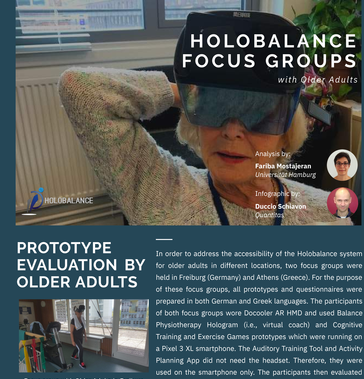


HOLOBALANCE FOCUS GROUPS

with Older Adults

Analysis by:
Fariba Mostajeran
University of Hamburg

Infographic by:
Duccio Schiavon
Qwantani

PROTOTYPE EVALUATION BY OLDER ADULTS

Focus group with Older Adults in Freiburg

In order to address the accessibility of the Holobalance system for older adults in different locations, two focus groups were held in Freiburg (Germany) and Athens (Greece). For the purpose of these focus groups, all prototypes and questionnaires were prepared in both German and Greek languages. The participants of both focus groups wore Docooler AR HMD and used Balance Physiotherapy Hologram (i.e., virtual coach) and Cognitive Training and Exercise Games prototypes which were running on a Pixel 3 XL smartphone. The Auditory Training Tool and Activity Planning App did not need the headset. Therefore, they were used on the smartphone only. The participants then evaluated these prototypes using several standard questionnaires.

RESULTS

The questionnaire was composed of a demographic questionnaire, the System Usability Scale (SUS) [4],User Experience Questionnaire (UEQ) [6],NASA Task Load Index (NASA_TLX) [5], and the perception of the virtual coach (Godspeed [3] and Social Presence [1] questionnaires). A summary of the results is presented here.

Focus group with Older Adults in Freiburg

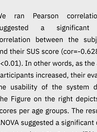
PART ONE

DEMOGRAPHIC

We interviewed in total 24 older adults (16 from Athens, 8 from Freiburg, 21 females) between 60 and 84 years old (M=71.42, Mdn=72, SD=7.99).

Eight of them were interviewed in Freiburg and the rest in Athens. Half of the participants had balance disorder and reported an average of 0.56 (SD=.82) falls and 0.85 (SD=1.32) near falls during the past 6 months.

Our participants were also physically active as more than 70% of them reported daily (37.5%) or weekly (33.33%) performance of exercises. The rest (25%) stated that they never do exercises. To perform exercises, about 70% of the participants indicated that they have 3m² or more free space at home. 13.04% of them had 2 m² and the rest (17.39%) had 1 m² available space at home.



Moreover, about 2/3 of these seniors had broadband internet at home and used computer (including smart phones, iPad, tablet, etc.) on a daily (58.33%) or weekly (4.16%) basis. On a scale of 1 (very bad) to 5 (very good) they rated their confidence in technology in average 2.33 (SD=1.34, Mdn=2.5). In addition, the majority of the participants (91.66%) did not have prior experience with AR or VR.

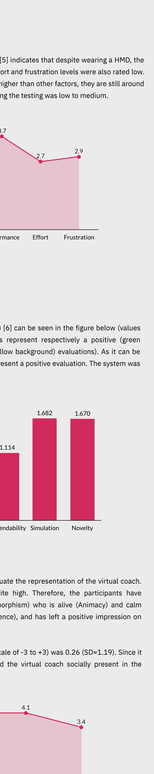
PART TWO

SYSTEM USABILITY SCALE (SUS)

The System Usability Scale (SUS) [4] is a standardized questionnaire to assess the usability of a given product or service. It consists of 10 statements which are scored on a 5-point Likert scale. The total score of the SUS can range from 0 to 100, where higher scores indicate better usability.

In order to calculate the SUS score we had to remove four not available responses. The mean SUS score for the remaining 20 responses was calculated as 57 (min=10, max=92, Mdn=60, SD=24.97).

We ran Pearson correlation which suggested a significant negative correlation between the subjects' age and their SUS score ($cor=-0.628$, $df = 18$, $p<0.01$). In other words, as the age of the participants increased, their evaluation of the usability of the system decreased. The Figure on the right depicts the SUS scores per age groups. The results of the ANOVA suggested a significant difference between the age groups ($F(4,15)=5.57$, $p<0.01$). The pairwise comparisons with Bonferroni corrections showed that the SUS score of the participants over 80 years old was significantly lower than the participants aged 60-64 ($p<0.01$) and 70-74 ($p<0.05$) years. Therefore, the usability of the system depends on the participants' age: the usability for 65-79 is OK to Good and for 80+ is the worst imaginable [2].



PART THREE

NASA-Task Load Index (NASA_TLX)

The results of the NASA-Task Load Index (NASA_TLX) [5] indicates that despite wearing a HMD, the system was not rated as physically demanding. The effort and frustration levels were also rated low. Although the mental and temporal demand was rated higher than other factors, they are still around the average score. Therefore, the overall workload during the testing was low to medium.

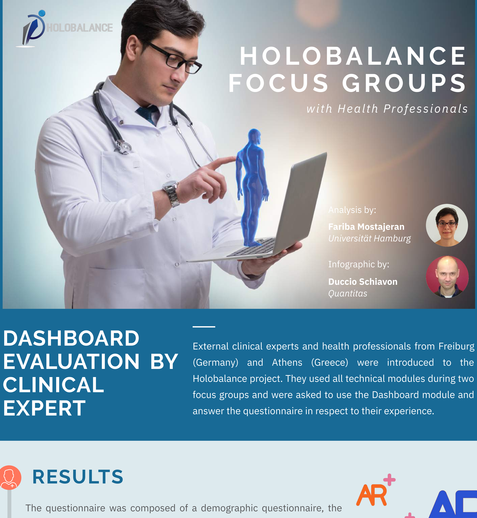
User Experience Questionnaire (UEQ)

The results of the User Experience Questionnaire (UEQ) [6] can be seen in the figure below (values above 0.8, below -0.8, or between these two values represent respectively a positive (green background), negative (red background), or neutral (yellow background) evaluations). As it can be seen, all scales have a mean value above 0.8 which represent a positive evaluation. The system was mostly perceived as novel, stimulating, and attractive.

Perception of Virtual Coach

The Godspeed questionnaire [3] was employed to evaluate the representation of the virtual coach. Different aspects of the virtual coach are rated quite high. Therefore, the participants have perceived the virtual coach like a human (Anthropomorphism) who is alive (Animacy) and calm (Perceived Safety), has intelligence (Perceived Intelligence), and has left a positive impression on them (Likeability).

In addition, the mean Social Presence [1] score (on a scale of -3 to +3) was 0.26 (SD=1.19). Since it is positive, it suggests that the participants perceived the virtual coach socially present in the physical room.



HOLOBALANCE FOCUS GROUPS

with Health Professionals

Analysis by:
Fariba Mostajeran
Universität Hamburg

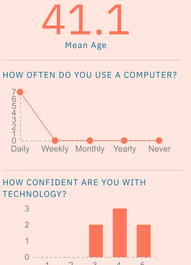
Infographic by:
Duccio Schiavon
Quantitas

DASHBOARD EVALUATION BY CLINICAL EXPERT

External clinical experts and health professionals from Freiburg (Germany) and Athens (Greece) were introduced to the Holobalance project. They used all technical modules during two focus groups and were asked to use the Dashboard module and answer the questionnaire in respect to their experience.

RESULTS

The questionnaire was composed of a demographic questionnaire, the System Usability Scale (SUS) [4], and the Technology Acceptance Model 2 (TAM2) [8]. A summary of the results is presented here.



Focus group with clinical experts in Athens

PART ONE

DEMOGRAPHIC

Five female and two male clinical experts and health professionals (average **age** 41.14, SD=15.6) filled out our questionnaire.

Five of them were located in Germany (71.43%) and the other two in Greece (28.57%).

All of them reported that they use computers (including smart phones, tablets, etc.) **daily**. On a scale of 1 to 5, they rated their level of confidence with technology in average 4.0 (SD=0.82). 71.43% (5 participants) did have broadband Internet at work.

57.14% (4 participants) had prior experience with virtual reality (VR) and/or augmented reality (AR).

All of them stated that they had **no previous experience using remote patient monitoring systems**.

PARTICIPANTS

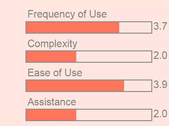
7 people

5 from Deutschland

2 from Greece

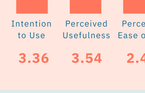
41.1
Mean Age

HOW OFTEN DO YOU USE A COMPUTER?



HOW CONFIDENT ARE YOU WITH TECHNOLOGY?

HAVE YOU EVER USED A VR/AR HEAD-MOUNTED DISPLAY?



DO YOU HAVE BROADBAND AT WORK?

DO YOU HAVE EXPERIENCE USING PATIENT MONITORING SYSTEMS?

"No" = 100%

PART TWO

SYSTEM USABILITY SCALE (SUS)

The System Usability Scale [4] is a standardized questionnaire to assess the usability of a given product or service. It consists of 10 statements which are scored on a 5-point Likert scale. The total score of the SUS can range from 0 to 100, where higher scores indicate better usability.

The total score of the Dashboard was calculated as 69.29 (SD=28.2) which is considered above average. The median SUS score was 82.5 which indicates an excellent usability score. The responses to the individual items are depicted on the right.

SUS SCORE

69.29

PART THREE

TECHNOLOGY ACCEPTANCE MODEL (TAM)

The extended Technology Acceptance Model (TAM2) [8] explains perceived usefulness and usage intentions. It has 26 questions which are grouped under 9 factors. The first three factors were selected as the most relevant to the Holobalance project. Thus, the technology acceptance processes which were evaluated are:

- Intention to Use
- Perceived Usefulness
- Perceived Ease of Use

The results of this evaluation can be seen on the right.

TAM SCORE

[1] J.N. Bailenson, J. Blascovich, A.C. Beall, J.M. Loomis, "Interpersonal distance in immersive virtual environments", Personality and Social Psychology Bulletin, vol. 29, no. 7, pp. 819-833, 2003.

[2] Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies, 4(3), 114-123.

[3] C. Bartneck, D. Kulić, E. Croft, S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots", International Journal of Social Robotics, vol. 1, no. 1, pp. 71-81, 2009.

[4] Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[5] Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage publications.

[6] Laugwitz, B., Held, T., & Schrepp, M. (2008, November). Construction and evaluation of a user experience questionnaire. In Symposium of the Austrian HCI and Usability Engineering Group (pp. 63-76). Springer, Berlin, Heidelberg.

[7] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204.

HORIZON 2020
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769574. **https://ec.europa.eu/digital-single-market/en/policies/ehealth**