds was get from the aftereffects of our client research about.

After children brainstorming and getting ideas, developers start developing interactive design according to the children needs. After designing, implement it (Fig. 2).

Our unique object was to focus on the gameplay issues, we immediately saw that amusement ease of use is so firmly identify with the gameplay that in diversion assessment general ease of use angles cannot be overlook. These heuristics, in any case, are reasonable for assessing amusements in different stages also since gameplay and diversion ease of use are regular for all recreations. General ease of use of the versatile

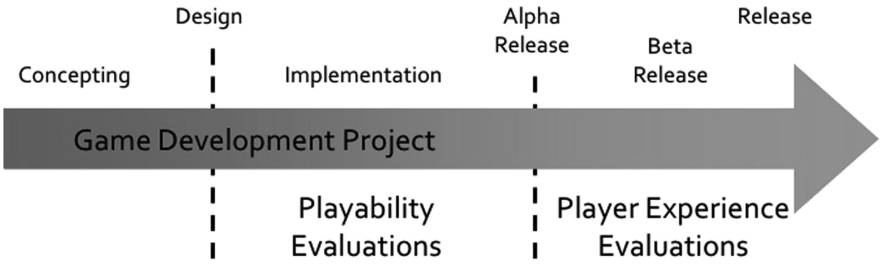



**Fig. 2.** Overview of reviews and heuristic

recreations is an imperative viewpoint and players would prefer not to battle with it, since players are not keen on the UI. The UI of the amusement and control keys that are utilize for controlling the diversion characters ought to be extremely characteristic and instinctive to utilize.

## 5 Results

### 5.1 Add Novelty

For kids, curiosity is important keeping in mind the end goal to have rich encounters that can prompt development. In the meantime, it needs to happen to not overpower them and cause tension.

### 5.2 Feeling of Control

It is specifically identify with dodging tension. There should be a base level of test with a specific end goal to stay away from exhaustion. Children can select the right level with ease and play the game.

### 5.3 Merging of Activity and Mindfulness

This includes having all consideration concentrated on the action, lost mindfulness, and a sentiment being unified with the action. These are largely pointers of an abnormal state of engagement and inundation with a movement. Specifically, the sentiment being unify with the action is probably going to require mechanical cooperation composed such that they feel regular and give a feeling of control.

### 5.4 Clear Objectives and Feedback

All together for the experience to be ideal, individuals need to know whether they are gaining ground. The utilization of the suitable measure of response time accurately. For this situation, regardless of what movement kids are doing, whether it is recounting stories, building, puzzle solving or performing music, obviously they require suitable criticism keeping in mind the end goal to learn ideally.

### 5.5 Transformation of Time

At the point when individuals take an interest in ideal encounters, they encounter a change of time, which appears to go significantly speedier. This ought to be recognizable to any individual who at any point encountered an ideal ordeal (e.g., talking with a dear companion, playing a computer game) and did not understand how much time had passed by.

### 5.6 Emotional Connection

The first thing is emotional attachment, which means how individuals feel about taking an interest in an action. The second is significance, or how much essential the action is to the individual taking an interest in it. In the examinations directed by O'Keefe and Linnenbrink-Garcia, members who had high passionate enthusiasm for action of high individual noteworthiness performed superior to anything the individuals who did not. Their examination recommends that collaboration fashioners need to discover courses for youngsters to associate sincerely with advances in exercises that issue to the kids, keeping in mind the end goal to draw in them in ideal encounters.

## 6 Conclusion

In this paper, we have concluded that if we apply UX methods to evaluate the game before hand then it will enhance the usability and UX of Children by playing games. As these methods can be implement at any phase of game development, so this would be helpful for learning and cognition of children. Developers should also involve children before starting development on any product. Instead of launching many applications, they should focus on these aspects too. As Children is an important part of our society, we should stop developed games that does not involve children as their tester.

The idea exhibited in this paper has not tried for influence yet. In any case, in our current proceeding with examines we are directing trials to test the influential energy of such amusements. Future work includes testing this game for influence that will give us more experiences on the human-centered approach to deal with games.

**Acknowledgements.** We authors acknowledge with thanks assistance rendered by Prof. Dr. Javed Anjum Sheikh, University of Lahore, Gujrat Campus for providing crucial insight during the course of the research work which greatly improved the manuscript.

## References

1. Read, J.C., Bekker, M.M.: The nature of Child Computer Interaction. Paper presented at the HCI2011, Newcastle, UK (2011)
2. Iversen, O.S., Smith, R.S.: Scandinavian participatory design: dialogic curation with teenagers. Paper presented at the Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany (2012)

3. Hengeveld, B., Voort, R., Caroline Hummels, C., Overbeeke, K., de Moor, J., van Balkom, H.: LinguaBytes. Paper presented at the Proceedings of the 7th International Conference on Interaction Design and Children, Chicago, Illinois (2008)
4. Marco, J., Cerezo, E., Baldasarri, S., Mazzone, E., Read, J.C.: User-oriented design and tangible interaction for kindergarten children. Paper presented at the 8th International Conference of Interaction Design and Children, Como, Italy (2009)
5. Malaka, R., Porzel, R.: Design principles for embodied interaction: The case of ubiquitous computing. In: Bärbel, M. et al. (eds.) Proceedings of KI 2009. LNAI 5803, pp. 711–718. Springer, Heidelberg (2009)
6. (2018). https://www.interaction-design.org/literature/article/the-cognitive-principles-of-gerhardt-powals-ace-your-capacity-to-understand-human-behavior
7. Ponnada, A., Ketan, K.V., Yammiyavar, P.: A persuasive game for social development of children in Indian cultural context. In: IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction, Kharagpur, India, 27–29 December 2012 (2012)
8. Hourcade, J.P.: Interaction design and children. Found. Trends Hum.-Comput. Interact. **1**(4), 277–392 (2007). Published in 17 June 2015
9. Read, J.C., Markopoulos, P.: Child-Computer Interaction. Int. J. Child-Comput. Interact. **1** (1), 2–6 (2013)
10. Lopes, A.: Using research methods in human computer interaction to design technology for resilience. J. Inf. Syst. Technol. Manag. **13**(3), 363–388 (2016)
11. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, New York City (2005)
12. Desurvire H., Caplan M., Toth, J.A.: Using heuristics to evaluate the playability of games. In: Proceedings of Computer-Human Interaction 2004, pp. 1509–1512. ACM Press, New York (2004)
13. Proceedings of the 2nd Workshop on Child, Computer and Interaction. ACM Press, New York (2009)
14. Nijholt, A.: Playful User Interfaces. Springer, Singapore (2014)

# User Engagement Through Multimodal Feedback and Involvement in Game Design with a Wearable Interface

Carlos Arce-Lopera(✉) and Arturo Gomez

Universidad Icesi, Calle 18 No. 122-135 Pance, Cali, Colombia
{caarce, arturo.gomez}@icesi.edu.co

**Abstract.** A wearable interface was designed, built and tested as a prototype to evaluate users' engagement in a game setting. The wearable interface consisted of a light sensor and three different types of actuators: a RGB led, a buzzer and a vibe motor. Evaluations of the interface performance were conducted focusing on two main components: the multimodal feedback system and the involvement in game design over several rounds. Experimental results showed that our wearable interface is comfortable and imperceptible. Moreover, our wearable interface is simple enough for users to be able to add game modes based on social agreements. Our results demonstrate how meaningful multimodal feedback and self-directed involvement in game design can address several human factors challenges faced by user engagement designers.

**Keywords:** Human factors · Game design · User engagement
Wearable interface · Multi-modal feedback

## 1 Introduction

User Engagement (UE) is a measure of the quality of the user experience when interacting with an interface. However, its practical measurement has been proven to be challenging to define, design for, and evaluate [1]. To understand the complexity of user engagement, several theories have been described [2]. For example, the self-determination theory, classifies motivation, which can be directly related to UE, into two categories: intrinsic and extrinsic motivation. The intrinsic motivation is driven by the user´s feelings of autonomy, competence and relatedness. On the other hand, the extrinsic motivation is directed by external factors, such as rewards or threats [3]. This example shows the multidimensionality of the concept of user engagement and the necessity for specific and diverse measuring tools.

One popular measuring tool is the User Engagement Scale (UES), a 31-item questionnaire, that classifies UE in six different dimensions [4]. However, by recognizing the difficulty to measure such a long scale for each subject and for each trial, several researchers have improved and reviewed the UES. The short versions include only four dimensions: Focused attention, Perceived usability, Aesthetic appeal and Reward or Satisfaction factor [1, 4]. Focused attention refers to the feeling of been absorbed in the interaction by losing the track of time. The Perceived usability is the


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 410–415, 2019.
https://doi.org/10.1007/978-3-319-94619-1_41

perceived degree of control and effort experienced. The Aesthetic appeal is the attractiveness of the interface according to our aesthetic perception using all senses. The last dimension, the Reward factor, is a combination of the overall success of the interaction, the sense of having fun, and the curiosity and interest for the interaction.

To be able to test for UE of a novel wearable interface in a game setting, we decided to test the importance of the self-directed, meaningful involvement with game design as a powerful tool to increase user engagement [5]. Additionally, meaningful game elements such as multimodal feedback, can give new insights and understanding on ways to retain the attention of users and customers.

## 2 Methods

### 2.1 Apparatus

The wearable interface consisted of a light sensor and three different types of actuators: a RGB led, a buzzer and a vibe motor. All components were connected to a Lilypad Arduino using a conductive thread (see Fig. 1). Therefore, the multimodal feedback stimuli was designed as a combination of audio, tactile and visual information.

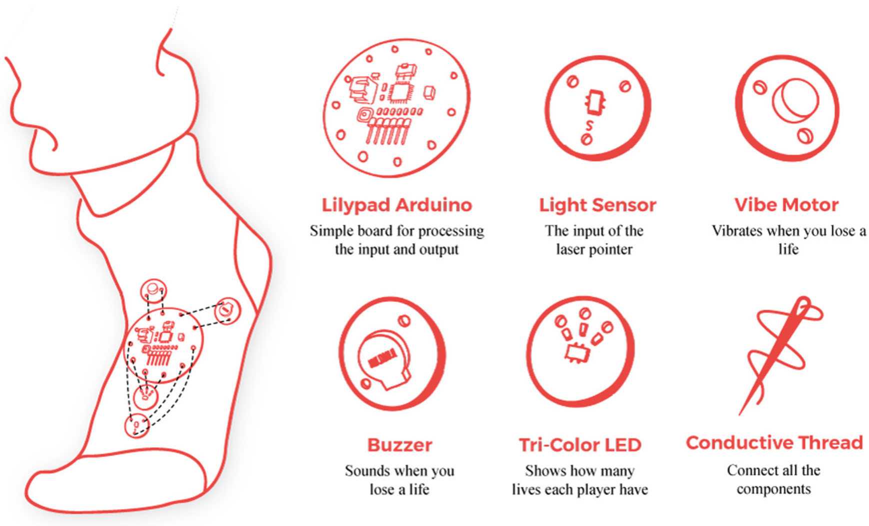



**Fig. 1.** Wearable interface components

The electronic components were sewed into socks that the users have to wear in one foot. A conventional laser pointer was given to be used in conjunction with the light sensor to activate the interaction. When the laser pointed to the light sensor, the feedback system was activated (see Fig. 2).

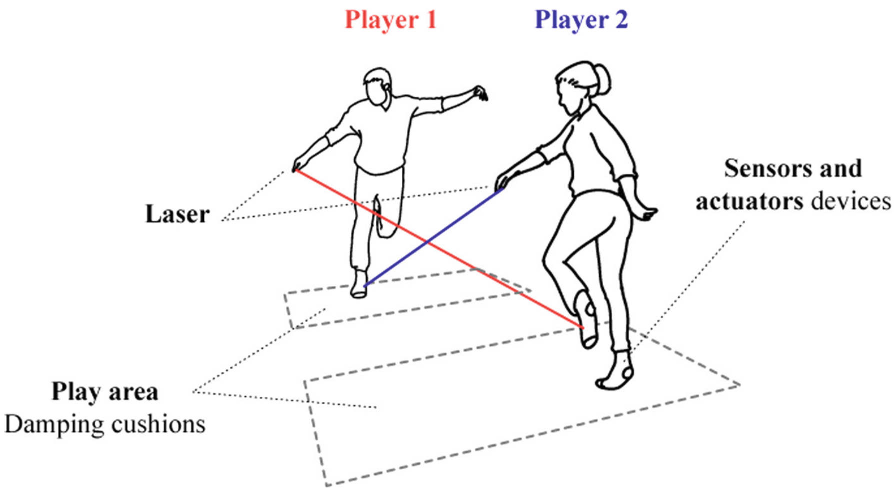



**Fig. 2.** Wearable interface with multi-modal feedback

### 2.2 Subjects and Task

Ten volunteers (three males and seven females; 19 years old in average) participated in the evaluation of the interface. All were naive to the purpose of the study and were instructed to wear the interface in their dominant foot. We asked participants to play a total of four rounds. In each round, they were instructed to abide by fixed rules or to change the rules for the next round. The order of the latter instruction was randomized for each game of four rounds always ensuring that players had two rounds with fixed rules and two rounds with socially mediated rules.

### 2.3 Evaluation

We designed and conducted evaluations of the interface performance focusing on two main components: the multimodal feedback system, and the involvement in game design over several rounds. The NASA-TLX was used as a tool to assess the perceived workload for each player on each round. Also, players' accuracy and response time were recorded. To evaluate user engagement, the game sessions were recorded on video and analyzed manually to extract the different game strategies but also human gestures associated with frustration or enjoyment. This approach is fundamentally different from the UES questionnaire where the responses are bias by the user's perceptions of their interaction. Here, the four dimensions of UE were deducted from a combination of user perceptions (NASA-TLX), quantitative objective data (response time from the interaction) and subjective interpretation by qualified evaluators (video analysis).

# 3 Results

## 3.1 Multi-modal Feedback System Evaluation

Experimental results show that the wearable interface is comfortable and imperceptible during the development of the game since the players understood the meaning of each multimodal feedback from the first round. The users were not instructed about how to interpret the feedback system nor were they aware of its multimodality. They only were told to point to the opponent light sensor to initiate the competitive interaction. All users associated the visual feedback with the status of the player in the game; the green light was associated with the start of the game, the yellow light was related to a warning to a possible loss of the round and the red light was associated with the end of the interaction. The color metaphor that resembles the traffic light rules was selected by the users for their interaction. The tactile feedback was given when the light sensor was activated; one vibration of two seconds meant a hit (light sensor activation) and two consecutive vibrations meant the end of the round. Lastly, the audio feedback indicated the start and ending of the round and the activation of the light sensor using different sounds. In summary, the wearable interface with a multimodal feedback that reinforces the meaningful aspects of the interaction was successful in conveying simple and concrete feedback that support the interaction without introducing confusion or frustration with the interface.

## 3.2 Involvement in Game Design

Involvement in game design as a strategy to increase UE is difficult to measure. For this reason, we used a mixed approach with a combination of self-reported user perception, quantitative recording of response time and events and a subjective video analysis of the interaction. By combining, these three sources of information, the four dimensions of UE can be derived and then contrasted between the two conditions, with or without fixed rules.

The involvement in game design, which were socially mediated instead of imposed increased user engagement. This skewed engagement became evident as the users perceived the game with socially mediated rules as more physically demanding and required more effort (see Fig. 3). The mental demand, performance perception and perceived frustration was not significantly different between the two conditions. However, the users perceived that the temporal demand was lower when the rules were self-imposed by them. Additionally, the time spend in the socially mediated rules condition was in average 20% longer than the time spend in the fixed rule condition.

In addition, the video recordings showed increased gestures of enjoyment when the game obeyed the self-imposed rules. The subjective video analysis (see Fig. 4) classified users in three different categories depending to expressions of frustration, laughter, self-reported satisfaction and commentary by users. The three classes were excellent, very good and good. The difference between the classes were the number of interactions, the number of laughs or frustration expressions interpreted in the video. The results showed that the socially mediated rules rounds were classified with greater likelihood in the top class than the fixed rules condition. The expressions of enjoyment,

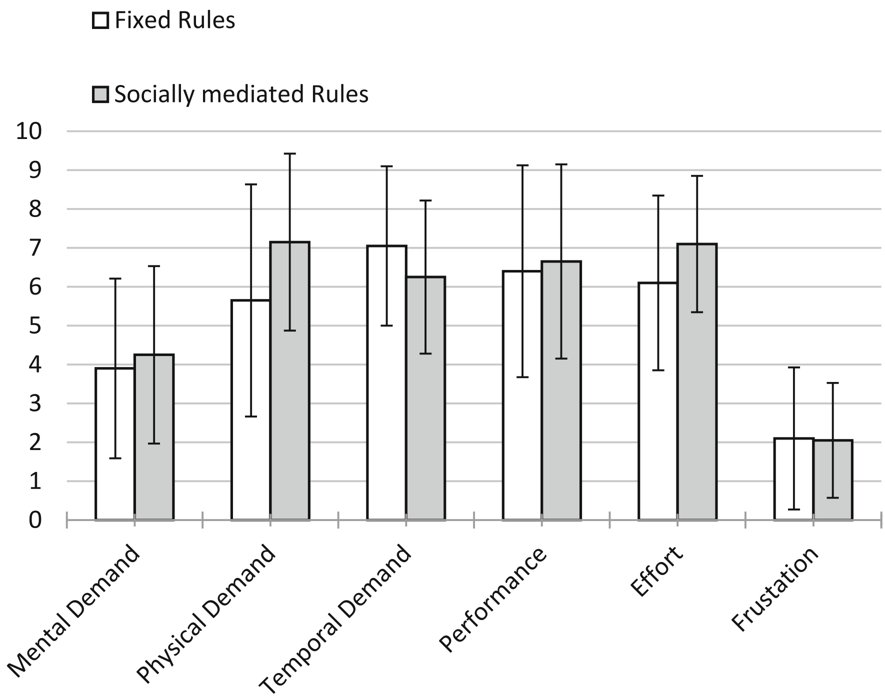



**Fig. 3.** Nasa TLX results.

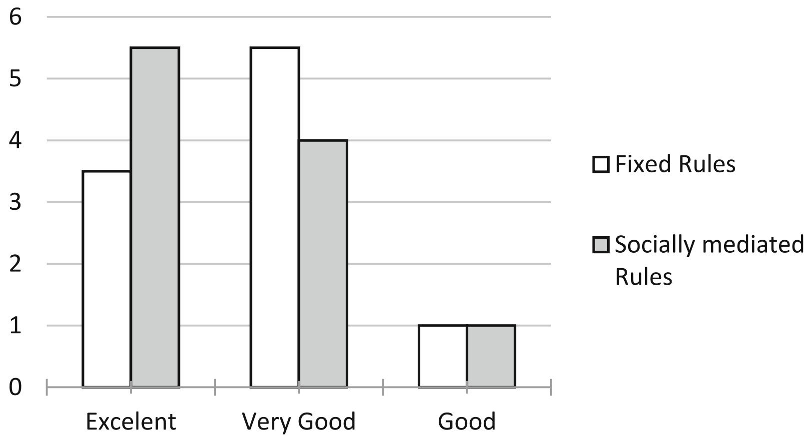



**Fig. 4.** Video analysis results.

such as laughter and jokes increased by 20% as well as interactions between players, such as speech and physical contact. This did not only happened during the rounds but also before and after each round. Although the physical and social effort was greater without fixed rules, the level of satisfaction was greater too.

The four dimension of UE, namely the Aesthetic appeal, Perceived usability, Focused attention, and Reward Factor, can be derived from our results. The Aesthetic appeal of our interface was evaluated by implementing a multimodal feedback system that ensured an imperceptible feedback with congruent multimodal information of meaningful interactions. The Perceived usability was evaluated for both, test and control conditions (with or without fixed rules) using the NASA-TLX instrument and video analysis. The results showed that the fixed rule condition had a lower perceived usability. The Focused attention dimension of UE was measured using a combination of recorded response time, NASA-TLX responses and video analysis. The results showed that when in the socially mediated rules condition, users spend more time but felt that the temporal demand was lower. Finally, the Reward factor was tested with the video analysis and commentaries from the users. Users responses agreed that self-directed rules were more engaging and satisfying.

## 4 Conclusion

We introduced a wearable interface that is comfortable and imperceptible. The interface is simple enough for users to be able to add game modes based on social agreements which demonstrates how meaningful multimodal feedback and self-directed involvement in game design can tackle several human factors challenges faced by user engagement designers.

## References

1. O'Brien, H.L., Cairns, P., Hall, M.: A practical approach to measuring user engagement with the refined User Engagement Scale (UES) and new UES short form. Int. J. Hum.-Comput. Stud. **112**, 28–39 (2018)
2. Silpasuwanchai, C., Ren, X.: A quick look at game engagement theories. In: Norman, K.L., Kirakowski, J. (eds.) The Wiley Handbook of Human Computer Interaction, pp. 657–679. Wiley, Hoboken (2018)
3. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. **55**, 68–78 (2000)
4. Wiebe, E.N., Lamb, A., Hardy, M., Sharek, D.: Measuring engagement in video game-based environments: investigation of the user engagement scale. Comput. Hum. Behav. **32**, 123–132 (2014)
5. O'Brien, H.L., Toms, E.G.: What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inf. Sci. Technol. **59**, 938–955 (2008)

# How Popular Game Engine Is Helping Improving Academic Research: The DesignLab Case

André Salomão(✉), Flávio Andaló, and Milton Luiz Horn Vieira

DesignLab, Universidade Federal de Santa Catarina, Campus Universitário, CCE, sala 101, Florianópolis, SC 88040-535, Brazil
andresalomao3d@gmail.com,
{flavio.a,milton.vieira}@ufsc.br

**Abstract.** In this article we'll be presenting some of the researches being made at DesignLab, inside the Federal University of Santa Catarina (UFSC) - Brazil, using the game engine and how it is helping we get things faster and more pleasant to the audience. DesignLab is a multidisciplinary research lab that works with professors, researchers, undergraduate and graduate students, mainly from Design and Animation programs. Our game engine of choice was Unreal Engine 4. We are using it to develop academic research on multiple topics, such as game design's cycle, serious gaming, 3D animation projects and brand awareness study in games.

**Keywords:** Game design · Game engine · Academic research Animation · Serious gaming

## 1 Introduction

Game engines are getting more popular than ever. The most popular engine by the number of developers is Unity with "over 5.5 million registered developers" [1] by 2016, followed by Unreal Engine, which has achieved "over 4 million developers" [2] in 2017.

This result in a lot of interest from students in learning and using these engines in university in different ways. Since they are very versatile, they can go beyond just learning game design, they are being used for producing short movies, TV series, Virtual Reality applications, Design Visualization and scientific research.

The engine of choice in our case was Unreal Engine in version 4, called UE4. It is a development tool made by Epic for anyone working with real-time technology, with plenty support from the most basic functions to the more advanced features [3].

The interest demonstrated by the students has resulted in a quite wide variety of projects utilizing the Unreal Engine 4, that includes projects on areas like game design's cycle, serious gaming for learning purposes, rendering 3D animations and brand study in games. All these project's details are in the following chapters.


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 416–424, 2019.
https://doi.org/10.1007/978-3-319-94619-1_42

## 2 Engines

When deciding a platform for our projects, we had to choose between the two most popular: Unity or Unreal Engine 4. After reading and testing both, we decided to use UE4 for our projects, to explain our choice we highlight three of the main core features that we decided it was of most importance.

The first is the photoreal rendering in real time *"Achieve Hollywood-quality visuals out of the box. Unreal Engine's physically-based rendering, advanced dynamic shadow options, screenspace reflections and lighting channels provide the flexibility and efficiency to create awe-inspiring content."* [3], this allowed primary the animation's project to have a better result aesthetic wise;

The second feature was the Extensive Animation Toolset [3], not only it was possible to create dynamic objects in the environment, but the tool allowed us to import the work done in other programs and make adjustments inside UE4 with its features.

The third is the main reason we were able to create playable games, the Blueprint system [3] allows the student to build behavior and interactive content without writing a single line of coding, which is important since learning programming languages like C++ are not part of the design's course, our focus.

In addition, in case of any problems during development, Epic has created a documentation website with information going from the basics of learning how to get started with UE4 to more advanced mechanics. This information has proved to be valuable especially to get new students started with UE4, help our students outside of class, guides to using all tools and keeping everyone updated with information on new updates [4].

Moreover, it is also open source and free to use in academic research, allowing our students to freely use its tools for their projects, while professionals need to pay a "5% royalty after the first $3000 per game per calendar quarter" [5].

Unreal engine 4 has a lot more positive points worth talking about, so that combined with an easy incorporation into our previous workflow, the wide support documentation and advanced features the platform offers, has made it the best choice to introduce as a viable platform for the students and academic research.

## 3 DesignLab

In 1994, DesignLab was created, a laboratory focused on Applied Research in Design, current and linked to the Graduate and Post-Graduation Program in Design at the Federal University of Santa Catarina (UFSC). Its purpose is to carry out research projects with an emphasis on technology, as well as an extension of this study for the development of products of social, economic and cultural character.

Our projects cover a wide range of subjects, we create and give support to the development like the research in the areas of 3D Animation, Games, Health Technologies and RFID (Radio Frequency Identification), among possible futures ones that can fit into our line of research. We are also a core of research in accessibility certified by CNPQ (National Council for Scientific and Technological Development), with development in researches with focus on technology that helps increase people's quality of life.

The workgroup that is active at the moment is composed of groups with professors, doctors, and masters, as well as undergraduate students, masters, doctorate and post-doctoral students.

In 2007, DesignLab began implementing a space, called TECMIDIA, and it was supported by several research institutions (FINEP, CNPq, and FAPESC). This new workplace is a multimedia environment that provides a coherent production flow with requirements originated in the researches in development.

The following projects were developed with the support and help from various professionals in each line of research that is being worked and support by the DesignLab.

## 4 Projects

### 4.1 Game Design: Prototype Development

One the first study with the support of the DesignLab was the development of a game prototype. This project had also the objective to try to work around a few problems our students were having during development of their projects at the university. *"we intend to prevent the problems that may appear later in the development of the project, by giving the students plenty of opportunity to test their work and providing them the necessary space to try things out. By the end of the project, a more confident and prepared student will be considerably more aware of their own capacity, which will certainly lead to a better result."* [6].

Because of the lack of security and space to make the harder decisions during the semester *"The students couldn't achieve the whole extension of the project like they have planned."* [6].

By combining MDA Framework [7], Jim Wallson's design cycle [8] with Schell's book [9] we tested a different approach to solving our problems. This project is the first to follow the author's methodology. Its main purpose was to study the cycle of developing and designing a game. The importance of this combination was to *"show to the undergraduate student when designing a game, that all the elements do not sustain by themselves; they are always being influenced by others and the coherence between all of them is one element that distinguishes a great design from a bad one."* [10] (Fig. 1).

Here the Unreal Engine 4 showed its strengths of being easy to use to test everything during the development, *"working in a loop going back and forward on each stage, moving forward just when it was tested and seems to be working"* [6].

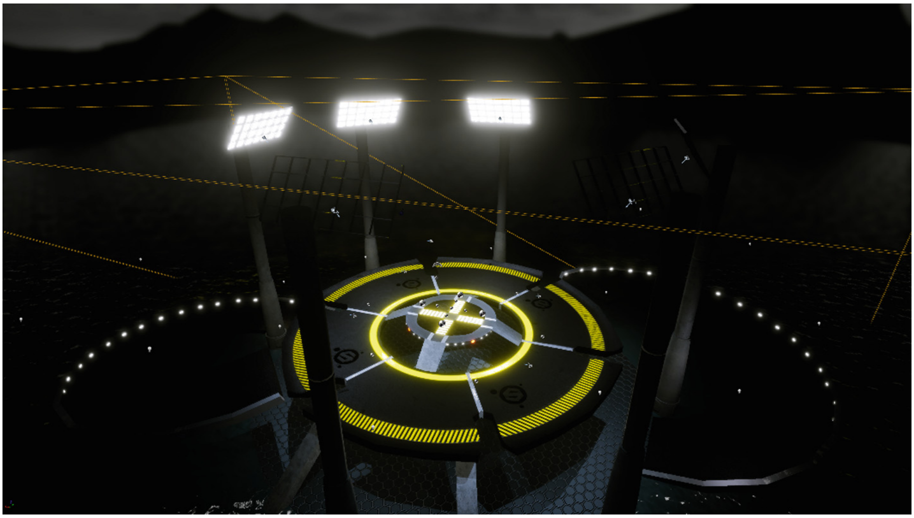


**Fig. 1.** Project's environment inside Unreal Engine 4 for testing.

### 4.2 3D Animation Projects

Parallel to that first project, we started researching the use of the game engine as a final render for animation students, where we could find the weakness and strength of this approach.

The main strength is that, as a game engine, UE4 uses what we call an online or real-time render. This means that it must render and display at least 24 frames every second to get the same smoothness of a film, meaning that every frame must not take longer than $1/24^{th}$ of a second, or 0.042 s. But games usually show 60 frames every second, this means that in this case a frame must be rendered in only 0.017 s. To get this number of frames rendered in such a short time it has to cut some corners on the real-world physics, limit the number of polygons shown on screen and also pre-compute a lot of things before start running.

On the other hand, an offline render is not running in real time, so it can simulate all the optical physics with raytracing, use a huge number of polygons and process everything at render time. The time to render a single frame using a modern raytracer can vary from a few seconds to hours depending on the size and detail of the scene and also the computer power.

Before beginning to utilize Unreal Engine 4 as the final render for the animation projects, we had Scanline, Quicksilver or Mental Ray as the offline renders available with the 3Ds Max free educational version, being the last a raytracer as the name implies. The main issue with using offline render is that the students would never see anything close to the quality of the final rendering without sometimes taking hours or days for a couple of samples made of seconds. Many of the students finished their semester's project with just a preview while leaving the final rendering to be processed during a vacation in our render farm.

Because of this limitation, all projects would work with very basic quality, having only small samples of what would be the final quality rendering. Since using offline renderer was a lot of the times a much slower process, the students would often see mistakes or problems much further in the development, making it hard to adjust them during the time frame they had to deliver it.

Since we started working if UE4 and gaining experience, we decided to try using UE4 as the final rendering solution in one student project develop during a semester. Even though UE4 was made for creating games, it has a very powerful system for edit and renders movies called Sequencer and while the student was working on light and material of their projects they see the results in real time. One problem with this solution though is that the students must learn how to use another program and also pass through the process of export all the assets created and animated in 3ds Max and import in UE4. But since many animation students show interest in game design, they enjoyed the fact of learning a popular game engine. In Fig. 2 there is a sample image of this first project.

**Fig. 2.** This first project was a short movie about a robot looking for a power outlet to charge his batteries.

In April 2017 3ds Max version 2018 was launched with a modern raytracer called Arnold, that has an interactive rendering mode that shows a quick preview of changes made on the scene. As soon as we got this version we adopted Arnold as our offline rendering with much improvement over the older renders available, but even though on another semester we let the students choose. We had the students divided into seven groups and six off them preferred to work with UE4 as a final render.

The results showed that, after understanding the limitations of a real-time rendering and learning the basics of importing assets, creating light and materials in UE4, the students were very pleased of working with real-time rendering and also about the fact that they wouldn't have to wait for over a week to get their final rendering. Below there are some images of the animations produced using UE4 at the design course in UFSC (Figs. 3 and 4).

**Fig. 3.** Frame of another student's short movie called "Cristal Cave" made in UE4.

**Fig. 4.** Frame of the short movie called "Pandora".

### 4.3 Measuring Brand Awareness in Games Using Eye Tracking Data

The next project under development in the DesignLab is a study based on the article about brand placement in video games [7]. Where we are modifying the game's template named *Vehicle Game* that was developed and published by Epic in the Unreal Engine 4 marketplace for free.

For the study of this project, we are modifying specifics assets included in the template and creating new ones where we can place local and international brands, adjusting them in specific locations where the player might pay attention to it, without sacrificing the gameplay's experience.

To keep track if the player is actually aware of the changes or is noticing the brands, we will be utilizing the technology of the eye tracking. This technology will enable us to keep track of what the player's eyes are focusing on during the course of playing the game.

The idea of this project is to check and test a few possibilities; the first being if the misplace of the marketing is harming the gameplay experience; the second is that if well placed, it's actually creating a positive effect for the brand; third is how we can properly balance the brand usage inside the game.

We intend to do cycles of tests, in each test we will modify the template as necessary by experimenting with each brand's location inside the track alongside any other necessary element. This way we can start have arguments on proper brand placement inside games and see how much aware the player are of those, by seeing which one is being too aggressive, or maybe being completely ignored by the player. That way we will have a better indication of how are the effects of brand awareness inside games, and how we might be able to use it more efficiently.

The template given by Epic has shown UE4's versatility and capabilities, by having a complete gameplay fully functional allowed us to focus on only editing the necessary parts of the game. This saved us as more than a couple of weeks of creating the map, testing gameplay until we were able to jump into the actual testing intended for this project. By having this easy to set-up template, we are able to make more than one cycle of tests without having much downtime between.

### 4.4 Serious Games

The next line of research inside the laboratory is the development of serious games, Crookall defines it as *"…games that make use of computer technology and advanced video graphics and that are used for the purposes of learning and training, as researched and talked about in countless conferences, publications, institutes, and websites."* [12]. The study of serious games started with the need for a line of research that helps to find new ways of teaching a target audience in subjects or functionalities that are in favor of our community.

At the moment of the writing of this paper, well underway and it's called "Bones Box". The goal of this project is to create a game that is going to be used as a tool to teach students about human-body bones. To create a more immersive experience, this project is using the technology of Virtual Reality, this helps the project to create a more interactive way of approaching the learning process.

In addition to the topic of serious games, a UAV (Unmanned Aerial Vehicle) simulator is under development. This project is still in early development and only in its design phase. The idea behind it is to create a simulator that enables the training of real-life situations. For example, it is possible to use UAV in favor of helping the society in monitoring areas of nature's preservation.

For this project, we will take the approach of Crookall where he defends that it is necessary the implementation of debriefing [12] while also utilizing learning theories, like Experiential Learning Theory [13, 14], and to add to both of this topic, it will be introduced to this project, the game design ideas of Schell [9]. Combining these topics for the design of a simulator that enables a learning and training environment in the form of a serious game.

To enable the development of this project, Unreal Engine 4 became the platform of choice, the reason is that most simulator platforms are heavy made with programming and not gaming-friendly, which means we can simulate something, but it's harder to turn it into a game. UE4 can do both, it might be harder to reach precise simulation, but it is easier to turn our simulator into a game and design levels of training's exercise.

The simulation will not start in such advanced subject; it will first create a very basic tutorial, similar to what games do, where we will introduce the basics mechanics behind maneuvering the UAV. After each test session, it is scheduled a debriefing, including but not only a questionnaire utilizing Schell's game design [9] questions for the person testing the simulator. The idea is to get the result Joe Wolfe means when he talks about the purpose of debriefing *"allows the individuals who were in the experience to share, cross-fertilize, and to generalize their learnings from and between all who participated in the same experience."* (Joe Wolfe, personal communication) [12].

This will allow us to meet the idea behind Experiential Learning Theory, *"…the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience"* (Kolb 1984) [13].

With the feedback, we will adjust the design of the levels inside the simulator, treating it like a game development while also taking into consideration proper learning theories and tools while in development of the game. As cited at the beginning of this project, the end goal idea is that the simulator can replicate real-life activities, like the one example of monitoring areas of nature's preservation. And also recreate real-life locations from the where the training activity could occur, with the help of the realistic aesthetic quality of the real-time renderer of UE4.

## 5 Conclusion

The use of Unreal Engine 4 by the students and academic research has resulted in a number of projects in topics like serious gaming, game design, animation, branding. The versatility showcased by UE4 has allowed these projects to perform in their respective areas by using from the most basic to the more advanced features [3].

The documentation of every single aspect made by Epic [4] and contributors has allowed design's students to create basic game functionalities and modify existing one without going through programming, speeding up the design and testing process. While animation projects had a real-time renderer available to them, making possible an easier visualization of each project.

Overall, the impact of the usage of one gaming engine for the projects has given us a different approach and solutions to problems we were having and created opportunities for new topics of research. This shows the potential of tools like game's engines, by combining its features with each theory used in each project, like the references [7–9, 11–14].

The combination of the development of each project has created a portfolio with a wide range of academic topic researches for the university. Following with the positive results, we will keep pushing game engines as a viable research tool. The DesignLab has the expectation to keep encouraging this alternative approach to the projects as a way to not only improve the results but to develop new propositions and methodologies. These results are an expansion of our academic research potential.

## References


1. Unity Technologies Looks at 1.7 Billion Devices Globally for its Q1 2016 Global "Games by the Numbers" Report. https://unity3d.com/company/public-relations/news/unity-technologies-looks-17-billion-devices-globally-its-q1-2016
2. Celebrating an Unreal 2017. https://www.unrealengine.com/en-US/blog/celebrating-an-unreal-2017
3. Unreal Engine 4 Product & Features Info. https://www.unrealengine.com/en-US/features
4. Unreal Engine 4 Documentation. https://docs.unrealengine.com/latest/INT/
5. Unreal Engine 4 Commercial Game Deployment Guidelines. https://www.unrealengine.com/en-US/release
6. Andaló, F., Salomão, A., Vieira, M.L.H., Mendes, B.: Game design for students: teaching as a whole context. In: Ahram, T., Falcão, C. (eds.) Advances in Human Factors in Wearable Technologies and Game Design. AHFE 2017. Advances in Intelligent Systems and Computing, vol. 608. Springer, Cham (2018)
7. Hunicke, R., LeBlanc, M., Zubek, R.: MDA: a formal approach to game design and game research. In: Proceedings of the AAAI Workshop on Challenges in Game AI, vol. 4, no. 1. AAAI Press, San Jose (2004)
8. Wallman, J.: It's only a game. In: Chestnut Lodge Wargames Group Fourth Annual Conference (1995)
9. Schell, J.: The Art of Game Design: A Book of Lenses, Second Edn. CRC Press Taylor & Francis Group, Boca Raton (2008)
10. Salomão, A., Andaló, F., Horn Vieira, M.L.: Understanding game design for the development of a game environment. In: Stephanidis, C. (ed.) HCI International 2017 – Posters' Extended Abstracts. HCI 2017. Communications in Computer and Information Science, vol. 713. Springer, Cham (2017)
11. Nelson, M.R.: Recall of brand placements in computer/video games. J. Advertising Res. [s.l.] **42**(2), 80–92 (2002). http://dx.doi.org/10.2501/jar-42-2-80-92. WARC Limited
12. Crookall, D.: Serious games, debriefing, and simulation/gaming as a discipline. Simul. Gaming **41**, 898–920 (2010). https://doi.org/10.1177/1046878110390784. First published on January 6, 2011
13. Kolb, D.: Experiential Learning: Experience as a Source of Learning and Development. Prentice Hall, Upper Saddle River (1984)
14. Kolb, A.Y., Kolb, D.A.: The learning way: meta-cognitive aspects of experiential learning. Simul. Gaming **40**, 297–327 (2009). https://doi.org/10.1177/1046878108325713. First published on October 10, 2008

# The Relation of Attention Between Player Profiles: A Study on the Eye-Tracking and Profile BrainHex

Victor Moreira[1](✉) and Maria Lúcia Okimoto[2]

[1] Department of Graduate Design, Federal University of Paraná, Curitiba, Brazil
victoremmoreira@gmail.com

[2] Ergonomics and Usability Laboratory (LabErg), Federal University of Paraná, Curitiba, Brazil
lucia.demec@ufpr.br

**Abstract.** The analysis of the profile of video game players predicts certain behaviors that the player will have during the game and, as result, can predict what the focus of attention of the players. In this research we used eye-tracking to measure the dynamic patterns of acquisition of visual information of the players, initially tracing the profile of the participants and subsequently collecting eye-tracking data. We assess the relationship of attention according to the player profile. Using the BrainHex questionnaire to profile the player, collecting eye movement data with Eye Tribe and using Ogama software to generate the attention of maps needed for the analysis. It was found that, as a rule, the focus of attention in certain actions corresponds to the player profile, while the rate of non-correspondence was insignificant.

**Keywords:** Player profile · Focus of attention · Eye-tracking

## 1 Introduction

The interfaces of digital games are rapidly changing, with the aim of making them more fluid and fun. The biggest challenge these interfaces face is to show information the players need in the moments in which they need it. Rogers et al. [1] explain that the interface is an exchange of knowledge and information that occurs between a user and a computer system. Therefore, the efficiency of an interface depends on the exchange of information between the user and the system. However, it is important to recognize that large amounts of information (e.g. visual pollution) cause a disservice to the interface, as well as the lack of information.

Research on interface seeks to understand what information the user needs and how to make this information available in a simple and intuitive interface. The eye-tracking sensors offer a chance to test aspects of multimedia learning theories and processing during learning [2]. Moreover, using this approach can help in understanding how players focus their attention during games [3]. However, there are other ways to use eye-tracking sensors in serious games, shown in the study by Almeida et al. [4] and Deng et al. [5]. Games can also be controlled by eye movement, which could be an


T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 425–434, 2019.
https://doi.org/10.1007/978-3-319-94619-1_43

affordable solution for people with reduced mobility or lack of motor control. The use of eye-tracking data can change interaction with games, producing new input experience based on attention.

On the other hand, categorization of user profiles is a tool commonly used to synthesize consumer groups. In the 90s, Richard Bartle conducted a study of MUD (Multi-User Dungeon) players in order to find motivation patterns. Bartle [6] describes the four profiles of players encountered (Killers, Achievers, Socializers and Explorers), improving the model [7, 8] by framing the player profiles under the lens of new gaming technologies. Since then, other authors also developed research on player profiles, among which stands out the study by Nacke et al. [9] with the model and questionnaire BrainHex.

By analyzing the models that describe the player profiles, it is notable that certain profiles have certain preferences for information on the interface. For example, players with social profiles look for ways to communicate, in which case interfaces with chats are more interesting for this public. On the other hand, players with the explorer profile need a map available on the interface. In this case, the player profile is directly linked to the focus of their attention on certain points of the interface.

The purpose of this research is to find the relation between the focus of attention of the player and her profile. Starting with the identification of the player profile, then recording the game session to later collect eye-tracking data, we intend to draw a parallel between the focus of attention and player profile. Accordingly, the following hypothesis was outlined: different profiles of players will have different areas of focus, namely, that there is a link between attention and player profiles. Simply put, the player with a predominantly explorer profile will focus her attention on the areas of interface that support this profile, such as maps, routes and new areas.

## 2 Method

The methodological framework of this work is divided into 3 parts (Fig. 1), the first participant selection is made using the BrainHex questionnaire. The second part details the game sectioning stage using eye-tracking tools, while the last part describes the process of data analysis.

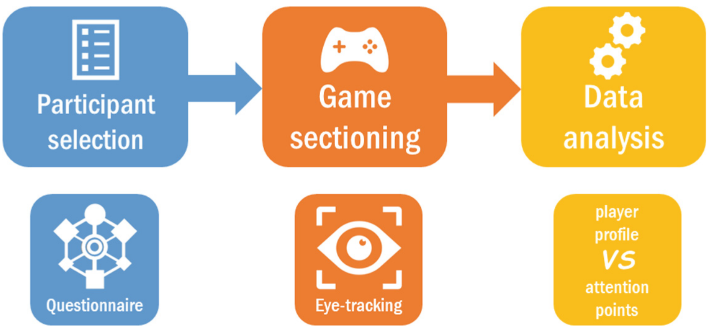



**Fig. 1.** Methodological procedures.

### 2.1 Selection of Participants

The players were invited to participate in the survey through social media and within UFPR (Federal University of Paraná). As a prerequisite, participants must be over 18 and consider themselves a gamer. To sort the player profile, participants answered the BrainHex questionnaire [9]. The questionnaire model combines concepts of neurobiology, psychology and demographics to assess the profile of each player.

There is still the possibility of using other models and questionnaires such as: The motivations of the players Yee [10] or the taxonomy of Bartle [6]. However, the BrainHex model was selected because of its large number of respondents [11], its use of public data, map relationships and its similarities with other models like: Bartle [7], Yee [10] and Lazzaro [12].

### 2.2 Game Section

After answering the questionnaire, the participants attended the Ergonomics and Usability Laboratory (LabErg) at the Polytechnic Campus at UFPR. During the meeting, participants read and signed the free and informed consent form and were introduced to data-collection instruments.

### 2.3 The Game

The game Watch Dogs was chosen because it is an open-world style game that gives players many ways to complete the challenges. By controlling a character in first person, one can understand the preference for actions in the game. These actions depend on the player to strategically fulfill them.

### 2.4 Eye-Tracking

To evaluate attention, we used the eye-tracking to assess the points at which the player's attention is directed. We used Eye Tribe hardware and Ogama software for data collection.

### 2.5 Data Analysis

The data analysis process occurred in two stages: first, without the knowing the player profile and subsequently knowing the player profile. Both analysis made use of Ogama software, and the main tools used were: heat map, paths and fixing, attention areas and attention focus. Bergstrom and Schall [13] describes these analysis tools as follows:

**Heat Map:** Consists of representing, through the use of color, a specific element analysis (or dimension). A color scale is used where the red is hot and green is the cold. This scale visually informs the plus and minus elements seen in the interface.

**Gaze Plot:** Summarizes eye behavior and displays fixings and paths taken by the eyes. It also indicates the sequence and order of an individual's eye movements. Using circles and lines to represent data, circles are used to represent fasteners connecting lines and circles represent these scan paths (saccadic movements) that occur between anchorages.

**Cluster:** It is automatically generated based on the intensity and concentration of viewing points. The main areas of attention of individuals summarize the focus of interest.

Bee Swarm: dynamically represents a set of points in an image that corresponds where individuals concentrated their attention during a time interval. Attention focus is most commonly used when you want to analyze a number of individuals, in order to relate the points of attention of an individual in relation to the other.

The analysis of the play sessions uses primarily data of attention areas and attention focus of the participants. These analyses are performed as follows:

1. Select a particular action in the game, for example, driving from point A or B, talk to a NPC, pick a point on the map, etc. The time of action may vary among participants.
2. Search for differences in attention among the participants. At this stage, compare the heat maps of participants who made the same kind of action.
3. Select one of the heat maps at random and compare it with the player profile.
4. Check that the attention focus coincides with the player profile, verifying points that either justify or do not justify the profile.

## 3 Results

Eight undergraduate and graduate students from UFPR participated in this study (6 men and 2 women). Data loss rate was 8% due to error during collection and gameplay video of the rescue. Each game section lasted 45 min, being recorded in 15 min intervals due to limitations of the collection program. We used a computer with good performance for gaming, aided by a GTX 980 TI video card, using a wi-fi joystick as input for the game (Fig. 2).


**Fig. 2.** Photo setup set for data collection.

The first step of the analysis sought to assess recurring actions between players, for example: walking the streets and sidewalk, driving a car, hiding and analyzing the space around oneself and others. Stretches of at least 10 s were compared so that the eye-tracking data was consistent. Each player took a different time for the same task, however, the attention map (Figs. 3 and 4) showed the same action of two individuals occurring with different attention focuses. In Figs. 3 and 4 participants are performing the same task—running from the cops. While participant A makes use of the mini map, seen on the bottom right-hand corner of the screen, participant B focuses her attention on the track.


**Fig. 3.** The participant A, makes great use of the mini map.


**Fig. 4.** Participant B, make little use of the mini map.

The results of the questionnaire BrainHex displays the player profile in primary and secondary. In the population of participants, the majority (62.5%) belonged to the Achiever profile, while the Seeker profile was 25%, and only 12.5% were in the Conqueror profile. Other key player profiles were not identified in this study. The secondary profile offers a wide range of profiles. Figure 5 shows graphs relating to the profile of primary and secondary participants (Fig. 6).

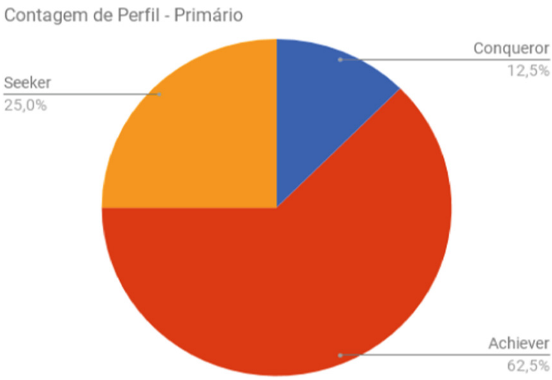



**Fig. 5.** Primary profile of the participants.

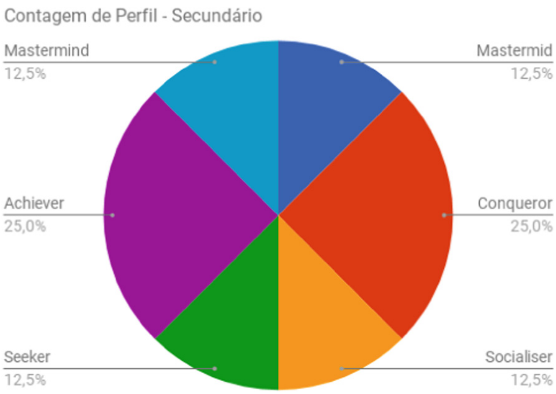



**Fig. 6.** Secondary profile of participants.

Analysis of the actions and attention focus of individuals in relation to player profile was built on a 3-level scale: matches the profile completely, partially consistent with the profile, and not consistent with the profile. This scale was used as described in the BrainHex profile [9], exemplified in Table 1.

The classification of player actions takes into account the primary and secondary profile, when one of the two fits with the actions, this is considered to be consistent. Adverse actions to the player profile were also classified; i.e. when a search player does

**Table 1.** Example of assigning values to the actions player profile.

| Profile *Conqueror*<br>Description: you like to defeat difficult enemies fighting until you achieve victory | Actions |
|---|---|
| Matches the profile completely | Fight vigorously to win the victory without worrying about the consequences |
| Partially consistent with the profile | Fight strategically to win a win worrying about not causing major damage |
| Not consistent with the profile | Struggling to analyze the location and the people, he is much more concerned with the consequences than with the mission |

not make a move that does not match her profile, this is also classified. The index of matches the profile completely reached 70%, partially consistent with the profile were 25%, and not consistent with the profile 5% (Fig. 7).

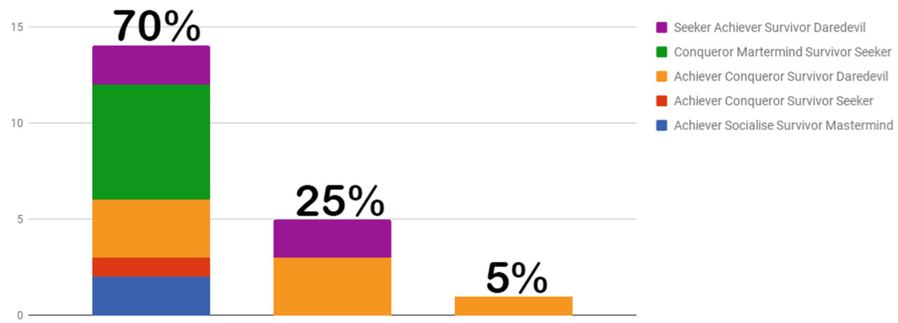



**Fig. 7.** Index of actions/attention matching profile.

## 4 Conclusion

This study has shown that the relation between player profile and the attention focus during the game are equivalent. Although some players pursue actions or focus attention on acts that do not match their profile, the data shows that most of the players act according to the descriptions of their profile. This reinforces that the BrainHex model can be used as a linear of actions and attention of the players during the game. From the analysis of the results, the following conclusions were made:

*1 - Understand what the players like is as important as understanding what them does not like.*

During the first analysis (in which researchers do not know the player profile), data collection was based primarily in actions which the player spent more time doing, or in actions that were coincident (e.g. driving a car). This type of observation generated images which describe the actions and where the player attention was focused. In the second analysis (in which researchers know the player profile), collection became more specific, sometimes seeking that which reinforced the player profile and others that which did not. At this point, the data from the primary and secondary profiles (see an example in Table 2) were repeatedly consulted; However, we realize that analyzing adverse profiles could broaden the collection strategy.

**Table 2.** Example of results of the questionnaire BrainHex

| | |
|---|---|
| Achiever | 20 (main profile) |
| Mastermind | 16 (main profile) |
| Conqueror | 14 |
| Seeker | 11 |
| Daredevil | 7 |
| Survivor | 6 |
| Socialiser | 3 (profile averse) |

*2 - Achiever BrainHex profile is broad, covering the main actions of one player.*

The description of the Achiever profile relates to several other player's profiles. Nacke et al. [14] describes the profile as:

While a Conqueror can be seen as oriented to the challenge, the archetype of the Achiever is more explicitly goal-oriented, motivated by long-term achievements. This distinction may be subtle, but it's important: the preference for Achiever style is routed in "ticking boxes" while the preference for conquering style is rooted in overcoming challenges. The satisfaction felt in the achievement of the objectives is sustained by dopamine (and therefore the pleasure center), but must be understood as being, ultimately, obsessive in its focus. The "Achievers" prefer games subject to conclusion. While the pleasure center is related to this preference, the decision center probably plays a role: the subjective reports of Achiever-style players show a compulsive attachment to achieve goals (Nacke et al. [9, p. 3]).

This behavior described by the author in more detail in International Hobo [11], makes this profile linked to several others that have as their objective focus. In this sense, the goal of the player can be surviving, linking her to the Survivor profile, or the goal can be to make the best strategy, linking her to the Mastermind profile. Seeing that the Achiever profile was found in 80% of profiles of the individuals in this research, the data may change significantly when applied in heterogeneous population profiles.

*3 - Error rate is between 30% (optimistic) and 5% (pessimistic).*

The population of participants in this study was too small to assess the same behavior for the entire range of players. However, the success rate being high shows that the hypothesis of the work is consistent. The error rate highlights the small number of participants, while the success rate underscores the possibility of this theory existing.

*4 - The 45 min this game is little for players inexperienced with this type of game.*

During the game sections, some players have had more difficulties than others to understand the game controls. Despite being a requirement that participants declare themselves as video game players as a prerequisite, some players showed more skill than others. For this reason, the allotted time of 45 min, for 2 people, was little to adapt to the game. In this case, it would be more fair to set a point of the game where the player should get to finish the collection and not limit the collection of data by time.

Making games more accessible with basis on your audience, be it by understanding its limitations or preferences, is a constant challenge of developing games. Therefore, understanding the relation between attention focus of the players and their profiles can provide other ways of understanding the public. In addition to this verification, there are still other possibilities, such as checking the attention of people with disabilities (e.g. deaf). Understanding the difficulties and limitations of the public player cannot be restricted only to people without disabilities. An example of this is the SOMA game, which used eye tracking as game mechanics, enabling people with physical disabilities to play without the need of a mouse.

From this study we propose that it is possible to apply eye-mapping data to player profiles, a statement which can be subject to tests, which may be performed as follows: a researcher informed on the player profile models can evaluate the gameplay and the result of the eye tracking and, without knowing the player profile, measure a player's profile based on their analysis. This test should be done with more than 3 researchers for possibility of the triangulation of data.

**Acknowledgment.** The authors thank Maria Lucia Okimoto Ph.D. for providing the use of LabErg (Ergonomics and Usability Laboratory) for the development of this study, as well as computers and Eye Track. And Wilk Oliveira dos Santos for support regarding the BrainHex questionnaire in Portuguese. And CAPES for financial aid, which enabled the realization of dedicated study.

## References

1. Rogers, Y., Sharp, H., Preece, J.: Design de Interação – Além da Interação Homem-computador. Bookman, Porto Alegre (2013)
2. Mayer, R.E.: Unique contributions of eye-tracking research to the study of learning with graphics. Learn. Instr. **20**, 167–171 (2010). https://doi.org/10.1016/j.learninstruc.2009.02.012
3. Sundstedt, V., Stavrakis, E., Wimmer, M., Reinhard, E.: A psychophysical study of fixation behavior in a computer game. In: Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization, APGV 2008, vol. 1, pp. 43–50 (2008). http://doi.acm.org/10.1145/1394281.1394288
4. Almeida, S., Veloso, A., Roque, L., Mealha, Ó.: The eyes and games : a survey of visual attention and eye tracking input in video games. In: Proceedings of SBGames, pp. 1–10 (2011). https://doi.org/10.13140/rg.2.1.2341.3527
5. Deng, S., Kirkby, J.A., Chang, J., Zhang, J.J.: Multimodality with eye tracking and haptics: a new horizon for serious games? Int. J. Serious Games **1**, 16–34 (2014). https://doi.org/10.17083/ijsg.v1i4.24

6. Bartle, R.: Hearts, clubs, diamonds, spades: players who suit MUDs. J. MUD Res. **1**, 19 (1996)
7. Bartle, R.: Virtual worlds: why people play. Massively Multiplay. Game Dev. 2 **2**, 3–18 (2005)
8. Bartle, R.A.: From MUDs to MMORPGs: the history of virtual worlds. In: Hunsinger, J., Klastrup, L., Allen, M. (eds.) International Handbook of Internet Research, pp. 23–39 (2010). https://doi.org/10.1007/978-1-4020-9789-8
9. Nacke, L.E., Bateman, C., Mandryk, R.L.: BrainHex: a neurobiological gamer typology survey. Entertain. Comput. **5**, 55–62 (2014). https://doi.org/10.1016/j.entcom.2013.06.002
10. Yee, N.: Motivations of play in online games. J. CyberPsychol. Behav. **9**, 772–775 (2007). https://doi.org/10.1089/cpb.2006.9.772
11. International Hobo: BrainHex. http://blog.brainhex.com/
12. Lazzaro, N.: Why we play games: four keys to more emotion without story. In: Game Developer Conference, GDC, pp. 1–8 (2004)
13. Bergstrom, J.R., Jonathan, A.S.: Eye Tracking in User Experience Design. Elsevier, Amsterdam (2014)
14. Nacke, L.E., Bateman, C., Mandryk, R.L.: BrainHex: preliminary results from a neurobiological gamer typology survey. In: Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6972, pp. 288–293 (2011). https://doi.org/10.1007/978-3-642-24500-8_31

# Author Index




T. Z. Ahram (Ed.): AHFE 2018, AISC 795, pp. 435–437, 2019.
https://doi.org/10.1007/978-3-319-94619-1