Investigating pedestrian-robot interaction in a context manipulation experiment

Kelian van Pernis

Supervisor: Dr. Baptist Liefooghe *

Master Thesis Artificial Intelligence, Utrecht University

February 14th, 2022

Abstract

In recent years a growing number of companies started pilot projects with delivery robots. Previous research on pedestrian-robot interaction has looked at, for instance, the effect of robot velocity and anthropomorphism on pedestrian behavior towards a robot, as well as people's ethical concerns towards these robots. However, little is known about what effect a clearly conveyed function of a robot has on behavior towards this robot. Therefore, a field study was done where observations of pedestrians encountering a stationary mobile robot without a clear function were compared to observations of pedestrians encountering a stationary Garbage cleaning robot. We found that, for pedestrians that encountered the Garbage bot, the chance of ignoring that robot was higher compared to pedestrians encountering the robot with no clear function. An additional finding was that women are more likely to ignore a mobile robot compared to men. The findings have emphasized that companies would benefit from ensuring that their robot conveys a clear function. The clear conveyance of the robot's function decreases the likelihood of pedestrians interfering with the robot's tasks. In addition, it decreases the chance of pedestrians getting distracted by the mobile robot, therewith increasing overall road safety. Future studies could conduct experiments with an autonomously driving robot, rather than a stationary one. These studies could aim to investigate the behavior of the demographic groups for which the current study did not provide enough data, as well as investigate how behavior towards mobile robots changes over time.

I. Introduction

In the past decades the field of automation and robotics has become increasingly prominent in society. In the automotive industry the focus has been set on developing (semi-)autonomous vehicles, with companies like Tesla gaining a lot of interest and popularity due to their advancements in autonomous vehicle software (Endsley, 2017). The advancements in the automation of vehicles have not been limited to passenger transport but are also seen in the field of autonomous mobile robots, which gained more interest in recent years. These robots can serve multiple purposes, but a prime example would be the last-mile delivery robots. The purpose of last-mile delivery robots is to employ a safer, environmentally friendlier, and more efficient method of delivering goods to populated urban areas. Many companies have started projects related to last-mile delivery tasks.

Figure 1: *From left to right the robot designs from DHL, Australia Post and Starship technologies.(USpostalservice, 2018)*

DHL has conducted pilots with a delivery robot named "PostBOT"; a robot that carries packages and closely follows the person delivering the mail. Australia Post has done similar pilots with a package delivery robot that does not need assistance from humans, but instead delivers packages from a central hub to people's homes. Besides delivery of mail,

*Thanks to Baptist for his excellent guidance in this project.

other companies have focused on food delivery. For instance, Starship has been partnering with food delivery companies to test out the effectiveness of food delivery by autonomous mobile robots (USpostalservice, 2018). Figure 1 shows these three delivery robot designs.

With the rapid rise of the global e-commerce market, advancements in the field of last-mile delivery robots become more relevant by the day. A 2019 e-commerce report revealed estimations that the total e-commerce revenue is expected to rise to an astonishing 2568.8 billion euros by 2023 (Striapunina, 2019). Naturally, this increase in online orders of goods will lead to an increase of packages being delivered worldwide every day. Deliveries in the last mile are not efficient and often slowed down by parking limitations, traffic congestion and environmental regulations (Akeb et al., 2018). Therefore, the introduction of these vehicles is a great opportunity for the e-commerce industry.

Besides last-mile delivery robots there are other applications for mobile robots. An example is the prototype garbage cleaning robot shown in Figure 2. This mobile robot uses image recognition to identify and collect garbage in places like parks where people often litter (Bai et al., 2018).

Figure 2: *A mobile garbage cleaning robot (Bai et al., 2018).*

Other examples of mobile robot applications include a bridge inspection robot aimed to replace human workers performing risky bridge inspection tasks (Sutter et al., 2018) and a recently developed COVID-robot designed to encourage groups of people in large public spaces to adhere to social distancing rules related to the, at the time of writing, COVID-19 pandemic (Sathyamoorthy et al., 2020).

The increasing number of real world projects involving mobile robots performing various tasks have raised ethical and legal concerns regarding the use of (semi-)autonomous mobile vehicles. An example of these concerns regard delivery robots collecting audiovisual information for purposes like accident documentation. The robots are often designed to collect information like photos, sound recordings and films in order to provide evidence in the case of accidents occurring that involved the robot. The collected data is seen as personal data according to the General Data Protection Regulation (Voigt and Von dem Bussche, 2017) and therefore should be handled with great care and in accordance with the rules and regulations regarding personal information.

Another concern relates to liability. When an autonomous mobile robot causes an accident there are several parties that can be blamed. The parties involved include the potential person or persons involved in the accident, the person monitoring the robot in an operator room, the company that deploys the robot and the manufacturer of the robot. The manufacturer should arguably be held responsible for any accidents caused by defective robots, with defective meaning that the robot did not work as specified by the manufacturer. However, the manufacturer can limit the risk of liability by clearly instructing the company deploying the robots on how to use them and inform the company of any risks involved in deploying them (Hoffmann and Prause, 2018). Liability would then rest with the company as a whole or the specific operator working during the accident. Making the operator a possible liable party would require recording everything that the operator monitors in order to have evidence that shows the operator can be held responsible, but this would circle back to the ethical concern of companies collecting and storing audiovisual information of people in traffic, mentioned in the previous paragraph. Thus, the examples have shown the complexity of dealing with ethical and legal concerns and show that regulative parties like the GDPR need to set clear rules regarding liability and the processing of personal information.

Before mobile robots can be deployed on a large scale it is vital to know how people would interact with these robots as knowledge about this can help with the optimization and design of the robots. Ideally, a robot should be designed and optimized such that pedestrians only interact with it when necessary, because any interactions that are unnecessary for the task of the robot, will ultimately lead to an unwanted delay in the execution of the robot's tasks. Taking DHL's PostBOT as an example, the task of the robot is to deliver packages. Therefore, DHL would only want pedestrians to interact with the robot when receiving packages meant for them. After all, any disturbance of curious pedestrians would slow the robot down and decrease the efficiency of the robot's delivery tasks. These mobile robots are sometimes also able to make use of deep learning models, so any knowledge about pedestrian behavior towards these mobile robots would be useful to acquire, because it could, in theory, be used to "feed" the models and teach the mobile robots how to deal with pedestrian interactions.

There have been a few studies that looked into the interaction between pedestrians and mobile robots. An example study was conducted by Chen and colleagues who ran a series of experiments to investigate the effect of robot moving speed on pedestrian velocity. They found that as the robot moved on faster, the pedestrians walked more slowly. In other words pedestrian's velocity was inversely correlated with the robot's velocity (Chen et al., 2018).

Besides the speed of the robot, other characteristics of the robot can have an effect on people's behavior towards it. In a study by Koay and colleagues a series of trials were conducted in which they looked at the effect of anthropomorphic features on pedestrian behavior towards the robot. They found that people feel more comfortable with the robot coming closer if it has a "mechanoid" appearance compared to when it has a "humanoid" appearance. However, over the course of five weeks in which several experiments were ran, the researchers noticed that the effect of the robot's appearance on pedestrian behavior diminished. This was explained by the concept of a "novelty effect", which entails that responses in the initial encounters differed significantly compared to following encounters, because during the following encounters the pedestrians had gotten used to the robots and their appearance (Koay et al., 2007).

A similar study looking at the novelty effect was conducted by Kim and colleagues who deployed small autonomous mobile robots on a university campus to investigate human interaction with these robots and, more specifically, to observe a potential change in perception towards these robots over time. They found that more encounters with the robot resulted in a decrease in perceived "mental capabilities" of the robot. This finding was explained by the "novelty effect" meaning that the expectations that pedestrians had of the robot changed after the first encounter (Kim et al., 2020).

Pedestrians can also occasionally have a negative attitude towards robots. Moore and colleagues coined the term "griefing" to describe unexpected or aggressive behavior towards automated vehicles (Moore et al., 2020). Several studies have looked into griefing behavior of pedestrians towards robots. One study reported griefing behavior towards service robots in a public plaza, which included people kicking and punching the robot (Salvini et al., 2010). In another study researchers built a hitchhiking robot and tested it by letting it walk through Canada and the Netherlands. This robot was at some point destroyed by people that encountered the robot (Smith and Zeller, 2017). Finally, a third study observed griefing behavior specifically with children damaging a robot in a shopping mall. The creators of this robot even designed an algorithm that, given an encounter with children and their parents, would attempt to move close towards the parents rather than the children in order to avoid any damage (Brščić et al., 2015).

Another field of research, that is relevant for this study, is gaze patterns of pedestrians. Knowing what pedestrians look at when traversing a road can give valuable insight into the effect of certain objects or vehicles and their design choices on pedestrian behavior. An example study is one done by de Winter and colleagues in which participants wearing an eye-tracker were asked to walk through a parking garage. The eye-tracker recorded all the data of the participants' gaze behavior. They found that while walking through the parking garage the participants looked at the ground only 20% of the time and gazed towards parked cars roughly 35% of the time, and more specifically looked at certain parts of the cars including the back, side and front as well as the tires of the car. Thus, pedestrian showed a general interest in the vehicles they passed (de Winter et al., 2021).

A recent study by Liu and colleagues looked at the effect of understanding an automated vehicle's (AV) intentions on the gaze behavior of pedestrians. They conducted a within-subject lab study where 10 participants interacted with an AV as well as a manual driven vehicle (MV) over the course of multiple trials. The participants were asked to walk a certain route through the lab, after which there would be an encounter with either the AV or MV. Afterwards the participants were asked to fill in a survey with questions about the extent to which the participants understood the intentions of the vehicle. They found that pedestrians' gaze duration would be longer for the AV compared to the MV. More importantly, they found that duration of pedestrians' gaze and their understanding of the intentions of the AVs were inversely correlated. In other words, not understanding what the AV was doing led to pedestrians gazing towards the AV longer on average compared to when the participants had a clear understanding of what the vehicle was doing (Liu et al., 2020).

The research discussed here gave some insight into pedestrian behavior towards robots and the effect of design choices like robot velocity and anthropomorphic features on the behavior of pedestrians towards mobile robots. The study by Liu and colleagues showed that people's perception and understanding of a mobile robot can influence how long they gaze at the robot. However, Liu and colleagues did not look at how design choices could contribute to a better understanding of the robot's function and intentions. If a robot has a clear design that conveys its function and intention, it is easier for people to understand the robot. Understanding the robot would then lead to shorter gaze durations.

The earlier discussed robot prototypes, which were shown in Figure 1, differ considerably in how their function is conveyed to pedestrians. DHL have decided to use their recognizable bright yellow color in combination with their company title and logo to clearly convey the mail delivery function of their "Postbot". Starship, on the other hand, has taken less effort into clearly conveying the function of their robot. Starship's robot has a generic white color and from the outward appearance of the robot it is hard to determine that it is a food delivery robot. This design, that does not convey the robot's function well, makes it harder for pedestrians to understand what the robot is doing which according to the findings of Liu and colleagues would lead to longer gaze durations. Pedestrians gazing at a mobile robot can cause multiple issues. Gazing at a robot for too long means the pedestrians are distracted and do not watch out for other traffic like bikes or cars coming their way. Besides safety there is also the issue of hindrance. A delivery robot needs to deliver its goods or food in a timely manner, and people gazing at the robot while standing in its way will inevitably slow it down. If conveying the function of the robot well could lower the chance of pedestrians bothering the robot while it is performing its duties, this would be valuable information to companies looking to deploy mobile robots on a large scale. Also, if conveying the robot's function well decreases the time spent gazing at the robot, this would lead to pedestrians paying better attention to the road thus improving overall traffic safety.

Therefore, to find out whether the extent to which the function of a mobile robot is conveyed influences how pedestrians interact with it, a between-subject observational study was conducted in Bunnik, the Netherlands, to investigate the behavior of pedestrians towards two different designs of a mobile robot. The first design was a robot that showed no clear function. The second design was a "Garbage bot" aimed to convey the function of transporting garbage. The latter design's function was clear and fit well in the context of the green area of Bunnik, where the local government is known for its effort of keeping the area clean. The two robot designs can be seen in Figure 3. In the present observational study, interactions between pedestrians and the garbage bot were compared to interactions between pedestrians and the robot that showed no clear function. During the observational sessions the robot was placed on the sidewalk after which the observer inconspicuously took place at a nearby picnic table and started taking notes of passing pedestrians' behavior towards the robot. The taxonomy for pedestrian behavior used in this study was adapted from a study by Usher and colleagues who created such a taxonomy for possible response behavior of pedestrians when interacting with a mobile robot (Usher et al., 2017). The observational study aimed to collect data which could provide insights into the behavior of pedestrians towards the robot. More specifically, the data from the observations was used to find an answer to the following research question:

How does a pedestrian's perception of a robot's function in its context influence their behavior towards this robot?

The expectation for this study was that pedestrians encountering the robot with no clear function would be more likely to ignore the robot compared to pedestrians encountering the Garbage bot. This hypothesis was mainly based on the findings of Liu and colleagues where pedestrians paid more attention to AVs when they did not understand what the robot was doing. Thus, a clearly conveyed function could help pedestrians understand what the robot is doing and decrease the attention given to the robot. It was also expected that a robot with an unclear function would cause curiosity of pedestrians potentially leading to them feeling more inclined to take a close look at the robot.

The following section will provide details on the method used in this study, including information on participants, materials, design, procedure and measures. Then in the results section the data from the observations is visualized after which it is analyzed further with a Bayesian Logistic Regression model. In the discussion section these results are reflected upon. This section will also elaborate on the limitations of the current study as well as potential future work, and at the end a conclusion of the present study will be drawn.

II. Method

Participants

In this study a dataset was created from observations of two groups of participants. The time frame of this research project allowed for observation of 483 participants (222 Male ; 261 Female). The groups of participants consisted of pedestrians that happened to walk past the robot in a naturalistic setting in Bunnik, the Netherlands. The estimated ages of the pedestrians ranged from 25 to 70 years old (M = 43.14, SD = 13.38).

For the data collection, any pedestrians that encountered the robot when walking in groups of 2 or more were seen as a single observation of an encounter. Thus for the data analysis the number of encounters was used, where an encounter included either a single person or a group of 2 or more people. This was done because whenever people walked together, the behavior of the person walking in front was usually copied by the other people. Naturally, whenever one person of the group for instance stopped to look at the robot the others stopped as well. If these groups would be split up during data analysis, their data could be misleading because if they had walked on their own they might not have stopped. In total 294 observations were made of interactions between pedestrians and the robot, 147 observations per participant group. This number was balanced by, on the last day of observations, ending the observation session slightly early when the number of observations for both participant groups were levelled out.

The experiment was approved by the ethics committee of Utrecht University. An informational flyer was made that the experimenter could give to the observed pedestrians after they encountered the robot in order to obtain consent for using their anonymous data in the current study. Giving the flyer (or verbally asking consent) after taking note of the observed behavior prevented the pedestrians from changing their behavior towards the robot due to knowing they were part of an experiment. If any of the participants did not want the data from their observations to be used, the data was removed from the study.

Materials

For the experiment a mobile robot was used to make observations. This robot was borrowed from a company called "More Work Less Carbon" (MWLC)*, which is based in Utrecht, the Netherlands. The robot designs can be seen in Figure 3. There are cameras attached on both the front and back part of the robot, which are normally used by the robot to gather visual information in order to navigate the road, but for this experiment the robot was standing still and thus the cameras were not used. The robot on the left in Figure 3 was used in the control condition of the experiment. For the Garbage bot condition this same robot was used but with a compartment mounted on top which was carrying (seemingly transporting) PMD waste. The components used for this design were:

- A transparent plastic box
- A PMD waste bag filled with plastic (mostly packaging material)
- Materials to mount the box on top of the robot, which included: screws, nuts, metal bars and a hex key.

In order to effectively observe pedestrians an inconspicuous setup was created for the observer to ensure that they would not draw the pedestrians' attention. The observer was placed at a picnic table near the robot and was having lunch during the observations. The observer wore headphones and had a mobile phone in their hands to take notes of observations. The notes were made in Google docs on the mobile phone[†]. After observations the notes were entered into a Microsoft Excel sheet for later use in data analysis. Data analysis was done on a pc with the latest versions of programming languages Python and R.

Figure 3: *Robot design in the Control condition (left) vs Robot design in the Garbage bot condition (right)*

As mentioned in the previous paragraph, a consent flyer was made that could be given to observed pedestrians to inform them of their anonymous data used in the study. The consent flyer is shown in Appendix A. Most of the time giving the flyer was not necessary because most observed pedestrians were enthusiastic about the study and had no problem with their anonymous data being used for research purposes.

Design

In order to find out whether a pedestrian's perception of a robot's function in its context influences their behavior towards this robot, an experiment was run over the course of 6 weeks at a pedestrian walkway in the green area of Bunnik, the Netherlands. Observations of this area prior to the study revealed that the area is popular for people to go on a walk around noon and that people from different age groups walk by, which was beneficial for the study. The intervals between passing pedestrians made this area more suitable for the current study compared to for instance a crowded shopping center street, because the intervals allowed for the observer to takes notes of the observed behavior as well as, where possible, approach the passing pedestrians to tell them they were observed for the study.

In the experiment the robot was standing still on a pedestrian walkway. Two designs were used for the robot in the experiment, as seen in Figure 3. The robot on the left was used in the control-condition. For this design the function of the robot was not clear. Therefore, the reactions to this robot formed a baseline to which the results from the Garbage bot design could be compared. The Garbage bot had a function that fitted well within the context. The robot was standing on a pedestrian walkway in an area that is known for its green environment. The local government makes a great effort to keep the area clean by often sending workers to gather and transport garbage. Therefore, a robot with a function of transporting garbage fits well in the context.

Given the naturalistic setting of the experiment, it was not possible to influence the type of people participating in the study. However, an attempt was made to create two homogeneous participant groups. Therefore, the observations made during this study were all made between 12:00 and 14:00 on weekdays, to increase the chance of the pedestrians encountered in one group being similar to those in the other group. Observations were only made on days without pre-

[†]The notes can be accessed here.

*Thanks to Hans Steuten and Jan-Willem van Bentum (MWLC) for their collaboration and for providing the robot used in this study.

cipitation, because this could influence pedestrians' behavior as well as decrease the number of pedestrians encountered during observations.

Figure 4 shows the layout for the experiment. The robot's location is shown with a cyan color, the observer with yellow and the pedestrians are shown in red. The observer's location had a clear overview of the robot and passing pedestrians. It was close enough for the observer to be able to see the pedestrians as well as intervene where necessary, and far enough to not draw any unwanted attention.

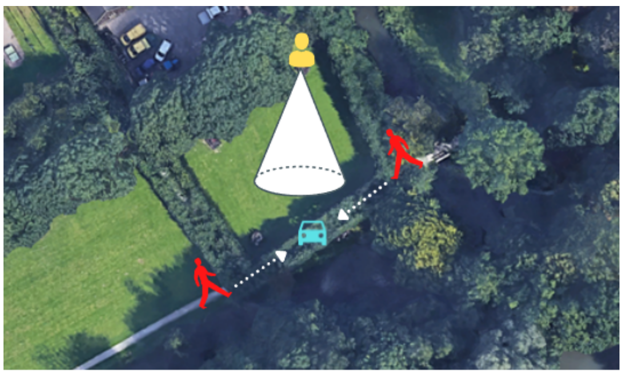

Figure 4: *The spatial layout of the experiment (Google, (n.d.))*

This experiment had a between-subjects design. The pedestrians interacting with the robot from the control condition formed participant group 1, and their behavior towards the robot was compared to the behavior of the pedestrians that interacted with the robot from the Garbage bot condition, who formed participant group 2. Random sampling was used in this study to ensure pedestrians were randomly assigned to either one of the two participant groups. On day 1 of observations the robot from the control condition was used, on day 2 the robot from the Garbage bot condition was used, on day 3 the robot from the control condition again and so on. At certain points during the study the number of observations per condition were far apart due to the fact that some days more pedestrians walked by the robot than on other days. Therefore, whenever the number of observations per condition were “drifting apart” the robot design that had fewer observations was used two days in a row to level out the number of observations in both conditions.

When conducting a study that includes observations it is crucial to take "observation bias" into account. Observation bias, or also known as the Hawthorne effect, refers to a change in behavior of participants in a study when they realize they are being observed. A strong method that can eliminate the observation bias is the covert observation method. This is an observational method in which participants have no idea they are being observed (Wu et al., 2017). Thus, in this study, pedestrians interacting with the robot were not told they were part of an observational study until after the interaction. The benefit of covert observation is that it results in a higher external, and more specifically, ecological validity. The higher ecological validity is due to the fact that the trials took place on a public road, where pedestrians were unaware that they were observed at the moment of interaction. Therefore, the study's findings give insight into the behavior of pedestrians towards small mobile robots in a real-life setting. The higher external validity stemmed from the random sampling of pedestrians that coincidentally encountered the robot, in an area of the Netherlands that is representative for at least a large portion of the country.

In order to test the experimental setup a pilot study was conducted. In the pilot experiment the setup as shown in Figure 4 was used. The main reasons for the pilot experiment was to test the chosen method as well as to see whether any unforeseen circumstances would occur that needed to be taken into consideration for the rest of the study. The robot from the control condition was used in the pilot study. The experimenter placed the robot on the pedestrian walkway and observed for 1.5 hours between 12:00-14:00 on a weekday. The pilot study confirmed a few assumptions as well as that it helped gain insight into certain factors that could influence the observations made.

First of all, with Bunnik being a very green and quiet area, it is a popular spot for runners. During the pilot experiment two people who were on a run encountered the robot. These people ran straight past it while barely looking at it (other than ensuring they would not run into the robot). Given the “goal setting” nature of this sport, it makes sense that in order to, for instance, maintain a steady running pace or run a certain distance within a certain amount of time, these people would refrain from stopping to look at the robot regardless of the way it looks. Therefore, in order to prevent data from runners skewing the results, it was decided to leave out any observations of runners encountering the robot.

Secondly, an assumption made in the design of the experimental setup was that the observer would not be easily noticed, let alone be seen as part of the experiment, due to the inconspicuous setting created. Based on reactions from observed pedestrians this assumption seemed correct. Often the pedestrians did not even notice the observer at all. According to the participants this was mainly due to the fact that it is common for people to sit down and have lunch in this area.

Lastly, an insightful observation during the pilot study was one where the pedestrian was upset about the fact that they felt like they were being recorded. The observer explained to the pedestrian that the cameras were not used for this experiment and explained the nature of the experiment. The pedestrian mentioned that “someone” might become angry with the presence of a robot fitted with cameras and could push it into the creek nearby. This encounter emphasized the need for the observer to be close enough to intervene in the case that a displeased pedestrian might show this type of griefing behavior.

Procedure

In table 1 information on the different observation sessions is shown. The observations were made on weekdays and exclusively on days without precipitation. In week 1 a pilot study was done to test out the experimental setup. On the day of the pilot study the robot was picked up from the storage location near the experiment location, and was carried to the experiment location by the observer. Then for 1.5 hours between 12:00 – 14:00 observations were made of interactions between passing pedestrians and the robot. At the start of the observation sessions the observer sat down at the picnic table and started up the Google Docs document on their phone to take notes of the observed behavior. Where possible the pedestrians were approached to explain the nature of the experiment and were asked consent for using their anonymous data. If necessary the consent flyer was given out. After observing for 1.5 hours the robot was carried back to the storage location and the session was ended. Due to the setup of the pilot study being successful the following observation sessions followed a very similar routine. At pickup the robot design was prepared, which meant that the transparent box with the PMD-waste bag either had to be mounted on top of the robot if the robot from the Garbage bot condition was to be used, or demounted if the robot from the control condition was to be used. The robot was then carried to the experiment location by the observer. Then, just like in the pilot study, observations were made of interaction between pedestrians and the robot for 1.5 hours within the time frame of 12:00-14:00. Notes were made on Google docs again by the observer, who was sitting at the picnic table. After observing, the robot was brought back to the storage location and the session was ended.

Table 1: *Data collection sessions*

Session	**Condition**	**Observations**‡
1 (Pilot)	Control	32
2	Garbage bot	26
3	Control	19
4	Garbage bot	18
5	Control	27
6	Garbage bot	14
7	Control	29
8	Garbage bot	38
9	Garbage bot	51
10	Control	17
11	Control	23
		Total: 294

Measures

For making observations of the interactions between pedestrians and the mobile robot, a coding scheme was required. Coding schemes are catalogs that classify the behavior of interest for the study being conducted (Kim et al., 2010). As mentioned in the introduction, Usher and colleagues created a taxonomy for possible response behavior of pedestrians when interacting with a robot. This taxonomy was used as a basis for a coding scheme that could be used for the observations in the current study. Table

‡A single observation in the experiment is the observed behavior of either an individual or a group of 2 or more people

2 shows the coding scheme used in this study. Some parts of the taxonomy were slightly adjusted to fit better with the conditions and environment of the experiment. The original taxonomy used distance to the robot as a factor, but this was irrelevant in this study because pedestrians would always come very close to the robot when passing it, due to narrow walkways. Thus, the behavior was not categorized on distance to the robot but solely on the action (or inaction) of the pedestrian towards the robot, as well as their walking behavior. A second adjustment that was useful in order to turn the taxonomy into an effective coding scheme was to number the response types.

Table 2: *Coding scheme used during observations (Adapted from (Usher et al., 2017)*

Response Category	Response Type	Response behaviour
Attraction	1: Interact	Stop and interact
Attraction	2: Watch	Stop and observe
Attraction	3: Curious	Slow down and observe
None	4: Ignore	No reaction to robot
Repulsion	5: Cautious	Avoid - small path deviation
Repulsion	6: Avoid	Avoid - large path deviation

Using a coding scheme in this experiment was beneficial as it allowed for quick notes of observations. This was especially useful when multiple people would walk past the robot in quick succession. For each pedestrian that walked past the robot, the observer noted down 6 characteristics, resulting in six data points. First of all the behavior was noted down according to the coding scheme. Then the group composition was noted down by writing down an ‘A’ for each adult and ‘C’ for each child. After this the gender of each pedestrian was noted down with either an ‘M’ or an ‘F’ for male/female. Then the estimated minimum age of the pedestrian was noted down. If a group of pedestrians were estimated to be roughly the same age only one age was noted down, but if pedestrians were estimated to have different ages the age of each pedestrian was noted down separately. If the pedestrians were walking their dog this was noted down with a "Y" for "Yes" or, if they were not, a "N" for "No" in the designated "Pet" column. If anything out of the ordinary was observed, like a pedestrian taking a photo, this was noted down in a dedicated comment section.

Each day, after data collection, the data from the observations was entered into an Excel sheet. When all data was collected the data of all observation sessions was aggregated into a single Excel sheet. The benefit of having the data in an Excel sheet was that it could be easily accessed with Python or R for (statistical) data analysis.

Statistical Analysis

1 Data Cleaning

Missing and incorrect values During the transfer of the data from the observer’s notes to the Excel data sheet, some data points were forgotten and thus a few cells in the data sheet had missing values. These values were later filled with the help of the observers notes. For the notes on age estimate, in some cases two people walking together were of a different age. In the data this would be denoted as for instance "35/65". For those data points the two ages were replaced by a single number representing the average of the two ages.

Outliers Due to the dependent variable in this study being a categorical variable, it is unconventional to talk of "outliers" as there can not really be an "extreme" value. However, if very few data points were collected for certain behaviour categories, this could be relevant in determining what data was useful to add to the Logistic Regression model.

In the current study the response types defined as: "Interact", "Cautious", and "Avoid", were rarely observed during the sessions, and could thus be seen as categorical "outliers". For those pedestrians that exhibited the "Cautious" and "Avoid" behavior it was hard for the observer to determine whether the robot was the cause of this behavior, because pedestrians could have just wanted to take a shortcut or have changed their mind on which route to walk, thereby coincidentally avoiding the robot by changing their walking direction. For the "Interact" response type there were only three occurrences, two of which occurred when two men were guiding two handicapped women along the path. The handicapped women naturally moved extremely slow, which led to the men guiding them spending a lot of time next to the robot, patiently waiting for the women to catch-up. This inevitably led to interaction with the robot. Thus, due to indications that these already rare occurrences of behavior were likely influenced by external factors, the data points of the three response types "Interact" , "Cautious" and "Avoid" were dropped from the data set.

For the Composition category, the group compositions of three or more people, as well as groups containing children, were rarely observed. Therefore, due to an inability of drawing strong conclusions due to low amounts of data for these group compositions, the data of groups larger than three people as well as the data of children was removed from the data set.

Finally, the number of observations of pedestrians walking their dog was very limited, therefore the data of pedestrians walking with pets was not analyzed.

Binning For some categories, like age, there were many groups within that category. This led to a low number of observations for those groups. For example, only 17 observations were made of people estimated to be 60 years of age. These 17 observations are not enough to conclude something about behavior of this age group. Therefore, data binning was applied to aggregate data into larger groups so that each group contains enough data to be able to form a conclusion on the potential difference in behavior between these groups. For the Age category participants were split into two groups: "Young Adults" and "Adults". The ages in the Young adults group ranged from [25-44] and the ages in the "Adults" group ranged from [45-70]. The Gender category was reduced from five groups (M, MM, F, FF, MF) to three groups (Male, Female, Mix), where the observations of individuals and duos were aggregated for groups "Male" and "Female". As mentioned in the previous paragraph, data of three response types of "Behavior" were removed from the data set. This left the responses types of "Watch", "Curious" and "Ignore". As seen in table 2 the first two response types belong to the "Attraction" category while the latter response type belongs to the "None" category. In order to effectively compare responses the data was grouped into two contrasting behavior types "Attract" and "Ignore", where the "Attract" group thus contained all "Watch" and "Curious" responses.

2 Bayesian Data Analysis

For the analysis of the data in this study a Bayesian approach was preferred over classical methods. Bayesian statistics focus on calculating a posterior distribution, normally using the following formula:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \tag{1}$$

In this formula θ contains the model parameters and Y contains the data. The posterior distribution is calculated using the likelihood, prior distribution and marginal likelihood. Firstly, The likelihood is defined as $p(Y|\theta)$ and refers to the distribution of the data given the parameters. Secondly, the prior distribution is defined as $p(\theta)$ and describes the distribution before any data is seen. Lastly, the marginal likelihood is defined as $p(Y)$ and has the purpose of normalization to ensure that the computed posterior distribution is a probability distribution (Bürkner and Vuorre, 2019).

Although Bayesian statistical models are often computationally expensive, this approach was preferred over classical approaches due to a higher model flexibility and overall more elaborate and informative results (Bürkner and Vuorre, 2019).

Besides the general advantages of Bayesian statistics mentioned in the previous paragraph, the preference for the Bayesian approach also stems from the subtle and slightly philosophical difference in interpretation of results between Bayesian statistics and classical frequentist methods. In a study like the one discussed in this paper where there are two conditions, classical statistical models would try to find a significant difference for these two conditions. If a significant effect would be found with a "Confidence Interval" of for instance 95%, this would be interpreted as an implication that repeating the experiment many times would result in finding the same results 95% of the time. Credible Intervals, as used in Bayesian statistics, indicate a 95% probability that the estimated parameter value lies within a certain interval.

The aspect of Bayesian models predicting an outcome fits well with the purpose of this particular study. Knowing the probability of pedestrians showing a certain behavior is useful knowledge for the robot because it can use this knowledge to anticipate pedestrian behavior and ensure that tasks are not slowed down or undesirably influenced in any way by pedestrian behavior (e.g. a mail delivery robot preventing pedestrians from unnecessarily slowing it down during delivery tasks).

III. Results

Before looking at the results of the data analysis it is worth mentioning some of the rarely observed behavior that was not categorized in the taxonomy of Table 2. A few pedestrians who stopped to look at the robot took out their phone and started taking pictures of the robot. Also, as mentioned in the method section, one pedestrian was upset because they felt like they were being recorded. They started a discussion with the observer about their displeasure regarding the robot. Two children that were observed during the study showed interesting behavior, with one child poking the robot with a wooden stick, and another child shouting about an intention to destroy the robot. The rarely occurring behavior mentioned here will be further reflected upon in the Discussion section of this report.

The next chapters will go into the descriptive statistics summarizing the data. After that an in depth description is given of the Binomial Logistic Regression model used to further analyze the data.

1 Descriptive Statistics

After data cleaning, the resulting data set contained information on 269 observations. As mentioned earlier, the response behavior data was grouped into the two contrasting behavior categories "Attract" and "Ignore". In the following paragraph the raw data counts of the two behavior categories are shown for both conditions in the study, as well as the difference in counts of the two behavior categories based on Gender, Age and Composition. The stacked bar charts visualize the proportion of pedestrians that ignored the robot. After discussing the count data, the findings of further analysis of the data using the Logistic Regression model is discussed.

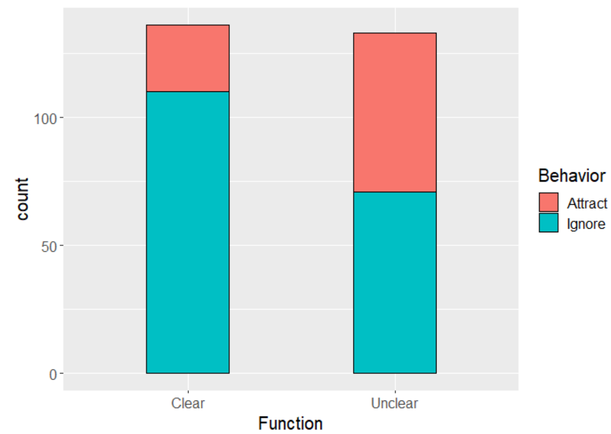

Figure 5: *Counts of behavior "Attract" and "Ignore" for the category Function*

Robot Function The data for the two participant groups is shown in Figure 5. In the participant group that encountered the robot that showed a clear function (Garbage bot), *19.1%* showed the "Attract" behavior and *80.9%* showed the "Ignore" behavior. For the participant group that encountered the robot that did not show a clear function, 46.6% showed the "Attract" behavior and 53.4% showed the "Ignore" behavior. Thus, the proportions show the expected effect of an unclear robot function causing pedestrians to show interest in the robot more frequently.

Gender For the attribute Gender there were three categories: Female, Male and Mixed. The data for Gender is shown in Figure 6. For the Female group 15.6% showed the "Attract" behavior and 84.4% showed the "Ignore" behavior. For the Male group this was 41.9% for "Attract" and 58.1% for "Ignore". For the Mix group this was 42.4% for "Attract" and 57.6% for "Ignore". The graph shows that the proportion of pedestrians ignoring the robot was higher in the Female group, compared to the Male and Mix groups. It should be taken into account here that the category "Mix" exclusively contains duos (a man and a woman) whereas the groups of Female and Male contain the aggregated data of individuals and duos. Thus, any difference between group Mix and the other groups could partially be attributed to the effect of Composition.

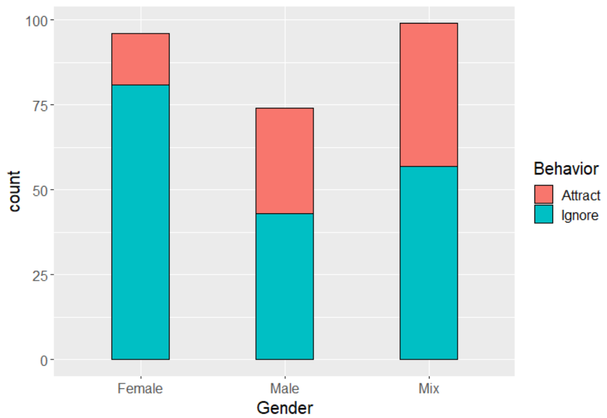

Figure 6: *Counts of behavior "Attract" and "Ignore" for the category Gender*

For the Age characteristic there were two categories: Young adults and Adults. Figure 7 shows the data of these two age groups. In the group of Young adults 34.0% showed the "Attract" behavior and 66.0% showed the "Ignore" behavior. For the Adult group this was 31.2% for "Attract" and 68.8% for "Ignore". The data showed that there was barely any difference in behavior between these two age groups.

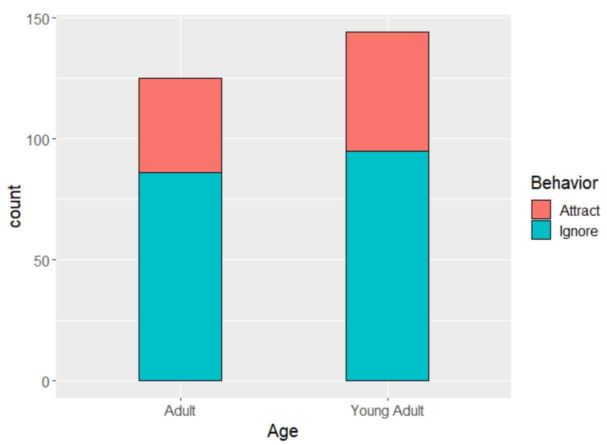

Figure 7: *Counts of behavior "Attract" and "Ignore" for the category Age*

The Composition characteristic had two categories: Individual and Duo. Figure 8 shows the data for this characteristic. In the group of individuals 27.5% showed the "Attract" behavior and 72.5% showed the "Ignore" behavior. For the group of duos this was 36.9% for "Attract" and 63.1% for "Ignore". The data shows that the proportion of pedestrians ignoring the robot was slightly higher with individuals compared to duos.

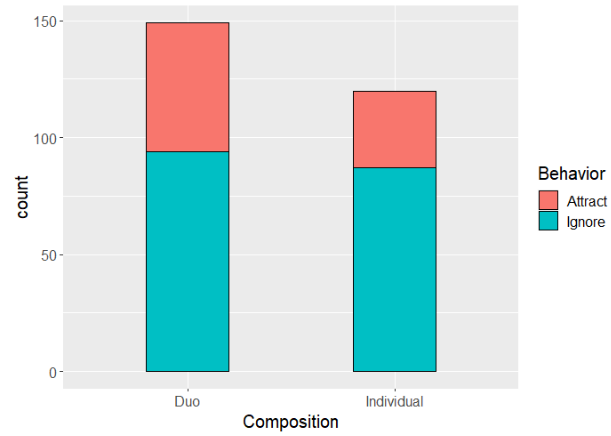

Figure 8: *Counts of behavior "Attract" and "Ignore" for the category Composition*

2 Binomial Logistic Regression

The proportions of the response behavior discussed in the previous paragraphs give an indication of the effect of both the function of the robot as well as characteristics of the participants on the pedestrians' behavior towards the robot. To further analyze the effect of the robot function and pedestrians' Gender, Age and Composition on the behavior exhibited towards the robot, a statistical analysis was done using a Binomial Logistic Regression model. Binomial Logistic Regression is an effective method of calculating predictions of a dichotomous outcome variable. This is a variable that has two outcomes, and the model predicts the probability of one of the outcomes (Schüppert, 2009). In the current dataset the dichotomous outcome is the dependent variable "Behavior", which is either "Ignore" or "Attract". In the following chapters the model selection process as well as the output of the chosen model are discussed in detail.

Brms To create the models an R package called "BRMS" was used. This package allows for creating models in R using the probabilistic programming language Stan. After creating a model in R the code is compiled to Stan code and run, after which the results are returned in an R object. The package has the advantage of being able to create the model in R while benefiting from the great modeling flexibility provided in Stan (Bürkner, 2017).

3 The Model

In this study pedestrians' characteristics were noted down during observations so that these characteristics along with the main manipulation of the study could be added to a model to determine which variables could predict pedestrians' behavior towards the robot. Multiple models were created and compared to find the model with the most predictive power. The following paragraph will discuss the model selection process.

Model Selection The first model created was one that incorporated all four predictor variables Function, Age, Gender and Composition. The second model used these same predictors but additionally looked at interactions between those predictor variables. The first model's output predicted an effect of Function and Gender on pedestrian behavior, but did not predict an effect of Age and Composition on pedestrian behavior. Therefore, a third model was created that only used Function and Gender as predictor variables. In order to compare the predictive power of these models the $loo()$ function of the BRMS package was used. This function applies the approximate Leave-one-out cross validation (ALOOCV) technique. The approximation lowers the computational cost of the Leave-one-out cross validation (LOOCV) technique (Beirami et al., 2017). Leave-one-out cross validation is a commonly used method for comparing Bayesian models on their estimated predictive performance on new unseen data (Sivula et al., 2020). Table 3 shows the result of applying the $loo()$ function to the three models.

Table 3: *Model comparison*

	elpd_diff	elpd_se
Model3	0.0	0.0
Model1	−1.0	1.1
Model2	−4.1	1.9

The output of the $loo()$ function in BRMS always puts the best performing model at the top and compares the other models to it. The "elpd_diff" column shows the difference in predictive power. The details of the calculations of the "elpd" scores are not relevant for the comparison, but the difference can be seen as the extent to which the models differ in their predictive power. As a rule of thumb, any difference lower than 4 can be seen as a small difference (Sivula et al., 2020). The difference between model 3 and 1 is low indicating only a slight difference in predictive power with that model. The difference between model 3 and 2 appears to be substantial as it is a difference of slightly more than 4. Based on the model comparisons, model 3 was chosen for further analysis of the data. The unused models 1 and 2 are shown in Appendix B. The following paragraphs will provide an in depth description of model 3 as well as discuss and visualize the predictions of this model.

Model Fitting The $brm()$ function of the BRMS package was used to fit the models. The following code snippet shows the code for model 3:

```
model3 <- brm(
  formula = Behavior ~ Function+Gender,
  data = obs,
  family = "bernoulli",
  prior = c(
    prior("normal(0,1)",
    class = "b"),
    prior("normal(0,1)",
    class = "Intercept")),
  warmup= 2000, iter = 5000,
  file = "model3"
)
```

The main argument $formula$ identifies the dependent variable on the left of the $\sim$ sign and the predictor variables on the right. The $+$ sign was put in between predictor variables to indicate no interaction between predictors was desired. The $family$ argument describes the distribution of the response variable. In this case *bernoulli* is used because this distribution is effective for a dichotomous dependent variable (Bürkner, 2017). The prior can be used to incorporate knowledge from previous studies. Given that the literature study did not yield any data on what to expect in this study, a weakly informative prior was used for every effect in the model. A weakly informative prior entails a prior that incorporates an assumption of the "extremeness" of an effect.

The alternatives are an informative prior that would incorporate prior knowledge or a non-informative prior that models no assumptions or prior knowledge whatsoever (Lemoine, 2019). The priors for the regression coefficients in this model were specified globally with a normal distribution. The normal distribution had a mean of 0 and a standard deviation of 1. This narrow distribution represents the assumption that the study would not result in "extreme' results. In other words, it was not expected that any of the predictor variables would cause the extreme effect of either "Ignore" or "Attract" to be near 100% of the response in one of the predictor variable categories.

Model output An abbreviated version of the output of the model can be seen in Table 4. The output contains the coefficient table that is based on the posterior distribution. For each term under "Population-level Effects" the table shows the posterior mean and the 95% credible interval. The first item under "Population-Level effects" is the Intercept. For each of the two predictor variables Function and Gender the Intercept takes one random group within that category and uses it as the reference category. This means the Intercept describes the posterior distribution for the group of pedestrians that fit in the reference categories: (Function:Clear, Gender:Female). In other words, the Intercept describes female pedestrians that encountered the Garbage bot. The other items in the list are compared to the Intercept to measure the effect of the particular predictor variables. The item "FunctionUnclear", for instance, describes the group of female pedestrians that encountered the robot with no clear function.

Table 4: *Summary of the Regression parameters in the model fitted to the pedestrian behavior data*

Population-Level Effects	**Estimate**	**95% Credible Interval**
Intercept	2.35^*	$[1.74; 2.99]$
FunctionUnclear	-1.36^*	$[-1.92; -0.82]$
GenderMale	-1.38^*	$[-2.07; -0.72]$
GenderMix	-1.21^*	$[-1.86; -0.58]$

* 0 outside 95% credible interval.

The values under "Estimate" are the log odds ratios of the response variable "Behavior". The odds ratios can be extracted from these log values by taking the exponent, following this formula:

$$P(y=1)/P(y=0) = exp(\mu_k) \tag{2}$$

In this formula, "y" corresponds to the "Behavior" category with the binary options 1 (Ignore) and 0 (Attract). To illustrate, taking the Intercept as an example, the odds ratio for that group would be $exp(2.35) = 10.48$. This means that the odds of the behavior "Ignore" are 10.48 times higher than the odds of the behavior "Attract" for the group described in the Intercept, which are female pedestrians encountering the robot that shows a clear function (Garbage bot).

The log odds ratios for each Population-Level effect gives an indication of the effect of particular predictor variables. The further away from 0, the bigger the effect. Negative estimates indicate a higher chance of the "Attract" behavior compared to the group described in the Intercept, and positives estimates indicate the opposite. In Table 4 the estimates for which 0 is not in the 95% credible interval are indicated with an asterix ($*$). When 0 is not in the 95% credible interval, we can conclude with at least 95% probability that the predictor variable had an effect on the pedestrian's behavior towards the robot. When looking at Table 4 it shows that this was the case for the function of the robot, which was the main manipulation of this study. According to the table it appears that pedestrians' gender also had an effect on the probability of them ignoring the robot.

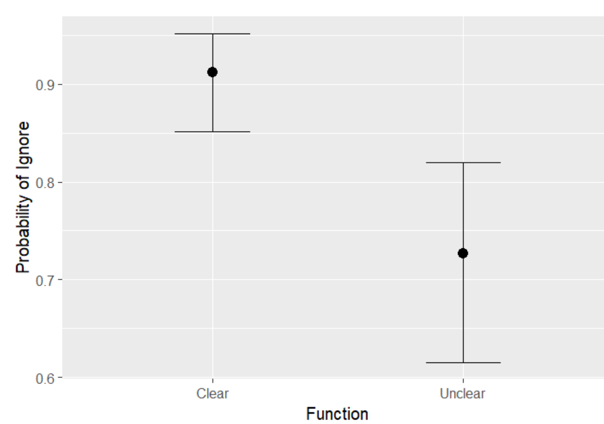

Figure 9: *Predicted probabilities of the behavior "Ignore" for predictor variable Function*

Using the $marginal\_effects()$ function in BRMS we can plot the effect of any of the predictor variables. Figure 9 shows the effect of the main manipulation Function on the probability of pedestrians ignoring the robot. The dots represent the predicted probability of behavior "Ignore", with the error bars representing the 95% credible intervals for the probabilities. The graph shows that according to model 3's predictions, pedestrians are more likely to ignore the robot with the clear function (Garbage bot) compared to the robot with the unclear function. The credible intervals are wide, especially for the predictions of the robot with an unclear function, indicating that the model's predictions (the dots) are not very accurate. However, given that the error bars do not overlap we can be confident that there is an effect of Function on pedestrian behavior towards a mobile robot.

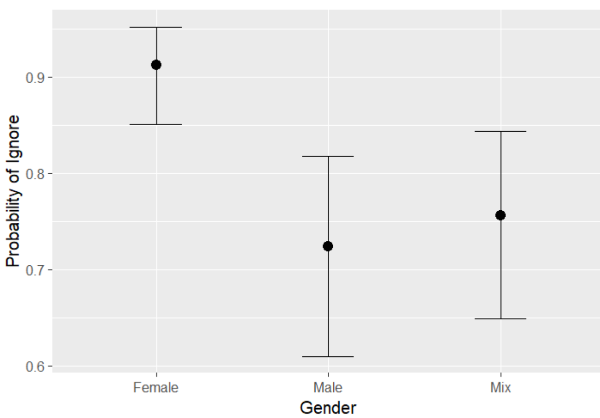

Figure 10: *Predicted probabilities of the behavior "Ignore" for predictor variable Gender*

Figure 10 shows the predicted probabilities of behavior "Ignore" for the category Gender. According to model 3's predictions, female pedestrians are more likely to ignore the robot compared to male pedestrians. When men and women are walking together they are less likely to ignore the robot compared to female pedestrians. It should be taken into account, as mentioned before, that the group "Mix" exclusively contains duos whereas groups Male and Female contain aggregated data of men and women. However, given that our first model was not able to predict any effect of Composition, it would be safe to assume that the difference between groups Mix and Female can, for the most part, be attributed to the effect of Gender.

IV. Discussion

This study aimed to find out whether pedestrians' perception of a robot's function in its context would influence their behavior towards that robot. The results have shown that when the function of the robot was clear, pedestrians became much less likely to stop and approach the robot compared to when the function of the robot was unclear, in which case pedestrians more often stopped to watch the robot or slowed down their walking speed to have a look at the robot.

Besides this main finding the results also showed that Gender had an effect on the likelihood of pedestrians ignoring the robot. Female pedestrians were more likely to ignore the robot compared to male pedestrians as well as men and women encountering the robot together. The results indicate that men tend to have more curiosity towards the robot resulting in men stopping to look at the robot more often than women. The data of the "Mix" category indicate that men show this same behavior when walking together with a woman.

For the results of the model discussed in the previous paragraphs the credible intervals between the compared groups did not overlap, indicating that we can be confident in concluding these effects were present. The first model that we discussed did not show any evidence of an effect of age or group composition on the behavior towards the robot. For the effect of age it should be taken into account that, if there had been enough data to compare smaller age ranges (e.g. age ranges like [25-34]) there could potentially have been differences between these smaller age groups. For group composition the data of groups larger than three were not taken into consideration due to a low number of observations of groups larger than three. Given that the model did indicate a slightly higher chance of ignoring the robot for individuals, it could be the case that the chance of pedestrians ignoring the robot is higher with groups larger than three compared to duos and individuals.

In the introduction several studies were discussed related to mobile robots. Research was done into the effect of design choices like robot velocity and anthropomorphism on people's behavior towards these robots, but not on the effect of conveyance of function. The current study provided new information on the effect of how well a robot conveys its function on pedestrian behavior towards the robot and therefore contributed to a more complete understanding of the effect of mobile robot design choices on pedestrian behavior towards these robots.

During the observation sessions pedestrians were assumed to encounter the robot for the first time, which means that there was likely a novelty effect in the encounters. This effect was researched by Koay and colleagues who saw a change in behavior towards mobile robots over the course of multiple encounters. Kim and colleagues found similar results in their study on the novelty effect. Although the novelty effect has an impact on the generalizability of the current study's findings, it does not undermine the results due to the fact that for both designs of the robot the novelty effect would have been present, allowing for an effective comparison of behavior towards the two robot designs.

The behavior displayed by the child that poked the robot with a stick and the other child that shouted about intentions to destroy the robot were in line with the findings of Brščić and colleagues regarding the griefing behavior of children towards robots. The pedestrian who got upset about the cameras attached to the robot showed that the ethical concerns discussed in the introduction are present in society and that these concerns should be addressed.

The current study effectively built upon the the findings of Liu and colleagues that showed evidence of pedestrians gazing at AV's longer when it was unclear what the AV was doing. A longer gaze essentially means more attention to the robot which is, in the context of for instance delivery robots, unwanted behavior of pedestrians. The findings of the current study have revealed that this amount of attention given to the robot can differ with the extent to which the robot's function is conveyed clearly. Ensuring that a mobile robot's function is conveyed well can reduce the amount of attention given to the robot by pedestrians and increase the likelihood of pedestrians ignoring the robot.

The evidence this study provided of the effect of the robot's function on pedestrian behavior implicates that a proper design of a robot that has the task of inspection or delivering goods can contribute to an efficient and smooth execution of these tasks. When the robot conveys a clear function, pedestrians that encounter the robot are likely to ignore it or at the very least not let their curiosity negatively impact the robot's ability to perform its tasks. For companies looking to deploy these robots this knowledge can be valuable because it can help with the efficiency of the robot's tasks. When taking Australia Post's robot as an example, ensuring that the robot conveys in a clear way that it is delivering mail should decrease the likelihood of pedestrians interfering with the robot's task. Increasing the efficiency of the robot in terms of number of packages delivered per hour could increase revenue as well as contribute to the satisfaction of customers. The same would apply to Starship's food delivery service. For food delivery it is arguably even more crucial that delivery times are kept to a bear minimum, given that the food gets colder the longer it takes and customers do not want to wait too long for their food to arrive when they are hungry. Thus these companies would benefit from spending time on the design of their robot and ensuring that it conveys its function well.

While the findings for the effect of Gender on behavior towards the robot might not be of great importance for the field of mobile robots, it does illustrate that specific characteristics of pedestrians can have an effect on their behavior towards mobile robots. Although the current study did not obtain enough data on elderly people and people under 25, it could be that these groups behave differently towards mobile robots. Companies deploying mobile robots could program them in such a way that they use pedestrians' characteristics to determine the likelihood of them approaching the robot. The robot could then try to avoid people with characteristics that indicate a higher chance of interference. The method of Brščić and colleagues who programmed a robot to stay close to adults and steer away from children perfectly illustrates such an implementation of pedestrian knowledge to allow the robot to perform its tasks more efficiently.

Limitations

Stationary Robot In this study observations were made of pedestrians who passed a stationary robot. The robot used in this study is designed to drive autonomously, but did not do so during the observational sessions due to limited human resources paired with safety concerns. Although the robot, when it is used in practice, might stand still at several occasions while performing its tasks, it will likely be driving at a slow pace for the majority of the time. Reactions to a driving robot could differ from the reactions to a stationary robot.

Participants The participants in this study were all adults with ages ranging from 25 to 70. During the observational sessions only a handful of children and adolescents were observed passing the robot. The data was too scarce to draw any conclusions on the behavior of children and adolescents towards the robot. Although the findings in this study are based on a large demographic age group, children and adolescents still make up for a significant proportion of the population in the Netherlands. More data of children and adolescents would have made it easier to generalize the results of the study to the general population. Besides the demographic limitation the method also allowed for participants to be counted twice in the study. If, for instance, a person encountered the robot on Day 1 and their behavior was observed and noted down, there is a slight probability that this same person encountered the robot again on another day, at which point their behavior was noted down again. The second time the participant encountered the robot they might have reacted differently than they would have if it had been the first encounter with the robot. This limitation should not have had a large impact on the findings regarding the research question, because due to random sampling the odds of a person encountering the robot for the second time would have been the same for the two participant groups.

Observer As described in the method section the observer was placed at a picnic table close to the robot. The observer had a clear overview of the robot and the pedestrians walking past the robot. While it was necessary for the observer to be close enough to the robot to be able to approach pedestrians when needed, it would have been best if the observer had been out of sight altogether during the interaction between pedestrians and the robot. A few pedestrians might have suspected that the observer placed the robot there and this suspicion could have influenced their behavior towards the robot.

External factors The field experiments in this study took place on the public road in a naturalistic setting. While the choice for a field experiment was justified by the benefit of the increased ecological validity, it should be taken into account that results from field experiments like these can be influenced by extraneous variables (e.g. variables that influenced the results but were not accounted for). The alternative to a field study would have been to create an "artificial setting" where the potential effect of extraneous variables could have been limited, but this would have had the unwanted effect of significantly decreasing the ecological validity of the study.

Generalizability Although the choice for a field study increased the ecological validity of the study, it is still hard to determine whether the results of this study can effectively be generalized to every situation. For instance, it could be that pedestrians in a busy shopping street behave significantly different towards mobile robots compared to pedestrians in the relatively quiet and green area where the current study was conducted. The observational sessions in this study took place in November and December which resulted in cold temperatures during those sessions. It could be that pedestrians walking outside when there are warm temperatures in summer would be in less of a hurry to get to their destination and be potentially more inclined to interact with the robot, compared to the pedestrians observed in the current study, who walked outside in cold temperatures. Another factor that could influence the generalizability of the study's findings is the novelty effect. It was assumed in the current study that most if not all pedestrians encountered the robot for the first time. Research by for instance Koay and colleagues indicated that behavior towards mobile robots can change after multiple encounters. Therefore, it could be that the observed behavior in the current study would be different from the behavior of pedestrians that encountered the robot multiple times.

Future studies

The findings of this study gave insight into the effect of pedestrians' perception of a robot's function on their behavior towards that robot. Besides this, it has laid some groundwork for future field studies on similar topics. The data collected during observational sessions in this study lacked information on the behavior of children and adolescents, people walking their dog and people walking in groups bigger than three. The field experiments were also conducted with a robot that was stationary for the entire duration of the experiments.

Future studies can aim to conduct experiments in which information can be gained on how adolescents and specifically children behave towards a mobile robot. The literature showed that children are known to be more likely to show griefing behavior compared to adults. Gaining more knowledge on the behavior of children could be beneficial due to the ability to incorporate this knowledge into the robots to "teach" them to watch out for the potential erratic behavior of children. Using a driving robot in future studies as opposed to a stationary robot will provide results with increased ecological validity due to the robot used in the study being a better representation of the robots that will likely be used in real settings. The literature also discussed the novelty effect when people encounter a mobile robot. Future studies could focus on experimenting with a mobile robot around the same group of people for a longer duration to investigate further how people's behavior towards mobile robots changes over time.

To gain more knowledge of how pedestrians will react to mobile robots in society it would be useful if researchers in future studies would attempt to cooperate with companies aiming to deploy mobile robots. The companies could then, together with researchers, create a prototype mobile robot that would be a good representation of a robot that would be deployed in the future, and have researchers use this prototype in their studies on pedestrian behavior towards mobile robots.

Conclusion

The current study aimed to find out whether pedestrians' perception of a robot's function in its context influences their behavior towards that robot. The findings found evidence supporting the hypothesis that a robot conveying a clear function increases the likelihood of pedestrians ignoring the robot. This finding has emphasized the importance of companies ensuring that the mobile robots they deploy show a clear function so that disturbance of the robot by pedestrians is minimized and road safety is maximized by reducing the chance of pedestrians being distracted by the robot. The findings also emphasized the possibilities for mobile robots to incorporate knowledge of pedestrian's characteristics to aid them in efficiently executing their tasks.

References

Akeb, H., Moncef, B., and Durand, B. (2018). Building a collaborative solution in dense urban city settings to enhance parcel delivery: An effective crowd model in paris. *Transportation Research Part E: Logistics and Transportation Review*, 119:223–233.

Bai, J., Lian, S., Liu, Z., Wang, K., and Liu, D. (2018). Deep learning based robot for automatically picking up garbage on the grass. *IEEE Transactions on Consumer Electronics*, 64(3):382–389.

Beirami, A., Razaviyayn, M., Shahrampour, S., and Tarokh, V. (2017). On optimal generalizability in parametric learning. *Advances in Neural Information Processing Systems*, 30.

Brščić, D., Kidokoro, H., Suehiro, Y., and Kanda, T. (2015). Escaping from children's abuse of social robots. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*, pages 59–66.

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, 80(1):1–28.

Bürkner, P.-C. and Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, 2(1):77–101.

Chen, Z., Jiang, C., and Guo, Y. (2018). Pedestrian-robot interaction experiments in an exit corridor. In *2018 15th International Conference on Ubiquitous Robots (UR)*, pages 29–34. IEEE.

de Winter, J., Bazilinskyy, P., Wesdorp, D., de Vlam, V., Hopmans, B., Visscher, J., and Dodou, D. (2021).

How do pedestrians distribute their visual attention when walking through a parking garage? an eye-tracking study. *Ergonomics*, 64(6):793–805.

Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the tesla model s. *Journal of Cognitive Engineering and Decision Making*, 11(3):225–238.

Hoffmann, T. and Prause, G. (2018). On the regulatory framework for last-mile delivery robots. *Machines*, 6(3):33.

Kim, J. H., Chang, S.-H., Schmidt, L., Otto, J., Buch, B., Seifert, U., and Coussement, P. (2010). playing_robot': An investigation of situated cognition in the context of (artistic) humanrobot interaction design. In *Proceedings of the 1st IROS 2010 Workshop on Robots and Musical Expressions (IWRME 2001), Taipei*, page 69.

Kim, M. J., Kohn, S., and Shaw, T. (2020). Does long-term exposure to robots affect mind perception? an exploratory study. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 1820–1824. SAGE Publications Sage CA: Los Angeles, CA.

Koay, K. L., Syrdal, D. S., Walters, M. L., and Dautenhahn, K. (2007). Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pages 564–569. IEEE.

Lemoine, N. P. (2019). Moving beyond noninformative priors: why and how to choose weakly informative priors in bayesian analyses. *Oikos*, 128(7):912–928.

Liu, H., Hirayama, T., Morales, L. Y., and Murase, H. (2020). What is the gaze behavior of pedestrians in interactions with an automated vehicle when they do not understand its intentions? *arXiv preprint arXiv:2001.01340*.

Moore, D., Currano, R., Shanks, M., and Sirkin, D. (2020). Defense against the dark cars: Design principles for griefing of autonomous vehicles. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 201–209.

Salvini, P., Ciaravella, G., Yu, W., Ferri, G., Manzi, A., Mazzolai, B., Laschi, C., Oh, S.-R., and Dario, P. (2010). How safe are service robots in urban environments? bullying a robot. In *19th international symposium in robot and human interactive communication*, pages 1–7. IEEE.

Sathyamoorthy, A. J., Patel, U., Savle, Y. A., Paul, M., and Manocha, D. (2020). Covid-robot: Monitoring social distancing constraints in crowded scenarios. *arXiv preprint arXiv:2008.06585*.

Schüppert, A. (2009). Binomial (or binary) logistic regression. In *Statistics Seminar, Spring*. Citeseer.

Sivula, T., Magnusson, M., and Vehtari, A. (2020). Uncertainty in bayesian leave-one-out cross-validation based model comparison. *arXiv preprint arXiv:2008.10296*.

Smith, D. H. and Zeller, F. (2017). The death and lives of hitchbot: The design and implementation of a hitchhiking robot. *Leonardo*, 50(1):77–78.

Striapunina, K. (2019). Ecommerce report 2019: Statista digital market outlook – market report.

Sutter, B., Lelevé, A., Pham, M. T., Gouin, O., Jupille, N., Kuhn, M., Lulé, P., Michaud, P., and Rémy, P. (2018). A semi-autonomous mobile robot for bridge inspection. *Automation in Construction*, 91:111–119.

Usher, J. M., McCool, R., Strawderman, L., Carruth, D. W., Bethel, C. L., and May, D. C. (2017). Simulation modeling of pedestrian behavior in the presence of unmanned mobile robots. *Simulation Modelling Practice and Theory*, 75:96–112.

USpostalservice (2018). Autonomous mobile robots and the postal service. pages 5–12.

Voigt, P. and Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10:3152676.

Wu, K.-S., Chen, Y.-S., Lin, H.-S., Hsieh, E.-L., Chen, J.-K., Tsai, H.-C., Chen, Y.-H., Lin, C.-Y., Hung, C.-T., Sy, C. L., et al. (2017). A nationwide covert observation study using a novel method for hand hygiene compliance in health care. *American Journal of Infection Control*, 45(3):240–244.

Appendix

A. Consent Flyer (Dutch)

Hallo!

Mijn naam is Kelian van Pernis en ik heb u zojuist geobserveerd voor een mens-robot interactie onderzoek van de UU. De data die ik noteer is:

- Uw reactie op de robot qua loopgedrag / interactie
- In welke leeftijdscategorie u ongeveer zit
- Uw geslacht (M/V)
- In welke samenstelling u liep (alleen/tweetallen etc)
- Wel/geen huisdier.

Deze anonieme data wordt enkel voor het onderzoek gebruikt. Als u (toch) bezwaar heeft tegen het gebruik van deze data kan het altijd nog verwijderd worden, stuur dan graag een email naar k.a.vanpernis@students.uu.nl.

B. Unused models

Model 1 code:

```
model1 <- brm(
  formula = Behavior ~ Function+Age
  +Comp+Gender,
  data = obs,
  family = "bernoulli",
  sample_prior = "yes",
  prior = c(
    prior("normal(0,1)",
    class = "b"),
    prior("normal(0,1)",
    class = "Intercept")),
  warmup= 2000, iter = 5000,
  file = "model3"
)
```

Model 1 output:

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	2.28	0.41	1.49	3.11
FunctionUnclear	-1.36	0.29	-1.92	-0.80
AgeYoungAdult	-0.20	0.28	-0.75	0.35
GenderMale	-1.51	0.37	-2.24	-0.81
GenderMix	-1.06	0.38	-1.82	-0.32
CompositionIndividual	0.36	0.38	-0.39	1.11

Model 2 code:

```
model3 <- brm(
  formula = Behavior ~ Function*Age
  *Comp*Gender,
  data = obs,
  family = "bernoulli",
  sample_prior = "yes",
  prior = c(
    prior("normal(0,1)",
    class = "b"),
    prior("normal(0,1)",
    class = "Intercept")),
  warmup= 2000, iter = 5000,
  file = "model3"
)
```

Model 2 output:

```
Population-Level Effects:
                                                             Estimate Est.Error l-95% CI u-95% CI
Intercept                                                        2.16      0.52     1.18     3.21
FunctionUnclear                                                 -0.87      0.54    -1.92     0.19
AgeYoungAdult                                                    0.01      0.53    -1.03     1.06
GenderMale                                                      -1.20      0.62    -2.39     0.04
GenderMix                                                       -0.87      0.55    -1.95     0.20
CompositionIndividual                                            0.08      0.59    -1.06     1.23
FunctionUnclear:AgeYoungAdult                                   -0.43      0.60    -1.61     0.75
FunctionUnclear:GenderMale                                      -0.37      0.73    -1.80     1.05
FunctionUnclear:GenderMix                                       -0.49      0.60    -1.65     0.69
AgeYoungAdult:GenderMale                                        -1.00      0.69    -2.37     0.35
AgeYoungAdult:GenderMix                                         -0.48      0.61    -1.67     0.67
FunctionUnclear:CompositionIndividual                           -0.40      0.65    -1.69     0.85
AgeYoungAdult:CompositionIndividual                              0.67      0.67    -0.62     1.99
GenderMale:CompositionIndividual                                 0.30      0.66    -0.97     1.61
GenderMix:CompositionIndividual                                 -0.02      1.00    -1.96     1.93
FunctionUnclear:AgeYoungAdult:GenderMale                        -0.10      0.80    -1.68     1.46
FunctionUnclear:AgeYoungAdult:GenderMix                          0.63      0.69    -0.72     1.98
FunctionUnclear:AgeYoungAdult:CompositionIndividual              0.03      0.72    -1.39     1.43
FunctionUnclear:GenderMale:CompositionIndividual                -0.20      0.77    -1.70     1.33
FunctionUnclear:GenderMix:CompositionIndividual                 -0.00      0.98    -1.93     1.93
AgeYoungAdult:GenderMale:CompositionIndividual                   0.21      0.76    -1.30     1.67
AgeYoungAdult:GenderMix:CompositionIndividual                   -0.01      1.01    -1.98     1.99
FunctionUnclear:AgeYoungAdult:GenderMale:CompositionIndividual   0.61      0.83    -1.01     2.21
FunctionUnclear:AgeYoungAdult:GenderMix:CompositionIndividual   -0.01      1.00    -1.97     1.98
```