HTI Minor Project

Health exercises with an avatar, what effect does it have on the performance of elderly.

Report

19 Februari 2013

Loes Wagter 0720099
Bart van Wezel 0740608
E-mail: l.n.wagter@student.tue.nl
E-mail: b.j.p.a.v.wezel@student.tue.nl

Index.

Introduction.

In the past years many studies indicated that physical exercise is important to keep the body fit, especially for elderly. Nevertheless, the elderly are often not very mobile and prefer staying at home. Also, they sometimes cannot join group sessions, because they cannot keep up with the tempo of the session or can't do all the exercises. Therefore, research is done in several physical exercise activities that can be done at home. These researches focus on three types (Buttussi, Chittaro, & Nadalutti, 2006[1]):

- Computer-supported physical games
- Virtual (embodied) trainers
- Mobile application and devices for physical activities

The research in this report is about the second focus; virtual (embodied) trainers. There are two types of virtual embodied trainers, embodied robots and virtual avatars. Both can be used to stimulate elderly to do physical exercises at home. The goal of both virtual avatars and embodied robots is to help people exercises and thus they should be able of several things:

- Explain the exercises: It should be clear what users have to do.
- Present the exercises: Users should be able to follow the exercises.
- Monitoring the users motion: Users should do the exercise correctly.
- Keep the user motivated

Elderly do not always have the opportunity to hire a personal trainer or to go to the gym. When they use a virtual avatar or an embodied robot they can do most exercises at home at their own tempo. Through this personal adaptation elderly will be encouraged and simulated to exercise more and stay in a better condition.

In our research we investigate how we can improve their stimulations for exercises. Both a trainer avatar and a trainer robot have some big advantages in this stimulation. Nevertheless, what are the differences between them and more important which gives the best training results? The biggest difference between these training assistants is that an embodied robot is 3D and a virtual avatar is 2D. This means that the virtual avatar is shown on a screen and an embodied robot can actually walk through the room. Of course, there are more differences.

Firstly we know several results on how a robot works as trainer. Investigations on robots as personal trainer are done within the KSERA project. (www.ksera.ieis.tue.nl)[2]. In one of these researches it has been already investigated what kind of feedback improves the physical results of elderly. The results of Eberts (2012)[3] research state that low exhaustive exercises are more precise imitated when positive verbal feedback is used, whereas the performance of high exhaustive exercises is not effected by feedback of the robot. Furthermore, an investigation in how the amplitude and speed of the robots' limbs affects the performance of the elderly is done. (Ebert, 2012)[3]. In this research, they measure the amplitude and speed of the limbs of elderly and try to find results on how well elderly can imitate robots' movements. All these researches help in the development of an effective embodied trainer.

Secondly, we also have some results on how a virtual avatar works as a trainer. Virtual avatars are usually 3D animated robots or taped humans, which are shown on a screen. This can be a on a television, a monitor, a laptop or even a mobile phone. These platforms give avatars the opportunity to do the exercises and give feedback. There has been done research to virtual avatars (Ruttkay and Welbergen 2008)[4]. In this research, they made users exercise for 3-5 minutes with the virtual avatar. The virtual avatar shows some exercises and gets information of the motion of the user. The virtual avatar gives feedback to the user if they perform the exercise incorrect. Users were very positive about the virtual avatar, but there also was some critical feedback: too general feedback, speech quality and bad embodiment of the virtual avatar.

We see that researchers are exploring both the possibilities for avatars and robots as personal trainers. Both types of trainer become more advanced after each research. Nevertheless, an interesting question could be what the differences in physical performance is of elderly when using either a robot or an avatar in their training.
We are already familiar with some similarities of the performance of elderly. For example, people were positive about training with an avatar (Ruttkay and Welbergen, 2008)[4] as they also were fore training with a robot (Ebert, 2012)[3]. Nevertheless, we do not know if the physical results of elderly are improving when using an avatar instead of a robot or the other way around.

Because of this uncertainty, we like to state the following research question: Is exercising with a robot better for the physical results of elderly people than exercising with the avatar of this robot?
We expect that users will be more exhausted when repeating robots' exercises than the repeating the ones of the avatar of that same robot, when they are both acting as health instructor. We think that copying movements from a physically 3D robot is easier. We estimate this because we think that being able to see a movement from an angle you like is easier. In addition to that we also estimate that a shape relatable and touchable very similar as your own feels more confidant and is easier to copy.

To find an answer to our research question and test if our estimations are true there are three hypotheses stated:
<u>Our first hypothesis (H1) is that users rate the exercises with the real robot as more exhaustive as the exercises with the avatar of the robot.</u>
We expect this to happen because of the estimation explained in the previous paragraph. We find results here by comparing the results of a test with an avatar with the results of Eberts (2012)[3] research. We will let participants do the same exercises as in Eberts research and let them answer the same questions about how exhaustive this exercise was.

<u>Our second hypothesis (H2) User can perceive and therefore copy robot's arm movements better than avatar's arm movement.</u>
We expect to find prove that the forward movement of arms is harder to copy from 2D screen that from a robot. We don't think that there will be much difference in amplitude with the side movements of arms. We estimate this because we think that forward movements look a little flat on screen and that our visual perception therefore finds it harder to detect the amplitude.
We hope to find prove for this by comparing the users' amplitude to the robots' amplitude with the amplitude of the user when having a virtual agent do the same as the robot did in the already existing investigation.

<u>Our last hypothesis (H3) is: Users rate the real robot just as confortable to work with as the avatar of the robot.</u>
Although this is not specifically about the physical performance of elderly it is about motivation to fitness. Motivation is a big stimulant to exercise well, if people are more positive about the avatar than about the robot, the motivation to fitness with the avatar would be easier to maintain. We have already some prove that they like working with both, when we compare the results of Ruttkay and Welbergen (2008)[4] with Ebert (2012)[3], as we have done earlier in this introduction. Nevertheless this comparison is not based on the same questions and situation; therefore we like to specifically to test the hypothesis and not combine research results.

Methods

Task

Participants were told to do the same exercises as the virtual avatar, in appendix 1 the precise text can be read. The virtual avatar performed 4 exercises these exercises the users should try to copy as accurate as possible. These exercises had 9 types of variations in both speed and amplitude of arms. The table below explains the 9 variations, these are the same variations ebert[3a] used in his research.

Table 1: 9 different amplitude-speed combinations (M1-M9) for every basic exercise (from eberts[3a] report)

Amplitude/Speed	**Slow**	**Medium**	**Fast**
Small	m1	m2	m3
Medium	m4	m5	m6
Large	m7	m8	m9

We measured the performance of participants with the amplitude movement. The participants were also asked to fill in a Godspeed survey[5] before the whole experiment and after the whole experiment they needed to fill it in again. In between each variation of one exercise participants were asked to fill in a short survey.

Participants

In total 16 people participated in the experiment. All the participants had a minimum age of 55. The reason we wanted a minimum of 55, is that the participation in the KSERA project requires participants to be at least 55 years old because we want to study how elderly people perceive exercises. The research of Ebert, which we compared our results to, involves the minimum of 55 years.

Gender distribution: 10 men and 6 women. The ages of the participants varies between 58 years and 77 years. (Average = 66, Standard deviations is 5,62919512)

All the participants did not have a mental or physical disability. They were able to do all the exercises. They were able to understand what the avatar does. They were able to hear and see the avatar.

Apparatus

For the research we used an apparatus that helped us monitoring the elderly and an apparatus that stimulated the elderly to exercise.

In the map below you can see that the participants were 2 meter from the avatar. The Kinect is above the avatar. The projector is behind the avatar, so the participants do not have trouble with the projector. The users could fill in the questionnaire at their right side, so the Kinect would not lose the participants.

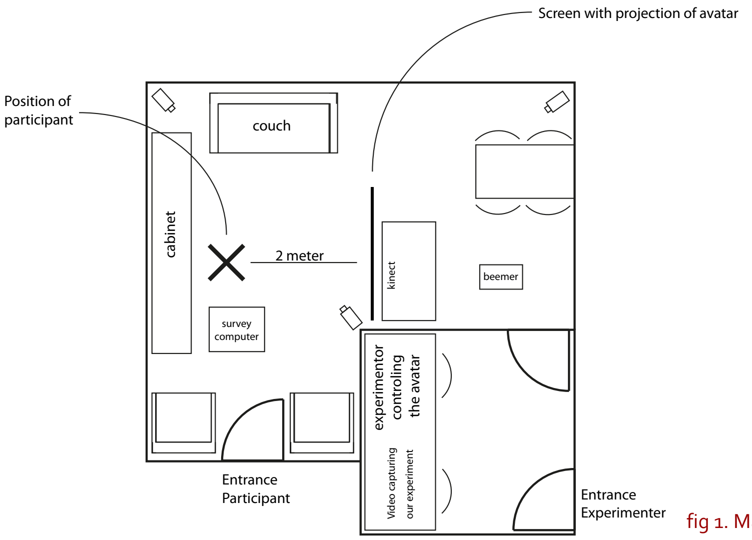


fig 1. Map of the living room set-up

The Kinect
When the elderly are stimulated to perform their exercises we need to monitor how well they perform these exercises. We measure this in the amplitude of the arms. This amplitude can be measured by monitoring the distance of the wrists during the exercises. To measure this we use a Xbox Kinect. This Kinect works with 2 cameras and an infrared camera. The two cameras use the image disparity as in the human eyes, together with the infrared image, to compute from the taped images one 3D image. With this 3D image we are able to detect the position of the hands with respect to each other and we are able to detect the position of the hand within respect to the torso. To process the Kinect measurement we use a Matlab script.

fig 2. Kinect

The Avatar:
We used the avatar to present the exercises to the participants. The participants were told to do exactly the same as the avatar, so the avatar is the stimulus. The participants see that the nao performs an exercise and the participant's copy that exercise.

Software:
A laptop with choreographe and python IDLE to drive choreographe. To the elderly only the choreographe robot will be shown. This avatar will perform the exercises and be the stimulation for elderly to perform their tasks. This will be shown by the use of a beamer that beams the choreographe nao on a screen. Choreographe is a program that controls the visual part of the robot. Python IDLE is used to tell choreographe what the robot should do.

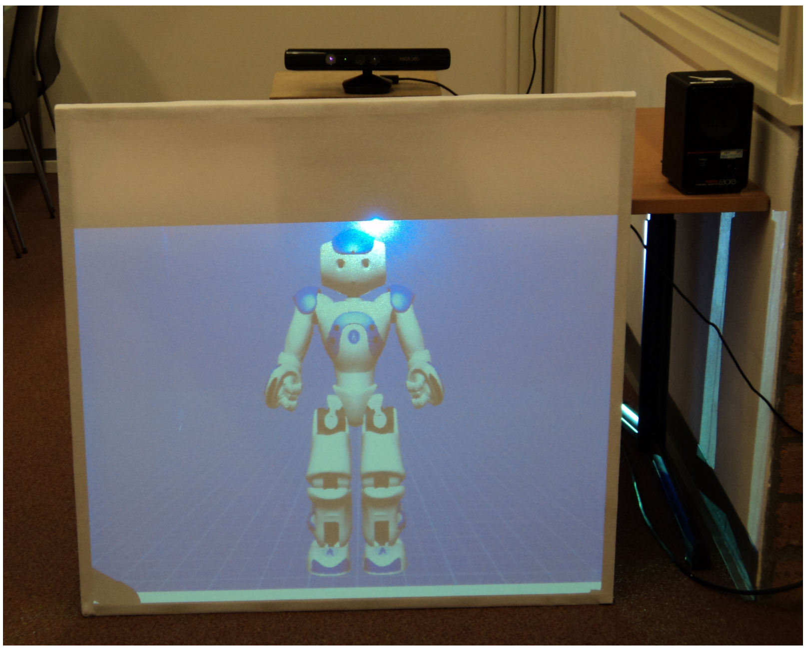


Fig 3. Picture of the robot.

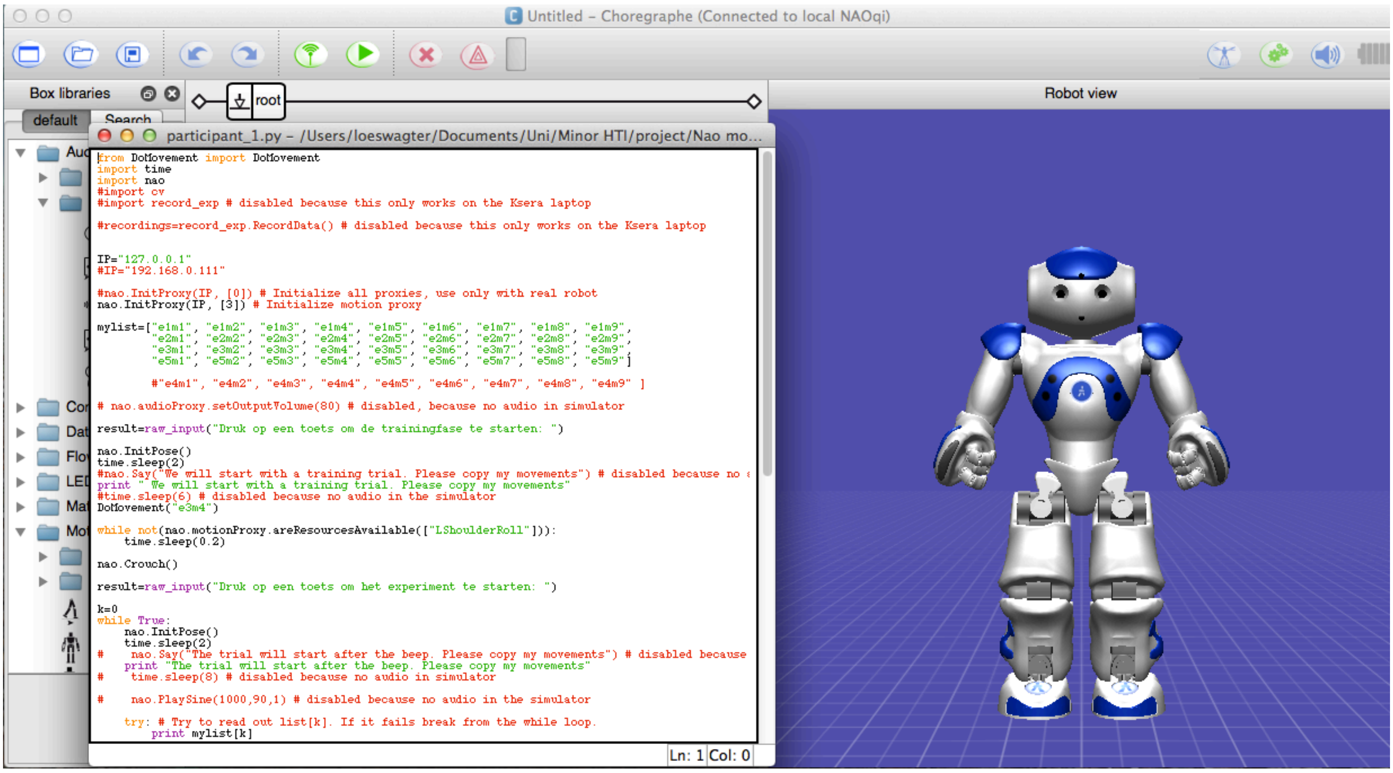
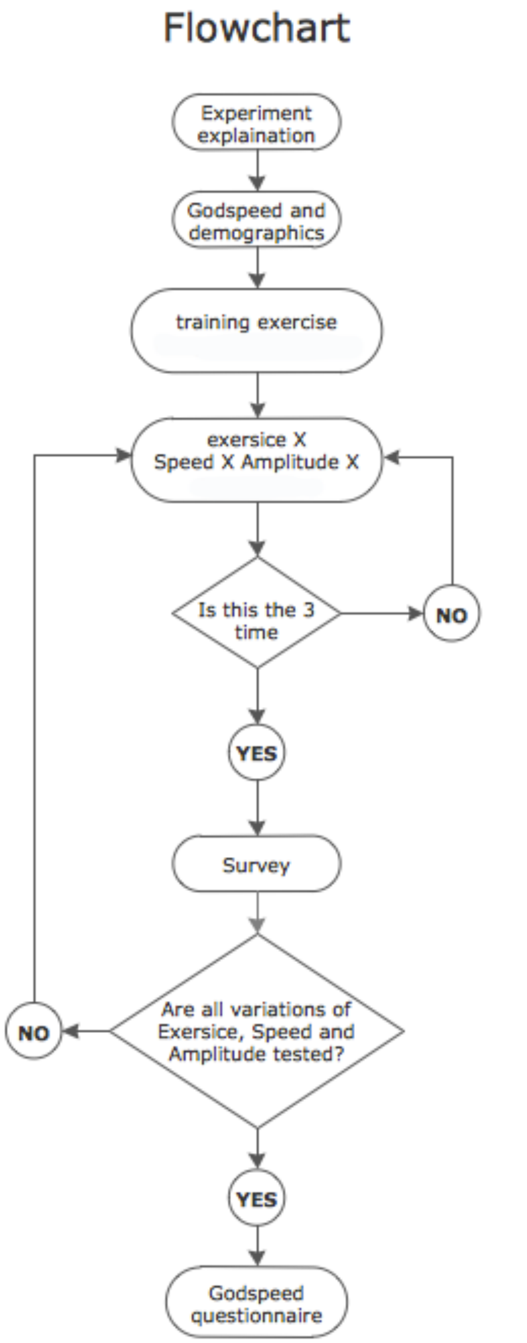


Fig 4. Screenshot of choreographe and IDLE. Only the right part with the avatar will be shown to participants.

Python idle sets the order of exercise. Each exercise involves 9 types this means that exercise 1 has 9 types of variation in speed and amplitude. Each of these variations is executed 3 times shortly after each other, so there is a 4 sec pause and the same exercise will be repeated up till 3 times. After each series of one variation the participant needs to fill in the small survey.



Pic 5. Flowchart of the experiment.

The survey:
For the survey we used a website called www.thesistools.com[6]. The participants were asked to fill in the Godspeed questionnaire (appendix 2). After this they did an exercise and then they filled in a short survey (appendix 3). With the surveys we could measure their opinions about avatars before and after the experiment and with the short surveys we could measure how they experienced the exercise.

Experiment design

We have three different independent variables. The first one is the amplitude of the movement of the robot. The amplitude differs with three levels: small, medium and large. The second independent variable is the execution speed of the exercises. We have three different levels of speed: slow, medium and fast. The last independent is the type of exercise. We have 4 different exercises.
For each exercise we do each combination of amplitude and speed. There are 9 combinations of speed and amplitude. So each participant has to do 4x9 = 36 exercises. None of the participants was able to complete 36 exercises within the time limit of the experiment. The time limit was 60 minutes.

We used both quantitative and qualitative dependent measurements. The qualitative measurements were surveys. They had to answer a Godspeed survey [5] before and after the experiment. The Godspeed survey is found in appendix 2. They also had to answer a short survey between each exercise. This short survey can be found in appendix 3.

With the Kinect we could measure the other dependent variable of the experiment. We measured the distance between both hands and the distance between the hand and the participants' torso. This is the quantitative data.

Experiment procedure

Since we compare our results to the results of Ebert's research it is very important that the experiment procedure is the same as it is in Ebert's experiment with a robot. Our only difference is that we have an avatar instead of a robot.
Therefore the experiment took place in a lookalike living room and had the avatar already showing on a screen when they enter.

We asked participants to come to the TU/e and to wait in the waiting area. We picked them up and guided them to our lookalike living room. After a participant entered, he/she got a general introduction to the experiment (appendix 1) and filled out the informed consent form (appendix 5). After filling the form, they were asked to fill in the Godspeed questionnaire (appendix 2). After they filled in the Godspeed questionnaire they were asked to hold their hand in the air so that the Kinect could recognize their body and calibrate.

When the Kinect had finished the calibration phase, the experiment started. Exercises were presented to participants in randomized order. After they finished an exercise, they filled in the in-between questionnaire (Appendix 3). After they filled in the questionnaire the next exercise would start. After they finished all the exercises they filled in the Godspeed survey again. When they finished the Godspeed survey they would sign the papers that they did the experiment and were offered a monetary reward. After they got their reward they left the room.

Results

We can divide the data collected in two types; data gained from questionnaires (qualitative) and data gained by the Kinect (quantitative). We only analyzed a partial set of the data that could be used to give an answer to H2:
User can perceive and therefore copy robot's arm movements better than avatar's arm movement.

.

Quantitative data: participant amplitude score for exercise 3 and 5.

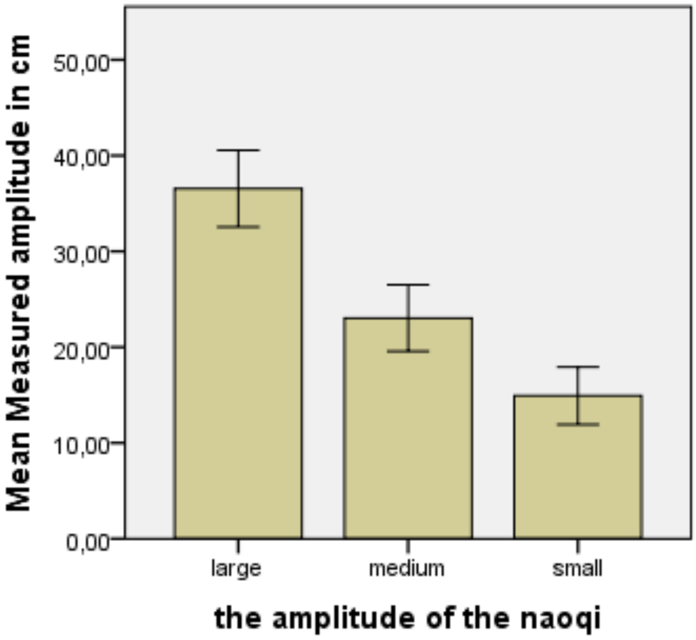


Figure 6: Amplitude scores registered by the Kinect on exercise 3 including the 95% confidence interval

Figure 6 shows the estimated means for the different levels of amplitude for exercise 3. The figure shows an increase of the level of amplitude results in a higher amplitude score. Therefore, as expected, the repeated measurement ANOVA showed a significant effect ($F(2,16)=130{,}065$, $p<0{,}001$)

Contrast analysis reveals a significant difference between all stimulus levels. People were able to correctly perceive small, medium and large avatar's amplitudes. Contrast between small and medium ($F(1,8)= 78{,}081$, $p= 0{,}000$) and the contrast between medium and large ($F(1,8)= 99{,}496$, $p= 0{,}000$).

Figure 7 shows the estimated means for the different levels of amplitude for exercise 5. The figure shows that a higher amplitude level does not always result in a higher amplitude score.

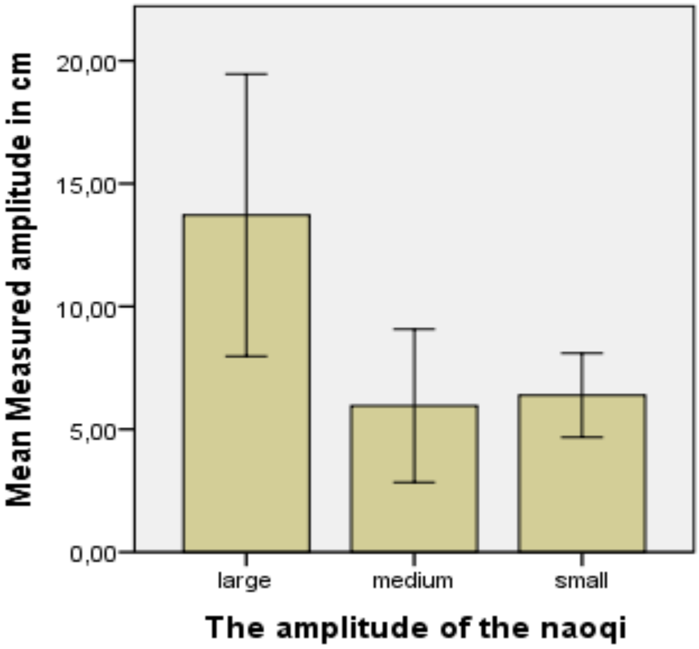


Figure 7: ***Amplitude scores registered by the Kinect on exercise 5 including the 95% confidence interval.***

A repeated measure ANOVA on the mean amplitude of movements for exercise 5 shows a significant effect of avatar's amplitude ($F(2,13)=11,653$ $p=0,002$),but contrast analysis reveals that the significant difference is only between a small amplitude and large amplitude and between a medium amplitude and large amplitude.. The small and medium amplitude are not significantly different ($F(1,1)=0,085$, $p=0,780$). This means that people were not able to distinguish significant differences between small and medium amplitude.

Qualitative data: participant perceived amplitude for exercise 3 and 5.

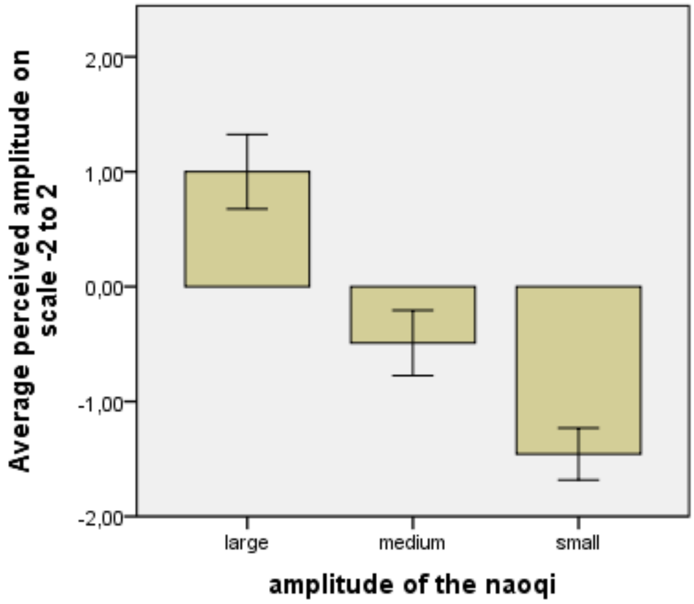


Figure 8*: Perceived amplitude by participants for exercise 3 including the 95% confidence interval:*

Figure 8 shows the perceived amplitudes by participants for exercise 3. The figure shows an increase of the level of amplitude results in higher perceived amplitude. Therefore, as expected, the repeated measurement ANOVA in shows a significant effect for amplitude. (F(2,18)=44,470., p< 0,001).

Contrast analysis reveals a significant difference between all stimulus levels. People were able to correctly perceive small, medium and large avatar's amplitudes. Small and medium amplitude:(F(1,9)=10,449, p = 0,001) Medium and large amplitude:(F(1,9)=35,564, p <0,001).

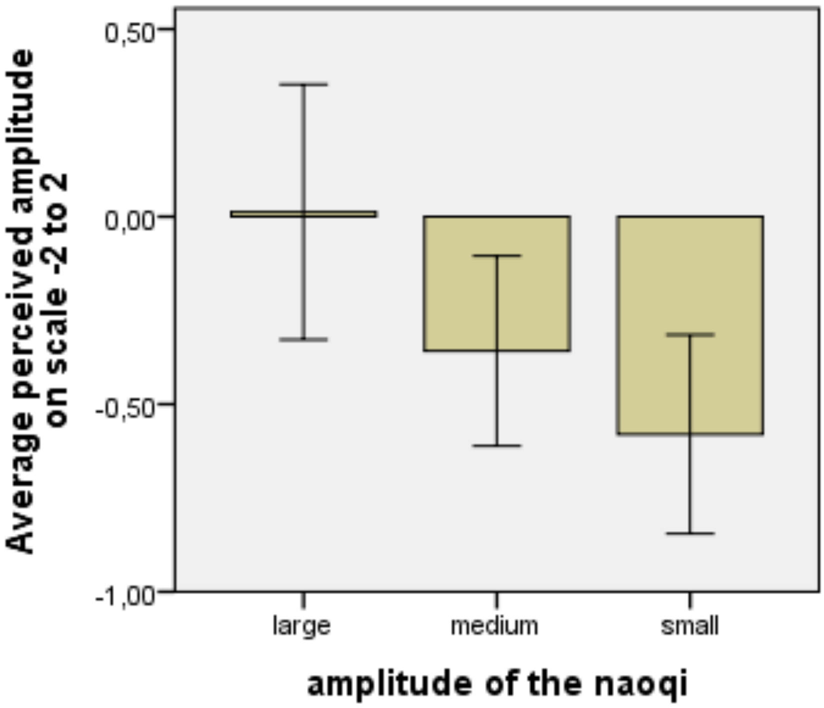


***Figure 9**: Perceived amplitude by participants for exercise 5 including the 95% confidence interval:*

Figure 9 shows the perceived means by participants for exercise 5. The figure shows an increase of the level of amplitude results in higher perceived amplitude. However the difference between each level is not very big. The repeated measurement ANOVA showed a small significant effect. (F(2, 16)= 3,220, p= 0,075) .

As the difference between each level is very small it is not expected to find a significant difference between a small and medium amplitude and neither between a medium and big amplitude. Contrast analysis reveals that the there is no significant between small and medium amplitude and between medium and large amplitude. Small and medium amplitude are not significantly different (F(1,8)=1,532, p=0,251). Medium and large amplitude are not significantly different (F (1,8)= 2,837, p =0,131).

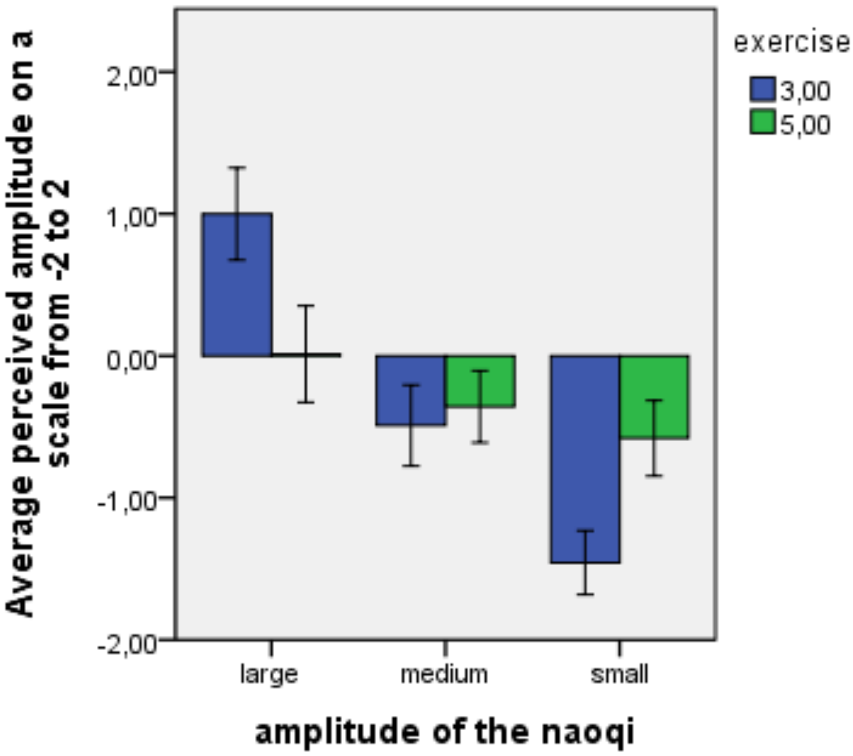


Figure 10: Perceived amplitude by participants for exercise 3 &5 including the 95% confidence interval.

Figure 10 shows the perceived amplitude from both exercise 3 and 5. It shows that people perceived the differences in amplitude 3 as large, but that the differences in amplitude 5 were perceived as small. This is also seen in the other data were no significant difference was found.

Discussion and conclusion.

This study shows how well people can perceive and copy exercises shown by an avatar. This study is closely related to the research project of Ebert[3a] who performed the same experiment but with a real robot. To learn about training with an avatar the preciseness of the execution was measured and the experience of participants was measured. There was measured how well participants executed the exercises. This was done by following their arm movement (amplitude) with a Kinect. Furthermore, the experience was measured by a questionnaire that provides insight in the experience of participants on the exercise; how tiring was it; how big was your arm movement and how fast did they find it. We tried to find the qualitative perception of the avatar's arm movement with respect to speed and amplitude. In this discussion the results are discussed on their own and compared to the findings of training with a real robot. (Ebert, 2012)[3]

The expectations for this research are that cues are found that show that people find it more exhaustive to fitness with a real robot than with an avatar. (See page 3 H1) Furthermore it is expected that people like working with a robot just as much as they like working with an avatar, no difference are expected to be found. (See page 3 H3) Differences in amount of arm movement (amplitude) are likely to be found. In this discussion only the aspect of differences in the amount of arm movement are discussed.

In data about the amplitude experienced and measured in the previous section some interesting results occur. These outcomes help to argue on the hypothesis of the arm movement amplitude of the users is bigger with the real robot than with the avatar of this robot. The outcome is that people cannot divine the amount of amplitude shown in exercise 5, this can be seen in the Kinect data gained and in the survey results. They rate level amplitude 2 higher than level amplitude 1, also the Kinect data shows that participants produced a bigger amplitude at level 2 then 1.However, in exercise 3 they do experience the right level of amplitude at the right moment. This shows that exercise 5 was 'hard to read'. This could be, because of the angle in which the avatar is shown, from the front, whereas exercise 5 is an exercise with the arms moving forward. The pictures below are screenshots of exercise 3 and exercise 5.

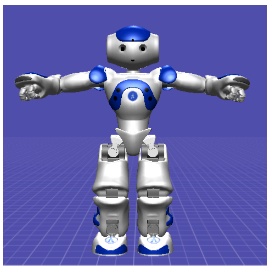


Fig 11 how exercise 3 looked.

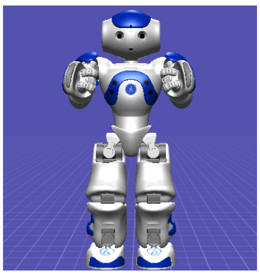


Fig 12 how exercise 5 looked.

A suggestion to solve this could be showing the avatar from the side during exercise 5. As in this figure:

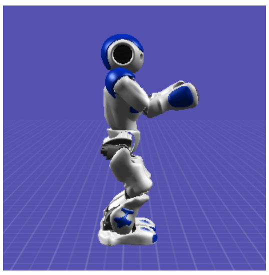


Fig13 A suggestion on how to show the avatar during exercise 5

Recommendations

In the future more research can be done in exercising for elderly. There are still many things unknown. It is still not sure whether participants rate the exercises with the robot as more exhaustive as exercising with the avatar. This can be very important, because this is a small indication that users have benefit from the exercises. Also unknown is if the users rather work with the robot or with the avatar. This can be a huge difference in people's motivation. This could be of use when users have to exercise weekly.

Also some techniques evolve fast. With 3D movies and 3D games, a 3D avatar for fitness is another possible solution for the elderly. The design of the avatars can make a big difference. Perhaps a more human look alike avatar is better for fitness for elderly. Some other improvements for both the robot and the avatar could be better feedback for the users. With different games that react on movements of the users like Kinect, can be used to make sure the users do the fitness exercises correctly.

References

[1] Buttusi, F., Chittaro, L., & Nadalutti, D. (2006, September 12-15) Bringing Mobile Guides and Fitness Activities Together: A solution based on an embodied virtual trainer. *MobleHCI'06*

[2] www.ksera.ieis.tue.nl

[3a] Ebert, N. (2012) Exercising more with a humanoid robot.

[3b] Ebert, N. (2012, September) Improving Heath Exercise Performance With An Embodied Agent Using Verbal Feedback.

[4] Ruttkay and Welbergen (2008) Elbows higher! Performing, observing and correcting exercises by a Virtual Trainer.

[5] Bartneck, Kullé, Croft & Zoghbi. (2008, November) Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots.

[6] *www.thesistools.com*

Appendix 1, Introduction text.

You will first fill in the questionnaire. You can do this on the laptop on the table next to you. After that we will set the body sensor. This is a camera that tracks the movement of your arms and legs. In order to calibrate the equipment you should stand up your arms as the experimenters will show you and maintain that position until the experimenter confirms the end of the calibration phase.
After that the exercises begin.
The avatar of the robot NAO will show you some exercises. You should replicate all the movements of the robot as good as possible. During the experiment you will perform 4 different trials consisting of the same series of health exercises. These exercises involve only arm movements.
A sound beep signalizes the beginning of each trial and at the end there is another sound beep that alerts you to fill in a questionnaire. Thus you have will have about 1 minute to answer the questions proposed for each trial and to bring you back to the initial position.
The end of the last trial is communicated through a dual beep sound. After that please go back to the table and fill in the questionnaire and post-questionnaire.

Appendix 2, Godspeed questionnaire

Please rate your impression of robots on these scales:

1	Artificial Kunstmatig	1	2	3	4	5	Lifelike Levensecht
2	Ignorant Onwetend	1	2	3	4	5	Knowledgeable Veel wetend
3	Mechanical Mechanisch	1	2	3	4	5	Organic Organisch
4	Fake Onecht	1	2	3	4	5	Natural Natuurlijk
5	Dislike Afkeer	1	2	3	4	5	Like Geliefd
6	Machinelike Lijkend op een machine	1	2	3	4	5	Humanlike Lijkend op een mens
7	Awful Afschuwelijk	1	2	3	4	5	Nice Mooi
8	Anxious Angstig	1	2	3	4	5	Relaxed Ontspannen
9	Inert Passief	1	2	3	4	5	Interactive Interactief
10	Unintelligent Onintelligent	1	2	3	4	5	Intelligent Intelligent
11	Quiescent Rustig	1	2	3	4	5	Surprised Verrast
12	Unkind Niet lief	1	2	3	4	5	Kind Lief
13	Unpleasant Onplezierig	1	2	3	4	5	Pleasant Plezierig
14	Unconscious Onbewust	1	2	3	4	5	Conscious Heeft een bewustzijn
15	Moving rigidly Houterige bewegingen	1	2	3	4	5	Moving elegantly Vloeiende bewegingen
	Incompetent	1	2	3	4	5	Competent

16	Onbekwaam						Bekwaam
17	Stagnant Stilstaand	1	2	3	4	5	Lively Levendig
18	Dead Dood	1	2	3	4	5	Alive Levend

Please rate your impression of robots on these scales:

19	Foolish Dwaas	1	2	3	4	5	Sensible Gevoelig
20	Artificial Kunstmatig	1	2	3	4	5	Lifelike Levensecht
21	Irresponsible Onverantwoordelijk	1	2	3	4	5	Responsible Verantwoordelijk
22	Agitated Opgewonden	1	2	3	4	5	Calm Kalm
23	Unfriendly Onvriendelijk	1	2	3	4	5	Friendly Vriendelijk
24	Apathetic Apatisch	1	2	3	4	5	Responsive Responsief

Appendix 3, between exercises questionnaire

During the last trial, the robot moved fast.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

I feel rested after this trial.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

During the last trial, the arm movements of the robot were small.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

I feel tired after this trial.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

The robot was quick during the last trial.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

During the last trial, the arm movements of the robot were big.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

During the last trial, the robot moved slowly.

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

This last trial was fatiguing (vermoeiend).

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

During the last trial, the arms movement of the robot were substantial (aanzienlijk).

Strongly disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

Appendix 4, Consent

PARTICIPANT NUMBER:

INFORMED CONSENT FORM

Title: Avatar Workout.
Experimenter: Bart van Wezel and Loes Wagter
(b.j.p.a.v.wezel@student.tue.nl and l.n.wagter@student.tue.nl)
Supervisor: elena torta (e.torta@tue.nl)

Description
You are invited to participate in an experiment on human-robot interaction. During the experiment the avatar of a humanoid robot will present a series of health exercises. You should replicate them as good as possible. You will be asked to rate the performance on several aspects.

Method
The experiment will consist of replicating health exercises shown by the avatar of a robot. You are asked to rate your impression of the robot before and after the trials.

Confidentiality
For analysis purposes, video recordings will be made during the experiment. Also, data obtained from the questionnaires will be used for this purpose. All data will be processed anonymously.
Your individual privacy will be maintained in all published and written data resulting from the study. The data will only be used for scientific publications and are not retraceable to individual persons.

Duration
The experiment will last approximately 45-60 minutes.

Voluntary participation
Your participation in this experiment is voluntary and you have the right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. You have the right to refuse to answer particular questions.
I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions I have been asked have been answered to my satisfaction. I consent voluntarily to be a participant in this study.

Name:
Signature:
Date: