Universiteit Leiden

Master Computer Science

A study into human interaction with dysfunctional social behaviour in robots.

Name: Taz Smit
Student ID: s1936875

Date: 30/08/2023

Specialisation: Artificial Intelligence

1st supervisor: Joost Broekens
2nd supervisor: Peter van der Putten

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Abstract

This paper focuses on human-robot-interaction, specifically with a robot showing socially dysfunctional behaviour. This research takes a first step in that direction by designing and evaluating robot behaviour and setting up a pilot experiment that can test the interaction of humans with such a robot. Interactions like this can help broaden our understanding of HRI, yet this part of HRI is critically under-researched. This paper aims to start closing the gap of the interaction spectrum where the robot is the one displaying dysfunctional behaviour.

Contents

1 Introduction

When it comes to human-human interaction, not all interactions are always positive or even constructive. This is part of human behaviour. Some people do not get along. A person can have a bad day. Some people have a disorder that makes it harder to socialise [1]. These different angles of the spectrum of human interaction have been well documented and researched [2] [3].

When it comes to human-robot-interaction (HRI) however, only one side of the spectrum is ever considered. Research in robot behaviour is done mostly to make robots as social and helpful as possible [4]. While there exists research on humans having trouble working with robots [5], to understand HRI when the robot is displaying dysfunctional behaviour. Considering the fact that a lot has been learned about human interaction by looking at dysfunctional behaviour in humans, one could consider the idea that the same could be said about HRI.

This thesis aims to start closing that gap in HRI studies by working towards an experiment that could be used to test a human's reaction to a robot that displays several types of dysfunctional social behaviour, like aggression and anxiety. In this thesis it is investigated what dysfunctional behaviour is best suited for modelling with a robot by looking at DSM-V criteria, then these behaviours are designed and tested. Based on this pilot it was decided that the behaviours are too difficult to express clearly due to basic interaction flaws such as the flow of the interaction. Finally, a second pilot was performed with a more basic approach where the robot is modeled as a "helpless" robot needing navigation help from users.

1.1 Motivation and Related Work

The idea to build a robot with multiple dysfunctional behaviours originated from an idea to make an antisocial robot. As stated in the previous section there is a lot of research done in HRI when it comes to robots learning to interact with humans in a positive way, so the idea here was to make a robot that flipped that scenario on its head. Little research has been done on robots with dysfunctional behaviour. Considering the fact that the spectrum of interaction is not limited to positive interactions where everyone wants to work together, this is critically under-researched.

Multiple studies have been done on robots in public needing help. One study used small cardboard robots called Tweenbots [6], which simply moved forward in a straight line. The robots had a little flag attached to them which told any passerby where the robot wanted to go. The Tweenbots relied on humans to turn them, help them over or around obstacles and keep them safe from traffic. None of the Tweenbots were ever damaged or lost and when they got stuck they would always be helped back on track by someone. Another study on this phenomenon is the case of hitchBOT [7]. As their name suggests, hitchBOT was a robot that would hitchhike around to get to their destination. This project took the concept of whether or not humans feel safe with robots and turned it on its head, instead asking whether or not robots are safe with humans. Contrary to the Tweenbots experiment however, hitchBOT was not as safe. While hitchBOT had no trouble hitchhiking through Canada, Germany and The Netherlands, when hitchBOT came to America they were vandalised.

Londoño et al. describe how robots can learn to work together with humans by learning which actions not to take and then making sure to avoid those actions [8]. This study helps show the importance of researching opposite interactions to those preferable, as these interactions can help learn robots later on how to behave more positively. Another study by Rosenthal and Veloso focuses on robots asking for help from humans when they cannot complete a task [9]. These robots are programmed to first wait around to see if there is anyone nearby to help and if not to go out and find someone willing to help them. When it comes to robots asking for help multiple studies have been done on how exactly the robots should ask for help to make sure they will actually receive it. This includes trying out different phrases to see which is most likely to get a response [10] and using nonverbal methods like lights and sounds [11]. Though all these studies considered some variation of annoyance as a variable in their studies, it was not the focus and was usually meant to be avoided as much as possible. This research aims to do the exact opposite.

There are studies done that look at the way robots are perceived by humans [12], which can help with the creation of a personality as well as give insight in what type of experiment works well to test such a personality. Different studies found that humans can detect personalities in robots [13] [14]. A research done by Salem et al. showed a difference between a human's perception of a robot and their willingness to cooperate with that robot [15]. The study involved a robot that displayed either correct or faulty behaviour. The human participants were asked by the robot to do different tasks, some normal and some unusual. The study found that participants scored the robot as more trustworthy and reliable, and anthropomorphised the robot more when the robot performed correct behaviour, while a faulty robot was anthropomorphised less. While subjective perception of the robot changed between correct and faulty behaviour, the participants' willingness to perform tasks did not. This suggests that the personality of a robot can indeed be tested with an experiment involving a simple task, as humans are likely to follow the robot's instructions as long as the task is not an unusual one.

2 Research Questions and Approach

2.1 Research Questions

This thesis revolves around the following research questions:

1. How to design dysfunctional social behaviour for robots?
2. How to study the influence of dysfunctional social behaviour on the human during the interaction and how to measure the human's experience of the interaction?

These questions revolve around two sequential parts. The first part is about implementing dysfunctional social behaviours. This part of the thesis involves research into what makes certain behaviours recognisable and how to implement that into a robot. The second part involves designing a test that will allow these personalities to be displayed and will make sure the human participants interact with the personalities to be able to measure their experience.

While research has been done on how to make robots display emotion and how humans interact with robots like that [4] [16], very little research is done specifically on observing humans as they interact with a robot that displays socially unacceptable behaviours. While this thesis will not be able to fully answer this problem, it proposes ways to implement dysfunctional behaviours in robots and proposes how to set up an experiment that will allow these behaviours to come to light and will allow humans to interact with these behaviours.

2.2 Overall Approach

The main research questions can be broken up into the following subquestions:

1. Which disorder types have externally portrayed dysfunctional social behaviours?
2. Which of these behaviours are the most promising candidates to be modelled in a robot?
3. What type of interaction or task can be used to make humans interact with these behaviours?
4. How can the chosen behaviours be modelled in a robot in such a way that they are displayed during that interaction?

To answer subquestion 1 the DSM-V [1] is used as a basis. It categorises disorders by type and provides symptoms, which can be used as a base. First, symptoms should be selected based on whether or not they manifest externally. Internal symptoms are hard to perceive and will therefore not be useful in this research.

Once a selection has been made, the chosen behaviours should further be selected based on whether or not they can be implemented in a robot. This will also answer subquestion 2.

The assumption for question 3 is that the task needs to be cooperative and the human needs to help the robot. A task that can be done without the robot's interference would make it hard for the intended personality to come across. An interaction that would force the human to work with the robot would be a task where the robot has to accomplish something it cannot do alone and needs the human's help for. This type of task requires active cooperation between the robot and the human, which will allow the human plenty of time to interact with the personality.

For subquestion 4 the following design perspective is taken. First a robot is created with a base personality that simply performs the chosen task without any social dysfunction. This way all personalities stem from the same base, making them easier to test. It also makes sure the robot is able to perform the chosen task.

This thesis is split up into three parts. The first part reports on different types of behaviours that can be modelled on a robot. The second part reflects on the design of a robot with these dysfunctional behaviours. The third part reports on a suitable task and study setup. Section 3 explains how the different behaviours were chosen and grouped together to form two distinct personalities. Section 4 shows the first approach to implementing these behaviours and the resulting problems. Section 5 goes into detail about how the pilot experiment was formed and conducted.

3 Identifying Dysfunctional Behaviours based on DSM-V

To be able to model behaviours on the robot, first suitable behaviours need to be chosen. As explained in section 2 these behaviours were found using the DSM-V. In the DSM-V disorders are categorised by type, so the first step is to choose appropriate disorder types that have externally displayed behaviours.

Disorder Type	**Disorders**	**External Behaviours**
Neuro-developmental Disorders	Communication Disorders	Reduced vocabulary
		Stuttering
	Autism Spectrum Disorder	Stimming
		Nonverbal communication
	Hyperactivity Disorder	Hyperactivity
	Motor Disorders	Insufficient motor skills
		Tics
Depressive Disorders	Disruptive Mood Dysregulation Disorder	Outbursts
	Major Depressive Disorder	Depressed mood
Anxiety Disorders	Separation Anxiety Disorder	Distress when person leaves
	Selective Mutism	Not speaking when expected to
	Social Anxiety Disorder	Avoiding social situations
	Panic Disorder	Panicked reactions
Disruptive, Impulse-Control and Conduct Disorders	Oppositional Defiant Disorder	Angry
		Argumentative
	Intermittent Explosive Disorder	Aggressive Destructive
Personality Disorders	Histrionic Personality Disorder	Attention seeking
	Narcissistic Personality Disorder	Arrogant

Table 1: Different disorder types as specified in the DSM-V, split up in their individual disorders and external behaviours

Looking through the different types, there are a few that immediately jump out as inappropriate for modelling on a social robot. These types include:

- Somatic Symptom and Related Disorders
- Feeding and Eating Disorders
- Elimination Disorders

- Sleep-Wake Disorders
- Sexual Dysfunctions
- Gender Dysphoria
- Substance-Related and Addictive Disorders
- Paraphilic Disorders

All these disorder types have to do with something a social is physically incapable of doing and are therefore not useful for this research.

Next let's look at the different disorder types that are left and zoom in on their different disorders and behaviours. Table 1 shows different disorders and their external behaviours categorised by type. Some types and specific disorders have been left out because they do not have clear external behaviours. The left out types include Schizophrenia Spectrum and Other Psychotic Disorders, Bipolar and Related Disorders, Obsessive-Compulsive and Related Disorders, Trauma- and Stressor-Related Disorders, Dissociative Disorders and Neurocognitive Disorders. Only the disorders with easily identifiable external behaviours are shown per type.

When looking at these behaviours a pattern appears to emerge. There are behaviours that have the person act and react strongly, aggressive or even destructive. Another set of behaviours appears much more timid, behaviours including anxiety, stuttering and depressed mood. With this in mind the behaviours seem to mostly fall into two types: an angry, aggressive and more extraverted type and a scared, hard to talk to and more introverted type, see table 2. These two types of robots are chosen because they both have a distinct set of dysfunctional behaviours which are hypothesised to provoke different reactions from the humans they will interact with.

Negative Approaching Type	**Negative Avoiding Type**
Argumentative	Stuttering
Arrogant	Avoiding
Aggressive	Panicked
Angry	Scared
Hyperactivity	Stimming
Outbursts	
Attention Seeking	

Table 2: The two types of dysfunctional robots with their behaviours

4 Modelling and Piloting Dysfunctional Behaviour

4.1 Robot Hardware and Software Platform

To help make the interactions more believable a humanoid robot was chosen to display the different behaviours. This robot is called Pepper (see figure 1), a humanoid robot with audio, visual and touch sensors, as well as a touch screen made by Aldebaran, previously known as SoftBank Robotics Europe [17]. The sensors can be accessed to gather input and allow the robot to interact with humans. Pepper has wheels instead of legs, which allow them to move around quickly and stable.

To interface with Pepper, a portal called RobotInDeKlas was used [18]. This portal allows the user to write external code which can be uploaded to Pepper remotely, allowing for much faster programming.

Figure 1: A picture of Pepper, taken from the Aldebaran website (credits: `https://www.aldebaran.com/en/pepper`)

4.2 Piloting Dysfunctional Behaviour

In order to test how humans react to robots with dysfunctional personalities, first the personalities had to be modelled. As explained in the previous section, the two personalities were built off of a base personality that could do a simple task. This task is further explained in section 5. The first iteration of this base personality was built on a dialogue system, where Pepper would wait for a verbal command to start the task, move, stop and finish. Figure 2 shows a state diagram of the interaction. States would change either based on a timer or on a voice command, usually in the form of words picked up from what participants would naturally say. For example, if Pepper were standing still (the Pause state) a participant might say "Come on Pepper!" or "Move Pepper!", causing Pepper to switch to the TryMove state and check if it was safe to move. The two dysfunctional personalities both used this base as a starting point. Each had their own version of the different interactions however, as can be seen in table 3.

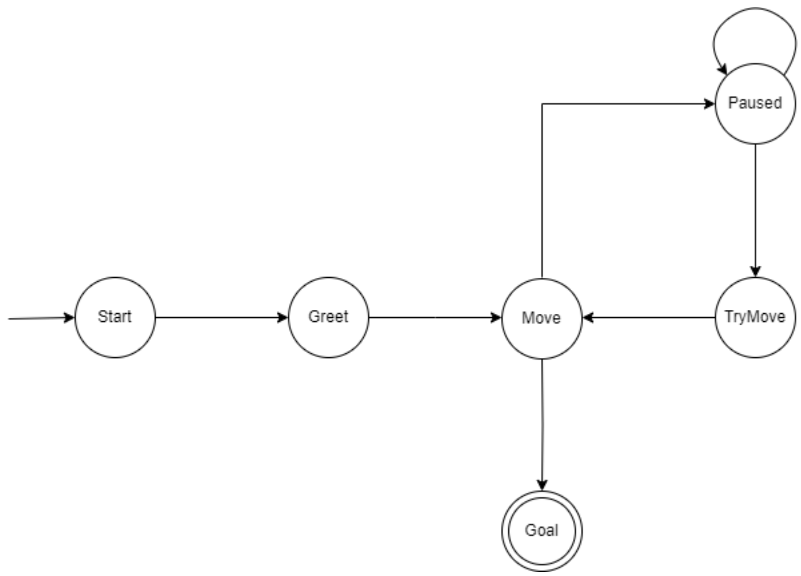

Figure 2: The state diagram for Pepper's original interaction. The different personalities had different reactions to the participant's comments and would display their own actions during the Paused state according to their personality.

4.3 Negative Avoiding Personality

The Negative Avoiding Type had a greeting that had them hesitate, accidentally ask the question too loud and startle away from their own sudden boldness. When helped, they would thank the participant in disbelief. This type would refuse to get close to the participant and would move very slowly. When idle, this personality had Pepper clench and un-clench their hands, look away from the participant or look around the room distracted. When trying to move this personality would always ask if it was alright or if it was safe to move.

4.4 Negative Approaching Personality

The Negative Approaching Type started off demanding to be helped, ordering rather than asking the participant to do the task at hand. This type would get right in the participants face, moving fast and decisive and looking the participant in the eyes the whole time. This personality would idle by crossing their arms or asking what was taking so long and would announce rather than ask that they were going to move again.

	Avoiding	**Approaching**
Start	"Uh hi? I uhh.. I need help! There are these things in the way and I have to get out of the room. Could you please help me?"	"What? Oh. Well, there is stuff in the way and I kinda wanna leave."
Greet	"Really? Thank you!"	"Thanks."
Move	*Pepper moves until an obstacle is detected*	
Paused	*Pepper clenches and unclenches hands* *Pepper slowly looks around* *Pepper looks at participant and quickly looks away*	*Pepper clenches and unclenches hands* *Pepper looks around*
TryMove	"May I move?" "May I?" "Is it safe?"	"Can I move yet?" "Why is this taking so long?" "I'm gonna go now."
Goal	"Thank you for helping me!"	"Finally."

Table 3: The different responses and actions of the two personalities per state (see figure 2)

4.5 Technical Implementations

As stated before, to program Pepper the portal RobotInDeKlas was used. This portal allowed easy access to Pepper's different functionalities. The programming language Python was used to program Pepper. To model the two behaviour types on Pepper first a base program was established. This program used Pepper's built-in facial recognition features to find a face and turn towards it. Pepper would then walk towards the human they found. This feature was established in the code by having Pepper find a face and then making Pepper turn until that face was in the center of their vision. To be able to look a human in the eye Pepper was made to move their head up or down, again until the found face was in the center of their vision. Another one of Pepper's built-in features is object detection, which will automatically stop Pepper if they are about to bump into something. This was taken advantage of by continuously giving Pepper a command to walk (or rather, drive) forward. Pepper could only execute that command if there was nothing in the way, so this allowed Pepper to move only when it was safe. These two features combined became the base of the experiment.

With the base completed, the two personalities could be modelled according to figure 2 and table 3. Every time Pepper stopped, instead of instantly giving a move command, Pepper would enter the Paused state for a while and there was a chance a random "idle animation" would play. These idle animations consisted of clenching hands, looking around the room or saying a sentence in line with the displayed personality, see the previous two sections. After a couple seconds Pepper would try to move again, which they would announce in a way consistent with the displayed personality. If they could move, they would, otherwise they would go back to the Paused state and try again in a couple of seconds. Whether or not Pepper could move was determined by if their wheels were rotating. Because Pepper could only move if it was safe, if their wheels stayed still it meant there was something in the way.

4.6 Prototyping

After finishing a prototype of both personalities, they were tested on two individuals. From this small pilot it became clear that there were a lot of problems when trying to model the personalities. First of all, the dialogue based interaction made it so interactions were slow, because Pepper was waiting for voice commands too often. Pepper could not always hear what the other person was saying and only certain phrases were able to be picked up. Pepper also tended to wait too long to try and move again, while the idle animations barely showed. This created an interaction where the participant kept waiting for Pepper to finally do something, which took them out of the experience.

4.7 Reflection and Redesign

After the first pilot it became clear that to be able to test the personalities, first the interaction itself had to be tested. Because the interaction was not smooth, the participant kept being focused on the flaws in the interaction itself and therefore did not pay any attention to the displayed personality. The personalities themselves also had a lot of moving parts that did not always work together well or clashed in unforeseen ways. This led to Pepper displaying undesired behaviour, which further derailed the experiment.

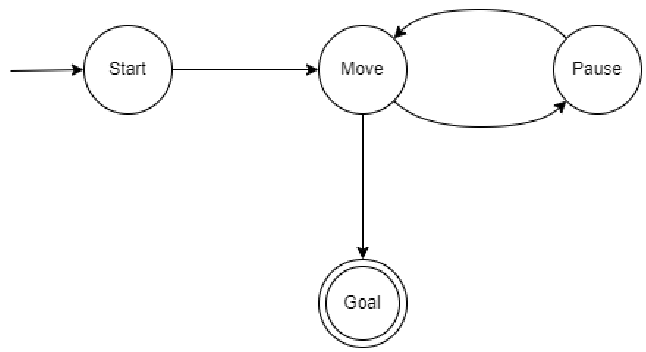

Figure 3: The state diagram for Pepper's base personality that was used in the experiment.

The pilot showed the importance of designing a proper interaction before adding personalities. Because of this, the focus was shifted to designing an interaction between Pepper and a human that went smoothly and without delays from interactions like voice commands. This new base personality was to be used in the actual experiment to see if the above requirements were met.

Figure 3 shows the state diagram of the new base personality. In contrast to the one in figure 2, this personality simply greeted the participant and then instantly started moving, only pausing when there was an obstacle in the way. Pepper would sometimes comment on the fact they were standing still for too long, saying “I can't move” or “Something is in the way”. To further simplify the interaction, some aspects of the interaction were controlled by the researcher. The interaction started and ended with the press of a button outside of the participant's vision. The researcher could also let Pepper say that they lost track of the participant's face, in case the participant stayed out of Pepper's vision for too long.

5 A Pilot Interaction Experiment with a Helpless Robot

5.1 Experimental Protocol

The experiment was a small pilot to test if humans would properly interact with Pepper and if this setup allowed the personalities to properly come across. The chosen setup placed Pepper into a room filled with obstacles that they could not get past on their own. A human participant was asked to enter the room and help Pepper reach the door by removing the obstacles and getting Pepper to move towards the entrance of the room. This setup was used to force the participants to interact with Pepper.

The experiment was done with the base personality to test if this type of setup was practical. The base personality would greet the participant, tell them they were stuck and could not leave the room without the participant's help and then ask the participant to please move the objects and guide them towards the door.

5.2 Measures

The experiment was measured using a survey with a 5-point Likert scale for each answer. The survey was based on two different questionnaires, one measuring different aspects of Pepper's behaviour and one grading Pepper's usability as a system.

The first part of the survey used the Godspeed questionnaire [19], which has five different parts: anthropomorphism, animacy, likeability, perceived intelligence and perceived safety. For this experiment, all of them were used. Each part consists of scales to rate the robot on, using the Likert scale. Together these scales provide insight in how Pepper is viewed in these five different aspects.

The second part of the survey was inspired by the System Usability Scale (SUS) [20]. This scale originally consists of ten questions which are scored and together provide insight in the usability of a system. It is not an absolute measurement, but rather a way to compare, just like the Godspeed questionnaire. While the goal of the experiment was not to measure whether or not Pepper was a usable system, some of the questions from the SUS gave inspiration for a way to measure how well Pepper was received by the participants. The questions that were used, albeit in a modified way, were questions 2, 3, 6, 8 and 9.

Besides the survey, the participants were also asked a couple of open questions in a more dialogue-oriented way to provide qualitative insight in the experiment. These questions were: "What did you think of the experience?", "Is Pepper alive?" and "Does Pepper have a personality and if yes can you describe it?". These questions were asked to learn more about the effectiveness setup of the experiment and to gain more insight in the answers to the survey.

5.3 Results

The experiment was performed with 5 participants of different ages and split genders. After the experiment each participant was given a survey containing Likert scale questions from the Godspeed questionnaire as well as 5 questions inspired by the SUS and were asked some open questions according to the previous section. The results of the survey can be seen in figures 4 and 5. The (paraphrased and translated) answers to the open questions can be found in table 4. For privacy reasons the answers are in no particular order.

All 5 participants were immediately willing to help Pepper as soon as Pepper told them what to do. Some started talking to Pepper, others did not, but every participant immediately started removing obstacles and help Pepper to the door. The participants took on average 3 minutes to get Pepper to their goal and all participants succeeded in helping Pepper.

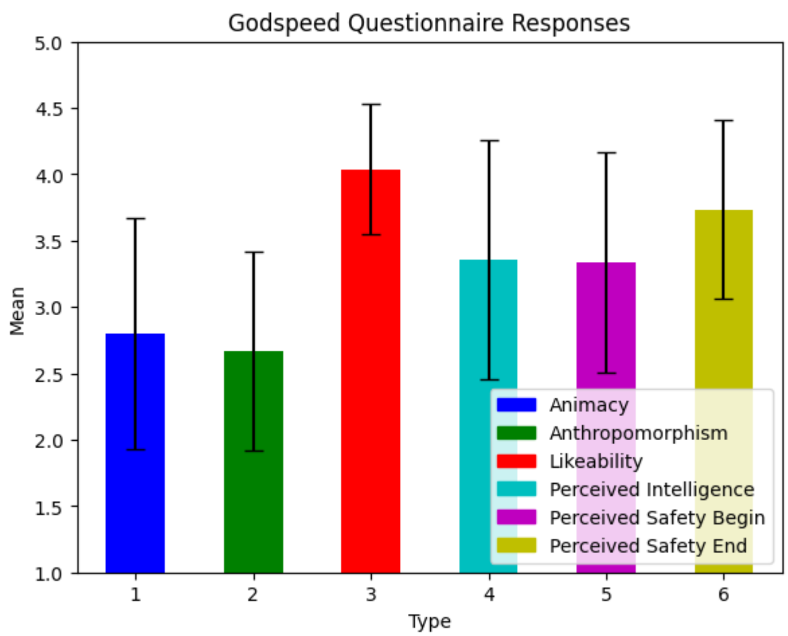

Figure 4: Mean responses to the Godspeed questionnaire categories taken from 5 participants.

Figure 4 shows how Pepper ranked on different aspects from the Godspeed questionnaire. It can be seen that Pepper ranked high in likeability, but lower on animacy and anthropomorphism. This is consistent with the fact that the version of Pepper the participants interacted with was only the basic program with very little personality. When asked about whether or not Pepper was alive, the response was mixed but leaning slightly towards the negative, as can be seen

"What did you think of the experience?"	Fun. Pepper doesn't go as fast. Having to wait that long is not very useful. But it's awesome to see a robot do this much. The facial recognition is pretty cool.
	At the start it was a little creepy, but it's pretty cool.
	Fun. You're slowly starting to wonder whether or not you're doing it right. Is also pretty fun. Was a little uncertain about it. "Why aren't you following me?".
	Pepper was standing still a lot. Did Pepper recognise my face properly?
	Amazing experiment! Did get a little impatient, because I didn't know how I could accelerate it. The robot was only taking small steps and I couldn't really figure out how to make those steps bigger. It was good that Pepper reminded me that I had to stay visible. Had forgotten that for a moment.
"Is Pepper alive?"	A little. Pepper was taking a bit. Sometimes I was thinking "are you real?".
	Yes. Was really creepy when Pepper started moving at the beginning.
	No. The response time was too low, so no. Did get the idea I was actually talking to something, but Pepper wasn't really responsive.
	Yes, because there was a little light in Pepper's eyes. Made me think of the twinkling in someone's eyes. Also because the hands moved. Didn't know for sure if I was allowed to touch the robot, so I didn't because I didn't know if it would mess something up, but I almost couldn't resist grabbing Pepper by the hands and leading Pepper towards the door. The robot was kind of cute.
	Muah. No. Very little complete movements.
"Does Pepper have a personality and if yes can you describe it?"	Cheerful, affectionate.
	There wasn't really one. A bit pushy.
	Friendly and a bit clueless.
	Dependent and grateful. A bit vulnerable like that.
	Nice. The question was clear, not forceful. Calm.

Table 4: Translated and slightly paraphrased responses to the open questions

in table 4, which is also consistent with the questionnaire responses. Overall the participants cited Pepper's slower response time and lack of movement or responses to attempts at conversation. The question on Pepper's personality had most participants rate Pepper as nice, friendly and slightly clueless, as well as vulnerable and affectionate. This response coincides with the high likeability score and the slightly higher score in perceived safety at the end of the experiment compared to the beginning. According to the results the participants also ranked Pepper as being slightly intelligent, which was backed by their answers to the open questions as multiple participants spoke about Pepper in a way that suggested Pepper was making their own decisions instead of being controlled by a program. However they still referred back to Pepper's programming if something was perceived to be wrong, such as when Pepper would not recognise their face or refused to move even when the path was cleared. Participants stated they would start to wonder if Pepper was working properly whenever they refused to move.

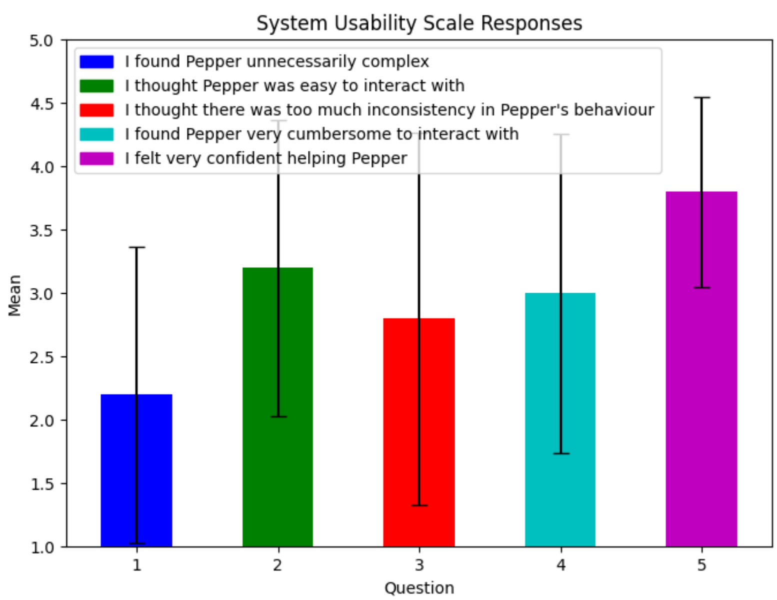

Figure 5: Mean results of the System Usability Scale taken from 5 participants. The scores are rated from 1 meaning strongly disagree to 5 meaning strongly agree.

Figure 5 shows that overall the participants were able to work well with Pepper. The participants felt confident in helping Pepper and had very little problems understanding what to do. The higher scores in inconsistency and Pepper being cumbersome to work with are consistent with how the participants described what they thought of the experiment. Participants mentioned sometimes being unsure if Pepper could see them or wondering why Pepper was moving slow or in short bursts instead of moving all at once.

During the experiments it became clear that once participants heard what they were supposed to do they would instantly clear the way for Pepper and would then try and coax Pepper forward towards the door. Most would try to talk to Pepper in a soothing voice, beckoning them forward and encouraging them when they moved. Some would ask Pepper to take bigger steps or otherwise try to get them to move faster. One participant mentioned afterwards that they were not sure if they were allowed to touch Pepper so they did not, but they had almost wanted to grab Pepper by the hand and guide them that way. Another participant was completely silent during the experiment, only moving obstacles out of the way and making sure to stay within Pepper's view.

6 Overall Discussion

From the low scores in animacy and anthropomorphism it can be concluded that this version of Pepper was not enough to convince the participants that Pepper was acting on their own or fully alive. This did not stop participants from helping Pepper though. During the experiment itself the participants treated Pepper as one might a small child, encouraging them to walk forward and slowly guiding them towards the door. They waited patiently as Pepper slowly made their way across the room.

As most participants saw Pepper as nice, friendly, vulnerable and helpless, it can be concluded that even this basic program which just consisted of Pepper looking at a face and moving towards it if able is enough to give people the impression of personality. The helplessness likely came from the fact that Pepper would stop and start seemingly randomly while they tried to move, instead of moving smoothly towards the participant. This is likely because of the programming, but during the experiment it made for an additional layer of personality that participants zoned in on. The simple explanation at the start and some small comments during the experiment were enough to help the participants understand what was going on, as can also be concluded from the Usability Scale questions.

The results show that this simple task where a human has to help Pepper get across the room is enough to have participants interact with Pepper. The participants described Pepper with a consistent personality of helplessness and friendly despite this version of Pepper only having the base interaction with no intended personality attached to them. The result that participants were able to discern a personality in Pepper is consistent with the hypothesis to subquestion 3.

This experiment with the base personality also starts paving the way towards an answer to subquestion 4, as section 4.7 showed that different personalities can be added onto an existing base personality, as long as that base interaction is sound. With the positive results from this current base personality, this next step could once again be attempted.

7 Conclusion and Future Work

In this paper an experiment was proposed and tested to help pave the way towards a new section of the HRI spectrum, the part where the robot is displaying dysfunctional behaviour. Two personalities were found to be ideal candidates for this test, a Negative Avoiding Type and a Negative Approaching Type, both created from external dysfunctional behaviour. To be able to properly test these personalities, it was concluded that first a base interaction had to be formed and perfected before the personalities were added in. This interaction became a setup where the robot had to leave a room filled with obstacles that they could not move on their own. The experiment showed that humans interacting with this robot were able to perceive helplessness within the robot, to the point that they described the robot with a personality even when there was none intended.

In a future experiment, this base interaction can be used to test the two proposed personalities and observe the effect on a human's interaction with a robot displaying dysfunctional behaviour. Future research should keep in mind to keep the interaction smooth and to not let the participants wait too long for the robot to give feedback. This research has shown that not much is needed for humans to discern a personality, therefore priority should be given to a consistent interaction over too many moving parts to make the personality seem more real. That way this gap in HRI can finally start to close, contributing to a more complete picture of the interaction between humans and robots.

References

[1] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders.* (5th ed., text rev.), 2022.

[2] Karen D Lincoln. Social support, negative social interactions, and psychological well-being. *Social Service Review*, 74(2):231–252, 2000.

[3] Ayline Maier, Caroline Gieling, Luca Heinen-Ludwig, Vlad Stefan, Johannes Schultz, Onur Güntürkün, Benjamin Becker, René Hurlemann, and Dirk Scheele. Association of childhood maltreatment with interpersonal distance and social touch preferences in adulthood. *American journal of psychiatry*, 177(1):37–46, 2020.

[4] Paul Marks. Antisocial robots go to finishing school. *New Scientist*, 191(2569):28–29, 2006.

[5] Nicolas Spatola and Olga Wudarczyk. Implicit attitudes towards robots predict explicit attitudes, semantic distance between robots and humans, anthropomorphism, and prosocial behavior: From attitudes to human-robot interaction. *International Journal of Social Robotics*, 13:1149–1159, 2021.

[6] Kacie Kinzer. Tweenbots, 2009. `http://www.tweenbots.com/`.

[7] David Harris Smith and Frauke Zeller. The Death and Lives of hitchBOT: The Design and Implementation of a Hitchhiking Robot. *Leonardo*, 50(1):77–78, 02 2017.

[8] Laura Londoño, Adrian Röfer, Tim Welschehold, and Abhinav Valada. Doing right by not doing wrong in human-robot collaboration. *CoRR*, abs/2202.02654, 2022.

[9] Stephanie Rosenthal and Manuela Veloso. Mobile robot planning to seek help with spatially-situated tasks. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, AAAI'12, pages 2067–2073. AAAI Press, 2012.

[10] Vasant Srinivasan and Leila Takayama. Help me please: Robot politeness strategies for soliciting help from humans. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 4945–4955, New York, NY, USA, 2016. Association for Computing Machinery.

[11] Elizabeth Cha and Maja Matarić. Using nonverbal signals to request help during human-robot collaboration. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5070–5076, 2016.

[12] Mark Scheeff, John Pinto, Kris Rahardja, Scott Snibbe, and Robert Tow. *Experiences with Sparky, a Social Robot*, pages 173–180. Springer US, Boston, MA, 2002.

[13] Hadi Beik Mohammadi, Nikoletta Xirakia, Fares Abawi, Irina Barykina, Krishnan Chandran, Gitanjali Nair, Cuong Nguyen, Daniel Speck, Tayfun Alpay, Sascha Griffiths, et al. Designing a personality-driven robot for a human-robot interaction scenario. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 4317–4324. IEEE, 2019.

[14] Kwan Min Lee, Wei Peng, Seung-A Jin, and Chang Yan. Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication*, 56(4):754–772, 2006.

[15] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. Would you trust a (faulty) robot? effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, pages 141–148, 2015.

[16] M. L. Lupetti. Shybo - design of a research artefact for human-robot interaction studies. *Journal of Science and Technology of the Arts*, 9(1):57–69, 2020.

[17] Amit Kumar Pandey and Rodolphe Gelin. A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3):40–48, 2018.

[18] Interactive Robotics. Robot in de klas, 2016. `https://www.robotsindeklas.nl/`.

[19] Astrid Weiss and Christoph Bartneck. Meta analysis of the usage of the godspeed questionnaire series. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 381–388. IEEE, 2015.

[20] John Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.