Hospital robots: an experimental study on communicating urgency using non-semantic speech and built-in LED's

MART SEIP, University of Twente, The Netherlands

1 ABSTRACT

During the next few years, robots might be used in hospitals to take over simple tasks from personnel, such as transporting sample material through corridors from one place to another. In this study we look at how the Harmony robot can communicate urgency in case the robot requires help, for example, because of a corridor block or closed door, using non-semantic sound and LED's. The Harmony robot aims to automate on-demand delivery tasks around the hospital and to automate bio-essay sample flow. This paper describes a 2x2 research which uses a combination of thematic qualitative analysis and quantitative analysis to analyse the responses of a group of 39 participants to different video recordings. Data was collected by using an online questionnaire with open questions to gather opinions of the participants on the different sound and light effects used, the appropriateness of the sound, how important they believe the cargo of the robot to be and the willingness to help the robot. The results are that when the robot is transporting important goods, participants considered it best to communicate using urgent, but not appropriate sounds. Using the LED pattern did have a positive effect on how appropriate participants believed the robot to communicate, but did not have a positive effect for communicating urgency. Further exploration of different communication modules might result in finding a proper balance for communicating urgency appropriately.

Additional Key Words and Phrases: urgency, robots, human-robot interaction, sound, warning, audio visual, hospital, communication, light, non-semantic speech, obstacles

2 INTRODUCTION

This research will be used for the Harmony project [1]. The Harmony project is a project between five universities, a research organization, a multinational company, an SME and two hospitals. The objectives of the Harmony project are twofold, automating on-demand delivery tasks around the hospital and automating bio-essay sample flow.

This research will primarily focus on the objective of automating on-demand delivery tasks around the hospital. Despite being largely invisible to patients, hospital logistics are important in order to give the best patient care [8]. With an increasing elderly population, it is important to have effective and efficient logistics. The Harmony robot that is used in this research is an early prototype of the end product. The exact looks of the Harmony robot prototype can be seen in Figure 3. It is about 1.60 meters tall and has a touch screen on top which can be used to show facial expressions. Next to that, it has wheels to drive around. It cannot speak and does not have any linguistic capabilities, but will have speakers in order to make sounds. Next to that, it has multiple LED's, on both sides of its head (used when making turns), four smaller LED's surrounding it on the bottom and a larger LED in the center of the bottom part of the robot.

When the word 'urgent' is used in this paper, an action that has a high priority and has to be done within as little time as possible is being meant. The higher the urgency, the more important an action is perceived to be and the faster it needs to be executed. This paper also uses the phrase 'non-semantic speech' several times. Non-semantic speech can be defined as a sound that is not part of any human language but does come across as a language. This kind of speech is often used in movies where robots cannot speak, but do communicate, such as R2D2 [2] from Star Wars and WALL-E [3].

2.1 Problem statement

The development of the hospital robot for the Harmony project requires the robot to smoothly move through hospital corridors, such that sample materials can be delivered to the right place in time. To achieve this goal, the robot will use the same corridors as personnel, such as nurses, doctors and cleaners, but also patients and visitors. As can be imagined, a robot is an interesting object for most people, which might cause people to interfere with the robot, stopping it to for example take a selfie or just to have a better look. Especially children seem to be interested in robots and might even hinder them by deliberately blocking their path or kicking it, as was researched by Nomura et al. [7]. Hospital personnel might also accidentally interfere the robot in its job, for example by blocking the corridor with an empty bed or cleaning trolley. In order to prevent this, the robot has to effectively communicate urgency, such that the problem is clear for bystanders and obstacles can be taken out of the way such that the robot can continue its journey.

3 RELATED WORK

There is a lot of literature available on the subjects of human-robot interaction, robots in the workplace, hospital robots, etc. An overview of important related work is given in this section.

3.1 Urgency

Hellier & Edworthy [6] researched how urgent a variety of speech warnings were perceived. In their research they identified three styles of communicating a warning: urgent, non-urgent and monotone. They found out that urgent words are spoken at a higher frequency and louder compared to the other two styles. An interesting observation was that the male voice was considered more urgent than a female voice. These outcomes were used as inspiration for the design of the non-semantic sound in this research.

[1] https://harmony-eu.org/project/about/

[2] https://en.wikipedia.org/wiki/R2-D2

[3] https://en.wikipedia.org/wiki/WALL-E

TScIT 37, July 8, 2022, Enschede, The Netherlands

3.2 Appropriateness

In order to communicate in an appropriate way, some related work can be considered. In a research describing the effects that robots have in the workplace, Dietsch [4] for instance mentions the need for people to get acknowledgement from robots in order to not be perceived as rude. The perception of people on the robot in this research might therefore be influenced by the response the robot has after the person decided to help out the robot. In case the robot for instance does not thank the person in any way, someone might not decide to help again, as the behaviour was possibly seen as rude.

Next to that, Srinivasan & Takayama [10] conducted some experiments in order to find out how robots can be designed to retrieve help in a more polite way. What was found in their research and could be useful in this one is that using positive politeness is most effective in human-robot interaction, that people are more likely to help robots with smaller requests and people would put more effort into helping an autonomous robot rather than one that is believed to be operated by a person.

3.3 Requesting help

By communicating urgency, the Harmony robot attempts to get help from participants. A paper from Daly & Bremmer [3] researched how help can be obtained from bystanders when a robot is stuck. They conclude that people are more likely to help a robot showing emotion, for instance happy or sad, compared to a robot that remained neutral. This insight can be considered for the computer-generated non-semantic sound, by including the emotion that the sound is communicating in the decision on which sound to use.

Second of all, Cha & Fong [2] also researched making robots retrieve help from participants. They have conducted a research on how to make non-humanoid robots using nonverbal signals in order to achieve collaboration. This research was not conducted in a hospital environment and did not necessarily focus on getting people to move (objects) out of the way, but is still useful since it uses light and sound signals in order to request help from bystanders. An interesting observation from this research is that the sound and light signal was perceived differently, they state sound was alerting the bystander, whereas the colour of the light was perceived as how urgently the help was needed.

Third of all, Fischer et al. [5] researched how interactions can best be initiated in order to retrieve help. Their comparison of the response to acoustic signals can be useful for conducting this research. They found out that a simple beep was not effective in order to get someone's attention.

Finally, research on asking for a favour is considered. The results of Saunderson & Nejat [9] in their research on how directness and familiarity influence the outcome when a robot asks for a favour were used for designing the methodology of this study. The results of this study are that a familiar robot is "more persuasive, trustworthy, and people are more willing to help it" compared to an unfamiliar robot.

3.4 Safety

Since this research includes human-robot interactions it is considered important to also take potential safety issues into consideration. Vasic & Billard [11] researched the safety issues that human-robot interactions might cause. In their paper they mention that it would be best to prevent the interaction between humans and robots as much as possible, but if not possible, one of the safety concepts is to avoid obstacles as much as possible. In small hospital corridors, however, this is not always feasible, definitely since the corridors in hospitals will be also used by personnel, patients and visitors.

4 GOALS AND RESEARCH QUESTION

The goals, as described below, have to be achieved using solely non-semantic speech and built-in LED lights. Non-semantic speech will be used because this is the main method the robot will use in order to communicate. The robot will not be equipped with voice recognition or other linguistic capabilities, therefore it was decided to use non-semantic speech. This research will only consider sound and LED patterns, but the robot also has the capabilities to communicate using movements, facial expressions and LED light colour.

Goal 1: Making the Harmony robot communicate urgency in a way that is perceived as urgent by a majority of the participants in the research.

Goal 2: Making the Harmony robot communicate urgency in a way that is perceived as appropriate by a majority of the participants in the research, in order to assure that people who interfere with the robot more often, such as hospital personnel, to remain helping the robot over time.

Based on the prior work on this topic, which can be read in the next section, two hypotheses were created.

Hypothesis 1: It is possible to make the Harmony robot prototype communicate urgency in a way that is effective and perceived as appropriate using non-semantic sound.

Hypothesis 2: The Harmony robot prototype will communicate urgency more effectively and appropriately when using built-in LED patterns next to non-semantic sound.

In order to achieve these goals and to be able to confirm the hypotheses, the following research question is being used.

RQ: How can the Harmony hospital robot effectively and appropriately communicate urgency using non-semantic speech and built-in LED patterns?

5 METHODOLOGY

This section describes the methodology that is being used in order to conduct this research. The research will be conducted through an online survey measuring the variables appropriateness, importance, and willingness to help based on different videos. It is a 2x2 experiment with a beeping or non-semantic sound and a blinking or not blinking light effect. Next to that, it is a between-subjects experiment, since it was expected people would consider responding to five different videos as to time intensive.

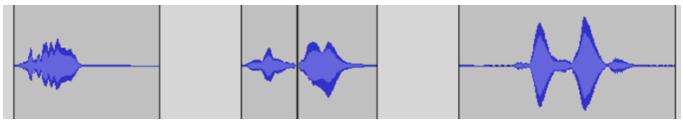

Fig. 1. Graphical representation of the non-semantic sound used

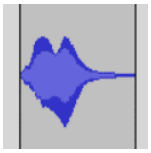

Fig. 2. Graphical representation of the beeping sound used

Fig. 3. Screenshot from one of the videos showing the experimental setup

5.1 Creating sounds

The first step in this research was to select the type of sounds used for this research. A combination of a non-semantic computer-generated sound and a beeping sound was chosen. For computer-generated non-semantic speech, inspiration was retrieved from movies such as Disney Pixar's WALL-E and R2D2 from Star Wars. WALL-E for instance includes a variety of scenes where robots use non-semantic sounds in order to clear the way and communicate urgency. The non-semantic sound is existing out of three different two-tone sounds. To create the non-semantic sound open-source robot sounds were downloaded from freesound.org [4] and edited using the computer program Audacity [5]. A few sounds were selected and then used in a pattern as can be seen in Figure 1.

At first, only one two-tone sound is used, if there is no response, two two-tone sounds are played shortly after one another, of which the first is equal to the sound initially used. If there is still no desired effect, after a short while three two-toned sounds are played. The idea behind this is to make it look like the robot is actually trying to communicate and do a request for help in an increasingly urgent way. In human language it could be considered 'help', then 'please help', followed by 'anyone please help!'. As can be seen when comparing Figure 1 and 2, the beeping sound is a small fragment from the non-semantic sound. The beeping sound will be used in order to identify the difference in perception between both situations. The beeping sound was played at a regular interval, like an alarm.

5.2 LED settings

The second step was to decide on the LED colour and pattern. It was decided to solely use white-colored LED's during this research. For the LED pattern it was decided to use a blinking pattern. The lights will be on if the robot speaks and off when it is silent. The goal is to find out what the difference in perception is between using lights in combination with the sound and solely using sound.

[4]https://freesound.org/people/ScreamStudio/sounds/397253/
[5]https://www.audacityteam.org/

5.3 Creating videos

The third step was to create videos of the robot driving towards an object. As soon as the robot gets to the object, the sound and LED signals will be enabled. As this is a 2x2 experiment, versions with a beeping or non-semantic sound, together with the enabled or disabled light pattern will be used. In total five videos will be used.

- Video **A1**: Beeping sound with no blinking LED's
- Video **A2**: Non-semantic sound with blinking LED's
- Video **B1**: Beeping sound with blinking LED's
- Video **B2**: Non-semantic sound with no blinking LED's
- Video **AB3**: No sound nor LED pattern

Except for these differences, the videos will be identical. Video AB3 is a control video, which can be used to compare the participants' responses to the other videos.

5.3.1 Materials. In order to create the videos a few materials were used. First of all the Harmony robot was used, controlled by a remote controller. Second of all a video camera was used. Third of all a chair was used to block the corridor. The experimental setup can be seen in Figure 3.

5.3.2 Video description. All videos used have a duration of about 28 seconds, with 15 seconds of time driving through the corridor, after which the chair was hit and the sound and light effects were used for 14 seconds. During this time there was 1 whole repetition of the non-semantic sound pattern and 11 repetitions of the beeping sound. It was decided to keep the time identical for both videos, but to use the sounds in the same way they would be actually used in this period. The 14 seconds are divided into two equal parts where the perspective halfway switches from looking at the back of the robot to looking at the front side of the robot. In the video, the speed of the robot was the slowest speed mode that is available on the robot.

5.4 Questionnaire

The fourth step was to create the questionnaire. The questions in the questionnaire are mostly open questions and aim for determining whether the participants think the cargo the robot is carrying is

important or not, based on the used sound and light signals. Participants received questions about their perception of the robot and its communication. Because this is a mainly qualitative research it was decided to use mostly open questions, such that participants are able to explain their reasoning, giving a larger insight into the motivation of the participants. Next to the open questions, one scale question was used. This question allows participants to say how important they believe the robot's cargo is for each video. These results were used to better compare how the same participant would rate different videos and identify the differences between the videos. The questions, which can be seen in Appendix A, were created in such a way that the main variables, as mentioned in the next subsection, could be analyzed. There were four versions of the questionnaire, one for version A and one for version B in both Dutch and English.

Next to a textual explanation, there was an introduction video for all participants to watch before watching the other videos. This introduction video explained the purpose of the robot and showed it moving through corridors. The footage used in the introduction video was different from the footage used in the other videos.

5.4.1 Variables. The questions in the questionnaire are aimed to find out what the opinion of the participant is on the two dependent variables mentioned in the research question, namely **urgency** and **appropriateness**. In order to determine how urgent a participant believes the robots' communication is for each video, three separate questions are being asked. First of all, how important does the participant believe the robot's cargo to be on a 1 to 10 scale? Second of all, whether the participant believes the robot's cargo to be important and why. Third of all, participants are asked whether they would decide to help the robot, in what way and why. Together these three questions answer whether a participant believes the robots' communication is urgent. Two questions are being asked in order to determine how appropriate a participant believes the robots' communication to be. The first question is simply asking for their opinion on the sound itself, whereas the second question asks whether the participant believes the sound to be appropriate in a hospital environment and why (not).

5.5 Participants

In total 39 people participated in this research, of which 20 are female and 19 are male. The participants were divided into two groups. Both groups were shown their own specific videos after watching an introduction video which was equal for both groups. Group A watched videos A1, A2 and AB3 whereas group B watched videos B1, B2 and AB3. After watching each video, questions were answered in a digital questionnaire which can be seen in Appendix A. The participants were being told up front that the robot is a cargo robot, but not what it was carrying.

The participants were recruited in several ways. First of all, a secondary school IT teacher was asked whether the questionnaire could be conducted in the fourth year of both havo [6] and vwo [7]. Pupils could decide themselves whether they wanted to participate or not. One class did survey version A, whereas the other class did version B. Next to that family and friends were asked to complete the questionnaire, a random selection was made between the people filling in version A and version B. By spreading the different versions in an equal matter, it was attempted to gather as many participants for both versions. Eventually, 22 people completed version A, whereas 17 people completed version B. The English version of the questionnaire was filled in 6 times and the Dutch version 33 times.

[6] https://en.wikipedia.org/wiki/Hoger$_{a}$lgemeen$_{v}$oortgezet$_{o}$nderwijs
[7] https://en.wikipedia.org/wiki/Voorbereidend$_{w}$etenschappelijk$_{o}$nderwijs

Table 1. Results on the three main variables measured in absolute numbers and percentages

Video	Appropriateness	Importance	Willingness to help	Total
A1	7 (31,82%)	9 (40,91%)	20 (90,91%)	22
A2	16 (72,72%)	7 (31,82%)	20 (90,91%)	22
B1	13 (76,47%)	8 (47,06%)	15 (88,24%)	17
B2	8 (46,06%)	9 (52,94%)	13 (76,47%)	17
AB3	23 (58,97%)	4 (10,26%)	16 (41,03%)	39

6 RESULTS

In this section, the results of the questionnaire are discussed using both a qualitative and quantitative analysis. Qualitative thematic analysis as described by Braun and Clarke [1] will be conducted. Next to that, a quantitative analysis will be conducted comparing the proportions between different videos of participants who think the robot is appropriate, important and are willing to help.

6.1 Quantitative analysis using two proportions

For conducting the quantitative analysis Table 1 has to be considered. Table 1 shows for each of the five videos the number of participants who think the video was appropriate, important and the willingness of the participant to help the robot. As the survey contained open questions, for this table the answers of all participants have been transformed into yes or no answers. Simply put, the number of times where the answer to the main closed question could be considered a 'yes' on whether the participant believed the robot communicated appropriately, was important and the participant would decide to help the robot were counted and added to a table.

For the quantitative analysis, a two-proportion hypothesis test was used. The null hypothesis in those tests was that both proportions are equal whereas the alternative hypothesis was that both proportions are not equal.

6.2 Appropriateness

First of all, appropriateness will be considered. When comparing videos A1 and A2 the null hypothesis of both proportions being equal is rejected at a 1, 5 and 10% significance level. The proportion of A2 is significantly larger. This means the video with non-semantic sound and blinking LED's is perceived to be more appropriate compared to the video with a beeping sound and no blinking LED's. When testing with the proportions of B1 and B2 the 5 and 10% significance levels are rejected, but at a 1% significance the test was not rejected. This means that B1 seems to be larger, but at a 1%

significance level this is not certain. This shows that the video with beeping sound and blinking LED's is perceived as more appropriate than the video that uses non-semantic sound and no blinking LED's.

When comparing B1 and A2 there is no significant difference, as the null hypothesis could not be rejected at 1, 5 and 10% significance levels. The same is applicable for videos A2 and AB3. Next to that, videos B1 and AB3 could not be rejected at a 1 and 5% significance level but could be rejected at a 10% significance level. Therefore B1 is a little more likely to be perceived as appropriate compared to video AB3, but there seems to be no difference between video A2 and AB3.

It can be concluded that both videos with blinking LED's are perceived to be more appropriate compared to those where the LED's are not blinking. No difference could be identified between using non-semantic and beeping sounds. Next to that, only at a 10% significance level there was a difference measurable between video AB3, without sound and blinking LED's, and video B1, with beeping sound and blinking LED's. At other significance levels, there was no difference.

6.3 Importance

Second of all, the variable importance will be considered. When comparing videos A1 and A2 and B1 and B2 no differences in how participants perceived the robot's cargo to be important could be determined. At a 1, 5 and 10% significance level the null hypothesis could not be rejected. Comparing videos A2 and AB3 however shows that at a 5 and 10% significance level the null hypothesis is rejected. The null hypothesis is also rejected at 1, 5 and 10% when comparing videos B1 and AB3, meaning they are likely not equal.

From these results, it can be concluded that participants did not perceive a difference in how important they believed the robot's cargo to be between the different sounds and LED patterns. All four videos were perceived as important. However, when comparing the video without sound and LED patterns with videos A2 and B1 they are likely not equal. Therefore, participants believe the robot's cargo to be more important when using a sound compared to when not using a sound.

6.4 Willingness to help

Last of all, when looking at the variable 'willingness to help', for videos A1 and A2 and B1 and B2 no difference could be identified. Both were not rejected at a 1, 5 and 10% significance level meaning they are likely equal. When comparing videos A1, A2, B1 and B2 with video AB3 it turns out they are all not equal, as the null hypothesis is rejected at a 1, 5 and 10% significance level. It appears that the participants are more willing to help the robot when it is making a sound compared to when not making a sound.

It can be concluded that there are no differences between using non-semantic or a beeping sound and between using blinking or not blinking LED's. However, participants are far more willing to help the robot when it is making a sound, regardless of which one, compared to when it is not making a sound.

Table 2. The number of participants who mention the sound to be 'annoying' for each video

Video	Total number of mentions	Total number of participants
A1	7	22
A2	5	22
B1	2	17
B2	7	17
AB3	0 (no sound used)	39

6.5 Theme 1: Participants who describe the sound used as annoying

One of the themes that could be identified is participants mentioning they think the sound used is 'annoying'. As Table 2 shows, a total number of 18 out of 39 participants mention in total 21 times that the sound used is annoying in their opinion. A few participants mention this for both sounds used. Table 2 shows the distribution of participants with this opinion on the different videos. It can be seen that slightly more people consider the non-semantic sound to be annoying (videos A2 and B2).

6.5.1 Beeping sound. One of the participants mentions the sound is too loud and aggressive. Next to that, another participant adds that this sound is not appropriate if there are very sick people laying nearby. A different participant thinks the sound is communicating its goal clearly but is still annoying. Multiple participants mentioned that it sounds more like an alarm, which is in their opinion not appropriate in a hospital environment. Also, a few participants explain that they think the sound is okay for a short period of time, but after a while they think it becomes too annoying.

A majority of the participants who mention the sound is annoying also think the sound used is not appropriate for the hospital environment. Next to that all participants except for two also don't think the robot is carrying something important. Despite their opinion of the robot not being all too important, all participants except for one would actually help the robot out, which is interesting because they don't believe it to be important. It seems like their opinion on how annoying the sound is has a larger influence on whether they would help or not rather than their beliefs of the robot being important. One participant mentions that if after removing the object the sound would not stop, she would try to turn off the robot.

6.5.2 Non-semantic sound. The opinions on whether the sound is appropriate differ a lot, with some participants mentioning they believe it to be appropriate, whereas others think it is more annoying and therefore less appropriate compared to the beeping sound. One of the participants mentions it to be less clear than the beeping sound. A different participant mentions that the non-semantic sound makes it more clear that the robot is there and needs attention.

Only three out of the 12 participants who mention the non-semantic sound to be annoying do think it is appropriate for the hospital environment. Seven of the participants who are in this group believe the robot to be important, a few participants mention this is "because it makes a more annoying sound" or that the sound

Table 3. The number of participants who mention to prefer the non-semantic over the beeping sound

Video	Total number of mentions	Total number of participants
A2	13	22
B2	1	17

is more obtrusive and therefore more important. All participants except for two say they would help the robot. One of both participants who would not decide to help the robot mentioned this is because "it is interesting to know what else the robot has to say", whereas the other says he would walk away from the robot.

6.6 Theme 2: Perceptions on non-linguistic communication

As Table 3 shows, in total a number of 14 participants mention that they prefer the non-semantic sound over the beeping sound. A large majority of 13 participants who mention this watched videos A1 (not blinking + beeping) and A2 (blinking + non-semantic).

The one participant who preferred the non-semantic sound in video B2 believes it sounds like it needs more help compared to the first video. He explains this is mostly because it makes a "more sharp" noise. He also believes the sound to be appropriate, important and would decide to help the robot.

The group of participants who preferred the non-semantic sound in video A2 over the beeping sound in video A1 mention that the sound is "more clear", "better as it looks less like an alarm" and believes people to "watch faster because of the different tones". A different participant says that "the sound is better as it is not too loud that will annoy people" whereas another participant says the sound is "cute" and "far less annoying". This same participant also thinks it "almost sounds like an animal and not as an alarm clock". On the contrary, this participant also mentions the sound not to be urgent and "therefore the cargo must not be too important", however this person would decide to help the robot and remove the chair.

Another participant mentions that the sound has "more emotion, which is handy". A few participants, including the participant I mentioned before, believe the sound to be very appropriate in a hospital setting as it cannot be confused with other sounds. A different participant tells it catches more attention whilst being less annoying. In general, multiple participants mention the robot sounds more confused.

Out of this group of 13, only 5 participants think the robot's cargo is important, however all participants say they would decide to help the robot. When comparing the rating the participants give to how important they think the robot is, 5 participants give the non-semantic sound a higher rating compared to the beeping sound. For 2 participants the rating remains the same and for 6 participants the non-semantic sound is less important than the beeping sound, despite the fact they do mention they prefer it over the beeping sound.

6.7 Other observations

This research focuses on just two variables, sound and light effect. For both variables two options are there, making this a 2x2 experiment. However, multiple things were noticed by participants that were not researched in this experiment.

First of all, multiple participants mentioned that "the robot must not be important as in that case it wouldn't be so slow". It seems like the speed of the robot matters for the perception of amongst others the importance of the cargo.

Second of all, the loudness of the sound was mentioned multiple times. Some of the participants thought it was too loud or too soft, which can be explained by the fact that this was online research and therefore the loudness of the sound was depending on the computer used. The two sounds used in this research were exactly as loud. From these remarks it can be said that the loudness of the sound is likely to affect the results.

Third of all, the frequency of the sound might play a role. For the beeping sound this frequency did not change, making some participants think of it as an alarm clock noise. The intervals between the non-semantic sounds however did deviate as they would with regular speech. This interval might play a role in the perception of the robots' communication.

Last of all, some of the participants mentioned that their opinion on the sound depends on how long the sound would play. It can be imagined that in some results the sound would be over soon as the robot is helped out quickly, but in other cases it might take more time before the robot gets help.

7 DISCUSSION

7.1 Results

The described related work has given interesting insights which helped to understand the results. For instance the work by Srinivasan & Takayama [10] which highlights that autonomous robots are more likely to receive help might have played a role in this research, as it appeared the robot was an autonomous robot. Also, the research from Daly & Bremmer [3] who conclude that emotions play a role when people decide on helping robots was applicable, as multiple participants mentioned the emotions expressed by the robot.

What follows from the results is that using a combination of non-semantic sound and built-in LED patterns does not work for communicating urgency both effectively and appropriately. This implies if only using these communication modules in the Harmony robot, a decision has to be made between the desire to communicate urgency effectively or appropriately. A different option is to further research a different set of communication methods, such as the use of facial expressions or movements. The results cannot tell whether the main underlying goal of participants helping the robot by moving objects or letting it pass would be achieved in an actual hospital environment. Further study is recommended keeping into account the limitations of this research and including the use of alternative communication methods.

7.2 Limitations

7.2.1 Participants. It is important to mention that the group of participants in this research were mostly in a specific age group. A majority of the participants were between 16 and 21, with a few outliers of participants with the age of 23, 25, 27, 48 and 49. In general, this group is most likely not representative of a regular hospital population, not personnel or patients. In total 39 participants took part in this research. It has to be noted this is a limited number of subjects and more information could be retrieved when using a larger group of participants. Due to the fact this was online research, there was no way to control the loudness of the sound the participants used. Some of the participants described the sound to be too hard or too soft, but this will depend a lot on the loudness of the speakers of the computer that was used to play the videos. Findings could also differ across different robotic platforms.

7.2.2 Harmony robot prototype. The colour of the LED's used for the robot might have an influence on the response and perceived urgency. This factor was however not used in this research, as the colour of the light did not change. Next to that, the robot used in this research, the Harmony robot, is a very specific and not yet widely used robot. Therefore it is likely that with a different robot, the results would differ as well. It is also important to consider that for this research an early prototype of the Harmony robot was used and not yet the final version. As the looks of the robot will likely change in future versions or features such as a robotic arm will be added, these changes would likely influence the outcome.

7.2.3 Location. For this research, it was not possible to move the robot outside of the university building in order to conduct an experiment at an actual hospital. Therefore all videos were recorded at the university. Despite the fact that a majority of the participants did believe the robot was in an actual hospital, there were participants who could see or knew this was not the case. This might have somewhat influenced the results.

7.2.4 Physical effort. It is important to consider that participants in this research only had to answer questions in a digital questionnaire. For instance regarding helping the robot it is assumed the participants gave a realistic answer on what they would do in this situation. However when participants mention they would help the robot, this can not be proven as no actual physical effort had to be done. It has to be taken into consideration there might be a difference between what participants say they would do and would actually do.

7.2.5 Technical limitations. A few technical limitations play a role as well. With the current version of the prototype it was not possible to control the smaller LED's surrounding the robot in combination with the larger LED which is placed in the center of the bottom of the robot. Despite the fact that a difference can be noticed when turning the smaller LED's on and off, the effect would be larger and better visible in case the larger LED would be blinking at the same rate. This was also the reason why white LED's were used, as a different colour of the smaller LED's were not visible on the video because of the larger LED being more bright. The current version of the robot does not include speakers, so all sounds used were added to the video afterwards by using video editing software.

8 CONCLUSION

A few important insights can be gathered from this research. First of all, from the quantitative analysis it can be said that the robot was perceived to be communicating more appropriately when using the blinking LED pattern, but no difference was measured between using a non-semantic or beeping sound. Participants however believed the robot communicated appropriately as well when it was not using sound or LED patterns.

Second of all, participants believed the robot's cargo to be far more important when using a sound compared to when not using a sound, but no difference could be identified between using non-semantic or beeping sound. There also appeared to be no difference between blinking and not blinking LED's.

Third of all, participants appear to be willing to help the robot as long as it makes a sound, regardless of whether it is non-semantic or beeping, compared to when the robot does not make a sound.

In general, from the qualitative data it can be said that the opinions on the sound used to differ a lot between participants. More participants thought the non-semantic sound was annoying compared to the beeping sound. There seems to be a connection between how annoying people think the sound is and how much they want to help it. There is consensus amongst participants that annoying sounds are not appropriate for a hospital setting, however annoying sounds do seem to be interpreted as more urgent.

It can be said that people who prefer the non-semantic sound over the beeping sound think the non-semantic sound shows more emotions and is better because it looks less like an alarm. From these results it seems like it is difficult to communicate in an appropriate manner whilst also communicating urgency. Depending on the cargo of the robot, a decision has to be made whether a sound that is urgent, but can come across as obtrusive or a polite sound which does not communicate the urgency of the robot as well has to be used. An important conclusion is that regardless of using an urgent or appropriate sound, a large majority of the participants in this research would decide to help the robot.

When considering all conclusions, it can be said that Goal 1 and Goal 2 are met apart from each other, but not together. Hypothesis 1 turned out to be false, as it was possible to communicate urgency effectively or appropriately, but not both. Hypothesis 2 turned out to be partly true, as using a blinking light pattern did increase the perceived appropriateness, but was not more effective. These conclusions also answer the main research question, showing that with the Harmony robot prototype, it was not possible to communicate urgency effectively and appropriately using the tried combination of non-semantic sound and LED pattern. It has to be taken into account that these results might not be generalizable. They might vary depending on the robotic platform, sounds or led patterns used.

9 ACKNOWLEDGEMENTS

I would like to thank my supervisor Hideki Garcia Goo for her great help, feedback and suggestions. I also want to thank Mariët Theune for organizing weekly track meetings, giving us a lot of tips and sharing her experience. Last but not least I want to thank everyone who participated in this research, as without them it wouldn't have been possible to conduct it.

REFERENCES

[1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[2] Elizabeth Cha, Maja Mataric, and Terrence Fong. 2016. Nonverbal signaling for non-humanoid robots during human-robot collaboration. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2016). https://doi.org/10.1109/hri.2016.7451876

[3] Joseph Daly, Ute Leonards, and Paul Bremner. 2020. Robots in Need. *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (2020). https://doi.org/10.1145/3371382.3378301

[4] Jeanne Dietsch. 2010. People Meeting Robots in the Workplace [Industrial Activities. *IEEE Robotics Automation Magazine* 17, 2 (2010), 15–16. https://doi.org/10.1109/mra.2010.936950

[5] K. Fischer, B. Soto, C. Pantofaru, and L. Takayama. 2014. Initiating interactions in order to get help: Effects of social framing on people's responses to robots' requests for assistance. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (2014). https://doi.org/10.1109/roman.2014.6926383

[6] Elizabeth Hellier, Judy Edworthy, Ben Weedon, Kathryn Walters, and Austin Adams. 2002. The Perceived Urgency of Speech Warnings: Semantics versus Acoustics. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 44, 1 (2002), 1–17. https://doi.org/10.1518/0018720024494810

[7] Tatsuya Nomura, Takayuki Kanda, Hiroyoshi Kidokoro, Yoshitaka Suehiro, and Sachie Yamada. 2016. Why do children abuse robots? *Interaction Studies* 17, 3 (2016), 347–369. https://doi.org/10.1075/is.17.3.02nom

[8] Ali Gurcan Ozkil, Zhun Fan, Steen Dawids, Henrik Aanes, Jens Klestrup Kristensen, and Kim Hardam Christensen. 2009. Service robots for hospitals: A case study of transportation tasks in a hospital. *2009 IEEE International Conference on Automation and Logistics* (2009). https://doi.org/10.1109/ical.2009.5262912

[9] Shane Saunderson and Goldie Nejat. 2021. Robots Asking for Favors: The Effects of Directness and Familiarity on Persuasive HRI. *IEEE Robotics and Automation Letters* 6, 2 (2021), 1793–1800. https://doi.org/10.1109/lra.2021.3060369

[10] Vasant Srinivasan and Leila Takayama. 2016. Help Me Please. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016). https://doi.org/10.1145/2858036.2858217

[11] Milos Vasic and Aude Billard. 2013. Safety issues in human-robot interactions. *2013 IEEE International Conference on Robotics and Automation* (2013). https://doi.org/10.1109/icra.2013.6630576

A APPENDIX: QUESTIONNAIRE

A.1 Section 1: Consent

In this study a robot that will transport objects from one place to another inside hospitals is considered. The study is meant to discover how the behavior of a hospital cargo robot is perceived. When choosing to voluntarily participate in this research, you will be shown three short videos and an introduction video and are asked to answer some questions regarding these videos. This will in total take approximately 10 to 15 minutes.

This research has been reviewed by the Ethics Committee Information and Computer Science. All answers you give will be gathered anonymously, please beware that it is therefore not possible to withdraw after participation, since it is unknown to us which data belongs to you. No personally identifiable information will be collected.

The data provided will be stored for the duration of this research and will not be used in future studies. In case you want to contact the researcher after you took part in this research, please send an email to h.m.c.seip@student.utwente.nl.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information Computer Science: ethicscommittee-CIS@utwente.nl.

A.1.1 Question 1. I have read and understood the study information or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction. I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason. I understand that taking part in the study involves digitally answering questions.

A.1.2 Question 2. I understand that information I provide will be used for publication, I agree that my information can be quoted in research outputs and I agree to joint copyright of the responses to Mart Seip.

A.2 Section 2: Demographics

A.2.1 Question 1. What is your age?

A.2.2 Question 2. What is your gender?

A.3 Section 3: Video A1 or B1

A.3.1 Question 1. What do you think the robot tried to communicate?

A.3.2 Question 2. How did the robot communicate this?

A.3.3 Question 3. What is your opinion on the sound the robot is making?

A.3.4 Question 4. How important do you think the robots cargo is on a scale from 1 to 10?

A.3.5 Question 5. Why do you think the robots cargo is important or not? Please explain.

A.3.6 Question 6. What would you do if you encountered the robot in this situation?

A.3.7 Question 7. What is your opinion on the appropriateness of the sound in this environment? Please explain.

A.4 Section 4: Video A2 or B2

A.4.1 Question 1. What difference(s) did you notice compared to the video before?

A.4.2 Question 2. What is your opinion on the sound the robot is making?

A.4.3 Question 3. How important do you think the robots cargo is on a scale from 1 to 10?

A.4.4 Question 4. Why do you think the robots cargo is important or not? Please explain.

A.4.5 Question 5. What would you do if you encountered the robot in this situation?

A.4.6 Question 6. What is your opinion on the appropriateness of the sound in this environment? Please explain.

A.5 Section 5: Video AB3

A.5.1 Question 1. What difference(s) did you notice compared to the video before?

A.5.2 Question 2. How important do you think the robots cargo is on a scale from 1 to 10?

A.5.3 Question 3. Why do you think the robots cargo is important or not? Please explain.

A.5.4 Question 4. What would you do if you encountered the robot in this situation?

A.5.5 Question 5. What is your opinion on the appropriateness of the sound in this environment? Please explain.

A.5.6 Question 6. Where do you think all three of these videos were recorded?