



AALBORG UNIVERSITET
STUDENTERRAPPORT

v/ Institut for
Electronic Systems
Produkt- og Designpsykologi
Fredrik Bajers vej 7
9220 Aalborg

Titel:

Subjektiv oplevelse af interaktionen med
en social robot i en dansk lufthavn

Projekt:

Investigation of Subjective Experiences

Projektperiode:

01. sept. til 18. dec. 2017

Projektgruppe:

Gruppe 782

Vejleder:

Dorte Hammershøj

Sidetal artikel: 8

Sidetal arbejdsblade: 183

Synopsis:

I dette studie undersøges danske rejsendes subjektive oplevelse af at blive betjent af en social robot i Aalborg Lufthavn (AAL), gennem to økologiske feltstudier. I begge tests blev de rejsende mødt af en *Double* robot, som tilbød at hjælpe dem med at finde vej. Efter en kort interaktion, beder robotten den rejsende om at følge efter sig. I første test kørte robotten de rejsende hen til et semi-struktureret interview om deres oplevelse. På baggrund af observationer og interviews blev 567 affinity notes kategoriseret i et affinity diagram med 10 kategorier, hvorfra vigtige parametre i forhold til brugeroplevelsen af sociale robotter udledes. På baggrund af disse parametre udvikles 24 skalaer, som efterfølgende er blevet anvendt i test nummer to. I anden test ledte robotten de rejsende hen til en testleder, hvor de på de udviklede skalaer vurderede deres oplevelse af interaktionen med robotten. Resultaterne fra anden test blev analyseret med Principal Component Analysis, som viste både positive og negative korrelationer mellem flere af skalaerne.

Deltagere:

Andreas Kornmaaler Hansen

Lucca Julie Nellemann

Juliane Nilsson

Emil Bonnerup

Sara Nielsen

Subjective Experience of Interacting with a Social Robot at a Danish Airport

Unpublished Conference Paper for SEMCON at Aalborg University, 2017

Andreas Kornmaaler Hansen, Emil Bonnerup, Juliane Nilsson, Lucca Julie Nellemann, and Sara Nielsen

School of Information and Communication Technology

Aalborg University, Aalborg, Denmark

Email: 17gr782@es.aau.dk

Abstract— This paper investigates the subjective experience of interacting with a social robot at Aalborg Airport (AAL) by conducting an ecological field study where 23 important attributes in Human-Robot Interaction (HRI) were elicited. Scales were developed from these attributes and tested at AAL. In both tests travellers were recruited by a remote controlled *Double* robot, which had an iPad with an interface asking if it may help the travellers with wayfinding at AAL. When the subjects had chosen the desired location they were asked to follow the robot. In the first test the robot led them to a semi-structured interview about their experience and in the second test to an experimenter who asked them to rate their experience on the developed scales. After the first test the observations and the subjects' statements were interpreted and coded using an affinity diagram. 567 affinity notes were sorted by a bottom-up procedure into 10 categories from which the 23 scale questions and scales were developed. The scales were used in a second test at AAL, where 43 subjects rated the robot and the interaction on the 23 scales. The ratings were analysed with Principal Component Analysis (PCA), which showed both positive and negative correlations within each of the three groups: *Robot's height, distance, and direction*.

I. INTRODUCTION

This study originates from a social robot research project at Aalborg University with the aim of developing and implementing robots in a variety of contexts. This raises technological, normative, and empirical questions as to how these robots are to behave and in which settings they might be useful. Several studies have already shown that both utilitarian and hedonic attributes influence the way users experience the interaction with social robots [1], [2]. Utilitarian attributes involves the utility, practicality, and usability of the robot, whereas hedonic attributes reflect the user experience, enjoyment, and acceptance while interacting with the robot [1]. Some gender and cultural differences are found to affect the interaction with a social robot. These are the willingness to accept robots [3], and the intimate distance, which is smaller for southern Europeans and Japanese compared to Americans and northern Europeans [4]. How one experience Human-Robot Interaction (HRI) also depend on one's gender, with females being less willing to accept and interact with robots than males, who also perceive the robots as more human-like compared to how females perceive robots [1].

It is unknown whether these results generally apply to Danish travellers. People have different experiences and might articulate them differently depending on both their culture and

the context. It is important to elicit attributes based on the user's own words and terms.

This study aims to elicit important attributes in HRI at an airport and develop relevant scales based on these attributes. Furthermore the study aims to analyse if some attributes have an influence on others. The developed scales might be used by robot designers for specific contexts and potentially for tailored user experience evaluations.

II. METHOD - ELICITATION OF ATTRIBUTES

The first test was conducted on Danish travellers who interacted with a social robot in a natural setting. The test was conducted at Aalborg Airport (AAL) after the security check and before the travellers reached the shopping and dining area. The travellers who interacted with the robot were asked to participate in a semi-structured interview about their first impressions while being observed during both the interaction and the interview. The *Wizard of Oz* method was used.

A. Materials

For the test a *Double* robot from Double Robotics Inc, [5], with an iPad Air 2, [6], was used. Based on a small pretest it was decided to change the head mount so that the iPad is angled upwards, to make the robot appear more welcoming. The modified *Double* robot is shown on Figure 1 and Figure 2. The *Double* robot was connected to a laptop via Wi-Fi connection and its own software. The *Double* application allows a web page to be superimposed on the iPad screen while the controller is still able to control the robot from the laptop. The web page presented on the iPad was a wireframe developed in Marvel, (www.marvelapp.com), aiming to depict the potential usage of the robot as a wayfinding tool at AAL. The entire wireframe was formulated in Danish.

B. Subject Recruitment

30 subjects from the age of 8 to 62 years ($M=37.9$, $SD=17.1$) distributed among 16 females and 14 males participated. The subjects were recruited by the robot itself, which was remote controlled by a present robot controller, to provide a more ecological and undisturbed interaction between robot and subject. The robot approached potential subjects and the wireframe on the iPad asked them if it might help them find their way around AAL and presented a "Yes/No" option.



Fig. 1: *Double's* front.



Fig. 2: *Double's* profile.

If "No" was pressed, the robot wished the traveller a pleasant journey and left. If "Yes" was pressed, the subjects were presented with four wayfinding options: Food, Shopping, Toilets, or Gate information, as shown on Figure 3. The subjects were then kindly asked to follow the robot after they had chosen their preferred option. The robot then led the subjects to an interviewer who shortly informed them of the study and received verbal consent to record audio before the semi-structured interview. In total 18 interviews were conducted of which 11 were on single travellers and seven were on groups of travellers.



Fig. 3: Options presented to the subjects after having pressed "Yes" and been informed of the robot being part of a research project.

C. Semi-structured Interview

The interview was a two part semi-structured interview. The first part consisted of probing the subjects for their first impression and experience of interacting with the robot in regard to their thoughts about the robot itself and what they think other travellers might think about the interaction. In addition to these conversation topics the subjects were asked about their opinion regarding the robots approach, usefulness, and reliability. Afterwards, the subjects were asked about

previous experiences and problems at airports, in order to gather potential use cases where the robot might be helpful.

The second part consisted of asking specific questions relating to the robots physical characteristics such as speed, height, distance, movements, appearance, and approach. These questions were asked because these attributes have been found to affect the experience of HRI [4].

The two parts were conducted in continuation of each other and the questions in the second part were only asked if the subjects had not previously answered them spontaneously.

D. Roles

Five researchers were present during the test at AAL. One controlled the *Double* robot, one conducted the interview, and the remaining three observed the travellers as they interacted with or walked past the robot. Since the robot controller had no way to know what happened on the screen, one of the observers signaled to the robot controller when to start leading the subjects to towards the shopping area.

E. Data Processing

The interviews were transcribed and coded along with observations into affinity notes. The purpose was to create an affinity diagram, which brings insights and issues into a hierarchical diagram based on subjects' statements and behaviour [7]. This affinity diagram is pivotal in eliciting the attributes that affect HRI, and thereafter in creating the scales to be used for further work. To include more of the gathered insights, it was decided to use both the spontaneous answers from the conversation topics and the answers from the specific questions in the affinity diagram. The two types of answers are not differentiated.

III. RESULTS - ELICITATION OF ATTRIBUTES

In total 567 affinity notes were sorted into 10 overall categories with individual subcategories. The 10 superordinate categories were appearance, trust, behaviour, approach, problems with the touch screen, avoidance of interaction, personal interest, positivity towards the robot, usefulness, and tech-experience.

Based on the 10 categories, attributes were elicited according to the criterion of a) being an influencing attribute and b) the possibility of formulating the attributes as a scale question. The field study was conducted on Danish speaking test subjects, and the attributes are therefore listed in both English and Danish. The scale questions are all presented on a *Visual Analogue Scale* (VAS) with closed anchor points and are either bi- or unipolar. If a scale is bipolar a midpoint will be marked either with or without a label. A bipolar scale without a label will be noted with *No label*, whereas an unipolar scale, which does not contain a midpoint, will be noted with a -.

In the following the 23 derived scales (noted *S*), along with the specific scale question (SQ), will be presented in corresponding order as presented for the test subjects in the second test and sorted by each page. An example of each set of scales is given, to illustrate how they were presented in the test.

A. Page 1

SQ01: How do you think the screen on the robot reacted?

SQ02: How did you experience the robot?

SQ03: How was it to use the robot?

SQ04: How did you experience the movements of the robot?

See Table I and Figure 4 for how the scales were presented.

TABLE I: Scale labels page 1

S	Left label	Midpoint	Right label
1	Extremely bad (Ekstremt dårligt)	No label	Extremely well (Ekstremt godt)
2	Extremely unwelcoming (Ekstremt avisende)	No label	Extremely welcoming (Ekstremt imødekommennde)
3	Extremely difficult (Ekstremt svært)	No label	Extremely easy (Ekstremt nemt)
4	Extremely wild (Ekstremt vilde)	No label	Extremely calm (Ekstremt rolige)

Hvordan synes du, at skærmen på robotten reagerede?

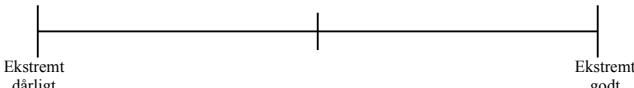


Fig. 4: Example of a bipolar scale relevant for the scale question: *How do you think the screen on the robot reacted?*, rated from *Extremely bad* to *Extremely well*.

B. Page 2

SQ05: I think that the robot stopped...

SQ06: I think that the speed of the robot is...

SQ07: I think that the height of the robot is...

See Table II and Figure 5 for how the scales were presented on page 2.

TABLE II: Scale labels page 2

S	Left label	Midpoint	Right label
5	Way too close (Alt for tæt på)	No label	Way too far (Alt for langt fra)
6	Way too slow (Alt for langsom)	Fine (Fin)	Way too fast (Alt for hurtig)
7	Way too low (Alt for lav)	Fine (Fin)	Way too high (Alt for høj)

Jeg synes, at robotten stoppede...

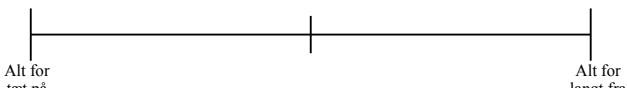


Fig. 5: Example of a bipolar scale relevant for the scale question: *I think that the robot stopped...*, rated from *Way too close* to *Way too far*.

C. Page 3 and 4

SQ08: I feel that the robot can help me

SQ09: I think that the robot was obstructing me

SQ10: I feel safe around the robot

SQ11: The robot startled me

SQ12: I like to be served by the robot

SQ13: I relied on the robot to lead me to the location I chose
See Table III and Figure 6 for how the scales were presented on page 3 and 4.

TABLE III: Scale labels page 3 and 4

S	Left label	Midpoint	Right label
8-13	Completely disagree (Helt uenig)	No label	Completely agree (Helt enig)

Jeg føler, at robotten kan hjælpe mig

A horizontal bipolar scale with two vertical tick marks. The left tick mark is labeled "Helt uenig" and the right tick mark is labeled "Helt enig". There is a small gap between the tick marks and the scale line.

Fig. 6: Example of a bipolar scale relevant for the scale question: *I feel that the robot can help me*, rated from *Completely disagree* to *Completely agree*.

D. Page 5

SQ14: How personal did you experience the help of the robot?

SQ15: How surprised were you by the robot's approach?

For each of the two scale questions the chosen labels are listed in Table IV. Also see Figure 7 for an example of the scale used.

TABLE IV: Scale labels page 5

S	Left label	Midpoint	Right label
14	Not at all personal (Slet ikke personlig)	-	Extremely personal (Ekstremt personlig)
15	Not at all surprised (Slet ikke overrasket)	-	Extremely surprised (Ekstremt overrasket)

Hvor personlig oplevede du robottens hjælp?



Fig. 7: Example of an unipolar scale relevant for the scale question: *How personal did you experience the robots help?*, rated from *Not at all personal* to *Extremely personal*.

E. Page 6

SQ: What do you think about the robot?

This question covers scales 16-19. The scale used was similar to the one shown in Figure 7. See Table V for the labels used.

TABLE V: Scale labels page 6

S	Left label	Midpoint	Right label
16	Not at all annoying (Slet ikke irriterende)	-	Extremely annoying (Ekstremt irriterende)
17	Not at all elegant (Slet ikke elegant)	-	Extremely elegant (Ekstremt elegant)
18	Not at all exciting (Slet ikke spændende)	-	Extremely exciting (Ekstremt spændende)
19	Not at all cute (Slet ikke sød)	-	Extremely cute (Ekstremt sød)

F. Page 7

Q: What else do you think about the robot?

This question covers scales 20-23. The scale used was similar to the one shown in Figure 7. See Table VI for the labels used.

TABLE VI: Scale labels page 7

S	Left label	Midpoint	Right label
20	Not at all cool (Slet ikke dej)	-	Extremely cool (Ekstremt dej)
21	Not at all intrusive (Slet ikke anmassende)	-	Extremely intrusive (Ekstremt anmassende)
22	Not at all funny (Slet ikke sjov)	-	Extremely funny (Ekstremt fun)
23	Not at all human (Slet ikke menneskelig)	-	Extremely human (Ekstremt menneskelig)

Based on the affinity diagram a 24th attribute was derived, this attribute will not be presented along side the aforementioned scales but will be included in a separate demographic page as it does not directly concern the robot. The attribute is formulated in the scale question: *How fond of technology are you?*, which will be evaluated on an unipolar scale, similar to the other unipolar scales, with anchor points: *Not at all fond* (slet ikke glad) and *Extremely fond* (ekstremt glad).

When comparing the attributes for HRI found in this study with attributes for HRI from previous conducted studies on social robots [1], [2], [3], [8], [9], [10], [11], attributes such as distance, anthropomorphism, height, speed, movement, trust, enjoyment, technological knowledge, and usefulness reoccur. New attributes were found compared with previous mentioned studies, these are: Elegance, cuteness, coolness, startling, excitement, welcoming, obstructive, whether the robot help is seen as personal, and whether the encounter with the robot was surprising. However, these might be measured indirectly in the aforementioned studies.

According to [12] *social distance* has an effect on HRI. The subjects in AAL did not mention it specifically, which might be due to *power distance*, where the subjects feel in control as they were the dominant part and the robot was the subordinate in the interaction. Furthermore this could also be due to the given task being a cooperative task, as the robot's purpose was to help the subject to a specific location of their choosing. *Task distance* in cooperative tasks might incline subjects to act more friendly, intimate, and involved in the interaction with the robot, which was supported by all the positive comments from the subjects.

IV. PART II - SCALE TESTING

This section describes the second part of the study where the scales were tested at AAL. The purpose was to get the travellers to rate their experience with the robot on the developed scales.

V. METHOD - SCALE TESTING

The test was conducted over the course of two days and in the same way as the first test described in section II.

However, in this test instead of the robot leading them directly to an interviewer, it led them in the direction of their chosen location. After a short distance, an experimenter politely stopped them and informed them about the ongoing study. This was done to give the subjects a more natural experience, where the robot appeared to lead them in the right direction. After they accepted to participate, they were led to a PC nearby to rate their experience on the developed scales. The experimenter was ready to take notes while they rated. The order in which the scales were presented is described in section III. The physical variables of the robot were varied over the course of the test. This included the angle of approach, the robot's height (and speed), and the distance to the subjects. As a consequence of the study being ecologically performed in an actual airport it was not possible to precisely control the distance to the subjects or the angle of approach. These had to rely on subjective judgement from the researchers. Height was measured using a measuring tape. The robot height was set to either of the following five heights: 118 cm, 123.5 cm, 129 cm, 140 cm, or 151 cm. The first and last mentioned heights are dictated by the robot's minimum and maximum height, whereas 129 cm was chosen according to what the subjects in previous study mentioned as fine, which roughly corresponds to human elbow height according to the mean height of the Danish population.

A. Materials

The same materials were used as those used in the first test described in section II along with the same *Double* and tilted head mount. The same wireframe was used except with a few adjustments. Additionally, software was developed in Processing 3.3.6 (www.processing.org) to be able to collect data from the scales. The program was presented on a *Microsoft Surface Pro* (5) with a wireless mouse. The 24th scale was presented on a sheet of paper along with a few other questions to collect demographic information.

B. Scale Program

The scale program consisted of 23 scales distributed on seven pages. The number of scales presented on a single page were maximum four and minimum two. The subjects were instructed to set a marker representing their response on the scales using the provided mouse. The scales were organised to be as consistent on each page as possible e.g. the same type of scales with a midpoint were presented on one page and internally the scales were organised such that two similar attributes did not appear right after each other.

C. Subject Recruitment

Over the course of two days 43 subjects participated. They were aged from 10 to 72 ($M=40.1$, $SD=13.4$) and distributed among 16 females and 27 males. Subjects reported that they travel between 1 and 100 times per year ($M=15.3$, $SD=18.1$). The subjects were again recruited by the robot with the same wireframe as in the previous test.

D. Roles

Four researchers were present during the test at AAL. One controlled the *Double* robot, one instructed the subjects to answer the scales, one observed when the robot should start leading the subjects towards the shopping area and one noted the physical variables of the robot such as height, direction of approach, and distance to the subjects. As with the previous test the robot controller had no way to know what happened on the screen, one of the observers signaled to the robot controller when to start leading the subjects to towards the shopping area.

E. Data Processing

Data was gathered in Processing 3.3.6 and for each subject data were saved in a .csv format and analysed using MATLAB, Excel, and RStudio. Data were analysed using boxplots and Principal Component Analysis (PCA), which were further analysed in order to compare correlated attributes. The PCAs were conducted on robot's height, distance, and direction separately in order to assess how the subjects responded to the physical changes of the robot.

VI. RESULTS - SCALE TESTING

Due to a programming error the dataset contains several zero-answers which could happen when subjects did not answer on a scale. It was decided to remove the missing answers. Two criteria were used to determine when to exclude some of the zero-points that occurred because of a non-answered scale. 1) If all scale ratings presented on one of the pages are zero, it was probably because subjects skipped a page and 2) If there is a strong tendency for high ratings on a scale and a single zero-answer seems unlikely, it was denoted as a missing answer.

The different ratings on the scales in the second test are visualised on Figure 8.

From Figure 8 it is notable that there exists much greater variation in some scales compared to others. For some SQs, a big part of the scale is used (1, 4, 8, 12, 14, 19), where for others a smaller part of the scale is used (2, 3, 5, 7, 9, 11, 13, 16). Furthermore some of the scale ratings appears normal distributed (1, 2, 6, 17, 20, 21) and others appears skewed (8, 9, 10, 11, 13, 18). In some SQs the data accumulates around midpoints or anchor points (5, 7, 9, 10, 11, 13, 16). SQ05 and SQ07 are centralised around the midpoint, which for SQ07 can relate to the label *fine* being too broadly describing. The closed anchor points used on the scales can induce accumulation of data points around these. Also the use of bipolar scales poses the risk of picking two words which are not each others opposite.

When looking at gender differences it seems that women rated SQ04, SQ11, and SQ21 lower than men. It suggests that females experienced the movements of the robot more calm compared to males, but also that they find the robot less intrusive and startling.

When conducting a PCA on all the scale results it showed that only 53.1 % of the total variance is explained when using the first three dimensions, which is why PCA relating to different groups such as the robot's height, distance, and direction are conducted in order to reduce the dimensions and to be able to explain more of the variance. Figure 9a shows a Scree plot from PCA relating to the robot's height, where 94.1 % of the variance is explained using three dimensions. Figure 9b shows the Biplot relating to the robot's height. Similar Biplots were made for direction and distance. The attributes which loads the most on the components are:

Height: SQ19 appears to contribute the most to PC1. SQ01 the most to PC2. SQ06 the most to PC3.

Distance: SQ01 appears to contribute the most to PC1 along with SQ15. SQ10 contributes most to PC2.

Direction: SQ10 contributes most to PC1. SQ22 and

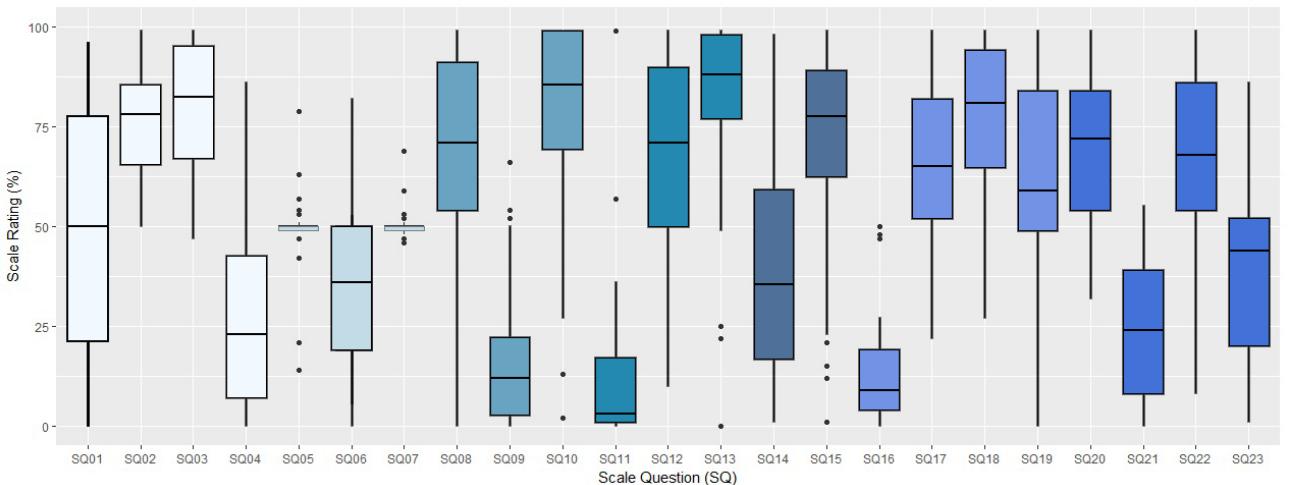


Fig. 8: Boxplot showing the 23 attributes. The boxplot shows the median, where the box is ranging from 25-75 % of the data, and the whiskers from 0-25 % and 75-100 % of the data. Each colour represents a new page of scales.

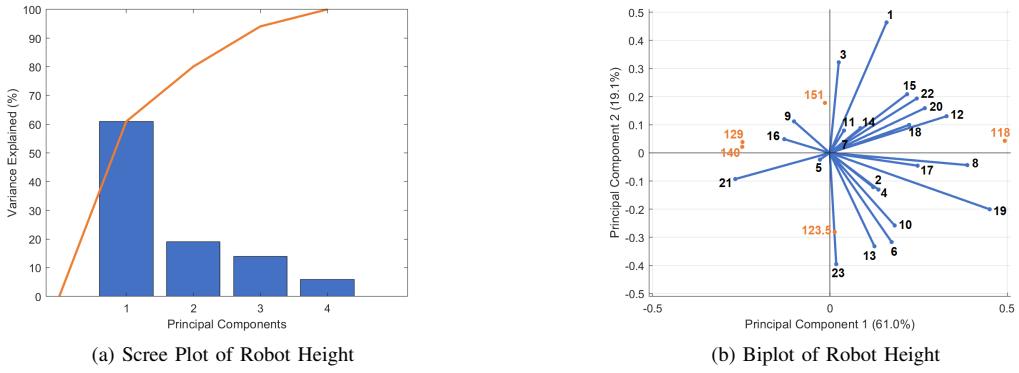


Fig. 9: (a): Scree plot showing the connection between the number of Principal Components and Variance Explained (%); PC1 (61 %), PC2 (19.1 %), PC3(14 %), PC4(5.9 %). (b): Biplot showing how the different attributes contributes to components and which attributes that correlates. The black numbers denotes SQ and the red to the different heights in cm.

SQ23 contributes almost equally to PC2. SQ13 and SQ20 contributes almost equally to PC3.

2D and 3D Biplots from the different PCAs shows positive and negative correlations between different attributes. These correlations are presented in Table VII.

TABLE VII: Correlations from PCA

PCA	Positive correlation	Negative correlation
Height	SQ08 + SQ17	SQ02 + SQ09
	SQ10 + SQ13	SQ04 + SQ12
	SQ12 + SQ18	SQ12 + SQ21
	SQ14 + SQ15	SQ16 + SQ19
	SQ20 + SQ22	SQ18 + SQ21
Distance	SQ01 + SQ12	SQ02 + SQ09
	SQ07 + SQ17	SQ05 + SQ21
	SQ08 + SQ21	SQ10 + SQ13
	SQ10 + SQ22	SQ13 + SQ22
		SQ14 + SQ16
		SQ19 + SQ20
Direction	SQ05 + SQ07	SQ01 + SQ12
	SQ08 + SQ10	SQ06 + SQ23
	SQ09 + SQ14	SQ09 + SQ10
	SQ18 + SQ20	SQ10 + SQ14
		SQ13 + SQ21

To investigate these correlations further, ratings from correlating SQs are compared, as seen in Figure 10. Even though correlations occur when doing PCA on groupings, these plots of comparison do not take groupings into consideration. This might be why some correlations are not found when comparing SQs, even though the Biplots showed a correlation. From Figure 10 it can be seen that the robot is perceived more exciting when subjects like to be served by the robot. Attributes that have a clear correlation from this kind of comparison are shown in Table VIII.

TABLE VIII: Correlations from graphs

Positive correlation	Negative correlation
SQ04 + SQ09	SQ02 + SQ09
SQ08 + SQ10	SQ09 + SQ10
SQ08 + SQ17	SQ12 + SQ21
SQ10 + SQ13	SQ13 + SQ21
SQ12 + SQ18	SQ18 + SQ21
SQ18 + SQ20	SQ20 + SQ22

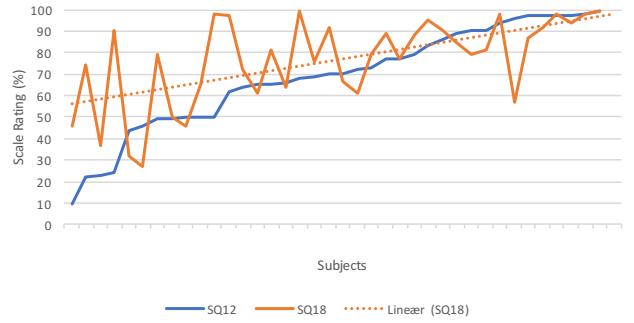


Fig. 10: Comparison between ratings on SQ12 and SQ18 based on 41 subjects. Two were removed due to incomplete datasets.

The physical variables can be compared with SQs as well as the demographic parameters in the same manner as comparing SQs with each other, to see how these variables and parameters affect the ratings of the attributes. Figure 11 were SQ06 is compared with the robot's height shows that when the robot's height increases the robot is perceived as being less annoying. When comparing SQ06 with age, the younger a subject is, the more inclined one is to rate the robot as moving too slow. Also, a small negative correlation was noted in relation to how exciting and funny the robot seemed. The older the subjects were, the less exciting and funny they rated the robot. Within height the smaller the robot was the faster and more wild it was rated. This was expected, as the robot automatically moves more slowly when the height increases and faster when the

height decreases. Additionally it was rated more cute, elegant, and trustworthy, regarding to wayfinding, when it was small.

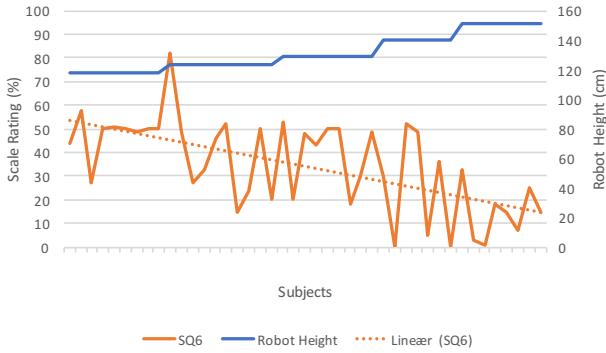


Fig. 11: Comparison between the robot's five heights (cm) and ratings (%) on SQ06 based on all subjects.

VII. DISCUSSION

A. Equipment

Due to the Marvel wireframe running inside the *Double* application, the touch screen was less responsive than expected. This resulted in some travellers not wanting further interaction with the robot. It also affected the experience for those who chose to interact with it anyway.

The scaling software written for the second test had some response issues as well, which ultimately led to incomplete datasets because the program often failed to notice mouse-clicks. This is not ideal for a scale which relies on an immediate impulse response without much thought. This have almost definitely changed some subjects' responses in either direction of the scale due to them clicking in another location on the scale than initially chosen.

B. Interview

The interview in the first test was supposed to be very open and allow the subjects to talk freely. In reality, however, it ended up being a bit more structured, with the subjects directly answering the questions they were asked and waiting for the next one. This introduced some interviewer bias, since the predetermined questions dragged the conversation in a certain direction, and meant subjects talked a lot about a specific aspect of the interaction, because they were asked directly about it.

C. Treatment of data

In the first test spontaneous answers, answers to specific questions, and observations were all analysed together in an affinity diagram. It could be argued that these answers are two different categories, where only the spontaneous answers are the subjects' own words and areas of interest. The subjects did not provide as many spontaneous answers as hoped, hence all answers and observations were analysed together to gain more insight.

If a scale was not answered in the second test the software returned 0, which also was a valid scale response for the subjects. After looking through the data it was decided to remove some of the zeros, which seemed to stem from the unresponsive software bug. This might have led to real zero-responses being excluded from the dataset and for missing answers being included.

D. First impression

The results of this study were based entirely on the subjects' first impression of the robot. The long term experience of interacting with the robot remains unknown. Several subjects interacted with the robot, because it was new and exciting and made them curious, which might be explained by the novelty effect. Their user experience and focus points might change relative to their familiarity with the robot and therefore different attributes might be elicited when conducting a study not only regarding first impression.

E. Context

Some of the attributes which were elicited might only appear because the study was conducted at AAL, where the *Double* robot helped with specific tasks. These attributes are considered to be important for travellers at AAL and can be used when evaluating a social robot in this specific context. However, these are not attributes which can easily be adjusted by a designer such as how the subjects perceive the robot (elegant, cute, welcoming, etc.). Though this might be true, it is found that the robot was rated more cute, elegant, and trustworthy, regarding to wayfinding, when it was small. On the basis of these correlations the subjective experience might be improved by changing physical, easily adjustable characteristics such as height.

F. Labels and scale questions

It appears that there was some confusion regarding some of the scales e.g. where SQ05 was misinterpreted. The question was meant as to how close the robot stopped in relation to the subject, but some subjects thought it was meant in relation to their chosen location which they obviously could not answer. A few also commented on the attributes funny and cute. Some subjects expressed that the robot was funny, but not humorous funny and therefore had trouble rating the scale question because of this duality. Another subject commented that she refused to rate the robot as cute, because she regarded it as a human quality and that she refused to humanise the robot.

For the scale rating to SQ02, none of the subjects rated under 50 %. Half of the scale, regarding unwelcoming, might therefore be redundant. Instead an unipolar scale regarding welcoming could be developed and this would give a wider range to answer on the scale to the attribute welcoming.

The low variance and centering around the midpoint on some of the scales might be that the subjects did not find the scale questions relevant or did not understand them and

therefore chose a neutral rating. It might also be that the selected stimuli was not varied enough. This could be reinforced by the midpoints showing the subjects where the middle is. If the ratings are constantly located at the same point, they do not contain much information about how different variables and parameters affect the experience of interacting with the robot.

G. Redundancy

Some of the attributes might be redundant if they correlate with other attributes having similar overall characteristics. This relates to attributes such as cool and funny, which might be redundant when also asking subjects about how cute and exciting they found the robot. This is due to the correlations, but also the word funny, which can be widely used, e.g. as humorous funny or weird funny. Attributes like how safe subjects felt using the robot could be combined with how much they relied on the robot leading them to the chosen location under a shared trust attribute. Asking subjects how trustworthy they found the robot and how much they felt it could help them could then give an impression on the three attributes by only asking two questions. Furthermore, attributes such as SQ12 (like being served by the robot) might be redundant when asking subjects how exciting and intrusive they found the robot.

VIII. CONCLUSION

The research conducted in this study revealed 10 main categories relating to HRI. Attributes were found for each of these 10 categories which describe the user experience of interacting with a social robot at Aalborg Airport. The observations and the 30 subjects' statements were interpreted and coded using an affinity diagram. 567 affinity notes were sorted by a bottom-up procedure into 10 categories which revolved around appearance, trust, behaviour, approach, problems with the touch screen, avoidance of interaction, personal interest, positivity towards the robot, usefulness, and tech-experience.

New attributes within the 10 categories were discovered compared to previous mentioned studies: Elegance, cuteness, coolness, startling, excitement, welcoming, obstructive, whether the robot help is seen as personal, and whether the encounter with the robot was surprising. Besides the newly found attributes some attributes were found to match previously found attributes. These attributes regards distance, anthropomorphism, height, speed, movement, trust, and usefulness.

23 attributes were elicited and scales were developed from these. After the second test, 10 instances of correlations were found for both height and distance and 9 for direction based on PCA with different groupings. 12 instances of correlation were found when comparing attributes that correlates when conducting PCA.

Based on the results in this study, it is suggested that the scale questions: *I relied on the robot to lead me to the location I chose* and *I feel safe around the robot* are combined to one attribute named: Trust. Also funny and cool are described by

exciting and cute and thus seem redundant. Finally, the scale question: *I like to be served by the robot* is also described by intrusive and exciting, thus also seem redundant.

The different types of scales used in this study could explain the difference in variance between the scales. The closed endpoints might affect how much the results accumulate at the ends. The label *fine*, the use of midpoints, and scale questions such as *I think the robot stopped...* should be reconsidered to improve the user's understanding of the scale and minimize misunderstandings.

ACKNOWLEDGMENT

We would like to thank Karl Damkjær Hansen, postdoc at Aalborg University for giving professional insights and feedback, helping with technical support, and lending the *Double* robot to the authors. We would also like to thank Professor Dorte Hammershøi for supervising the study. Lastly we would like to thank Aalborg Airport for providing access to the wanted user group and facilities by letting the authors conduct their field studies at the airport.

REFERENCES

- [1] M. M. de Graaf and S. B. Allouch, "Exploring influencing variables for the acceptance of social robots," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1476 – 1486, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0921889013001334>
- [2] M. M. de Graaf, S. B. Allouch, and T. Klamer, "Sharing a life with harvey: Exploring the acceptance of and relationship-building with a social robot," *Computers in Human Behavior*, vol. 43, no. Supplement C, pp. 1 – 14, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0747563214005536>
- [3] J. E. Katz, D. Halpern, and E. T. Crocker, In *The Company of Robots: Views of Acceptability of Robots in Social Settings*, J. Vincent, S. Taipale, B. Sapiro, G. Lugano, and L. Fortunati, Eds. Springer, 2015.
- [4] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, E. A. Sisbot, R. Alami, and T. Siméon, "How may i serve you? a robot companion approaching a seated person in a helping context," pp. 172–179, 2006.
- [5] D. Robotics. (2017) Work from anywhere. [Online]. Available: <https://www.doublerobotics.com>
- [6] Apple. (2017, 11) ipad air - technical specifications. [Online]. Available: https://support.apple.com/kb/SP692?viewlocale=en_US&locale=da_DK
- [7] J. B. Wendell, K. Holtzblatt, and S. Wood, *Rapid Contextual Design*. Elsevier LTD, Oxford, 2005. [Online]. Available: http://www.ebook.de/de/product/3268220/jessamyn_burns_wendell_karen_holtzblatt_shelley_wood_rapid_contextual_design.html
- [8] D. Halpern and J. E. Katz, "Close but not stuck: Understanding social distance in human-robot interaction through a computer mediation approach," vol. 1, pp. 17–34, 2013.
- [9] V. Bogicevic, M. Bujicic, A. Bilgihan, W. Yang, and C. Cobanoglu, "The impact of traveler-focused airport technology on traveler satisfaction," pp. 351–361, 2017.
- [10] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, "Human-robot embodied interaction in hallway settings: a pilot user study," pp. 164–171, 2005.
- [11] M. Shiomi, K. Shinozawa, Y. Nakagawa, T. Miyashita, T. Sakamoto, T. Terakubo, H. Ishiguro, and N. Hagita, "Recommendation effects of a social robot for advertisement-use context in a shopping mall," vol. 5, pp. 251–262, 2013.
- [12] Y. Kim and B. Mutlu, "How social distance shapes human-robot interaction," *Int. J. Human-Computer Studies*, vol. 72, pp. 783–795, 2014.

Læsevejledning

Arbejdsbladene bør læses kronologisk, da nogle afsnit antager, at læseren har kendskab til tidligere afsnit i rapporten. Derudover er rapporten struktureret således, at resultater og viden løbende diskuteres og konkluderes.

Henvisninger

Alle henvisninger i rapporten fungerer som hyperlinks, som illustreret i eksemplet her: Figur 4.11.

Kildehenvisninger

Kildehenvisninger angives enten som en del af teksten eller i parentes. Et eksempel på de to kildehenvisningsmetoder: Taipale m.fl. (2015, s. 13) eller (Taipale m.fl. 2015, s. 13). Såfremt der refereres til en bestemt del af kilden, angives dette med sidetal, eksempelvis; s. 13 for én bestemt side eller ss. 1-3 for flere sider.

Afsnitshenvisning

Afsnitshenvisninger angives med et afsnitsnummer efterfulgt af et afsnitsnavn. Et eksempel på en afsnitshenvisning: Afsnit 1.1 (*Karakterisering af social robot*). Samme gør sig gældende for kapitler.

Figurhenvisning

Henvisninger til figurer angives med et decimaltal, som først gengiver kapitlets nummer efterfulgt af figurnummeret i det pågældende kapitel. Et eksempel på en figurhenvisning: Figur 1.1, der svarer til figur 1 i kapitel 2.

Henvisninger til elektronisk bilag

Henvisninger til det elektroniske bilag angives med en afsnitshenvisning, hvor stien til det pågældende bilag fremgår i det pågældende afsnit.

Arbejdsblade

Del I Introduktion

Kapitel 1 Sociale robotter	3
1.1 Karakterisering af social robot	3
1.1.1 Eksisterende sociale robotter	4
1.2 Projektsamarbejde	5
1.3 Teknologier i lufthavne	7
1.4 Interaktion med sociale robotter	8
1.4.1 Generelle tendenser	8
1.4.2 Parametre, der har indflydelse på accept og interaktion med sociale robotter	9
1.4.3 Bevægelsesmønstre og udseende	14
Kapitel 2 Projektafgrænsning	21
2.1 Udfordringer ved sociale robotter	21
2.2 Problemformulering	22

Del II Parametre, der har indflydelse på HRI

Kapitel 3 Feltundersøgelse	25
3.1 Metodeovervejelse	25
3.1.1 Contextual Inquiry	25
3.1.2 Laddering	26
3.1.3 Overvejelser til interviewer	27
Kapitel 4 Testdesign	29
4.1 Testens omfang	29
4.2 Testpersoner	29
4.3 Brugsscenarier	30
4.4 Samtaleemner	33
4.4.1 Løbende samtaleemner	33
4.4.2 Afsluttende samtaleemner	34
4.5 Rollefordeling	34
4.6 Testlokation og udstyr	35
4.7 Fremgangsmåde	39
4.8 Pilottest	40
4.8.1 Styring af robotten	41
4.8.2 Brugergrænseflade	41
4.8.3 Testdesignet	41
Kapitel 5 Databehandling	43

5.1	Ændringer af testdesign	43
5.2	Testpersoner	44
5.3	Databehandlingsmetode: Affinity diagram	44
5.3.1	Fortolkningssession	45
5.3.2	Udviklingen af <i>affinity diagram</i>	46
5.4	Valg af skala type	47
5.4.1	Endepunkter	47
5.4.2	Bipolær og unipolær	47
5.4.3	Labels	48
5.5	Resultater fra Affinity diagram	48
5.5.1	Interagerer ikke med R	49
5.5.2	Skærmens virker ikke	50
5.5.3	R kan assistere mennesker	51
5.5.4	R's væremåde	53
5.5.5	Henvendelse	54
5.5.6	R's udseende	56
5.5.7	Interesse for R	57
5.5.8	Positiv overfor R	58
5.5.9	Kendskab til teknologi	59
5.6	Udvælgelse af skalaer	63
5.7	Sammenligning mellem fundne parametre og relaterede studier	72
Kapitel 6	Diskussion	75
6.0.1	Observationer og resultater fra feltundersøgelsen	75
6.0.2	Skalaer	77
Kapitel 7	Delkonklusion	79
Del III Videreudvikling af skalaer til evaluering af HRI		
Kapitel 8	Tilpasning af skalaer	83
8.1	Tilpasning og rækkefølge på skalaer	83
8.2	Program til skalaer	91
Kapitel 9	Testdesign	95
9.1	Testens omfang	95
9.2	Brugsscenarier	96
9.3	Robottens bevægelse	96
9.4	Rollefordeling	98
9.5	Demografi	100
9.6	Testlokation og udstyr	101
9.7	Fremgangsmåde	101
9.8	Pilottest	102
9.9	Ændringer efter pilottest	104
Kapitel 10	Databehandling	107
10.1	Ændringer af testdesign	107

10.2 Testpersoner	108
10.3 Observationer i AAL	108
10.3.1 Observationer fra d. 01-12-2017	109
10.3.2 Observationer fra d. 05-12-2017	111
10.4 Skala Oversigt	114
10.5 Frasortering af manglende data	115
10.6 Fordeling af besvarelser	117
10.6.1 Standardisering af besvarelser	120
10.7 Principal Component Analysis	121
10.8 Principal Component Analysis: Robottens højde	125
10.8.1 Tendenser i forhold til robottens højde	130
10.9 Principal Component Analysis: Afstand	135
10.9.1 Tendenser i forhold til robottens afstand	138
10.10 Principal Component Analysis: Indgangsvinkel	139
10.10.1 Tendenser i forhold til robottens indgangsvinkel	144
10.11 Sammenligning af korrelerede parametre	145
10.11.1 Korrelerede parametre fra højde	145
10.11.2 Korrelerede parametre fra afstand	151
10.11.3 Korrelerede parametre fra indgangsvinkel	155
10.12 Demografisk indflydelse	159
10.12.1 Kønsforskelle	164
Kapitel 11 Diskussion	165
11.1 Skalaerne	165
11.2 Testpersoner	166
11.3 Fysiske parametre	167
11.4 Redundans	168
Kapitel 12 Delkonklusion	171
Kapitel 13 Elektronisk bilag	175
13.1 Transskriberede data	175
13.2 Affinity notes	175
13.3 Program til VAS	175
13.4 Skala Oversigt	175
13.5 Rådata i Excel	175
13.6 Hisogram og normalfordelings plot	175
13.7 MATLAB Scripts til PCA	176
13.8 3D Bi-plots	176
13.9 Tendenser	176
13.10 Korrelationsgrafer	176
Bibliografi	179

Del I

Introduktion

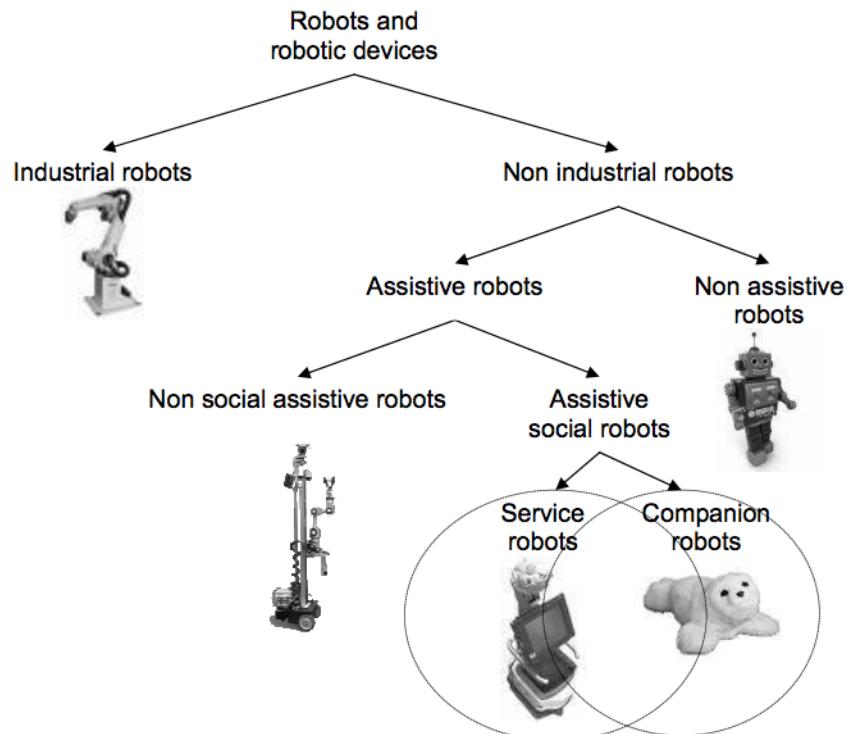
I takt med at flere og flere mennesker benytter sig af lufttransport, lægger det større og større pres på lufthavnene om at fornye sig, (Bogicevic m.fl. 2017, s. 351). Fornyelserne bør særligt foretages omkring integrering af nye, innovative teknologier, som assisterer og informerer de rejsende omkring deres flyafgang samt muligheder i lufthavnen, (Bogicevic m.fl. 2017, s. 352). Da dette projekt er et samarbejde med Karl Damkjær Hansens, der blandt andet arbejder med sociale robotter og som har indgået i et samarbejde med Københavns Lufthavn, vil det blive undersøgt, hvordan en social robot kan integreres i Københavns Lufthavn.

Sociale robotter 1

I det følgende kapitel vil der først blive specificeret hvad der karakteriserer en social robot og herunder blyses det hvilke eksisterende sociale robotter, der allerede er på markedet. Dernæst vil projektsamarbejdet blive yderlige specificeret. Efterfølgende fokuseres der på teknologier i lufthavne, hvorefter interaktion og udfordringer ved sociale robotter blyses. Afslutningsvist ledes der over i en problemformulering, som danner grundlag for det fremadrettede projektarbejde.

1.1 Karakterisering af social robot

Sociale robotter adskiller sig fra industrielle robotter, da det overordnede formål er, at de skal tage sig af mennesker, (Taipale m.fl. 2015, s. 13). Sociale robotter skal særligt tage sig af de svage i samfundet; ældre, handicappede, syge og børn, (Taipale m.fl. 2015, s. 14). Formålet med industrielle robotter er, at de kan udfører farligt og gentagende arbejde, hvorfor de potentielt kan rede menneskeliv, (Taipale m.fl. 2015, ss. 12-13). På Figur 1.1 illustreres en kategorisering af robotter, hvor det blandt andet fremgår at industrielle robotter og sociale robotter ikke tilhører den samme robottype.



Figur 1.1. Kategorisering af robotter fremsat af Heerink (2010, s. 13).

Da projektet fokuserer på sociale robotter, *Assistive social robots*, som skal indgå i en kontekst blandt mennesker, afgrænses der fra industrielle robotter, *Non assistive robots* samt *Non social assistive robots*. Sidstnævnte robottype er en form for fysisk teknologisk hjælpemiddel, der kan anvendes til rehabilitering, det kan eksempelvis være proteser eller intelligente kørestole, (Heerink 2010, s. 12).

Ifølge Graaf, Allouch og Klamer (2015, s. 1) kan sociale robotter karakteriseres ved, at de har forståelse og evne til at kommunikere på en menneskeagtig måde, hvilket gør det muligt for deres brugere at forstå dem. I tillæg argumenterer Breazeal (2003, s. 168) for, at når mennesker enten observerer eller interagerer med en autonom robot tildeles den en social model. For at en autonom robot kan kategoriseres som en social robot skal den, ifølge Breazeal (2003, s. 168), percipere omgivelserne, tage selvstændige beslutninger og udføre koordinerede handlinger for at løse deres opgave. Da sociale robotter designes til at interagere med mennesker, via simpel kommunikation, vil det øge brugerens accept af robotten (Graaf og Allouch 2013, s. 1476). Ydermere forudser Graaf og Allouch (2013, s. 1476), at sociale robotter i højere grad vil blive en del af vores hverdagsliv, og det er derfor nødvendigt at undersøge, hvordan mennesket perciperer sociale robotter og hvorfor de enten accepterer eller afviser dem (Graaf, Allouch og Klamer 2015, s. 1).

1.1.1 Eksisterende sociale robotter

Der findes et stort udvalg af sociale robotter når det kommer til børnelegetøj, blandt disse er den nok bedst kendte Sony's *AIBO*, som er en robot hund, (SONY 2017). *AIBO* er designet til at udvikle sin personlighed fra hvalp til voksen, afhængigt af interaktionen med dens ejer og omgivelser, (SONY 2017). *AIBO* illustreres på Figur 1.2. Selvom *AIBO* ikke er en virkelig hund, afspejler den hundes naturlige bevægelsesmønstre og behov, (Arkin m.fl. 2003, ss. 191-198). Det resulterer i, at ejeren tildeler *AIBO* hunde-agtige egenskaber, hvorved den behandles som en social ledsager med mentalkapacitet, (Graaf, Allouch og Klamer 2015, s. 2). Dette kommer blandt andet til udtryk ved en analyse af kommentarer i et forum, hvor 47 % tildeler *AIBO* en biologisk essens, 42 % vurderer at *AIBO* har forsættlig adfærd, 38 % vurderer at *AIBO* har følelser og 39 % vurderer at *AIBO* kan opdrages, udvikles og modnes, (Katz m.fl. 2015, s. 26). En anden social robot, som ligeledes anvendes blandt børn er *NAO*, designet af S. Robotics (2016a). *NAO* er en robot, der kan bruges til undervisning. *NAO* illustreres på Figur 1.2.



Figur 1.2. Oversigt over et udvalg af eksisterende sociale robotter.

I forbindelse med autistiske børn tyder det ligeledes på, at sociale robotter kan anvendes som en del af barnets terapi, (Dautenhahn og Billard 2002, s. 180). Ifølge Dautenhahn og Billard (2002, s. 185) er en af fordelene ved at anvende en robot, *Roberta*, at den er mindre kompleks end mennesker, hvilket gør det muligt for robotten at hjælpe barnet med at udvikle sociale kompetencer.

Ydermere kan sociale robotter ligeledes indgå som en form for terapi for ældre, som eksempelvis lidet at demens eller oplever ensomhed, (Vincent 2015, s. 110). *PARO*, som er en robot sæl, anvendes blandt andet i disse situationer. Ifølge Robots (2014), reducerer *PARO* patientens stress, gør dem mere afslappet og motiveret og forbedre sociale egenskaber. *PARO* illustreres på Figur 1.2. Udoer *PARO*, findes *ROMEO*, som er designet til at hjælpe ældre og andre der har mistet en del af deres mobilitet, (S. Robotics 2016b). *ROMEO* illustreres på Figur 1.2. I forhold til de før nævnte robotter er *ROMEO* forholdsvis stor; 140 cm høj. Ved at introducere sociale robotter, designet til at hjælpe ældre, kan det være med til at sænke belastningen på sundhedssystemet i takt med at den ældre har mulighed for et længere uafhængigt og sundere liv i sit eget hjem, (Graaf, Allouch og Klamer 2015, s. 1).

Double tillader at medarbejdere kan arbejde hvor som helst og gennem robotten være til stede på arbejdspladsen samtidig, (D. Robotics 2017). *Double* er illustreret på Figur 1.4. En lignende form for telekommunikation opleves med *HUGVIE*, som er en menneskeformet pude med hovede, arme, krop og ben, (Kaerlein 2015, s. 78). *HUGVIE* illustreres på Figur 1.2. I *HUGVIE*s hovede er det muligt at lægge en mobiltelefon i en lomme og dermed fører en samtale, eksempelvis med ens partner, samtidig med at brugeren krammer *HUGVIE* og mærker dens hjertebanken, som afspejler den transmiterede lyd.

Derudover anvendes sociale robotter som turguider og receptionister på museer, (Halpern og Katz 2013, s. 22). Ved robotter såsom *Rhino* og *Minerva*, har fokus primært været på at udvikle navigation og undvigelsesmanøvre, (Karreman m.fl. 2015, s. 318). Ifølge Karreman m.fl. (2015, ss. 318-319) fokuserer tidligere undersøgelser, med forskellige sociale robotter brugt på museer, på et enkelt aspekt og ikke på evalueringen af hele brugeroplevelsen.

1.2 Projektsamarbejde

Projektet dækker over et samarbejde mellem projektgruppen, inklusiv vejleder, og Karl Damkjær Hansens Ph.d., postdoc på Aalborg Universitet. Karls arbejde involverer udviklingen af sociale robotter, som skal indgå i en menneske-robot interaktion, *Human-Robot Interaction*, HRI. Dertil fokuseres der på, hvordan den sociale robot skal henvende sig til interaktionspartneren(e), hvor Karls primære arbejdsområde dækker det tekniske aspekt. Da sociale robotter skal kunne indgå i HRI er det en fordel at robotten afspejler menneskelig adfærd, eksempelvis i forhold til bevægelsesmønstre, hvor målet dels er at få robotten til at udføre dynamiske bevægelser og dels tillade robotten friheden til at bevæge sig 360°. Ved at opfylde de to mål bør det være muligt at få robotten til at gengive menneskelige bevægelser såsom vægtfordeling under en stående samtale samt øge samtalekredsen for at inkludere én eller flere samtalepartnere.

For at muliggøre dette arbejder Karl med at udvikle en robot, som balancerer på en kugle med en diameter på 25 cm, hvilket tillader 360° bevægelse, og i tillæg benyttes

inerti til at gengive menneskelige bevægelser. Robotten illustreres på Figur 1.3, dog er robotten stadig under udvikling og fotoet er taget den 27. oktober 2017. Ved at balancere robotten på en kugle bliver det muligt for robotten ikke blot at bevæge sig frem og tilbage, men også sidelæns, hvilket eksempelvis ikke er tilfældet ved *Double*, (D. Robotics 2017), som balancerer på en cylinder, hvilket er illustreret på Figur 1.4. Fælles for Karls robot og *Double* er, at de begge benytter sig af inertii, for at gengive menneskelige bevægelser, såsom vægtfordeling. Kuglens store diameter tillader robotten at køre ubesværet og naturligt på de fleste overflader og i tilfælde hvor robotten skal køre over et dørtrin, vil det fremstå flydende og ikke skramlende, som er tilfældet ved en mindre kugle diameter.



Figur 1.3. Fotografi af Karls robot.



Figur 1.4. Illustration af *Double*, (D. Robotics 2017).

Formålet med robotten er, at den skal indgå i Københavns Lufthavn, hvor den eksempelvis skal hjælpe de rejsende med at finde frem til deres gate i tide. At robotten skal indgå i Københavns Lufthavn udspringer fra samarbejdet mellem Karl, Combine og Københavns Lufthavn. De udviklere og designere, som Karl har været i kontakt med, giver udtryk for ikke at være interesserede i, hvordan robotten bevæger sig. De foretrækker at robotten har nogle foruddefinerede indstillinger for den grundlæggende måde hvorpå robotten bevæger sig, hvilket skal stemme overens med, hvad der forventes af en social robot. Derfor handler projektet for Karl om, at få robotten til at bevæge sig på en hensigtmæssig måde, så det er muligt for den at indgå i HRI. Det er med udgangspunkt i dette, projektsamarbejdet med Karl udspringer.

Det overordnede formål med projektet er derfor at undersøge hvordan den rejsendes subjektive oplevelse af interaktion med robotten er.

For at udvikle den sociale intelligens i robotten; hvordan robotten skal agere og reagere

blandt mennesker, er det nødvendigt at undersøge hvilke parametre, der har indflydelse på den subjektive oplevelse af HRI. I følgende afsnit undersøges eksisterende sociale robotter, der allerede er på markedet, interaktion med sociale robotter, herunder hvilke parametre der har indflydelse på denne og efterfølgende undersøges udfordringer ved sociale robotter. Selvom det på nuværende tidspunkt ikke tyder på, at diverse lufthavne har integreret sociale robotter i lufthaven, vil det bliver undersøgt om der er et grundlag for at denne teknologi potentiel kan introduceres.

1.3 Teknologier i lufthavne

Ifølge Kalakou m.fl. (2015, s. 203) tyder det på, at lufthavne konstant skal fornye sig og adoptere ny teknologi for at forbedre rejseoplevelsen. Dette kan ske ved at gøre det hurtigere, mindre stressfuldt, mere sikkert og effektivt at rejse. Det fremgår af Kalakou m.fl. (2015, s. 203), at rejsende ofte oplever flyrejser som værende tidskrævende, ukomfortable, stressende og dyre. Ifølge Bogicevic m.fl. (2017, s. 351) har amerikanske rejsende haft et dårligt forhold til lufthavnsindustrien, da den ikke lever op til deres forventninger. Derudover oplever de rejsende, at de ansatte er inkompentente, at der er for lange ventetider, dårlige navigationssystemer, manglende information samt ubelejlige flyoversigter.

Dette lægger ekstra pres på lufthavnene, da antallet af rejsende vil stige i fremtiden, (Kalakou m.fl. 2015, s. 203). I tillæg pointer Triebel m.fl. (2016, s. 609), at antallet af førstegangsrejsende, rejsende med begrænsede fremmedsprogskundskaber samt rejsende med særlige behov, ligeledes vil stige. En af de nuværende teknologiske løsninger luftenhaven gør brug af en applikation til smartphones, så de rejsende konstant kan modtage information omkring deres rejse, eller finde information omkring diverse butikker, (Kalakou m.fl. 2015, s. 203). En af fordelene ved at benytte lufthavnsapplikationer er, at omkring 76 % af de rejsende er i besiddelse af en smartphone, (Kalakou m.fl. 2015, s. 203). En anden måde at forbedre rejseoplevelsen er ved at gøre endnu mere brug af selvbetjening, så alt der vedrører check-in udelukkende består af selvbetjening, (Kalakou m.fl. 2015, s. 205). Ifølge Bogicevic m.fl. (2017, s. 351) forventes det, at inden for de næste par år vil alt, der vedrører check-in af passager og bagage foretages udelukkende ved brug af selvbetjening.

Ved at investere i nye innovative teknologier, der hjælper med check-in, bagageaflevering samt lokalisering af gaten, vil det, ifølge Bogicevic m.fl. (2017, s. 352), forbedre rejseoplevelsen. Det kan derfor være en fordel, at investere i sociale robotter, som er i stand til at interagere med de rejsende og hjælpe dem med lige netop deres behov, hvad end det vedrører boarding, gate information eller information omkring indkøbsmuligheder. I henhold til sociale robotter er der foretaget et EU-støttet projekt omkring udviklingen af *SPENCER*, der er en social robot, som har til formål effektivt, at guide rejsende fra deres ankomst gate til paskontrol og efterfølgende til transitten, (Triebel m.fl. 2016, s. 609). Baseret på resultaterne fremsat af Joosse og Evers (2017, s. 150), angav 12 ud af 16 testpersoner, at en social robot vil øge kundetilfredshed. Udo over det europæiske tiltag, har Sydkoreas største lufthavn: Incheon International Airport, ligeledes ambition om at investere i sociale robotter, hvis formål enten er at gøre rent eller at informere rejsende om afgangstidspunkter og vejvisning, (Vincent 2017).

I en undersøgelse foretaget af Brida m.fl. (2016) fokuseres der på den rejsendes

perciperede oplevelse af service kvaliteten i Chiles største lufthavn; Santiago de Chile Airport. Baseret på resultaterne konkluderer Brida m.fl. (2016, s. 213), at den vigtigste parameter er *Image perception of airport*, hvilket dækker over hvor innovativ lufthavnsterminalerne er, sikkerhed, passager opmærksomhed, terminal vedligeholdelse samt handicap faciliteter. Derudover finder Brida m.fl. (2016, s. 213) ydermere at information- og kommunikationsløsninger har en indflydelse på den perciperede service kvalitet. Denne parameter dækker over fly- og lufthavnsinformation via skærme eller skiltning. I tillæg forventer Brida m.fl. (2016, s. 210), at de rejsende værtsætter integration af nye teknologier i lufthavne.

Hvis det antages at disse resultater er generaliserbare samt forventningen om at rejsende værtsætter integration af nye teknologier er korrekt, så de ligeledes gør sig gældende for Københavns Lufthavn tyder det på, at det bør være muligt at introducere den sociale robot i Københavns Lufthavn. Integrationen af den sociale robot bør derfor øge den perciperede service kvalitet, da det er ny og innovativ teknologi, som er i stand til at levere informationer og kommunikere med de rejsende. Dertil kommenterer Triebel m.fl. (2016, s. 609), at *SPENCER*-projektet adresserer et yderst relevant forretningspotentiale, som kan få indflydelse på hele luftfartsindustrien, da der et voksende behov for at assistere de rejsende og mindske antallet af missede fly.

Når nye teknologier, som skal indgå i en lufthavn, skal designes er der nogle parametre, der bør tages højde for. Ifølge Bogicevic m.fl. (2017, s. 352) skyldes en stor del af de rejsendes frustration usikkerhed, derfor bør nye teknologier designes så de reducerer usikkerhed og øger den rejsendes selvtillid. Den rejsendes selvtillid har stor indflydelse på følelser såsom sikkerhed, forventninger og hvor komfortabel den rejsende føler sig ved at interagere med en service teknologi, (Bogicevic m.fl. 2017, s. 353). Sådan en teknologi skal ifølge Bogicevic m.fl. (2017, s. 353), udstråle integritet og pålidelighed, for at opnå den rejsendes tillid, som er med til at øge selvtilliden.

Da formålet med dette projekt er at undersøge, hvordan en social robot kan integreres i en lufthavn er det relevant at undersøge interaktionen med denne type robot, særligt i forhold til om de fornævnte parametre kan inkluderes.

1.4 Interaktion med sociale robotter

Inden der dykkes ned i hvilke parametre, der har indflydelse på hvordan interaktionen med sociale robotter percipieres og accepteres, vil nogle mere generelle tendenser blive diskuteret. Det dækker blandt andet over køns-, alders- samt kulturelle forskelle i henhold til synet på sociale robotter. Dernæst gives der nogle eksempler på, hvordan relaterede studier har målt brugerens perception og accept af sociale robotter.

1.4.1 Generelle tendenser

I den vestlige verden er vi i langt mindre grad villige til at acceptere sociale robotter end hvad der eksempelvis er tilfældet i Japan, (Katz m.fl. 2015, s. 28). Det skyldes, ifølge Katz m.fl. (2015, s. 28), at der er en indgroede frygt for maskiner og følelsen af manglende kontrol, hvilket ikke er tilfældet i den Japanske kultur, som åndeliggører robotter. At mennesker i den vestlige verden frygter robotter kan, blandt andet, retfærdiggøres med

at i 2012 angav 87 % af borgerne i Europa, at de aldrig har været i kontakt med en robot, hverken i hjemmet eller på ens arbejdsplads, (Höflich og Bayed 2015, s. 40). I en undersøgelse, beskrevet af Höflich og Bayed (2015, s. 41), fremgår det, at synet på robotter ikke har ændret sig de sidste 35 år. Når robotterne, ved brug af tegninger, visualiseres, så minder de i høj grad om robottypen: *Non assistive robots*, illustreret på Figur 1.1. Endvidere tyder det på, at Europær er frygter, at de vil miste deres job til robotten og at robotten er til for at erstatte mennesket, (Taipale m.fl. 2015, s. 22).

Derudover er der en tendens til, at mænd i højere grad perciperer robotten som menneskeagtig sammenlignet med kvinder, som i langt højere grad perciperer robotten som en maskine, (Katz m.fl. 2015, s. 28). Ifølge Graaf og Allouch (2013, s. 1479), perciperer mænd robotter som værende mere brugbare, de har større intention om at bruge dem og de er mere villige til at acceptere robotter end kvinder er.

Ydermere argumenterer Graaf, Allouch og Klamer (2015, s. 2) for, at den ældre population i højere grad accepterer sociale robotter, sammenlignet med den yngre population. Det antages, at en af årsagerne til det formentlig skyldes, at der anvendes sociale robotter i ældreplejen eksempelvis til mindske følelsen af ensomhed, se Underafsnit 1.1.1 (*Eksisterende sociale robotter*). I mere praktiske situationer, som eksempelvis ved rengøring, tøjvask og lignende, foretrækker ældre robotter fremfor mennesker, hvorimod hvis opgaverne er omsorgsrelateret foretrækkes mennesker, (Taipale m.fl. 2015, s. 22).

1.4.2 Parametre, der har indflydelse på accept og interaktion med sociale robotter

Graaf, Allouch og Klamer (2015, s. 1477) opstiller tre forskellige problemstillinger, der bør overvejes når brugerens accept af en social robot undersøges. I de følgende afsnit undersøges hvilke parametre, der har indflydelse på brugerens accept i forhold til de tre problemstillinger:

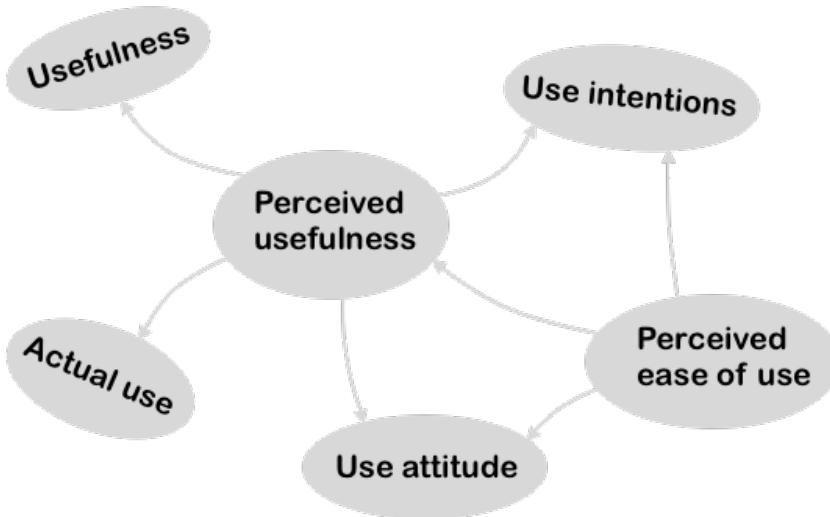
1. *the likely positive or negative consequences of the behavior*
 2. *the approval or disapproval of the behavior by respected individuals or groups, and*
 3. *the factors that may facilitate or impede performance of the behavior,*
- Graaf, Allouch og Klamer (2015, s. 1477).

Overvejelserne afspejler henholdsvis brugerens evaluering af robotten, hvilke sociale normative overbevisninger, der tilskrives robotten ved brug samt hvilke kontekstuelle faktorer, der spiller ind ved brug, (Graaf, Allouch og Klamer 2015, s. 1477). Baseret på Graaf, Allouch og Klamer (2015, ss. 1477-1478) tyder det på, at der er to overordnede kategorier af parametre, som har indflydelse på den første problemstilling: Utilitaristiske og hedoniske parametre. Førstnævnte dækker over det praktiske og anvendelige aspekt ved at interagere med en social robot, hvor sidstnævnte relateres til brugerens oplevelse af at anvende den sociale robot, (Graaf, Allouch og Klamer 2015, s. 1476).

I følgende to afsnit vil de utilitaristiske og hedoniske parametre undersøges nærmere, hvorefter de sociale normative overbevisninger og efterfølgende hvilke parametre, der har indflydelse på præstationsevnen, undersøges.

Utilitaristiske parametre

Som nævnt dækker utilitaristiske parametre over det praktiske og anvendelige aspekt ved at interagere med en social robot. Derudover har disse parametre ligeledes indflydelse på den specifikke adfærd. I følgende afsnit belyses hvilke parametre, der har indflydelse på det praktiske såvel som det anvendelige aspekt. Der tages primært udgangspunkt i Graaf, Allouch og Klamer (2015, s. 1477).

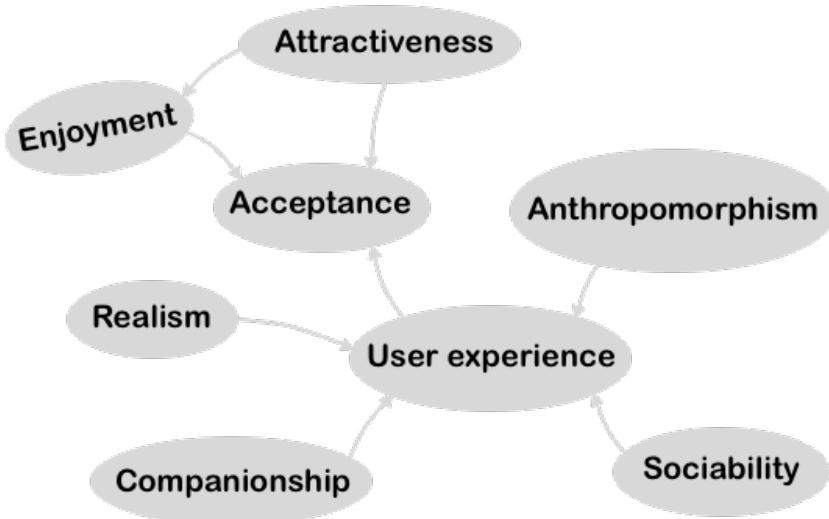


Figur 1.5. Sammenhængen mellem de seks forskellige parametre, der har indflydelse på det praktiske og anvendelige aspekt. Pilene indikerer hvilken retning indflydelsen er mellem to parametre.

Baseret på Graaf, Allouch og Klamer (2015, s. 1477) defineres *usefulness* som værende brugerens overbevisning om, at robotten vil forbedre de daglige aktiviteter. Ifølge Graaf, Allouch og Klamer (2015, s. 11) tyder det på, at *usefulness* er en vigtig parametre, der kan bidrage til langtidssigtet forhold mellem bruger og robot. *Ease of use* defineres som værende brugerens overbevisning om, at det er nemt at anvende robotten. Derudover argumenterer Graaf, Allouch og Klamer (2015, s. 1477) for, at i situationer hvor robotten skal indgå i en social interaktion med mennesker, er det nødvendigt at robotten gengiver menneskelige træk, for at brugeren føler sig komfortabel nok til at indgå i interaktionen. Robottens evne til at tilpasse sig den sociale kontekst afhængigt af brugeres behov, defineres som *perceived adaptability*. Ifølge Graaf, Allouch og Klamer (2015, s. 1477), har *perceived adaptability* indflydelse på *perceived usefulness*, *use attitude*, *use intentions*, som er repræsenteret på Figur 1.5. Derudover har *perceived adaptability* indflydelse på *enjoyment*. Ydermere tyder det på at *intelligence*, har en effekt på hvor realistiske robotten percipieres, (Graaf og Allouch 2013, s. 1477).

Hedoniske parametre

Som nævnt dækker hedoniske parametre over brugerens oplevelse af at anvende den sociale robot. Derudover har disse parametre ligeledes indflydelse på den specifikke adfærd. I følgende afsnit belyses de parametre, der har indflydelse. Der tages primært udgangspunkt i Graaf, Allouch og Klamer (2015, ss. 1477-1478). På Figur 1.6 illustreres de forskellige parametre, samt deres indbyrdes forhold.



Figur 1.6. Sammenhængen mellem de otte forskellige parametre, der har indflydelse på brugerens oplevelse. Pilene indikerer hvilken retning indflydelsen er mellem to parametre.

Baseret på Graaf og Allouch (2013, s. 1477) har både *enjoyment*, der defineres som værende følelsen af fornøjelse eller glæde forbundet med brug, og *attractiveness*, der defineres som værende den positive evaluering af robottens udseende, indflydelse på flere af de parametre gengivet på Figur 1.5. *Enjoyment* har indflydelse på *ease of use*, *use attitude* samt *use intentions*. Derimod har *attractiveness* indflydelse på *usefulness* og *ease of use*.

Antropomorfisering defineres som værende evnen til at tilskrive naturfænomener, guder, overnaturlige væsner og dyr menneskelige egenskaber, såsom følelser og motiver, (Hansen 2017). Dog defineres antropomorfisering i HRI sammenhæng, som værende evnen til at tildele og beskrive objekter med menneskelige egenskaber, for at rationalisere objektets adfærd, (Graaf og Allouch 2013, s. 1478). Ifølge Graaf og Allouch (2013, s. 1478) har anthropomorfisme indflydelse på *usefulness*, *use attitude* samt *use intention*, som er repræsenteret på Figur 1.5. Derudover har *anthropomorphism* ligeledes indflydelse på *attitude toward robots*, *social influence* samt *companionship*. Ifølge Halpern og Katz (2013, s. 19) resulterer antropomorfisering i, at mennesket betragter en social robot som en social enhed, og derfor behandler robotten som et menneske.

Der er forskellige årsager til at mennesker antropomorfiserer sociale robotter. Antropomorfisering kan forekomme i situationer hvor mennesket oplever en manglende kontrol og usikkerhed, hvor mennesket i højere grad har en tendens til at antropomorfisere sociale robotter, for at kunne forstå, kontrollere samt forudse robottens adfærd, (Graaf og Allouch 2013, s. 1478). Dette hører under *effectance motivation*, der defineres som værende ønsket om effektivt at kunne interagere med ens omgivelser, (Eyssel m.fl. 2011, s. 62).

Derudover har mennesker med en teknologisk baggrund en tendens til at tildele robotten sin egen personlighed, hvilket ikke er tilfældet med mennesker uden teknologisk baggrund, (Halpern og Katz 2013, s. 19). I tillæg argumenterer Graaf, Allouch og Klamer (2015, s. 2) for, at desto mere antropomorfisering brugeren oplever, desto bedre er evalueringen af robotten, desto mere fornøjet er de med interaktionen og desto større er chancen for at de oplever robotten som en ledsager.

Ifølge Eyssel m.fl. (2011, s. 61) har ensomme mennesker en stærk tendens til at antropomorfisere kæledyr og teknologiske objekter, såsom robotter. Det skyldes, at

mennesket har et behov for både tilknytning og et tilhørsforhold. Dette hører under *sociality motivation*, der defineres som værende ønsket og behovet for at skabe en social relation til andre, (Eyssel m.fl. 2011, s. 61). I det henseende vil en robot med et mere livagtigt udtryk perciperes som værende en venlig ledsager, (Graaf og Allouch 2013, s. 1478). Dette gengives som *companionship*, der defineres som brugerens perciperede mulighed for at opbygge et forhold til robotten. *Companionship* er illustreret på Figur 1.6 og uover at have indflydelse på *user experience*, har *companionship* indflydelse på den vedvarende interaktion med robotten.

Baseret på Graaf og Allouch (2013, s. 1478) tyder det på, at *realism* kan forbedre HRI og har dermed indflydelse på *user experience*, Figur 1.6. En robots *realism* afspejler i hvilken grad brugeren tror på, at robotten reagerer og opfører sig realistisk. Desto mere realistisk robotten perciperes, desto mere intelligent perciperes den, (Graaf og Allouch 2013, s. 1478).

Sociability defineres som værende brugerens overbevisning om hvorvidt robotten besidder de sociale, emotionelle samt kognitive færdigheder, der er nødvendige for en succesfuld tilvænning af robotten, (Graaf og Allouch 2013, s. 1478). Denne parameter har, ifølge Graaf og Allouch (2013, s. 1478), ligeledes indflydelse på *use attitude* og *usefulness*.

I henhold til teknologier i lufthavne argumenterer Bogicevic m.fl. (2017, s. 352) for, hvorfor der bør tages højde for *hedonic* parametre. En stor del af den teknologi, som kan findes i lufthavne bygger primært på *utilitarian* parametre, hvilket formentlig har været årsagen til en dårlig brugeroplevelse. For at forbedre dårlige oplevelser, blandt andet ved lange ventetider ved sikkerhedskontrolle, har lufthavne forsøgt at foretage nogle tiltag, der skal gøre oplevelsen bedre, hvilket eksempelvis kommer til udtryk ved et øget brug af smartphone applikationer. Derudover understreger Bogicevic m.fl. (2017, s. 352), at det er nødvendigt at skabe fornøjelige lufthavnsoplevelser. Fornøjelighed gengives på Figur 1.6, som *enjoyment*.

Sociale normer

I henhold til problemstilling 2: *the approval or disapproval of the behavior by respected individuals or groups*, fremsat af Graaf, Allouch og Klamer (2015, s. 1477), vil følgende afsnit undersøge hvilke parametre, der har indflydelse på det.

De sociale normer dækker over de synspunkter mennesket har i relation til deres egen adfærd og dækker ydermere de synspunkter og regler, der er gældende for en gruppe for hvorvidt en bestemt adfærd er passende eller upassende, (Graaf og Allouch 2013, s. 1478). Ifølge Graaf og Allouch (2013, s. 1478) dækker de sociale normer over *social influence* og *image*. *Social influence* er karakteriseret ved brugerens perception af hvad andre tænker omkring brugen af robotten. Denne parameter har indflydelse på *usefulness*, *ease of use*, *use attitude*, *use intention* samt *actual use*, (Graaf og Allouch 2013, s. 1478), jævnfør Figur 1.5. Det er særligt mennesker i ens tætte omgangskreds; familie, partner og venner, hvis mening har indflydelse på hvorvidt en bestemt teknologi vil blive brugt, i dette tilfælde en social robot, (Graaf og Allouch 2013, s. 1478).

Image karakteriseres ved brugerens overbevisning om at interaktionen med robotten kan lede til større anerkendelse og social status blandt ens omgangskreds, (Graaf og Allouch 2013, s. 1478). Derudover har *image* indflydelse på *perceived usefulness*, illustreret på

Figur 1.5.

I relation til sociale normer og tilnærmelsesvis *social influence* undersøger Kim og Mutlu (2014), hvilken indflydelse *social distance* har på HRI. *Social distance* referer til hvilken grad mennesker perciperer manglende intimitet grundet forskellige egenskaber såsom etnicitet, race, religion og beskæftigelse, (Kim og Mutlu 2014, s. 784). *Social distance* dækker over to forskellige parametre; *social structural distance* og *physical distance*.

Social structural distance referer til mennesket perception af og adfærd rettet mod andre, som er afhænger af hvordan mennesket kategoriserer andre, særligt i forhold til deres kompetance og varme, (Kim og Mutlu 2014, s. 784). Derudover kategoriseres mennesker afhængigt af deres indbyrdes forhold, om der er forskel i status, om interaktionen foregår som et samarbejde eller som en konkurrence. *Social structural distance* kan inddeltes i yderligere to kategorier; *power distance* og *task distance*. Førstnævnte har stor indflydelse på den individuelles adfærd og perception afhængigt af forholdet til interaktionspartneren, som afhænger af personligheder, roller og deres individuelle sociale klasse, (Kim og Mutlu 2014, s. 784). *Task distance* referer til at mennesker afhænger af hinanden for at opnå deres mål. Dette kan enten komme positivt til udtryk, når mennesker samarbejder for at opnå deres fælles mål, eller negativt, når mennesker konkurrerer mod hinanden og forhindrer hinanden i at opnå deres individuelle mål, (Kim og Mutlu 2014, s. 784).

Physical distance kan også beskrives som *proxemic distance* og referer til den fysiske adskillelse mellem mennesker, (Kim og Mutlu 2014, s. 784). Ved denne afstand opstår der bedre kommunikation mellem mennesker, hvor den individuelle vil være i stand til at løse egne opgaver bedre, (Kim og Mutlu 2014, s. 785).

Kim og Mutlu (2014, s. 794) undersøger forholdet mellem *power distance* og *proxemic distance* i forhold til HRI. *Power distance* udtrykkes ved at robotten enten agerer som *supervisor* eller som *subordinate*, hvor *proxemic distance* udtrykkes ved at den fysiske afstand mellem robot og menneske er lille eller stor. Baseret på disse resultater tyder det på, at *user experience* blev bedre når interaktionen foregik med en *supervisor* robot tæt på og når interaktionen foregik med en *subordinate* robot langt væk, (Kim og Mutlu 2014, s. 785). Dog finder Kim og Mutlu (2014, s. 785), at præstationen i opgaven blev forværet desto tætter robotten var på testpersonen, uafhængigt af *power distance*.

Derudover finder Kim og Mutlu (2014, s. 785), at *user experience* blev forbedret når robotten konkurrerede med testpersonen tæt på og når robotten samarbejde med testpersonen langt væk, hvilket var imod forventningen, (Kim og Mutlu 2014, s. 785).

Indflydelse på præstationsevnen

I henhold til problemstilling 3: *the factors that may facilitate or impede performance of the behavior*, fremsat af Graaf, Allouch og Klamer (2015, s. 1477), vil følgende afsnit undersøge hvilke parametre, der har indflydelse på præstationsevnen.

Control beliefs referer til brugerens overbevisning om hvilke ressourcer, muligheder og forhindringer, der er tilstede eller fraværende og som vil have en positiv eller negativ indflydelse på præstationsevnen, (Graaf og Allouch 2013, s. 1478). *Control beliefs* kan inddeltes i tre kategorier, som alle har indflydelse på *user acceptance*: *Perceived behavioral control*, *anxiety* og *experience*, begge rettet mod robotter, (Graaf og Allouch 2013, s. 1478).

Perceived behavioral control defineres som værende brugerens perciperede oplevelse af hvor let eller svært det var at interagere med robotten, (Graaf og Allouch 2013, s. 1478). *Perceived behavioral control* har indflydelse på *perceived ease of use, use intentions* samt *actual use*. *Anxiety* rettet mod robotter defineres som værende ængstelige eller andre negative emtioner forbundet med HRI og som forværres afhængigt af tidligere oplevelser, (Graaf og Allouch 2013, s. 1478). *Anxiety* har en negativ effekt på *perceived ease of use*, men *anxiety* kan reduceres med *enjoyment*, (Graaf og Allouch 2013, s. 1478).

Experience med robotter defineres som den opnåede oplevelse med robotter både direkte, i form af HRI, og indirekte via et medie, såsom nyhedsartikler og science-fiction film, (Graaf og Allouch 2013, s. 1479). Som nævnt i Underafsnit 1.4.1 (*Generelle tendenser*), har 87 % af borgerne i Europa aldrig indgået i en interaktion med en robot, hvorfor det antages at den direkte oplevelse med robotter er begrænset, hvert fald i Europæiske lande.

Tidligere erfaring med robotter har, ifølge Graaf og Allouch (2013, s. 1479), en positiv effekt på: *Usefulness, ease of use, attitude towards robots, use intention* samt *actual use*. I henhold til tidligere erfaringer med robotter, undersøger Halpern og Katz (2013) hvorvidt *online community* og erfaring med *avatars*, har en indflydelse på evnen til at antropomorfisere og acceptere robotter. *Community* defineres som forholdet mellem individet og den sociale struktur de tilhører, (Halpern og Katz 2013, ss. 20-21). Det essentielle ved *community* dækker over support, *sociability*, information, dannelse af social identitet og tilhørersforhold, (Halpern og Katz 2013, s. 21). Det eneste der adskiller *community* fra et *online community* er, at sidstnævnte foregår virtuelt. Halpern og Katz (2013, s. 25) fandt de testpersoner, som enten var en del af *online community* eller mere involveret med *avatars*, i højere grad var i stand til at genkende menneskelige træk ved robotten, sammenlignet med testpersoner uden disse erfaringer. Ydermere fandt Halpern og Katz (2013, s. 26) at testpersoner, som enten var en del af *online community* eller mere involveret med *avatars*, begge var mere villige til at acceptere robotter, som en del af deres sociale og fysiske miljø.

1.4.3 Bevægelsesmønstre og udseende

Som nævnt i Underafsnit 1.4.1 (*Generelle tendenser*) forekommer der store kulturelle forskelle i forhold til hvordan sociale robotter percipieres. Ikke nok med det, forekommer der ligeledes store kulturelle forskelle i hvordan robotter bør bevæge sig, blandt andet i forhold til hvor tæt de må komme på. Først vil afstanden mellem robot og menneske blive diskuteret, og derefter hvilken hastighed robotten bør bevæge sig med.

Afstand mellem robot og menneske

Ifølge Dautenhahn, Walters m.fl. (2006, s. 178) er den intime grænse for hvor tæt et andet menneske, tætte venner og familie, må komme på en mellem 20 cm og 30 cm for syd europærere og japanere. Hvorimod denne grænse for amerikanere og nord europæer er 46 cm til 122 cm. Ydermere har mennesker, der er opvokset i et landdistrikts ligeledes behov for en øget intim grænse, modsat mennesker, der er opvokset i bymiljøer, Dautenhahn, Walters m.fl. (2006, s. 178). Endvidere tyder det på, at kvinder normalvis er tættere på hinanden og vendt mere direkte ansigt-til-ansigt, sammenlignet med mænd, (Dautenhahn, Walters m.fl. 2006, s. 178). Det må derfor antages at kvinder, har en mindre intim grænse end mænd.

I undersøgelsen foretaget af Dautenhahn, Walters m.fl. (2006), undersøger de blandt andet hvor tæt en robot må komme før den overskridet den intime grænse, afhængigt af indgangsvinkel: Frontal, højre eller venstre. Robotten holdte en afstand svarende til 50 cm \pm 10 cm. I den første undersøgelse foretaget af Dautenhahn, Walters m.fl. (2006, s. 174), angav 76 % testpersonerne at afstanden mellem dem og robotten var *about right*, 19 % angav at afstanden var for stor og 5 % angav at robotten kom for tæt på. I den anden undersøgelse foretaget af Dautenhahn, Walters m.fl. (2006, s. 175), angav 53 % af testpersonerne, at når robotten nærmede sig frontalt, at den kom for tæt på, 27 % angav at afstanden var *about right* og 20 % angav at afstanden var for stor. Da robotten nærmede sig fra venstre, angav 80 % af testpersonerne at det var *about right* og 20 % angav at afstanden var for stor. Da robotten nærmede sig fra højre, angav 60 % af testpersonerne at afstanden var *about right* og 40 % angav at afstanden var for stor. I alle tilfælde foregår interaktionen mellem robotten og et siddende menneske.

I henhold til hvilken retning robotten skal nærme sig testpersonerne med fremgår det, at 59 % foretrækker at robotten nærmer sig fra højre, 28 % foretrækker venstre og 13 % foretrækker frontalt, (Dautenhahn, Walters m.fl. 2006, s. 175). Lignende tendens går igen ved hvor praktisk og komfortabelt testpersonerne vurder interaktionen, (Dautenhahn, Walters m.fl. 2006, ss. 175-176). Der skal dog tages højde for kønsforskelle, da en del af kvinderne faktisk foretrækker at robotten nærmer sig frontalt, hvilket formentlig hænger sammen med at kvinder i højere grad interagerer med hinanden ansigt-til-ansigt, sammenlignet med mænd, (Dautenhahn, Walters m.fl. 2006, s. 178). Dog argumenterer Dautenhahn, Walters m.fl. (2006, s. 178) for, at den frontale indgangsvinkel opleves som værende ukomfortabel, upraktisk, truende og konfronterende, hvorfor det bør undgås. I det henseende kommenterer Dautenhahn, Walters m.fl. (2006, s. 178) at i situationer, hvor der er en 45°s vinkel mellem to samtale partnere, vil følelser som aggression og konfrontation være reduceret. Derudover skal der tages højde for at 94 % af testpersonerne er højrehåndede, hvilket formentlig er årsagen til at størstedelen af testpersonerne foretrækker at robotten nærmer sig fra højre, (Dautenhahn, Walters m.fl. 2006, s. 175).

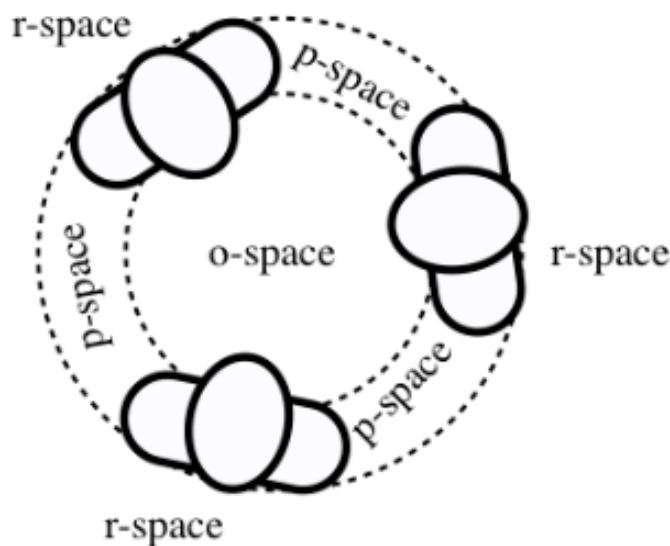
Det lader til at indgangsvinklen hvorved robotten nærmer sig er kontekstafhængig. Ball m.fl. (2015, s. 41) giver udtryk for at mennesker, der er alene, er mest komfortable, hvis robotten nærmer sig frontalt, da robotten befinner sig i synsfeltet. Derudover finder Ball m.fl. (2015, s. 48), at hvis der er et andet menneske til stede vil det ikke føles lige så ukomfortabelt at robotten nærmer sig bagfra, som hvis det andet menneske ikke var til stede. Derudover tyder det på, at tilstedeværelsen af en anden person resulterer i færre præferencer for hvilken indgangsvinkel robotten skal nærme sig ved, (Ball m.fl. 2015, s. 48).

En anden undersøgelse, der blandt andet fokuserer på afstanden mellem robot og menneske, er foretaget af Pacchierotti m.fl. (2005). Den differentieres mellem fire forskellige afstande, som er gældende for menneske-menneske interaktion. *Intimate distance* denne afstand går direkte fra kroppen til omkring 45 cm fra kroppen, og generelt beregnet til direkte fysisk kontakt eller privat interaktion (Pacchierotti m.fl. 2005, s. 165). Dernæst er *personal distance*, som går fra 45 cm til 1.2 m, denne afstand gør sig gældende for interaktion med familie og venner eller ved organiserede interaktion, som opstår når mennesker står i kø, (Pacchierotti m.fl. 2005, s. 165). Den tredje afstand er *social distance*, som går fra 1.2 m til 3.5 m og som er beregnet til mere formelle og arbejdsrelateret interaktioner, interaktioner

med bekendte. Ydermere fungerer *social distance* også som en form for separering mellem mennesker på offentlige steder, eksempelvis på stranden, (Pacchierotti m.fl. 2005, s. 165). Den sidste afstand er *public distance*, som starter fra 3.5 m og anvendes i situationer, hvor der er envejskommunikation, eksempelvis mellem publikum og den optrædende, (Pacchierotti m.fl. 2005, s. 165).

Baseret på Pacchierotti m.fl. (2005, ss. 169-170) anbefales det, at når en robot skal passere et mennesker bør robotten signalerer det med en afstand på 4 m til 6 m fra mennesket, formentlig er en afstand på 3.5 m også acceptabelt, da det overholder *public distance*. Såfremt robotten i tide signalerer dels at den nærmer sig og dels dens intention, som er at passere mennesket, er det ikke et problem at passagen foregår indenfor *personal distance* og endda kan en mindre afstand også accepteres, (Pacchierotti m.fl. 2005, s. 170). I tillæg kommenterer Pacchierotti m.fl. (2005, s. 170), at afstanden mellem robot og menneske under passagen er mindre vigtig så længe robotten signalerer sin intention i tide, så mennesket ligeledes kan reagerer på interaktionen, eksemplvist ved at bevæge sig til siden.

Da robotten skal indgå i en interaktion med rejsende i en lufthavn, skal robotten nødvendigvis være i stand til at interagere med mere end én rejsende af gangen. Det er derfor nødvendigt at tage højde for, hvordan interaktion mellem to eller flere mennesker udfører sig. Mennesker danner F-formationer ved ansigt-til-ansigt interaktion, (Marshall m.fl. 2011, s. 445) og disse formationer bygger på helt særlige bevægelsesmønstre.



Figur 1.7. Illustration af en F-formation dannet af tre deltagere samt dens inddeling af områder, (Marshall m.fl. 2011, s. 446).

På Figur 1.7 illustreres en F-formation bestående af tre mennesker. Mennesker danner et område kaldt *transactional segment*, hvor både opmærksomheden dirigeres hen og hvor objekter manipuleres. Størrelsen på dette område afhænger både af aktiviteten og af underkroppens position, (Marshall m.fl. 2011, s. 446). Når to eller flere menneskers *transactional segment* overlapper hinanden dannes der et fælles *transactional segment*, kaldt *o-space*, hvor alle har lige tilgang, (Marshall m.fl. 2011, s. 446). *O-space* illustreres på Figur 1.7. Området mellem menneskerne kaldes *p-space* og området omkring F-formationen

kaldes *r-space*, jævnfør Figur 1.7. *R-space* er det område hvor en robot vil befinde sig før den inkluderes i de rejsendes interaktion og derfra selv indgår i F-formationen. Ifølge Ball m.fl. (2015, s. 41) vil mennesker interagere med robotter i deres *p-space*. Størrelsen af *o-space* afhænger derfor af, hvor mange der deltager i interaktionen og hvordan dette område tilpasser sig efter spatial og holdningsmæssig adfærd blandt de deltagende, (Marshall m.fl. 2011, s. 446). Udover den cirkulære F-formation afbilledet på Figur 1.7, kan formationen ligeledes gengive en halvcirkel eller en rektangel, såfremt der er minimum tre interaktionspartnere involveret, (Marshall m.fl. 2011, s. 446).

Robottens hastighed

Ifølge Pacchierotti m.fl. (2005, s. 165) er den normale ganghastighed for et menneske mellem 3.6 km/t og 7.2 km/t, hvorfor det må forventes at en social robot ikke bør overskride det interval ved HRI. I undersøgelsen foretaget af Dautenhahn, Walters m.fl. (2006, s. 175) varierede robottens hastighed sig mellem 0.9 km/t og 1.44 km/t, hvor 60 % af testpersonerne angav at hastigheden var *about right* og 40 % angav at hastigheden var for langsom. Ifølge Dautenhahn, Walters m.fl. (2006, s. 178) bør en robot, efter en tilvænningsperiode eller efter behov, bevæge sig hurtigere end 1.44 km/t.

I undersøgelsen foretaget af Pacchierotti m.fl. (2005, s. 169) svinger robottens gennemsnitlige hastighed mellem 0.9 km/t og 1.4 km/t. Årsagen til at robottens gennemsnitlige hastighed ikke er højere skyldes, at når robotten passerer et menneske sænkes hastigheden, (Pacchierotti m.fl. 2005, s. 169). Baseret på testpersonernes vurdering af robottens hastighed ville en højere hastighed være at foretrække, (Pacchierotti m.fl. 2005, s. 169). Ydermere tyder det på, at robottens lave hastighed blev perciperet som værende mindre sikker og tilmed irriterende, (Pacchierotti m.fl. 2005, s. 169), hvorfor der bør anvendes en højere hastighed. For at afgøre hvilken hastighed robotten bør bevæge sig med argumenterer Pacchierotti m.fl. (2005, s. 167) for, at robotten bør være i stand til at tilpasse sin hastighed afhængigt af menneskets hastighed, da dette vil medføre en mere dynamisk interaktion mellem robot og menneske. Ifølge Althaus m.fl. (2004, s. 1897) bør robotten sænke hastigheden når den nærmer sig en gruppe af mennesker, ligesom et menneske ville gøre det, hvis det nærmrede sig en gruppe af mennesker.

Butler og Agah (2001, ss. 192-103) undersøger tre forskellige hastigheder, fordelt på to måder robotten kan nærme sig et menneske på. Robotten kan enten nærme sig direkte eller indirekte, ved direkte er hastigheden enten 0.91 km/t eller 3.66 km/t og ved indirekte er hastigheden 1.83 km/t, (Butler og Agah 2001, ss. 192-103). En indirekte tilgang gengiver, at robotten først drejer mod venstre og derefter mod højre, hvilket giver en fornemmelse af at robotten bevæger sig en anelse til højre for mennesket, (Butler og Agah 2001, s. 193). Med udgangspunkt i robottens hastighed fandt Butler og Agah (2001, s. 196), at den direkte hastighed på 3.66 km/t var ubehagelig og at den indirekte hastighed på 1.83 km/t var at foretrække, da den var mest behagelig.

Baseret på de tre undersøgelser tyder det på, at hastigheder til og med 1.44 km/t i nogen grad kan accepteres, men det vil være, at foretrække hvis hastigheden var højere. Eftersom 1.44 km/t er langommere end menneskets normale ganghastighed bør det være muligt, at øge robottens hastighed uden det påvirker interaktionen med mennesket negativt. Det lader til, at en øget hastighed potentielt kan forbedre interaktionen mellem robot og menneske, hvorfor der bør tages højde for dette. Dog tyder det på at robottens hastighed ikke må

overstige mennesket normale ganghastighed.

Robottens udseende og personlighed

Ifølge Severinson-Eklundh m.fl. (2003, s. 226) behøver en social robot ikke at have en menneskelignende krop, da bare det at robotten bevæger sig på en ikke-forudsigelig måde kan få mennesker til at percipere robotten, som havende en intention og en personlighed, jævnfør antropomorfisering. Dette hænger formentligt sammen med hypotesen omkring *Uncanny Valley* fremsat af Mori (2012). Hypotesen bygger på hvilken effekt robottens udseende har på det tilhørsforhold mennesket oplever, hvor tilhørsforholdet kraftig forringes hvis robotten afspejler alt for menneskelige træk, både i forhold til udseende og bevægelse, Mori (2012). Ydermere argumenterer Triebel m.fl. (2016, s. 610) for, at hvis robottens udseende er alt for menneskeligt vil det øge brugerens forventninger til systemet vedrørende robottens kognitive egenskaber. Dette kan lede til at brugeren bliver skuffet eller nægter at interagere med robotten, (Triebel m.fl. 2016, s. 610). I tillæg pointerer Severinson-Eklundh m.fl. (2003, s. 226), at robottens størrelse, form, farve og bevægelsesmønstre har en indflydelse på hvordan robotten percipieres.

Butler og Agah (2001, s. 189) præsenterer deres testpersoner for en cylinder-formet robot med og uden krop. Højden med krop er omkring 170 cm, hvor højden uden kroppen er 35 cm, (Butler og Agah 2001, s. 189). De finder, at når robotten har en krop får det testpersonerne til at føle sig mere ukomfortable, både i forhold til hvordan robotten skal nærme sig, direkte eller indirekte, og hvor kort afstanden mellem robot og menneske må være, sammenlignet med robotten uden krop, (Butler og Agah 2001, s. 196). Dog kan det ikke tydes ud fra resultaterne, hvorvidt det udelukkende er robottens højde eller robottens menneskeagtige krop, der får testpersonerne til at føle sig ukomfortable.

En undersøgelse, som forsøger at holde udseendet konstant og varierer højden, er udført af Shiomi m.fl. (2013, s. 255). Den lille robot er omkring 30 cm høj, hvor den store robot er omkring 120 cm høj. Ifølge Shiomi m.fl. (2013, s. 255) har robottens størrelse indflydelse på *attractiveness* og *ease of initiating interaction*. Begge robotter implementeres i et Japansk storcenter og har til formål at printe rabatkuponer ud til de besøgende samt agere vejviser, Shiomi m.fl. (2013, s. 252). Baseret på resultaterne henvendte flere besøgende sig til den lille robot for at starte interaktionen, sammenlignet med den store robot, (Shiomi m.fl. 2013, s. 260). I tillæg kommenterer Shiomi m.fl. (2013, s. 260), at den store robot kan virke skrämmende særligt for børn.

Ifølge Severinson-Eklundh m.fl. (2003, s. 226) har robottens personlighed indflydelse på hvor meget mennesker stoler på dens instruktioner. Robottens personlighed skal endvidere stemme overens med hvilke opgaver den skal løse, hvori situationer hvor robotten skal anvendes som underholdningsmiddel, er det en fordel at den udstråler glæde og humor, modsat situationer hvori opgaven er mere alvorlig, hvor robotten bør udstråle seriøsitet, (Severinson-Eklundh m.fl. 2003, s. 226). Selvom mennesker foretrækker en glad robot er de mere tilbøjelige til at følge instruktioner, hvis de leveres af en seriøs robots, (Severinson-Eklundh m.fl. 2003, s. 226). En måde hvorpå robottens personlighed kan komme til udtryk er via gestikker, som kan få robotten til at afspejle menneskelige følelser, såsom at være genert, tilbageholdende eller beskeden, (Severinson-Eklundh m.fl. 2003, s. 227). Ved hjælp af gestikker er det muligt for robotten at give interaktionspartneren feedback, eksempelvis

i form af at nikke med hovedet for at indikere, at en kommando er forstået, (Severinson-Eklundh m.fl. 2003, s. 228).

Med udgangspunkt i Chee m.fl. (2012, s. 272) er de to vigtigste, ud af *the big five*, personlighedskarakteristikas *extroversion* og *agreeableness*, eller venlighed og dominans, i forhold til sociale interaktioner. Dominans kommer blandt andet til udtryk ved en ret holdning, modsat en mere underdanig personlighed, som blandt andet kommer til udtryk ved en mere krum holdning, Chee m.fl. (2012, s. 273). Baseret på resultaterne fremgår det, at robotter med et mere menneskeligt udseende; bestående af to øvre kropslegemer, en overflade som afspejler en form for hud, en blanding af varme og kolde farver samt en form for beklædning, perciperes som værende mere venlig end robotter som kun har et øvre kropslegeme, en mere metallisk overflade, kolde farver og uden beklædning, Chee m.fl. (2012, s. 275).

Projektafgrænsning 2

I følgende kapitel fokuseres der på hvilke udfordringer, der kan opstå ved at designe sociale robotter. I det henseende vil der undervejs foretages afgrænsninger for efterfølgende at udarbejde en problemformulering.

2.1 Udfordringer ved sociale robotter

Sociale robotter er en relativ ny teknologi, eksempelvis blev den første version af *AIBO* lanceret i 1999, SONY (2017), og i 2012 angav 87 % af borgere i Europa, at de aldrig har været i kontakt med en robot, jævnfør Underafsnit 1.4.1 (*Generelle tendenser*). Selvom teknologien har udviklet sig en del igennem de sidste fem år, antages det, at procentdelen af europæiske borgere, der ikke har været i kontakt med en robot stadig er høj. En af årsagerne til at sociale robotter ikke har floreret på det europæiske marked i lige så høj grad som de eksempelvis har på det asiatiske, skyldes formentlig den europæiske frygt for robotter, (Katz m.fl. 2015, s. 28). I tillæg er der en frygt for, at robotter vil overtage ens job, (Höflich og Bayed 2015, s. 42). Det tyder på, at frygten blandt europærer enten skyldes uvidenhed omkring hvilke muligheder, der er med sociale robotter eller at de har et forældet syn på robotter. En måde at reducere frygten og fornye synet på sociale robotter, kan være ved at inddrage potentielle brugere i designfasen. Ved at involvere potentielle brugere i designfasen, kan de få indflydelse på, hvordan robotten skal designes for at de stoler på den og ikke frygter den.

I forhold til frygten for at sociale robotter tager ens arbejde, bør det ikke være et problem i denne sammenhæng, hvor den sociale robot implementeres i en lufthavn og som primært varetager opgaver, som normalt ikke udføres af mennesker. Ifølge Bogicevic m.fl. (2017, s. 352) vil de rejsende opleve, at der i de fleste lufthavne ikke vil være en menneske-menneske interaktion før de boarder flyet, da de nødvendige opgave håndteres ved hjælp af en form for teknologi.

Baseret på de undersøgelser, der er inddraget i det forrige kapitel er der kun én, som strækker sig over mere end en dag; Graaf, Allouch og Klamer (2015, s. 3), som strækker sig over tre perioder på hver 10 dage. Nogen undersøgelser, såsom Höflich og Bayed (2015, s. 273) og Halpern og Katz (2013, s. 23), præsenterer testpersonerne for billeder, hvorefter de skal vurdere robotterne ud fra nogle foruddefineret parametre. Höflich og Bayed (2015, s. 62) præsenterede deres testpersoner for et videoklip, som enten skulle få testpersonerne til at percipere robotten som værende forudsigtig eller uforudsigtig, når de efterfølgende skulle interagere med robotten.

Hvor i undersøgelser foretaget af Graaf og Allouch (2013, s. 1480), Butler og Agah (2001, ss. 190-191), Dautenhahn, Walters m.fl. (2006, s. 173) samt Kim og Mutlu (2014, ss. 786-787) præsenteres testpersonerne for en social robot. I undersøgelsen foretaget af Graaf

og Allouch (2013, s. 1480) skal testpersonerne deltage i en samtale med robotten, som styrer samtalen. Butler og Agah (2001, ss. 190-191) undersøger hvordan testpersonernes komfortabilitet påvirkes af hvordan robotten nærmer sig testpersonen samt robottens hastighed og afstand. Lignende undersøges af Dautenhahn, Walters m.fl. (2006, s. 173), hvor testpersonerne skal interagere med en social robot. Hvor testpersonerne i sidstnævnte enten skal samarbejde eller konkurrerer med en robot, som enten agerer *supervisor* eller *subordinate*.

Med udgangspunkt i de foregående afsnit omkring teknologier i lufthavne, jævnfør Afsnit 1.3 (*Teknologier i lufthavne*), interaktion med sociale robotter, jævnfør Afsnit 1.4 (*Interaktion med sociale robotter*), samt dette afsnit vedrørende udfordringerne ved sociale robotter, er det muligt at udarbejde en problemformulering.

2.2 Problemformulering

Ulempen ved de fornævnte undersøgelser er dels, at HRI foregår over en kort periode, hvorfor interaktionen med robotten er kortvarig og dels, at undersøgelserne som udgangspunkt foretages i laboratorier. Ingen af tilfældene er det muligt for testpersonerne at danne et decideret tilhørsforhold til robotten, da forholdet stort set kun bygger på førstehåndsintrykket. Derudover præsenteres robotterne i specifikke kontekster, som ikke nødvendigvis er den kontekst robotten er designet til at indgå i. Som tidligere diskuteret er der både store kulturelle-, alders- og kønsforskelle, hvorfor de parametre testpersonerne i de fornævnte undersøgelser vurderer robotten ud fra, ikke nødvendigvis er de parametre, der har betydning for danske brugere. På baggrund af dette kan følgende problemstillinger opstilles:

Ud fra hvilke parametre beskriver danske rejsende interaktion med en social robot i en dansk lufthavn?

Hvordan kan de fundne parametre anvendes til at evaluere HRI?

For at besvare den første problemstilling er det nødvendigt, at foretage en form for interview med danske rejsende i en dansk lufthavn, hvor de skal gengive deres subjektive oplevelse af HRI. Baseret på de rejsendes respons skal der foretages en kvalitativ analyse for at kategorisere den givne respons og dermed udlede, hvilke parametre, der har indflydelse på de danske rejsendes interaktion med en social robot. Disse parametre kan derefter sammenholdes med tidligere foretaget undersøgelser, for at afgøre om interaktionen mellem danske rejsende og en social robot er påvirket af samme parametre, som gennemgået i Afsnit 1.4 (*Interaktion med sociale robotter*).

For at besvare den anden problemstilling vil der blive udarbejdet skalaer, hvorpå de fundne parametre præsenteres. Efterfølgende forsøges det at anvende de udviklede skalaer til at evaluere interaktionen med robotten. Dette gøres ved at definere nogle brugsscenerier samt bevægelsesmønstre for robotten, som de rejsende præsenteres for. Baseret på de rejsendes evalueringer bør det være muligt, at undersøge hvilke parametre korrelerer med hinanden.

Del II

Parametre, der har indflydelse på HRI

For at besvare den første problemstilling: *Ud fra hvilke parametre beskriver danske rejsende interaktion med en social robot i en dansk lufthavn?*, vil der foretages en feltundersøgelse i en dansk lufthavn. Formålet med denne undersøgelse er, at udlede hvilke parametre, der ifølge danske rejsende, har indflydelse på interaktionen med en social robot. Disse parametre vil ydermere sammenholdes med de parametre, som allerede har vist sig at have en indflydelse på interaktionen med sociale robotter, jævnfør Afsnit 1.4 (*Interaktion med sociale robotter*).

Feltundersøgelse 3

Formålet med feltundersøgelsen er, at få danske rejsende til at sætte deres egne ord på oplevelsen af interaktionen med en social robot. Baseret på den kvalitative respons vil det forsøges at udlede hvilke parametre, der er essentielle for at designe en brugervenlig interaktion med en social robot. I det henseende er det ikke nødvendigvis vigtigt at undersøge, om en given interaktion fungerer, men nærmere hvordan den skal være for at fungere. Da konteksten hvori HRI foregår har stor betydning for hvilke parametre, der har indflydelse på oplevelsen, vil feltundersøgelsen blive udført i en dansk lufthavn. Det vælges at feltundersøgelsen udføres i Aalborg Lufthavn, da kontakt og aftaler med Københavns Lufthavn er tidskrævende og ikke nødvendigvis kan efterkommes inden for projektperioden. Ydermere er det nemmere at transportere robotten til Aalborg Lufthavn, hvor det allerede er aftalt at udføre undersøgelsen. Tilmed er Aalborg Lufthavn mindre end Københavns Lufthavn, hvorfor der vil være mere ro til at udføre undersøgelsen.

Feltundersøgelser, hvor potentielle brugere mødesude i den virkelige verden samt i den rette kontekst, kan ofte give en større indsigt i, hvordan interaktionen med produktet, i det her tilfælde en social robot, faktisk foregår. Endvidere kan det afsløre nogle af de problemer eller holdninger til interaktionen med en social robot, som ikke nødvendigvis vil blive udtrykt i en laboratorieundersøgelse. I denne sammenhæng vælges det, at de rejsende skal interagere med robotten under specifikke brugsscenerier, hvorefter de vil blive bedt om at beskrive deres oplevelse, hvilket uddybes i følgende afsnit.

3.1 Metodeovervejelse

Formålet med undersøgelsen er at fastlægge hvilke parametre, der vurderes at have betydning for *Human-Robot Interaction* (HRI) for danske rejsende ved interaktion med en social robot i lufthavnen. Det er en eksplorativ undersøgelse, hvor det ønskes, at undersøge om der er ny viden at finde og ikke nødvendigvis kun at bekraefte allerede fundne resultater. Den indsamlede data forventes derfor at være kvalitativ, da der primært vil blive undersøgt hvilke ord danske rejsende bruger, når de skal beskrive interaktionen med en social robot.

Følgende afsnit beskriver, hvilket metodeovervejelser, der er gjort i forbindelse med udarbejdelsen af testdesignet til feltundersøgelsen. Der er undersøgt forskellige metoder, for at kunne fastlægge en passende tilgang og opstilling af undersøgelsen. De undersøgte metoder er beskrevet i det følgende.

3.1.1 Contextual Inquiry

Den kvalitative metode *Contextual Inquiry* kan anvendes til at indsamle data i den kontekst, hvor brugeren agerer og metoden undersøges derfor i forbindelse med metodevalg

til undersøgelsen. I undersøgelsen, hvor det ønskes at udlede hvilke parametre, der har indflydelse på HRI, er produktet, der interageres med robotten. Da robotten endnu ikke er implementeret i lufthavnen er det ikke muligt at udføre *Contextual Inquiry*, efter de foreskrevne retningslinjer. Brugerens har ikke kendskab til robotten i den aktuelle kontekst før selve undersøgelsen, som ellers er en af forudsætningerne for *Contextual Inquiry*. Selvom metoden som helhed ikke anvendes, er der elementer og tilgange fra den der bruges.

Testlederens rolle i undersøgelsen tilstræbes, at være som intervieweren er ved *Contextual Inquiry*, dog med visse forbehold. Ved *Contextual Inquiry* taler intervieweren smed testpersonen, mens interaktionen foregår. Dette vil ikke nødvendigvis være tilfældet i feltundersøgelsen, da det ønskes at robotten selv rekrutterer og interagerer med testpersonerne uden forstyrrelser fra en testleder.

Efterfølgende ønskes det at foretage et interview med testpersonen, hvor interaktion med robotten stadig kan foregå, hvor nogle af elementerne fra *Contextual Inquiry* medtages. Testlederen og testpersonen skal være på samme niveau, testlederen kommer derfor ikke til at fungere som en decideret leder, som titlen ellers indikerer. I stedet vil testlederen føre en samtale med testpersonen efter interaktionen og undervejs spørge ind til relevante interaktioner eller problemstillinger, samt uddybende kommenterer, hvis der er behov for det. Der opstilles ikke konkrete spørgsmål til selve interaktionen, men i stedet opstilles der samtaleemner, som testleder og testperson efterfølgende kan snakke om. Disse samtaleemner vil dække over overvejelser og problemstillinger i forhold til robotten, hvilket er inspireret af tilgangen der er ved *Contextual Inquiry*.

3.1.2 Laddering

Efter testpersonerne har interagerede med robotten, vil de blive bedt om at deltag i et interview. Under interviewet fokuseres der på at lade testpersonerne tale så meget som muligt, for at undgå at dreje samtalen væk fra de ting, som har optaget testpersonen. For at få uddybende information fra testpersonerne omkring de betydende parametre ved en social robot, ønskes det at stille spørgsmålene på samme måde som ved brug af *UX Laddering*, der er en tilpasset *Laddering* metode designet til undersøgelser af brugeroplevelse, (Abeele og Zaman 2017, ss. 3-4). Ved brug af *UX Laddering* stilles testpersonen eksempelvis et spørgsmål om hvordan de oplevede interaktionen med robotten. Når testpersonen svarer, bliver deres svar overvejet, hvorefter der stilles uddybende spørgsmål ind til det. Sådan fortsættes det op ad trappen, indtil eventuelle værdier omkring deres holdning til den sociale robot er blevet fastlagt. Når *UX Laddering* skal bruges, er kontekst vigtig, (Abeele og Zaman 2017, s. 3), hvorfor det er positivt, at testpersonerne er blevet præsenteret for robotten i en kontekst, som den reelt set kunne indgå i.

Når en testperson starter med at besvare spørgsmål, er det sandsynligt at det foregår ved at opremse funktioner i stedet for specifikke egenskaber, (Abeele og Zaman 2017, s. 3). I det henseende er det vigtigt at forsøge at få testpersonen til at træde et par skridt ned ad trappen igen, så spørgsmålene og svarene i denne undersøgelse starter ved robottens specifikke egenskaber. Det kan opnås ved at stille spørgsmålet: *Hvad skyldes det?*, frem for at spørge hvorfor testpersonen mener noget. Ydermere er det vigtigt, at interviewet ikke tager for lang tid, (Abeele og Zaman 2017, s. 4). Testlederne skal naturligvis spørge videre ind, så testpersonen kommer til at tale om de bagvedliggende værdier, men kun hvis det virker naturligt i samtalen. Det er ikke alle oplevelser, der trigger virkelige værdier,

hvorfor det kan være nødvendigt at stoppe med at stille spørgsmål, hvis det ikke længere virker naturligt.

UX Laddering anvendes i feltundersøgelsen til at forme interviewet og forstå værdierne bag testpersonens udtalelser om robotten, hvorfra de vigtige parametre forventes at kunne udledes.

3.1.3 Overvejelser til interviewer

Det kan være relevant at fortælle testpersonerne om den beskrevne *Laddering* interviewmetode, så de ikke bliver irriteret over de konstant uddybende og opklarende spørgsmål. Derudover kan det være relevant at fortælle testpersonerne, at det ikke er projektgruppen, der har designet *Double*-robotten, hvorfor de frit kan fortælle hvad der falder dem ind, når de interagerer med robotten. Hvis testpersonerne begynder at snakke om generelle oplevelser eller hvordan andre vil forstå robotten, er det vigtigt at spørge dem ind til deres egne oplevelser frem for andres.

Ved afviklingen af interviewet er det vigtigt, at testlederen fremstår venlig og tilstedeværende, i stedet for at følge et manuskript. Derudover kan det være en god idé at "spille dum", for virkelig at få testpersonerne til at beskrive hvad de mener, frem for selv at tolke på det de siger. Det gør sig også gældende, selvom det de siger kan virke åbenlyst. Ydermere skal testlederen tænke over kropssprog og mimik, da det er nemt for testpersonen at fornemme, hvornår de har sagt noget, som testlederen havde håbet på og omvendt. Alle forudindtagelser og erfaringer fra tidligere tests skal derfor glemmes, så der på den måde kun fokuseres på testpersonens svar.

Testdesign 4

Følgende afsnit vil belyse de forskellige aspekter af feltundersøgelsen fra testens omfang, hvilke brugsscenerier og samtaleemner, der gøres brug af, rollefordelingen samt hvilke testpersoner, der tilstræbes at rekruttere til hvor testen afvikles og med hvilket udstyr samt fremgangsmåden.

4.1 Testens omfang

Med udgangspunkt i og for at besvare problemstillingen: *Ud fra hvilke parametre beskriver danske rejsende interaktion med en social robot i en dansk lufthavn?*, vil omfanget af denne test primært indebære en feltundersøgelse af, hvilke parametre danske rejsende tilskriver interaktionen med en social robot i en dansk lufthavn. Der vil derfor udelukkende indsamles kvalitativ data, som efterfølgende analyseres for, at udlede de parametre, der har indflydelse på interaktionen med en social robot. Formålet med feltundersøgelsen er derfor hverken at evaluere de rejsendes oplevelse af interaktionen med robotten, teste specifikke parametre, som potentelt kan have indflydelse på interaktionen eller at designe robottens bevægelsesmønstre. Dog vil det tilstræbes, at opnå et så naturligt brugsscenario som muligt, hvorfor testpersonerne skal løse en opgave ved hjælp af robotten i Aalborg Lufthavn.

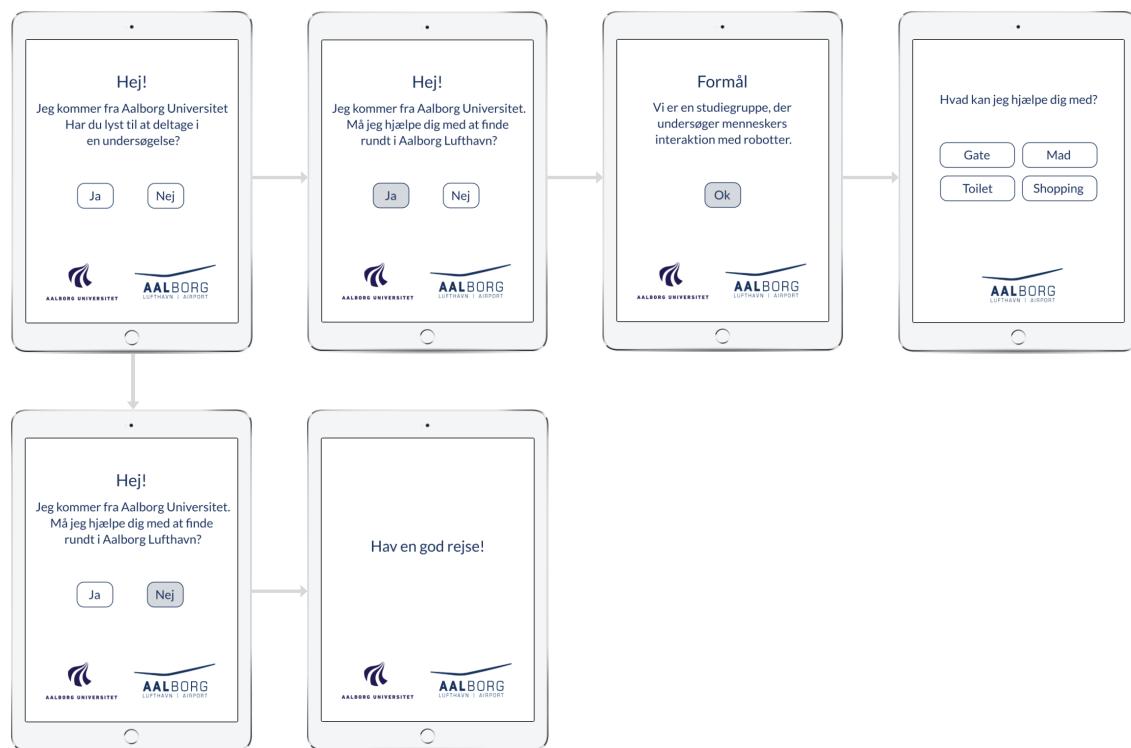
4.2 Testpersoner

Testpersonerne, de rejsende, bliver rekrutteret direkte i Aalborg Lufthavn i området efter sikkerhedskontrollen. Det er som udgangspunkt robotten, der står får rekrutteringen ved at henvende sig til de rejsende med spørgsmålet: *Jeg kommer fra Aalborg Universitet. Må jeg hjælpe dig med at finde rundt i Aalborg Lufthavn?*, har den rejsende lyst til at deltage vil robotten føre dem igennem nogle foruddefineret brugsscenerier. Fordelen ved at få robotten til at rekruttere testpersoner er, at testpersonerne får et upåvirket førstehandsindtryk af robotten, som det vil være tilfældet første gang de oplever robotten i lufthavnen.

Det tilstræbes, at foretage underøgelsen på både kvinder og mænd, gerne med forskellige aldre og rejseformål; forretningsrejse eller ferierejse, hvis muligt vil det ydermere bestræbes at inddrage førstegangs rejsende. Derudover er det et krav, at den rejsende er dansktalende for at undgå, at vigtige pointer går tabt i oversættelsen, når data efterfølgende skal behandles. Antallet af testpersoner er ikke forudbestemt, da undersøgelsen i stedet afsluttes, når der opnåes mætning, hvorved der ikke indsamles ny viden. Dette vurderes af de tilstede værende gruppemedlemmer.

4.3 Brugsscenerier

Da det tilstræbes at få et så naturligt scenario, som muligt vil der ikke blive stilt en decideret opgave som testpersonerne skal løse men for at få de rejsende til, at sætte deres egne ord på oplevelsen af interaktionen med en social robot defineres der fire brugsscenerier, som anses for at forekomme naturligt i en lufthavn. De fire brugsscenerier henvender sig enten til at den rejsende skal finde gate information, toiletfaciliteter, indkøbsmuligheder eller forplejning. Det er derfor op til testpersonen selv at vælge, hvilket af de fire brugsscenerie, der gemmengåes. De fire brugsscenerier gengives i *wireframe* designet i *Marvel*, som præsenteres på *Double's* skærm. *Wireframed* fremgår af Figur 4.1.



Figur 4.1. Oversigt over de forskellige skærmbilleder i det samlede *wirefram*, før der vælges brugsscenerie.

På Figur 4.1 illustreres de første skærmbilleder inden testpersonen vælger hvad robotten skal hjælpe dem med, jævnfør brugssceneriet. Derudover illustreres skærmbillederne i situationer hvor en rejsende ikke har lyst til at få hjælp af robotten. De fire forskellige brugsscenerier, som præsenteres afhængigt af testpersonen respons til spørgsmålet: *Hvad kan jeg hjælpe dig med?*, indgår ikke på Figur 4.1 men vil i de følgende afsnit blive præsenteret.

Gate information

Vælger testpersonerne at finde information om deres gate vil følgende skærbilleder blivet præsenteret på skærmen, jævnfør Figur 4.2.

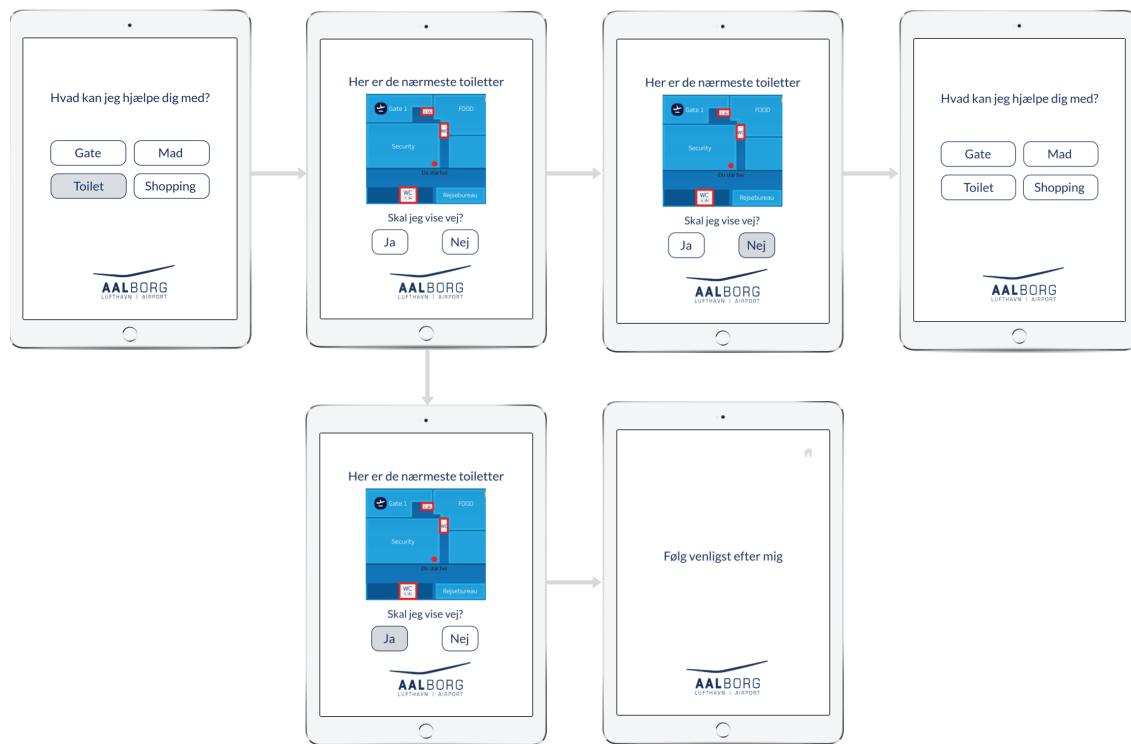


Figur 4.2. Oversigt over de forskellige skærbilleder, der anvendes i brugsscenariet: Gate information.

Figur 4.2 er blot et eksempel på hvilke skærbilleder testpersonen præsenteres for, da den valgte destination er *Gran Canaria*, vælges en af de andre destinationer vil navnet på den destination blive markeret og efterfølgende vil skærbillederne være tilpasset denne destination. Vælger testpersonen at trykke på *Udnyt tiden* bliver de spurgt om hvad de har lyst til at lave mens de venter. Vælger testpersonen *Shopping* vil skærbillederne være de samme som illustreret på Figur 4.4, hvis testpersonen vælger *Mad* vil skærbillederne være de samme som illustreret på Figur 4.5.

Toiletfaciliteter

Vælger testpersonerne at finde information om toiletfaciliteter vil følgende skærbilleder blivet præsenteret på skærmen, jævnfør Figur 4.3.



Figur 4.3. Oversigt over de forskellige skærbilleder, der er specifikke for brugsscenariet: Toiletfaciliteter.

Indkøbsmuligheder

Vælger testpersonerne at finde information om indkøbsmuligheder vil følgende skærbilleder blivet præsenteret på skærmen, jævnfør Figur 4.4.

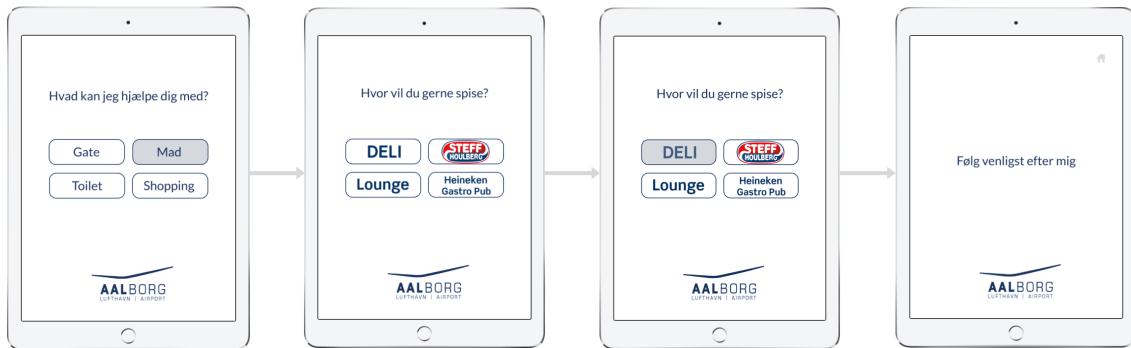


Figur 4.4. Oversigt over de forskellige skærbilleder, der anvendes i brugsscenariet: Indkøbsmuligheder.

Figur 4.4 er blot et eksempel på hvilke skærbilleder testpersonen præsenteres for, da den valgte indkøbsmulighed er butikken *Belsac*, vælges en af de andre muligheder vil navnet på den valgte butik blive markeret.

Forplejning

Vælger testpersonerne at finde information om forplejning vil følgende skærmbilleder blivet præsenteret på skærmen, jævnfør Figur 4.5.



Figur 4.5. Oversigt over de forskellige skærmbilleder, der anvendes i brugsscenariet: Forplejning.

Figur 4.5 er blot et eksempel på hvilke skærmbilleder testpersonen præsenteres for, da den valgte forplejning er *DELI*, vælges en af de andre muligheder vil navnet på den valgte forplejning blive markeret.

4.4 Samtaleemner

Mens undersøgelsen står på er det vigtigt, at holde samtalen i gang for at indsamle data, der gengiver testpersonernes oplevelse af interaktionen med robotten. For at sikre at testlederen har relevante emner at spørge ind til er det valgt, at opstille samtaleemner for undersøgelsen. Samtaleemnerne er inddelt i to grupper: løbende og afsluttende samtaleemner. Alle samtaleemner er vejledende, hvorfor de ikke nødvendigvis bliver formuleret ordret af testlederen.

4.4.1 Løbende samtaleemner

Samtaleemner henvender sig primært til emner, der forventes at være relevante for testpersonernes oplevelse af interaktionen med robotten. Disse samtaleemner udgør interviewet med testpersonen, som skal afsløre hvilke parametre, der er vigtige for testpersonen ved en social robot. Samtaleemnerne er som følger:

- Førstehåndsindtryk af robotten - fra rekrutteringen
- Måden hvorpå robotten henvender sig
- Hvad testpersonen synes om robotten
- Hvad testpersonerne tror andre rejsende tænker om interaktionen
- Robottens relevans
- Robottens pålidelighed
- Normal oplevelse i en lufthavn uden hjælp fra en robot

Da der gøres brug af elementer fra *laddering*, jævnfør Afsnit 3.1 (*Metodeovervejelse*), vil de listede samtaleemner udgøre grundlaget for samtalen, men ikke nødvendigvis være det eneste der bliver snakket om.

4.4.2 Afsluttende samtaleemner

De afsluttende samtaleemner vil blive anvendt som debriefing af testpersonen og afrundning på undersøgelsen. Ydermere vil opklarende spørgsmål omkring parametre, som testpersonen ikke selv har nævnt blive stillet, hvis testlederen føler behovet opstår. Samtaleemnerne er som følger:

- Hvad synes du om..
 - Robottens hastighed?
 - Robottens højde?
 - Robottens afstand til dig?
 - Robottens generelle bevægelse?
 - Robottens udseende?
 - Den retning robotten henvendte sig til dig fra?
- Hvor gammel er du?
- Hvor ofte flyver du?
- Spørgsmål og/eller kommentarer til undersøgelsen
- Spørgsmål og/eller kommentarer til vores projekt
- Afrunding og ønsk testpersonerne god rejse

4.5 Rollefordeling

For at udføre feltundersøgelsen er det nødvendigt at definere nogle roller, som relaterer sig til specifikke dele af undersøgelsen. Der vil i alt blive defineret tre roller.

Robotstyrer

Da de rejsende skal interagere med en *Double*-robot, som ikke kan forudprogrammeres er det nødvendigt, at have én til at styre robotten. For at de rejsende får et indtryk af at robotten er autonom og har en form for social intelligens er det favorabelt, at den rejsende ikke registrerer, at det er en person, som styrer robotten. For at efterkomme det vil personen, som styrer robotten være placeret så de rejsende ikke direkte kan se hvordan robotten styres, dog skal personen stadig have mulighed for at overvære og høre interaktionen mellem robot og den rejsende, så robotten kan styres derefter. Udover placeringen af personen, som styrer robotten, bliver de rejsende fortalt af testlederen, at personen vil tage supplerende noter.

Testleder

Testlederen har flere opgaver, først og fremmest er det testlederen de rejsende kommunikerer med, udover robotten, hvorfor det er vigtigt at testlederen sørger for at den rejsende føler sig tryg. Testlederen gengiver formålet med undersøgelsen: At finde ud af hvordan danske rejsende beskriver interaktionen med en social robot i en lufthavn. I den forbindelse nævner testlederen at der vil blive optaget lyd og såfremt at testpersonerne giver mundtligt samtykke til det, starter testlederen lydoptageren

Testlederen har desuden til opgave at styre samtalen ind på de nævnte samtaleemner beskrevet i Underafsnit 4.4.1 (*Løbende samtaleemner*). Afslutningsvist har testlederen til

opgave, at debriefe den rejsende omkring undersøgelsen samt at afslutte forløbet. Da sam-talen mellem testleder og testperson varierer udarbejdes der ikke specifikke instruktioner. Dog sørger testlederen for at introducere testpersonen til hvem projektgruppen er og at der vil blive taget noter undervejs. Fordi der anvendes elementer fra *laddering* vil testlederen informere testpersonerne om at der vil blive spurt en del ind til testpersonens udsagn.

Når robotten ikke interagerer med en testperson er det testlederens opgave, at holde øje med at robotten ikke kører ind i noget eller at andre rejsende ikke går ind i den.

Observatør

Der vil i alt være tre observatører til stede, hvor minimum én af dem vil observere interaktionen mellem testperson og robot før robotten leder testpersonen over til testlederen. Da det tilstræbes at interaktionen mellem testperson og robot er så naturlig som muligt, vil observatøren holde afstand til dem. De to andre observatører sidder med forskelligt udsyn til både interaktionen mellem robot og testperson, og ved interviewet. Formålet med at have tre observatører er for at være sikker på, at der indsamlas så meget data som muligt og fordi det forventes at observatørerne bemærker forskellige ting. Derudover vil der ikke optages video, hvorfor det er vigtigt at have gode notater. Ydermere vil det tilstræbes at én observatør fokuserer på testpersonernes mimik, én anden observatør fokuserer på kropsholdning og den sidste observatør fokuserer på decideret pointer koblet til en robot aktion. Er det muligt for observatørerne, at notere flere ting så gør de bare det. Observatørens noter vil blive nedskrevet på notespapir.

4.6 Testlokation og udstyr

Eftersom at robotten fortrinsvist skal indgå i en lufthavn og da formålet med undersøgelsen er, at undersøge ud fra hvilke parametre danske rejsende beskriver interaktionen med en social robot i en dansk lufthavn, er det favorabelt at udføre undersøgelsen i en dansk lufthavn. Som nævnt har det ikke været muligt at udføre undersøgelsen i Københavns Lufthavn, hvor robotten er tiltænkt, hvorfor undersøgelsen foretages i Aalborg Lufthavn. Lufthavnsvalget har formentlig indflydelse på testpersonernes oplevelsen og dermed den indsamlede data, da der er stor forskel på de to lufthavne. Det kommer særligt til udtryk ved antal rejsende, hvor Københavns Lufthavn, tiltrods for et fald på 3000 rejsende siden sidste år, servicerede 2,672,710 rejsende i september måned, (K. Lufthavn 2017a). Det er en del flere rejsende end hvad Aalborg Lufthavn oplevede i samme måned, hvor der blev serviceret 155,547 rejsende, (A. Lufthavn 2017a). Derudover er der stor størrelsес mæssig forskel på de to lufthavne, hvor Aalborg Lufthavn har én terminal med 11 gates, (A. Lufthavn 2017b), har Københavns Lufthavn to terminaler med 102 gates, (K. Lufthavn 2017c). Tiltrods for størrelsesforskellen, er der stadig stor forskel mellem Københavns Lufthavn og nogen af verdens største lufthavne, eksempelvis Europas største lufthavn: London Heathrow Airport, som i 2016 servicerede 75.7 millioner rejsende, (Heathrow 2017). Til sammenligning servicerede Københavns Lufthavn 29 millioner rejsende i 2016, (K. Lufthavn 2017b).

Alene baseret på antallet af terminaler forventes det, at det tilmed bliver sværere at lokalisere sin gate dels på oversigtstavlerne og dels i forhold til gatens fysiske lokalitet, hvorfor det ligeledes forventes, at der er større behov for den sociale robot

i Københavns Lufthavn sammenlignet med Aalborg Lufthavn. Dette kan potentielt begrænse testpersonernes forståelse for, hvorfor det kan være en fordel at interagere med robotten, fremfor at finde informationen selv, da det ikke er sikkert at de oplever behov for assistance.

Selvom der kan være behov for en social robot i indgangshallen, hvor de rejsende blandt andet skal igennem check-in og aflevere bagage, vælges det at udføre testen på den anden side af sikkerhedskontrollen. Det skyldes blandt andet antagelsen om, at de rejsende har mindre travlt og er mindre stresset efter de har passeret sikkerhedskontrollen og nu kun venter på at boarde flyet. Derudover er det begrænset, hvilke muligheder der er i Aalborg Lufthavn eksempelvis i forhold til shopping og café besøg, hvorfor det forventes, at det er nemmere at rekruttere testpersoner, da de formentlig bruger det meste af deres ventetid på at sidde og vente, fremfor at besøge lufthavnens butiksområde. Dog kan en ulempe være, at de rejsende ankommer senere til lufthavnen fordi den ikke er så stor, at det er nødvendigt at ankomme de anbefalede to timer i forvejen.

I henhold til det praktiske aspekt ved at udføre undersøgelsen vil der ikke foretages nogle synelige markeringer af testområdet, da det formentlig vil have indflydelse på, hvor realistisk situationen opleves af testpersonerne. Derimod vil testområdet være kendt dels af projektgruppen og dels af personalet i lufthavnen.

Til at udføre undersøgelsen er der behov for følgende udstyr:

- *Double*-robot
- iPad Air 2
- Vinklede hovedbeslag
- Computere med internetadgang
- Notespapir og kuglepen/blyant
- Lydoptager
- Guide til samtaleemner
- Bord og stol til observatør og robotstyre
- Højbord til testleder og testperson

Årsagen til at der gøres brug af en *Double*-robot er, at Karl's robot stadig er under udvikling, hvorfor det ikke er muligt, hverken at foretage en decideret interaktion eller undersøgelse med robotten. *Double*-robotten stilles til rådighed af Karl. Fordelen ved at anvende *Double* er, at det er et færdig udviklede produkt, som bevæger sig nogen lunde på samme måde, som Karl's robot er tiltænkt. Derudover er der flere udseendsmæssige træk, der er ens for de to robotter, dels at robotten enten balancere på en cylinder eller en kugle, dels at den øvre og nedre del af robotten forbinder med en stang og dels at robottens hovede består af en form for tablet, eksempelvis en iPad. Ydermere er *Double* let at styre, da den styres med piletasterne på ens tastatur samtidig med at det er muligt, at se hvor robotten bevæger sig, da den gør brug af det indbyggede kamera i iPad'en. Ulempen ved at anvende *Double* er, at det er et færdigt produkt, hvorfor det potentielt kan påvirke testpersonerne til, at tilbageholde eventuelle negative holdninger i frygt for at såre os.

Selvom der er flere udseendsmæssige ligheder mellem de to robotter er det stadig to forskellige robotter med hver deres udtryk, hvorfor det ligeledes kan være en ulempe at

undersøgelsen ikke foretages med Karl's robot. En af forskellene mellem de to robotter er hvad de balancere på cylinder kontra kugle, hvor Karl's robot, som balancerer på en kugle, har mulighed for at bevæge sig i alle retning, hvilket ikke er tilfældet med *Double*. Det kan have betydning for, hvor godt robotten egner sig til at indgå i F-formationer, jævnfør Figur 1.7, hvor det netop er naturligt at bevæge sig sidelæns for at inkludere flere i samtalen.

Hovedet på *Double*-robotten består af en iPad Air 2, med internetadgang. Det er dog muligt at anvende andre udgaver og versioner af Apple's iPad. Da testpersonerne interagerer med robotten via det udviklede *wireframe* er det ikke muligt, at se hvilket brugsscenario testpersonen vælger. For at løse det problem vælges det at lave skærmdeling fra iPad'en til en MacBook Pro gennem programmet *5kPlayer*, det kræver dog at både iPad, computeren hvorfra robotten styres og computeren hvorfra skærmdelingen gengives er forbundet til det samme netværk. Ved skærmdelingen er det ikke muligt at se hvad testpersonen trykker på men det er muligt at se hvilket skærmbillede de præsenteres, hvorfor det er muligt at styrer robotten derefter.

Baseret på et miniprojekt fra kurset: *Applied Experimental Psychology and Psychophysics*, hvori det blev undersøgt hvor indbydende robotten blev perciperet afhængigt af dens hovedposition, besluttes det at anvende et vinklede hovedbeslag. Dette hovedbeslag har til formål at vinkle hovedet på *Double*-robotten, så den fremstår indbydende og klar til at interagere med testpersonerne. Hovedbeslaget er designet af Karl Damkjær Hansen, som har 3D printet beslaget, og er illustreret på Figur 4.6.



Figur 4.6. Illustration af det 3D printede vinklede hovedbeslag designet af Karl Damkjær Hansen i forbindelse med projektsamarbejdet.

Ved at erstatte det originale beslag, som medfølger robotten, med det vinklede beslag, illustreret på Figur 4.6, vil det vinkle robottens hovede med 25° fra lodret. Den modificerede *Double*-robot illustreres på Figur 4.7, Figur 4.8 og Figur 4.9.



Figur 4.7. Front.

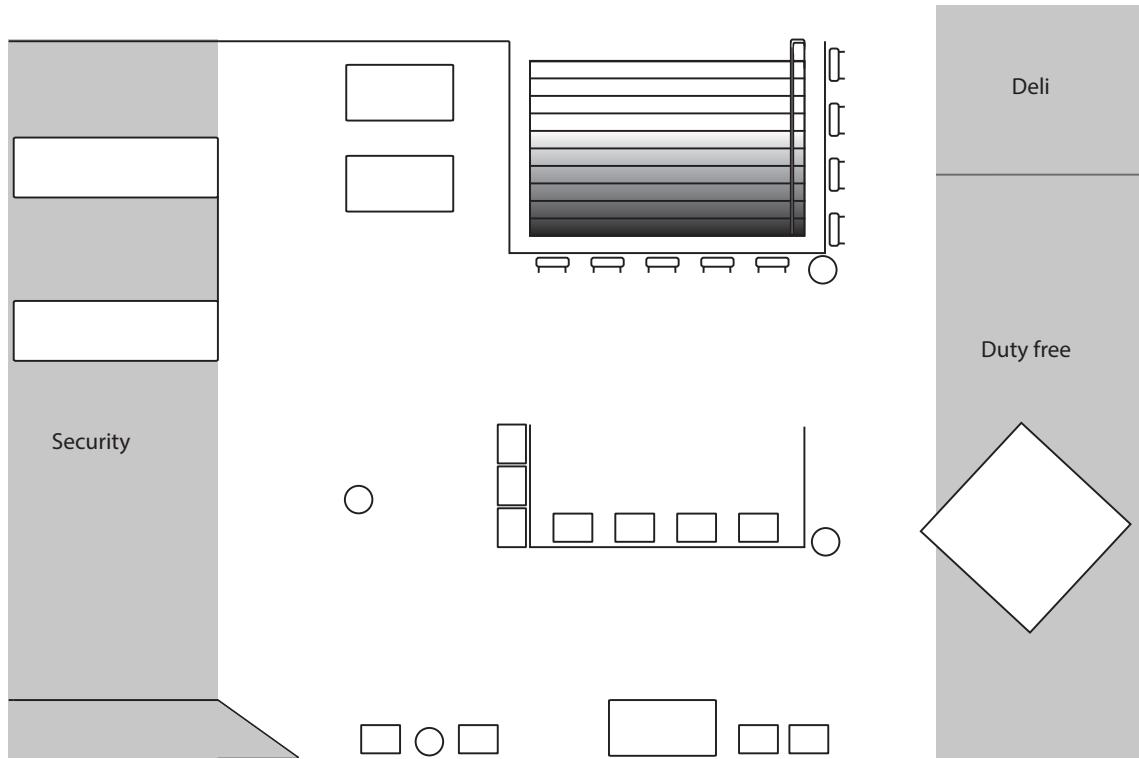


Figur 4.8. Profil.

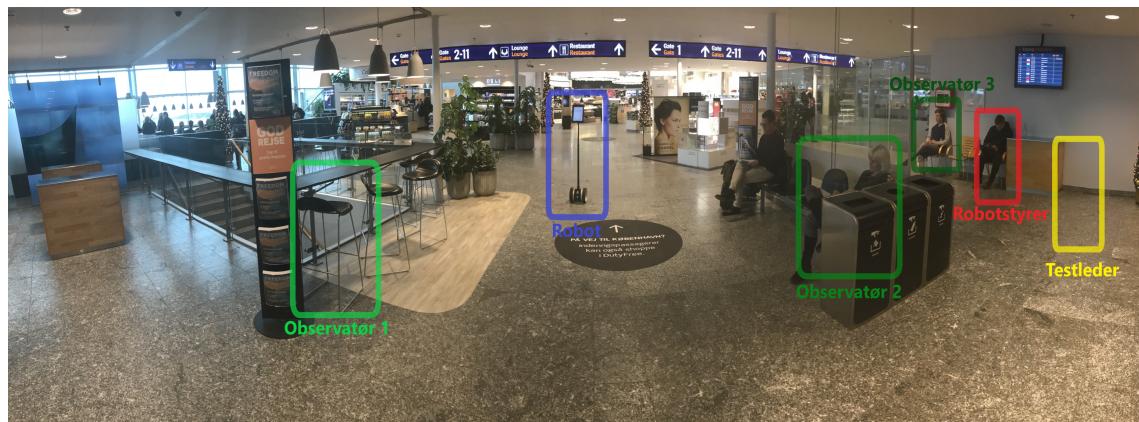


Figur 4.9. Profil nært.

Robotstyren får tildelt en computer med interadgang. Det er særligt vigtigt at robotstyrens computer har interadgang da den skal kommunikere med robotten, hvorfor det kan være nødvendigt at oprette et lukket netværk, eksempelvis ved at lave internetdeling fra ens egen telefon. Igennem *Doubles* egen hjemmeside er det muligt at uploadet linket til *wireframen*. Observatørerne får tildelt notespapir og kuglepen eller blyant. Lydoptageren anvendes af testlederen til at optage den auditive respons fra testpersonerne, der vælges at lydoptagelsen foretages på en af projektgruppens telefoner. Guide til samtaleemner printes for, at den pågældende testleder altid ved hvilke samtaleemner testpersonen og testlederen skal igennem. Bord og stole til observatører og robotstyre placeres så de ikke er i vejen for testlederen og testpersonen, men så det samtidig er muligt for både observatører og robotstyre at overvære interaktionen og høre samtalen. Højbordet indgår for at testlederen og testpersonen har en form for base, hvor testlederen introducere testpersonen til undersøgelsen og fører samtalen. Derudover vil der på højbordet blive sat forplejning til rådighed. På Figur 4.10 fremgår en plantegning af testopstillingen i AAL, hvor Figur 4.11 afspejler den reelle testopstilling.



Figur 4.10. Plantegning over testopstillingen i AAL. De objekter, der fremgår på plantegningen er steder hvor robotten ikke kan køre. Objekterne er enten pakkeborde, skraldespande, stole eller standere. Cirklerne indikerer de planter eller reklamestandere, der er i lufthavnen.



Figur 4.11. Panorama af testlokationen, hvor testopstillingen afspejler de muligheder, der var i Aalborg Lufthavn. Rollerne er markeret med grøn for hver af de tre observatører, blå for robotten, rød for robotstyreren og gul er for testlederen.

4.7 Fremgangsmåde

Så snart en af de rejsende befinder sig i det aftalte område vil robotten henvende sig ved at køre hen til den rejsende og spørge om den kan hjælpe dem med at finde rundt i Aalborg Lufthavn. Svarer testpersonen *Ja* vil der efterfølgende ganske kort stå, på skærmen, hvad formålet med testen er: At undersøge menneskers interaktion med robotter.

Derefter har testpersonen mulighed for selv, at vælge et af de fire brugsscenerier beskrevet i Afsnit 4.3 (*Brugsscenerier*). Uanset hvilket brugsscenario, der bliver valgt vil det ende med at robotten opfordrer testpersonen til at følge efter. Istedet for at følge testpersonen det rigtige sted hen følger robotten testpersonen hen til testlederen, som tager over. Testlederen har til opgave at small-talke med testpersonen for at skabe en behagelig stemning og undervejs ledes small-talken over i samtaleemnerne, jævnfør Underafsnit 4.4.1 (*Løbende samtaleemner*). Small-talken vedrører eksempelvis hvor testpersonen skal rejse hen, om de har rejst fra Aalborg Lufthavn før og lignende. I tillæg vil testpersonen introducere projektgruppen og at der sidder nogen og tager noter.

Inden samtalen drejes ind på samtaleemnerne skal testlederen sørge for at give testpersonen tilstrækkelig information omkring testen til at testpersonen kan give et mundtligt samtykke, som tillader at der kan optages lyd. Som nævnt vil der undervejs blive taget noter både før testpersonen er i kontakt med testlederen og efter når de to konverserer.

Da det er robotten, som har stået for rekrutteringen og der ikke har været en menneske-menneske interaktion før robotten følger testpersonen hen til testlederen, vil det første samtaleemne vedrøre testpersonens førstehåndsintryk dels af hvordan robotten henvendte sig til testpersonen og dels af hvordan det var at interagere med robotten. Da der anvendes elementer fra *laddering* er det en fordel at testlederen gør det klart, at der vil blive stilt en del spørgsmål, som skal få testpersonen til at reflektere over sin oplevelse, så det ikke virker påtrængende at der bliver stilt alle de spørgsmål. Undervejs i samtalen vil robotstyren sørger for at robotten køre rundt på forskellige måder, varierer afstanden til testpersonen samt højden på robotten. Det forventes at det er med til at få testpersonerne til at kommentere på netop robottens bevægelse.

Når alle samtaleemner er blevet inddraget og diskuteret er det testlederens opgave at afslutte undersøgelsen ved at debriefe testpersonen og ønske dem en god rejse. Under debriefingen vil testpersonen blive spurgt om alder og hvor ofte de rejser, hvor kønnet noteres af en af observatørerne. Der vil derudover bliver spurgt ind til nogle forskellige parametre, såsom robottens højde, hastighed, afstand, generelle bevægelse, udseende og indgangsvinkel, jævnfør Underafsnit 4.4.2 (*Afsluttende samtaleemner*).

Trykker den rejsende derimod *Nej* for ikke at deltage i undersøgelsen, ønsker robotten personen god rejse og forlader stedet og venter til at en ny rejsende befinner sig i området.

Efter hver testperson vil observatørerne konferere med hinanden og dele observationer eksempelvis i forhold til, hvad der kan være værd at fokusere på, men også i forhold til hvordan forskellige ting noteres, så der er en form for konsensus om noteringen. Ydermere vil der efter to eller tre testpersoner blive roteret roller blandt projektgruppen.

4.8 Pilottest

Ved udviklingen af testdesignet er de forskellige elementer i testen løbende blevet testet, primært på projektgruppens medlemmer. I følgende afsnit vil de forskellige elementer blive beskrevet.

4.8.1 Styring af robotten

Robotstyreren skal øve sig i at styre robotten, derfor kørte vedkommende robotten rundt på Aalborg Universitet, Frederik Bajers Vej 7, mens robotstyreren sad i grupperummet med en computer forbundet til *Double*. På den måde fik robotstyreren erfaring med at styre robotten ved brug af dens kamera. Robotten kørte rundt blandt studerende, hvorfor navigation blandt mennesker blev øvet. Det blev testet at få robotten til at dreje og køre i alle retninger og da gulvet er forskelligt rundt på Universitet, blev det også testet at køre robotten på forskellige underlag og over samlinger i gulvet.

Da det kan være svært at vurdere hvor tæt robotten kommer på personer omkring den, blev dens henvendelse testet og øvet ved at lade robotstyreren køre robotten hen til personer, for derefter at sammenligne afstanden mellem robotten og den person den henvender sig til, med hvad robotstyreren havde forventet at afstanden ville være.

4.8.2 Brugergrænseflade

Brugergrænsefladen, som er udviklet i *Marvel*, blev testet ved at lade et gruppemedlem, som ikke havde været med til at designe brugergrænsefladen, gennemgå alle mulige scenarier. Testningen af brugergrænsefladen er foretaget som en iterativ proces, hvor tilføjelser og ændringer er foretaget med det samme, hvorefter brugergrænsefladen er testet igen. Dette gøres indtil projektgruppen vurderer, at der ikke længere kan findes væsentlige fejl og mangler ved det.

4.8.3 Testdesignet

Efter de forskellige elementer af testen er på plads udføres den første sammenhængende pilottest. Her er det med gruppens vejleder, Dorte, som testdeltager. Formålet med at bruge gruppens vejleder som testperson i pilottesten er, at det både giver mulighed for at teste designet og efterfølgende evaluere på testdesignet i samarbejde med vejlederen. Denne evaluering resulterede i, at det blev fastlagt, at der ikke skal være nogen interaktion mellem interviewer og testperson før efter robotten har fulgt testpersonen hen til intervieweren. Fordelen ved at undgå kontakt mellem interviewer og testperson indtil selve interviewet er, at testpersonens interaktion med robotten derved ikke bliver forstyrret og det bliver derfor en mere økologisk oplevelse for testpersonen.

Efter den første sammenhængende pilottest udføres der i alt tre pilottests, hvor det er forskellige medlemmer fra projektgruppen der er testpersoner og interviewere. Disse pilottests har primært til formål, at give intervieweren erfaring med at udføre undersøgelsen og interviewet samt få tilpasset det således at undersøgelsen udføres som det er besluttet. Det er valgt at der ved udførelsen af undersøgelsen skal være to gruppemedlemmer, der skiftes til at være interviewer, derfor skal begge interviewere også prøve at have denne rolle i en pilottest. Udbyttet af disse pilottests er at gruppen får forventningsafstemt, hvordan udførelsen skal være for at det stemmer oversens med det der er besluttet.

Ved den sidste pilottest, der udføres, rekrutteres et gruppemedlem fra en anden gruppe end projektgruppen, til at deltage i testen. På den måde er testpersonen i denne pilottest ikke bekendt med formålet med testen og overvejelserne bag testdesignet, så forventningen er at feedbacken ikke vil være påvirket i samme grad som tilfældet er når

det er projektgruppen eller vejlederen, der gennemgår testen.

Næste step i forhold til pilottestning burde være at udføre en pilottest i det miljø som undersøgelsen reelt skal foregå i. Da projektgruppen ikke har ubegrænset adgang til Aalborg Lufthavn er det ikke sikkert at der er mulighed for, at udføre flere undersøgelser i lufthavnen over flere dage. Det vælges derfor at udføre undersøgelsen i lufthaven, hvor der ved de første testpersoner er fokus på om testdesignet fungerer som forventet. Det vælges at accepterer, at der kan forekomme ændringer i testdesignet undervejs i interviewet, som er nødvendige for at få det optimale udbytte af undersøgelsen. Desuden vil der naturligt forekomme variationer fra testperson til testperson, da undersøgelsen udføres i et naturligt miljø, hvor det ikke er muligt, eller ønsket, at kontrollere variabler, der påvirker interviewet og testpersonernes interaktion med robotten. Det anslåes at interviewet varer omkring fem minutter, hvor interaktionen med robotten forventes at være kortere.

Databehandling 5

I dette kapitel er der fokus på at behandle det indsamlede data fra feltundersøgelsen. Først beskrives hvilke ændringer, der er foretaget i forbindelse med testdesignet som resultat af testlokationen. Dernæst blyses det hvilke testpersoner, der har deltaget i undersøgelsen, hvor det efterfølgende vil blive beskrevet, hvordan den indsamlede data behandles. Efterfølgende vil kapitlet handle om at analysere det udarbejdede *affinity diagram*, hvorfra det bør være muligt at udlede, hvilke parametre danske rejsende tilskriver interaktionen med en social robot i en dansk lufthavn.

5.1 Ændringer af testdesign

De ændringer, der er foretaget i forbindelse med at udføre undersøgelsen i Aalborg Lufthavn vil være relateret til udstyr og fremgangsmåde.

Grundet placeringen i lufthavnen, efter sikkerhedskontrollen, var det ikke muligt at opsætte bord og stole til observatører og robotstyre som de kunne sidde ved. Dog var der tre stole til rådighed, én til robotstyren, to til to af de tre observatører. Ellers var det muligt at sidde andre steder, hvor der stadig var udsyn til både testpersonens interaktion med robotten og til interviewet. Ved at have spredt observatørerne over et større areal har det været muligt at indsamle observationer fra flere vinkler og med forskellige formål. Derudover blev der ikke opsat et højbord til testleder og testperson, dette blev dog ikke betragtet som et problem, da testlederen alligevel ikke skulle tage noter undervejs i interviewet. Det var ydermere ikke muligt at foretage skærmdelingen mellem hvad der præsenteres på robotten og en anden computer end den hvorfra robotten styres, fordi det krævede et lukket netværk. Det blev derfor besluttet at minimum én af observatørerne skulle være placeret et sted, hvorfra det var muligt at se hvad den rejsende trykkede på skærmen og signalere det til robotstyren, så robotten kunne styres derefter.

Forud for undersøgelsen var der ikke taget nogle beslutninger om, hvordan det skulle håndteres hvis mere end én rejsende interagerede med robotten og efterfølgende blev interviewet. Det blev derfor besluttet i lufthavnen, at det ikke ville være et problem og såfremt der var mere end én, der interagerede med robotten og efterfølgende fulgte robotten hen til interviewet, blev de alle sammen inddraget i interviewet. Dog har det indflydelse på den indsamlede data, da flere testpersoner i ét interview nødvendigvis ikke alle svarer på det samme spørgsmål. Derudover tages der ikke højde for mulige påvirkning blandt testpersonerne, som deltog i det samme interview. Fordelen ved at lade mere end én testperson interagere med robotten og deltagte i det efterfølgende interview er, at det vil afspejle potentielle naturlige situationer i en lufthavn, hvor de rejsende ikke nødvendigvis rejser alene.

Under interviewet blev der ikke lagt så meget vægt på at fortælle testpersonerne

om måden, hvorpå de blev interviewet. Laddering metoden blev ydermere ikke brugt så grundigt, da det ikke altid faldt naturligt at spørge uddybende ind under interviewet. I stedet holdt testlederen sig til de nedskrevne samtaleemner for at få indsigt i testpersonernes oplevelse med robotten.

5.2 Testpersoner

Der er i alt 30 testpersoner fordelt på 18 interviews, der har deltaget i undersøgelsen. Ud af de 18 interviews er der foretaget 11 enkeltpersonsinterviews, hvor de resterende syv interviews er foretaget med mere end én testperson. Testpersonerne udgør 16 kvinder og 14 mænd i et aldersspænd fra 8 til 62 år, ($M=37.9$, $SD=17.1$). Der er dog fire testpersoner, hvor alderen ikke er opgivet, hvorfor deres alder er et estimat. Det drejer sig om én mand, der estimeres til at være i slutningen af trediverne hvorfor hans alder er medregnet som 39 år, hvilket ligeledes gør sig gældende for én kvinde. Derudover er der en kvinde hvis alder er estimeret til at være i midten af halvtredserne, hvorfor hendes alder er medregnet som 55 år. Den sidste testperson hvis alder ikke kendes er en mand, der estimeres til at være mellem 40 og 45 år, hvorfor hans alder er medregnet som 43 år.

Foruden de testpersoner, som har interagerede med robotten og efterfølgende er blevet interviewet, blev der også noteret observationer for de rejsende, som enten interagerede med robotten og ikke deltog i interviewet eller som ikke interagerede med robotten men som reagerede på robottens tilstedeværelse. For de rejsende hvor der enten er noteret køn eller barn er der i alt 97 personer. Ydermere er der noteret 22 tilfælde, hvor der ikke er noteret køn, hvorfor der potentielt kan være mere end én rejsende per NA. Der er derfor ikke et specifikt antal rejsende, som enten har interagerede med robotten uden at deltage i interviewet eller som ikke har interagerede med robotten men som har reageret på dens tilstedeværelse.

I forhold til hvor ofte testpersonerne rejser svinger det mellem én gang om året til omkring én gang om ugen. Da det ikke er alle testpersoner, der har svaret på spørgsmålet og nogen har svaret, at de rejser tit, samt lignende uhåndgribelige svar, vælges det ikke at analysere hvilken betydning testpersonernes rejseerfaring har på hvordan de oplever interaktionen med robotten.

5.3 Databehandlingsmetode: Affinity diagram

Da formålet med undersøgelsen er at udlede hvilke parametre danske rejsende tilskriver interaktionen med en social robot, vil databehandlingen forsøge at opnå dette. Det kan blandt andet gøres med et *affinity diagram*, hvor data grupperes efter hvad testpersonerne giver udtryk for i forhold til problemer, behov og synspunkter fremfor forudbestemte kategorier. Hver gruppering får tildelt en label, som gengiver hvad testpersonerne gør eller tænker, (Holtzblatt m.fl. 2005a, s. 159). Fordelen ved at analysere det indsamlede kvalitative data ved hjælp af *affinity diagram* er, at hver eneste observation og pointe fra testpersonerne skal behandles, (Holtzblatt og Beyer 2015, s. 25). Ved at udvikle et *affinity diagram* for det indsamlede data bør det være muligt at udlede hvilke parametre danske rejsende tilskriver interaktionen med en social robot, hvorefter disse parametre kan anvendes i udviklingen af skalaer, som senere kan evalueres. Diagrammet giver således en

indsigt i hvad potentielle brugere oplever og hvad der kan forbedre eller forværre deres oplevelse af interaktionen med en social robot i en lufthavn. På baggrund af de fundne parametre samt en senere evaluering af oplevelsen bør det være muligt at designe robottens social intelligens.

Dette afsnit præciserer processen lige fra hvordan det indsamlede data fortolkes og gengives på *affinity notes* til hvordan det endelige *affinity diagram* bygges.

5.3.1 Fortolkningssession

Det tilstræbes så vidt muligt at følge reglerne og processen for fortolkningssessionen, som den er beskrevet i Holtzblatt m.fl. (2005b, ss. 101-122), dog er processen blevet tilpasset. Ifølge Holtzblatt m.fl. (2005b, s. 104) er det en fordel at gruppen, som deltager i fortolkningssessionen, har forskellig baggrund. Dette har ikke været tilfældet, da gruppen består af projektgruppens medlemmer, der alle har den samme baggrund. Der har været nogle afvigelser i forbindelse med gennemgangen af notater, hvorfra *affinity notes* formuleres. Testlederen, intervieweren, har under interviewet ikke taget noter men kun optaget lyd, hvorfor det blev besluttet at transskribere alle interviews, som intervieweren gennemgår i fortolkningssessionen. De transkriberede lydoptagelser fremgår af Afsnit 13.1 (*Transskriberede data*). Det bliver anbefalet at hvis fortolkningssessionen starter mere end to dage efter interviewet bør intervieweren genhøre lydoptagelsen. Da alle interviews er transskriberet af projektgruppen vurderes det at være unødvendigt at intervieweren genhører interviewet, da der kan snakkes ud fra transskriberingen. Fortolkningssessionen startede fem dage efter undersøgelsen blev afviklet.

Da der har været tre observatører til stede vil hver observatør bidrage med sine inputs til den specifikke testperson, hvilket gøres kronologisk. Observatørerne har ydermere noteret observationer af rejsende, som ikke har deltaget i undersøgelsen, disse observationer gennemgåes efter alle interviews. Når observationer, der ikke tilhører en af testpersonerne, noteres som *affinity notes* vil det blive noteret med *O* efterfulgt af køn: Mand noteres med *M*, kvinde noteres med *K*, børn noteres med *B*, hvorvis kønnet ikke er noteret vil det noteres med *NA*. Derudover er noterne nummeret efter kronologisk orden. *Affinity notes* udarbejdet til de 18 interviews vil blive noteret med: *IX-SX-XX*, hvor *IX* angiver interview nummer, *SX* angiver subjekt nummer, testperson, og *XX* angiver nummeret på *affinity note*. Indgår der mere end én testperson i interviewet er det noteret på to forskellige måder. Ved ét tilfælde er det noteret som: *S2*, *S2.1*, *S2.2* og *S2.3*, hvor det i de resterende tilfælde noteres som *SX*, som afspejler hvilket nummer testpersonen er. Derudover er indsigter noteret med *I* efter identificeringen, hvor design idéer og spørgsmål noteres som metoden fremsætter, henholdvis med *DI* og *Q*. Derudover bliver der brugt forkortelser til robot, lufthavn og Aalborg Lufthavn, som henholdvis forkortes til: *R*, *L* og *AAL*. For hver interview er der oprettet et dokument på Google Drev, hvori de tilhørende *affinity notes* noteres og projekteres på en skærm, som alle deltagere kan se.

Der er ikke blevet udarbejdet skilte til *rabbit holes*, men alle deltagere fik forklaret hvad formålet med *rabbit hole* er. Opstod der et *rabbit hole* kunne deltagerne til hver en tid sige ordene. Dette blev dog ikke konsekvent anvendt da projektgruppen oplevede situationer, hvor det var nødvendigt at diskutere et emne for at opnå enighed. Det tog omkring to arbejdsdage at udarbejde alle *affinity notes*. Sammenlagt blev der udarbejdet 567 *affinity notes*, som efterfølgende blev nedskrevet på gule *sticky notes*. Det skal dog nævnes at 99

af ud af de 567 *affinity notes* er foretaget på baggrund af observationer af rejsende, som ikke har deltaget i interviewet. De udarbejde *affinity notes* fremgår af Afsnit 13.2 (*Affinity notes*).

5.3.2 Udviklingen af *affinity diagram*

Følgende bygger på hvordan de udarbejdede *affinity notes* bruges til at udvikle et *affinity diagram*, hvor retningslinjerne for processen fremgår af Holtzblatt m.fl. (2005a, ss. 159-179).

Ligesom det anbefales at deltagerne i fortolkningssessionen har forskellig baggrund bliver det også anbefalet når et *affinity diagram* udarbejdes. I projektsammenhæng udarbejdes diagrammet af projektgruppen, som alle har den samme baggrund. Selvom det anbefales at udvikle diagrammet på én dag har det være nødvendigt strække processen over to dage. Den første dag blev alle gule og blå *sticky notes* placeret på væggen og på dag to blev de pink og grønne *sticky notes* placeret. Da alle de gule *affinity notes* skulle placeres og kategoriseres på væggen blev der taget udgangspunkt i de retningslinjer, som fremsættes af Holtzblatt m.fl. (2005a, s. 168). Hver deltager fik en bunke med 20 blandede *affinity notes*, hvor én deltager læser sin første note op og hænger den på væggen og de resterende deltagere gennemser deres bunke for en note, der kan placeres sammen med. Er det ikke muligt at finde en *affinity note*, der passer med den der allerede hænger på væggen, læser deltageren noten højt og hænger den på væggen for at danne en ny kategori. Efter de 20 *affinity notes* blev hængt på væggen, fik hver deltagere uddelt endnu 20 blandede *affinity notes*, men denne gang læste deltagerne kun deres note højt såfremt den startede en ny kategori. Denne proces fortsatte til at alle *affinity notes* var placeret på væggen. Reglen om at en kategori ikke må dannes ud fra en *DI* eller *Q* blev overholdt gennem hele processen.

Efter alle *affinity notes* blev kategoriseret startede det næste step i processen, som vedrører at tildele en blå label til hver kategori. Retningslinjerne for hvordan de blå labels tildeles blev tilnærmelsesvis fulgt, deltagerne blev inddelt i par og fik tildelt et område på væggen med *affinity notes*. Som retningslinjerne foreskriver skal parret starte med den længste gruppering, dette var ikke nødvendigvis tilfældet da en anden gruppering kunne fremstå mere overskuelig, hvorfor parret startede der. Efter alle grupperinger fik tildelt en blå label, blev der dannet nye par, som ligeledes fik tildelt et område på væggen, hvor der skulle formuleres grønne labels. Som metoden foreskriver, er de første grønne labels kun til for at danne en overordnet kategori med en initierende label, som efterfølgende kan omformuleres. Efter de initierende grønne labels er sat på væggen går parret igang med at formulere de pink labels. Efterfølgende gennemgik alle deltagerne, hele projektgruppen, diagrammet for dels at omformulere de grønne labels, tjekke at pink såvel som blå labels er korrekt formuleret. I forhold til formuleringen af labels blev det forsøgt, så vidt muligt, at holde formuleringerne i testpersonernes eget sprog. Det er dog langt fra alle labels og *affinity notes*, der er formuleret i førsteperson, derimod er flertallet formuleret direkte som udtalelser testpersonerne selv har udtrykt. Dog er de grønne labels formuleret mere generelt og abstrakt, hvilket er tilladt med denne type label.

For at gøre det nemmere at flytte *affinity notes* blev det besluttet, at når parret var igang med de initierende grønne labels, var det tilladt at samle de gule *affinity notes* så der under hver blå label hænger en samling af gule *sticky notes*. På den måde blev det nemmere at flytte dem rundt og det mindskede slid på limen. Det endelige *affinity diagram*

indeholder sammenlagt 149 blå labels, 47 pink labels og 10 grønne labels. Holtzblatt m.fl. (2005a, s. 174) fremsætter en tommelfingerregel om, at for hver pink label skal der være mellem to og seks blå labels under. I det udarbejdede diagram forefindes der kun én pink label, som har mere end seks tilhørende blå labels. Den pink label vedrører robottens højde og dækker over syv blå labels. I henhold til de grønne labels er tommelfingerreglen, at der må være fire til otte pink labels under, (Holtzblatt m.fl. 2005a, s. 174). Ud af de 10 grønne labels forefindes der tre kategorier, som har færre pink labels end tommelfingerreglen. De tre grønne kategorier vedrører situationer, hvor der ikke interageres med robotten, hvor skærmen ikke virker og hvor der opstår interesse for robotten.

5.4 Valg af skala type

Det vælges at anvende en *Visual Analogue Scale* (VAS) til skalaspørgsmålene. Fordelen ved at bruge en VAS er, at det giver mulighed for mere nuancerede besvarelser. Testpersonerne har mulighed for at svare hvor som helst på skalaen, modsat eksempelvis en *Likert*-skala, hvor svarmulighederne er afgrænsede til et bestemt antal bokse. Dog ifølge Lewis og Erdinç (2017, s. 73), er der ikke væsentlig forskel, når der er tale om UX research, på at bruge en 7-, 11-, eller 101-punkts skala, hvor 101-punkts skalaen er en VAS, der går fra 0-100.

5.4.1 Endepunkter

Ved VAS er der mulighed for at have lukkede eller åbne endepunkter. Fordelen ved at have åbne endepunkter er, at hvis samme testperson skal svare på det samme skalaspørgsmål gentagende gange, kan testpersonen få behov for at overgå tidligere afgivende besvarelser. Dette gør sig eksempelvis gældende, hvis testpersonen skal bedømme otte forskellige bolscher på de samme skalaer. Hvis de har bedømt et bolsche som værende stærkt, og der pludselig kommer et vanvittigt stærkt bolsche efterfølgende, så har de mulighed for at overgå deres første vurdering af, om bolschet er stærkt ved at sætte deres besvarelse udenfor endepunktet. Ved åbne endepunkter får testpersonen netop denne mulighed, da det er muligt at svare efter et endepunkt. Skalaerne i dette projekt skal testpersonerne kun svare på én gang for hver skala, hvilket gør at det ikke er så vigtigt at der er mulighed for at overgå tidligere besvarelser.

Det vælges at anvende skalaer med lukkede endepunkter da det vurderes, at der ikke er behov for at have åbne endepunkter. Desuden vurderes det, at det vil være nemmere for testpersonerne at bruge en skala med lukkede endepunkter, da der ikke skal tages stilling til, hvad der er efter endepunktet. Det kan ske, at der vil være ophobninger et stykke fra endepunkterne, da der muligvis vil være en tilbageholdenhed for at svare helt ude ved kanten. På den anden side er der også muligheden for at flere sætter ved endepunkter, fordi nogle måske kunne have ønsket at sætte længere ude på skalaen.

5.4.2 Bipolær og unipolær

Skalaerne kan opstilles som enten bipolær eller unipolær. Ved en unipolær skala vil det være samme ord for hele skalaen, hvor det er graden af dette ord, der bedømmes. Hvis skalaen er bipolær, vil det være to forskellige ord i hver ende af skalaen, hvor ordene er hinandens modpoler.

I forhold til dannelsen af skalaerne ønskes det for hvert skalaspørgsmål at tage stilling til, om der er mulighed for en bipolær skala, hvor de to ord er hinandens modpoler. Hvis det vurderes, at der ikke findes en naturlig og logisk modpol til ordet, vil der i stedet dannes en unipolær skala.

Årsagen til at det først undersøges om det er muligt at danne en bipolær skala fremfor en unipolær skala, er at der ved en bipolær skala dækkes over to udtryk, og det vil komme til udtryk hvor høj grad af modpolen, der er enighed med og ikke kun, at der ikke er enighed med den anden pol. Det er vigtigt, at der ved en bipolær skala er modpoler i hver ende, så der ikke opstår tilfælde, hvor det ikke er muligt at besvare skalaen, da personen måske er enig i begge poler og derfor ikke ved om de skal sætte besvarelseren i den ene eller i den anden ende. Derfor vil der kun blive dannet bipolære skalaer i tilfælde hvor der findes en naturlig og logisk modpol.

5.4.3 Labels

De labels, der sættes på skalaerne dannes ud fra testpersonernes egne ord og formuleringer. Det ønskes dog at der i hver ende af skalaen skal være ekstremer, så der opnås så stort et spænd som muligt mellem de to punkter, så der ikke er tilfælde hvor skalaen ikke er dækkende i forhold til det testpersonen ønsker at svare. Derudover hænger det også sammen med, at der ved opstilling af VAS, er retningslinjer for at endepunkterne skal være ekstremer.

Det vælges at bruge testpersonernes ord, men hvor der ved den laveste grad af ordet sættes *slet ikke* foran, og hvor det ved den højeste grad af ordet sættes *ekstremt* foran. Det betyder at ved et eksempel, hvor ordet er *glad*, så vil skalaens labels være *Slet ikke glad* og *Ekstremt glad*.

5.5 Resultater fra Affinity diagram

Formålet med projektet er at udlede, hvilke parametre danske rejsende tilskriver interaktionen med en social robot, for efterfølgende at udvikle skalaer, som kan bruges til at evaluere oplevelsen af HRI. Derfor bør der fremsættes nogle krav for hvordan parametrene udvælges, for dernæst at formulere dem til skalaer. Kravet er i første omgang, at det skal være muligt for en designer at ændre på parameteren, hvorfor parametre der eksempelvis relaterer sig til hvad brugeren foretrækker eller hvad brugeren har af tidligere erfaringer ikke medtages.

I det følgende afsnit vil hver af de 10 udlede grønne kategorier først blive præsenteret sammen med tilhørende potentielle skalaspørgsmål, efterfulgt af tilhørende skala labels. For hvert potentielt skalaspørgsmål vil det blive besluttet hvorvidt den specifikke parameter skal evalueres på en bipolær eller på en unipolær *Visual Analogue Scale*, (VAS) med lukkede endepunkter. I tilfælde hvor parameteren evalueres på en bipolær skala, vil der være markeret et midtpunkt, som enten kan være unavngivet eller navngivet. Derefter vil der fokuseres på udvælgelsen af de endelige skalaer.

For at udlede parametrene fokuseres der på én grøn kategori af gangen, hvorfra

parametrene udledes på baggrund af både pink og blå labels. Parametrene vil blive formuleret som udsagn, der potentielt kan evalueres på en skala, hvorfor parametre noteres med et S på en orange *sticky note*. De orange *sticky notes* kan derudover indeholde designidéer DI , indsigtter I samt idéer til det efterfølgende testdesign TD . Der er i alt dannet 50 orange *sticky notes*, hvorfra 42 er formuleret som udsagn, der kan omformuleres til potentielle skalaspørgsmål, hvis omformuleringen er nødvendig. Derudover forefindes der seks design idéer, én indsigt og ét foreslag til testdesign.

5.5.1 Interagerer ikke med R



Figur 5.1. Oversigt over den grønne kategori: *Interagerer ikke med R*, hvor er *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

På Figur 5.1 fremgår to orange *sticky notes*. Den ene er en idé til det næste testdesign, som relaterer sig til at de rejsende, der interagerer med robotten, men ikke ender med at følge efter robotten, stoppes og spørges ind til hvorfor de stoppede interaktionen. Den anden handler om at ignorere robotten. Da ordet *ignorere* dækker over flere aspekter vælges det at omformulere udsagnet så det bedst muligt afspejler hvad der sigtes efter at måle, nemlig om det er muligt at undgå robotten, hvorfor udsagnet omformuleres til følgende potentielle skalaspørgsmål:

1. Er det let at undgå R

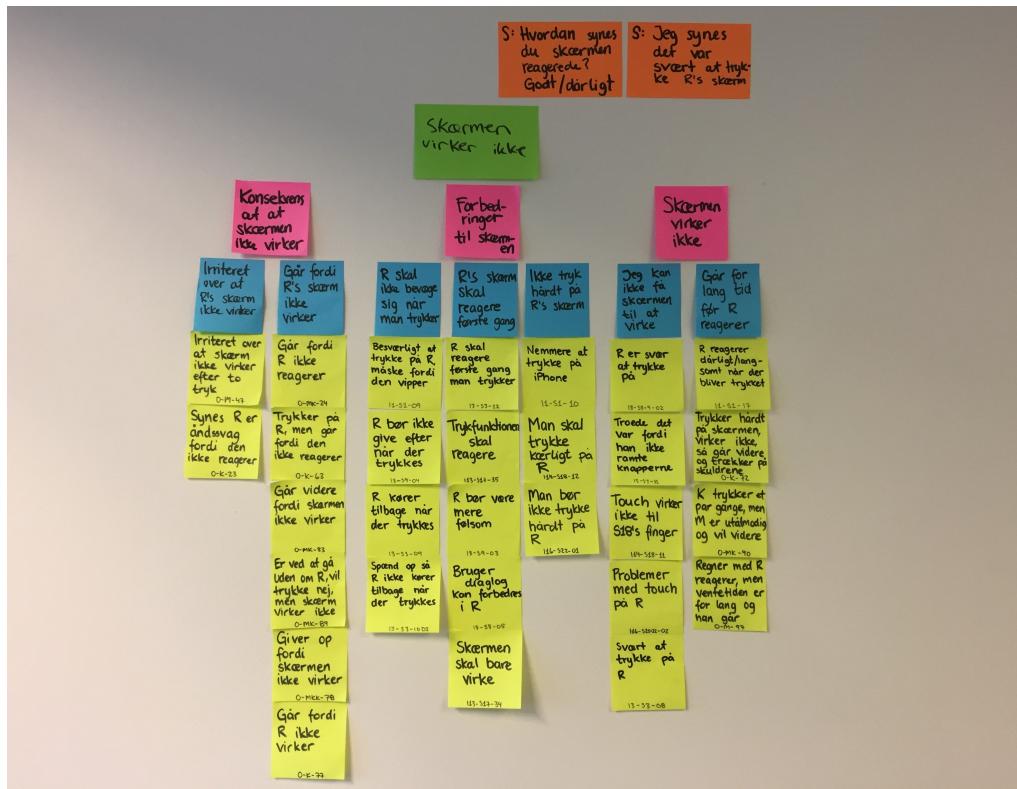
Til ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Ekstremt svært	-	Ekstremt let

Tabel 5.1. Skala labels vedrørende hvor let det er at undgå robotten. *SQ* er en forkortelse for skalaspørgsmål.

Baseret på det potentielle skalaspørgsmål samt tilhørende labels, der fremgår i Tabel 5.1, vil det udgøre en unipolær skala uden et midtpunkt.

5.5.2 Skærmens virker ikke



Figur 5.2. Oversigt over den grønne kategori: *Skærmens virker ikke* med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

Da de to orange *sticky notes* vedrører den samme problemstilling, at skærmens forvoldte problemer vælges det er sætte de to sammen og omformulere dem til følgende:

1. Hvordan synes du skærmens virkede

Til ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

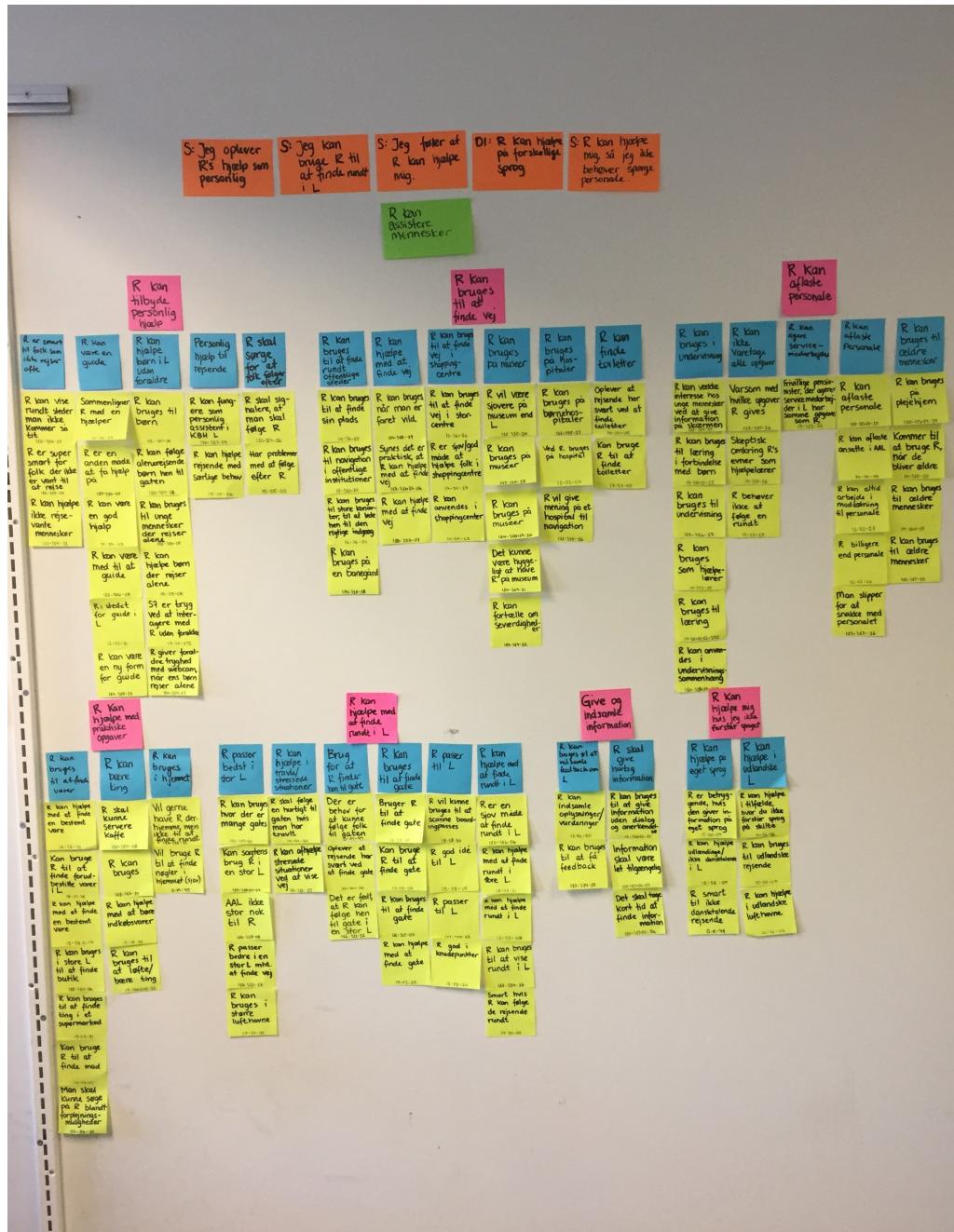
SQ	Venstre label	Midtpunkt	Højre label
1	Ekstremt dårligt	Intet label	Ekstremt godt

Tabel 5.2. Skala labels vedrørende skærmens på robotten. *SQ* er en forkortelse for skalaspørgsmål.

Det potentielle skalaspørgsmål vil blive evalueret på en bipolær skala med et unavngivet midtpunkt. Årsagen til at dette potentielle skalaspørgsmål ikke ekskluderes, selvom det

ikke er muligt for en designer at ændre på, er, at det antages, at skærmens reaktionsevne har indflydelse på hvordan testpersonerne evaluerer andre parametre. Derudover forventes det, at skærmen på den færdig udviklede robot ikke vil forårsage disse problemer, da brugergrænsefladen selvsagt ikke skal åbnes igennem *Doubles* applikation.

5.5.3 R kan assistere mennesker



Figur 5.3. Oversigt over den grønne kategori: *R kan assistere mennesker*, hvor *R* angiver robot og *L* angiver lufthavn, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

Baseret på de fire potentielle skalaspørgsmål vælges det at sammensætte: *Jeg kan bruge R til at finde rundt i L med Jeg føler at R kan hjælpe mig til et samlet potentielt*

skalaspørgsmål, som fremgår i nedenstående. Derudover fremgår der en design idé på Figur 5.3, som vedrører at robotten bør kunne hjælpe på forskellige sprog og ikke kun på dansk.

1. Jeg oplever R's hjælp som personlig
2. Jeg føler at R kan hjælpe mig
3. R kan hjælpe mig så jeg ikke behøver at spørge personale

Til hver af de tre ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Slet ikke personlig	-	Ekstremt personlig
2	Helt uenig	Intet label	Helt enig
3	Helt uenig	Neutral	Helt enig

Tabel 5.3. Skala labels vedrørende hvorvidt robotten kan assistere mennesker. *SQ* er en forkortelse for skalaspørgsmål.

Skalaen vedrørende hvorvidt robottens hjælp opleves som værende personlig vil blive evalueret på en unipolær skala. Hvor de to andre potentielle skalaspørgsmål begge vil blive evalueret på en bipolær skala med et midtpunkt henholdvist unavngivet og angivet med: *Neutral*.

5.5.4 R's væremåde



Figur 5.4. Oversigt over den grønne kategori: *R's væresmåde*, hvor *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejdede orange *sticky notes*.

Ud af de ni orange *sticky notes* er to designidéer hvor én relatere sig til, at det bør overvejes om robotten skal kunne tale og den anden relatere sig til at robotten skal tilpasse sine bevægelser afhængigt af brugerens bevægelser. Derudover er det valgt at sammensætte: *Jeg synes R's bevægelser er behagelige* med *Jeg synes R har rolige bevægelser* til ét potentielt skalaspørgsmål, listet som det første i nedenstående. Ydermere er det valgt at sammensætte: *Jeg oplevede R's hastighed som værende hurtig/langsom* med *Jeg synes at R's hastighed er for langsom/for hurtig* til ét potentiel skalaspørgsmål, der vedrører

robottens hastighed, listet som nummer to i nedenstående. De tre resterende bibeholdes:

1. Jeg synes at R's bevægelser er...
2. Jeg synes at R's hastighed er...
3. Jeg synes at R er irriterende
4. Jeg synes at R er levende
5. Jeg synes at R er anmassende

Til hver af de fem ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Alt for rolige	Intet label	Alt for vilde
2	Alt for langsom	Fin	Alt for hurtig
3	Slet ikke irriterende	-	Ekstremt irriterende
4	Slet ikke levende	-	Ekstremt levende
5	Slet ikke anmassende	-	Ekstremt anmassende

Tabel 5.4. Skala labels vedrørende robottens væremåde. *SQ* er en forkortelse for skalaspørgsmål.

De to første potentielle skalaspørgsmål vil begge blive evalueret på en bipolær skala med et midtpunkt henholdvist uden et label og med label: *Fin*. De tre resterende vil hver blive evalueret på en unipolær skala.

5.5.5 Henvendelse



Figur 5.5. Oversigt over den grønne kategori: *Henvendelse* med tilhørende *affinity notes* placeret under henholdvist blå og pink labels, samt de udarbejde orange *sticky notes*.

To ud af de syv orange *sticky notes* er angivet som designidéer og vedrører robottens position relativt til brugeren, hvor robotten fortrinvist skal vende front mod brugeren, samt være tæt på personen. Det vælges at ekskludere: *Jeg foretrækker at R/jeg henvender sig*, da det ikke er en parameter, der for en designer er mulig at ændre på. Derudover vælges det at omformulere: *Jeg synes R kom for tæt på til Jeg synes R er intimiderende*. Denne omformulering er baseret på dels de blå labels samt *affinity notes* kategoriseret under den pink label: *R skal respektere intimsfæren*. Det medfører at følgende potentielle skalaspørgsmål formuleres til:

1. Jeg synes at R er imødekommen
2. Jeg synes at R kom for tæt på
3. Jeg synes at R stod i vejen
4. Jeg blev overrasket over R's henvendelse
5. Jeg synes at R er intimiderende

Til hver af de fem ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Ekstremt afvisende	Intet label	Ekstremt imødekommen
2	Alt for tæt på	Intet label	Alt for langt fra
3	Slet ikke i vejen	-	Ekstremt i vejen
4	Slet ikke overrasket	-	Ekstremt overrasket
5	Slet ikke intimiderende	-	Ekstremt intimiderende

Tabel 5.5. Skala labels vedrørende robottens henvendelse. *SQ* er en forkortelse for skalaspørgsmål.

Baseret på Tabel 5.5 vil de to første potentielle skalaspørgsmål blive evalueret på en bipolær skala med et midtpunkt, som henholdvist er unavngivet og navngivet med *Tilpas*. De tre resterende potentielle skalaspørgsmål vil hver blive evalueret på en unipolær skala.

5.5.6 R's udseende



Figur 5.6. Oversigt over den grønne kategori: *R's udseende*, hvor *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

På Figur 5.6 fremgår der én designidé, som vedrører at robotten automatisk skal indstille sin højde afhængigt af brugerens højde. Det vælges at sammensætte: *Jeg oplevede R's højde som værende meget høj/meget lav med Jeg synes at R's højde er for høj/for lav til et enkelt potentiel skalaspørgsmål vedrørende robottens højde. Da det ikke er muligt for en designer at påvirke, hvad en potentiel bruger foretrækker vælges det at sammensætte: Jeg foretrækker at R ser menneskelig ud med Jeg synes at R ser menneskelig ud.* De potentielle skalaspørgsmål er som følger:

1. Jeg synes at R's højde er...
 2. Jeg synes at R er elegant
 3. Jeg synes at R ser menneskelig ud
 4. Jeg kan godt lide R's udseende

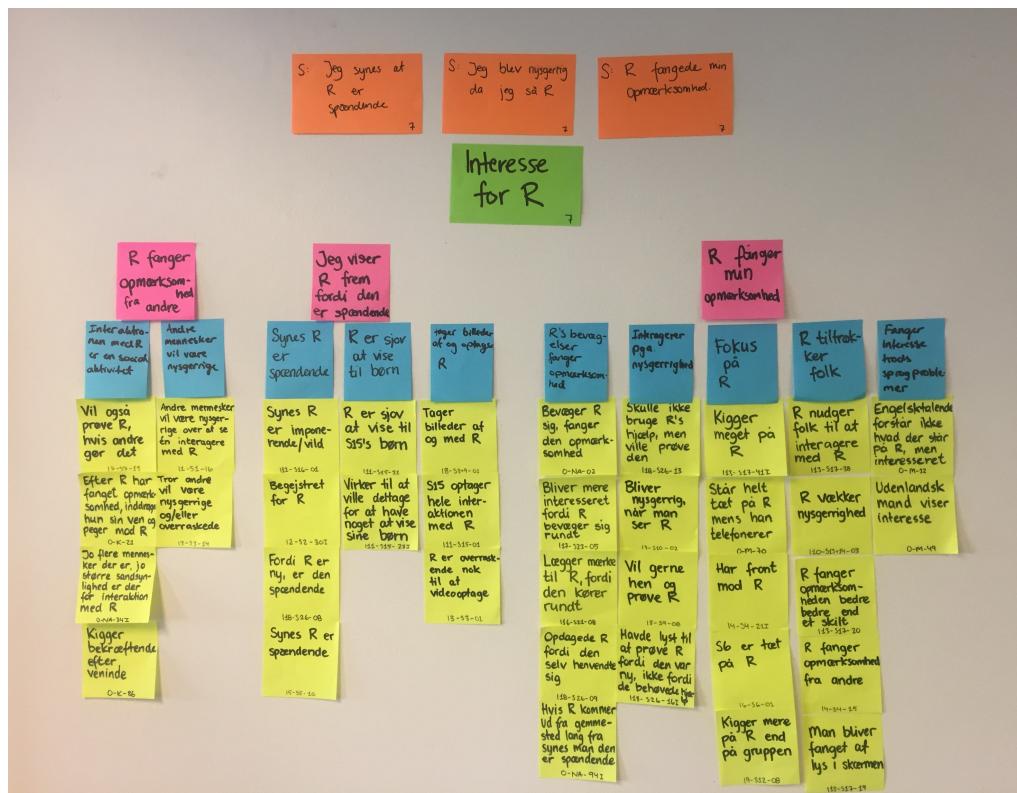
Til hver af de fire ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre	Midtpunkt	Højre label
1	Alt for lav	Fin	Alt for høj
2	Slet ikke elegant	-	Ekstremt elegant
3	Slet ikke menneskelig	-	Ekstremt menneskelig
4	Helt uenig	-	Helt enig

Tabel 5.6. Skala labels vedrørende robottens udseende. *SQ* er en forkortelse for skalaspørgsmål.

Baseret på Tabel 5.6 er det kun skalaen, som vedrører robottens højde, der evalueres på en bipolær skala med midtpunktet *Fin*. De resterende tre potentielle skalaspørgsmål bliver alle evalueret på en unipolær skala.

5.5.7 Interesse für R



Figur 5.7. Oversigt over den grønne kategori: *Interesse for R*, hvor *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

Det vælges at sammensætte *Jeg blev nysgerrig da jeg så R* med *Jeg synes at R er spændende* til ét potentielt skalaspørgsmål, da det antages at de vedrører det samme. De to potentielle skalaspørgsmål er som følger:

1. R fangede min opmærksomhed
 2. Jeg synes at R er spændende

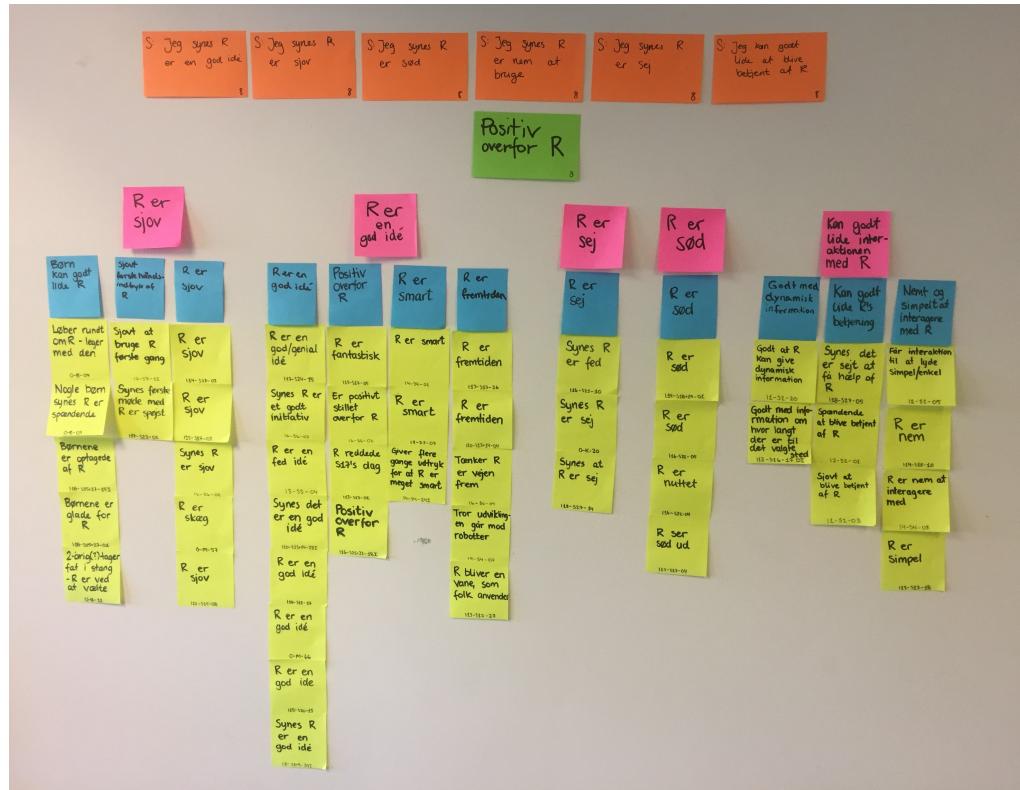
Til hver af de to ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Slet ikke	-	Ekstremt meget
2	Slet ikke spændende	-	Ekstremt spændende

Tabel 5.7. Skala labels vedrørende interesse for robotten. SQ er en forkortelse for skalaspørgsmål.

Baseret på ovenstående Tabel 5.7 vil de to potentielle skalaspørgsmål begge blive evalueret på en unipolær skala.

5.5.8 Positiv overfor R



Figur 5.8. Oversigt over den grønne kategori: *Positiv overfor R*, hvor *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

Det vælges at ekskludere *Jeg synes R er en god idé*, da det ikke er muligt for en designer at ændre på hvorvidt potentielle brugere betragter robotten som værende en god idé. Derudover vælges det at sammensætte *Jeg synes R er nem at bruge* med Jeg har brug for hjælp til at bruge R fra den grønne kategori: *Kendskab til teknologi*, da det vurderes at de to udtryk er modsætninger, som beskriver den samme oplevelse, hvorfor skalaspørgsmålet vil fremgå af Underafsnit 5.5.9 (*Kendskab til teknologi*).

1. Jeg synes at R er sød
2. Jeg synes at R er sjov
3. Jeg synes at R er sej

4. Jeg kan godt lide at blive betjent af R

Til hver af de fire ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Slet ikke sød	-	Ekstremt sød
2	Slet ikke sjov	-	Ekstremt sjov
3	Slet ikke dej	-	Ekstremt dej
4	Slet ikke	-	Ekstremt meget

Tabel 5.8. Skala labels vedrørende ens positive indstilling i henhold til robotten. *SQ* er en forkortelse for skalaspørgsmål.

Baseret på ovenstående Tabel 5.8 vil alle fire potentielle skalaspørgsmål blive evalueret på en unipolær skala.

5.5.9 Kendskab til teknologi



Figur 5.9. Oversigt over den grønne kategori: *Kendskab til teknologi* med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

Som nævnt sammensættes *Jeg synes R er nem at bruge* fra Underafsnit 5.5.8 med *Jeg har brug for hjælp til at bruge R* til ét enkelt potentielt skalaspørgsmål, som vedrører hvordan det var at bruge robotten. De resterende tre udsagn: *Jeg har erfaring med robotter*, *Min erfaring med teknologi har betydning for min oplevelse af R* samt *Jeg synes R minder om noget jeg kender* sammensættes. Det skyldes hovedsageligt at de to førstnævnte udsagn ikke er noget en designer har mulig for at ændre og selvom det potentielt er muligt at

designe robotten, som noget velkendt er det ikke muligt at vide hvad potentielle brugere da sammenligner robotten med. Det vurderes dog, at brugerens kendskab til teknologier og robotter kan have indflydelse på hvordan andre parametre evalueres, hvorfor det vælges at formulere dette som et skalaspørgsmål. Dette skalaspørgsmål vil ikke indgå på samme måde som de resterende, men vil indgå som en del af at indsamle demografisk information fra testpersonerne. De to potentielle skalaspørgsmål er formuleret som følger:

1. Hvordan var det at bruge R
2. Hvor meget kendskab har du til teknologi/robotter

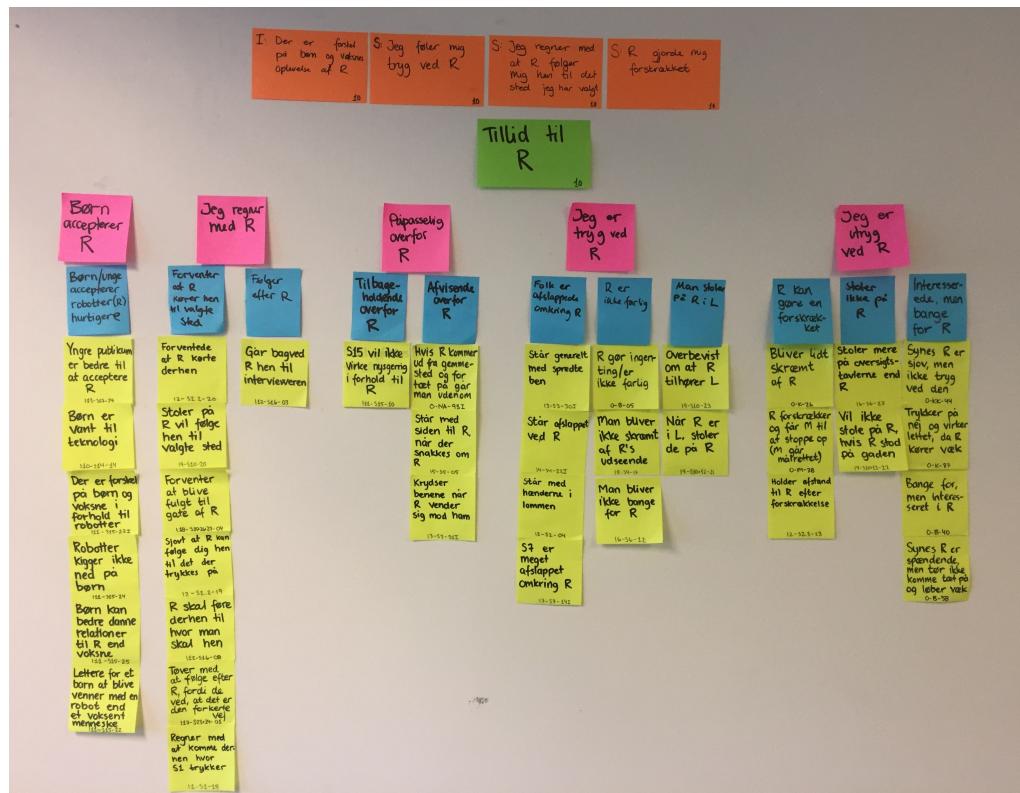
Til hver af de to ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Ekstremt svært	Intet label	Ekstremt nemt
2	Slet ingen kendskab	-	Ekstremt meget kendskab

Tabel 5.9. Skala labels vedrørende kendskab til teknologi. *SQ* er en forkortelse for skalaspørgsmål.

Baseret på Tabel 5.9 vil det første potentielle skalaspørgsmål relateret til hvordan den var at bruge robotten, blive evalueret på en bipolær skala med unavngivet midtpunkt. Det andet potentielle skalaspørgsmål, som vil indgå i indsamlingen af demografisk data, vil blive evalueret på en unipolær skala.

Tillid til R



Figur 5.10. Oversigt over den grønne kategori: *Tillid til R*, hvor *R* angiver robot, med tilhørende *affinity notes* placeret under henholdsvis blå og pink labels, samt de udarbejde orange *sticky notes*.

På Figur 5.10 fremgår der én orange *sticky note* noteret med et *I* efterfulgt af et udsagn. *I* står for indsigt, hvor det i dette tilfælde vedrører at der er forskel på, hvordan børn og voksne oplever interaktionen med robotten. Det er derfor hverken noget en designer vil kunne ændre på eller noget der skal være en del af det næste testdesign. De resterende tre orange *sticky notes* er dog formuleret som potentielle skalaspørgsmål:

1. Jeg føler mig tryg ved R
2. Jeg regner med at R følger mig hen til det sted jeg har valgt
3. R gjorde mig forskrækket

Til hver af de to ovenstående potentielle skalaspørgsmål kan følgende labels anvendes:

SQ	Venstre label	Midtpunkt	Højre label
1	Ekstremt uthyg	Intet label	Ekstremt tryg
2	Helt uenig	Neutral	Helt enig
3	Slet ikke forskrækket	-	Ekstremt forskrækket

Tabel 5.10. Skala labels vedrørende tillid til robotten. *SQ* er en forkortelse for skalaspørgsmål.

De to første potentielle skalaspørgsmål vil blive evalueret på en bipolær skala med midtpunkt henholdsvis unavngivet og navngivet med *Neutral*, jævnfør Tabel 5.10. Det sidste potentielle skalaspørgsmål vil derimod blive evalueret på en unipolær skala.

Igennem de foregående afsnit har det været muligt at reducere antallet af udsagn fra 42 til 30 potentielle skalaspørgsmål, hvoraf ét vil blive anvendt i forbindelse med indsamling af demografisk data. Dog vurderes det, at det er muligt med yderligere reducering af antallet af potentielle skalaspørgsmål for at testpersonerne, som skal evaluere interaktionen med robotten ikke skal forhold sig til 30 skalaer. Derudover antages det at nogen af de potentielle skalaspørgsmål reelt set måler det samme, hvorfor det er hensigtsmæssigt at sammensætte dem. I det følgende afsnit vil der derfor blive fokuseret på at udvælge og reducere antallet potentielle skalaspørgsmål samt omformulerer spørgsmålene, hvis nødvendigt.

5.6 Udvælgelse af skalaer

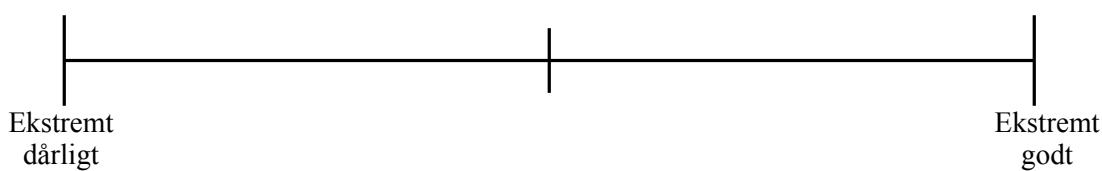
Ud fra de fundne parametre og de potentielle skalaspørgsmål beskrevet i forrige afsnit vil der i dette afsnit fokuseres på at udvælge de skalaer, der fremadrettet vil blive anvendt i forbindelse med at evaluere interaktionen med robotten. I det henseende kan det være at to eller flere tidligere foreslægt skalaspørgsmål bliver sammensat til ét i tilfælde af, at det vurderes, at de vedrører den samme parameter. Denne vurdering foretages blandt projektgruppen, hvorfor der er ikke fastsatte kriterier for hvornår to skalaspørgsmål sammensættes, uddover at de vedrører det samme.

Et krav for udvælgelsen af et skalaspørgsmål er, at det skal være muligt at ændre på den specifikke parameter. Derudover vil de udvalgte skalaspørgsmål blive præsenteret på deres respektive skalaer. Det tilstræbes at opstille den optimale skala for hvert enkelt parameter, hvorfor der ikke er fastsat én specifik skala type på forhånd.

I henhold til den første kategori: *Interagerer ikke med R*, hvor der kun er ét potentielt skalaspørgsmål: *Er det let at undgå R*, så vil dette skalaspørgsmål ikke fremgå på en endelig skala. Årsagen er, at spørgsmålet vil blive sammensæt med: *Jeg synes at robotten er anmassende*. Vurderingen bygger på, at hvis robotten er anmassende vil den formentlig ikke være let at undgå, hvorimod hvis robotten slet ikke er anmassende så er robotten formentlig let at undgå.

Skalaspørgsmålet: *Hvordan synes du skærmen virkede*, vil ikke behandles på lige fod med de resterende parametre, men det kan være med til at undersøge hvorvidt skærmenes reaktionsevne har en indflydelse på interaktionen. Grunden til at spørgsmålet ikke vil indgå på samme måde som de resterende parametre er, at det forventes at den færdig udviklede robot ikke vil have dette problem. Problemet med at skærmen reagerer dårligt skyldes formentlig, at *wireframed* åbnes igennem robottens egen software. Derudover er det observeret at flere testpersoner har problemer med at trykke på skærmen, hvor teorien/hypotesen er, at de trykker for hårdt. På Figur 5.11 fremgår skalaen, hvorpå det vurderes hvordan skærmen reagerede.

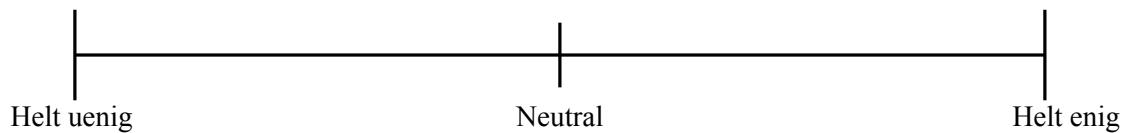
Hvordan synes du skærmen reagerede?



Figur 5.11. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Hvordan synes du skærmen reagerede?*

Årsagen til at der ikke er et label på midtpunktet på Figur 5.11 er, at det forventes, at der en logisk forståelse for hvad midten indikerer, hvor et label potentielt kan skævvrude responsen. De to potentielle skalaspørgsmål: *Jeg føler at R kan hjælpe mig* og *R kan hjælpe mig så jeg ikke behøver at spørger personale* bliver sammensat fordi de begge vedrører at robotten kan hjælpe dem. På Figur 5.12 fremgår den udvalgte skala for hvorvidt robotten kan hjælpe én.

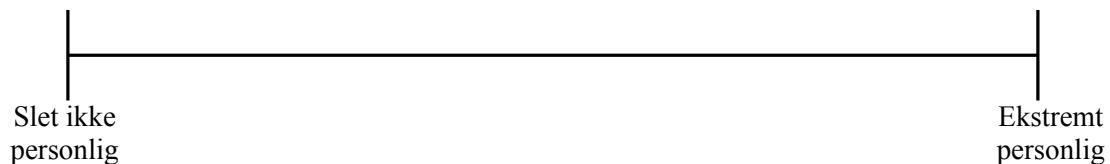
Jeg føler at robotten kan hjælpe mig



Figur 5.12. Bipolær VAS med lukkede endepunkter og navngivet midtpunkt til skalaspørgsmålet: *Jeg føler at robotten kan hjælpe mig*.

Årsagen til at der på Figur 5.12 fremgår et label på midtpunktet er for, at kalibrer skalaen omkring neutral i tilfælde af, at testpersonen ikke har en holdning om hvorvidt robotten kan hjælpe en. Fra samme kategori: *R kan assistere mennesker* udvælges skalaspørgsmålet: *Jeg oplever robottens hjælp som personlig*, hvilket fremgår på Figur 5.13.

Jeg oplever robottens hjælp som personlig

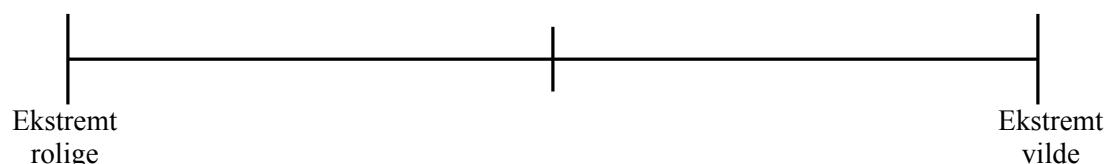


Figur 5.13. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg oplever robottens hjælp som personlig*.

Årsagen til at skalaspørgsmålet: *Jeg oplever robottens hjælp som personlig* ikke præsenteres på en bipolær skala skyldes hovedsageligt, at der ikke findes en naturlig og logisk modpart til personlig, i forbindelse med robotten.

I henhold til kategorien: *R's væremåde* vil skalaspørgsmålene vedrørende robottens bevægelse, hastighed samt hvorvidt robotten er irriterende vælges, som de fremgår i Underafsnit 5.5.4 (*R's væremåde*). Skalaen tilhørende robottens bevægelse fremgår på Figur 5.14, robottens hastighed fremgår på Figur 5.15 og hvorvidt robotten er irriterende fremgår på Figur 5.16.

Jeg synes at robottens bevægelser er

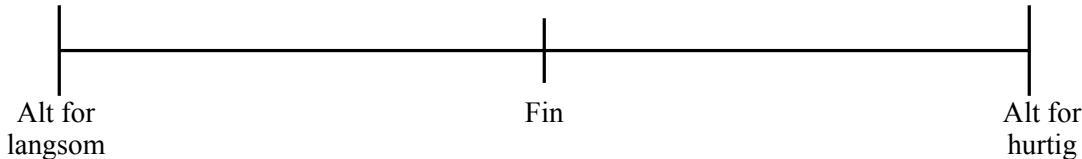


Figur 5.14. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Jeg synes at robottens bevægelser er*.

Årsagen til at Figur 5.14 er bipolær er, at det vurderes at ekstremt rolige bevægelser er det modsatte af ekstremt vilde bevægelser. Hvor årsagen til at der ikke er et label på midtpunktet er, at det vurderes, at der er en logisk adskilles mellem de to yderpunkter.

Derudover er det svært at finde et label, som ville kunne kobles til midtpunktet mellem rolige og vilde bevægelser.

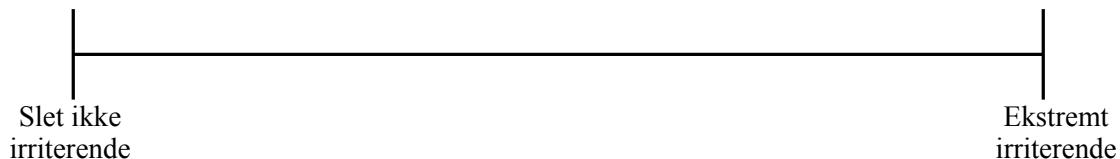
Jeg synes at robottens hastighed er



Figur 5.15. Bipolær VAS med lukkede endepunkter og navngivet midtpunkt til skalaspørgsmålet: *Jeg synes at robottens hastighed er*.

Årsagen til at Figur 5.15 præsenteres på en bipolær skala er, at de to labels på endepunkterne gengiver testpersonernes egne formulering af robottens hastighed. Derudover er midtpunkt angivet med: *Fin*, fordi det afspejler testpersonernes udsagn. Ydermere vurderes det, at der ikke findes en naturlig og logisk midte mellem de to yderpunkter, hvorfor der kalibreres omkring at robottens hastighed er fin.

Jeg synes at robotten er irriterende

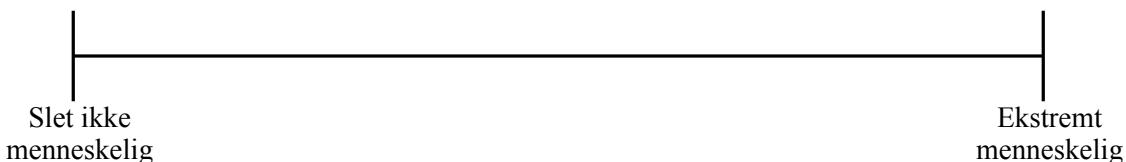


Figur 5.16. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er irriterende*.

Årsagen til at Figur 5.16 præsenteres på en unipolær skala er, at det vurderes, at der ikke forekommer en naturlig og logisk modpart til irriterende.

Derudover vælges det at sammensætte: *Jeg synes at R er levende* fra kategori: *R's væremåde* og *Jeg synes at R ser menneskelig ud* fra kategori: *R's udseende*, til ét enkelt skalaspørgsmål præsenteret på Figur 5.17.

Jeg synes at robotten er menneskelig



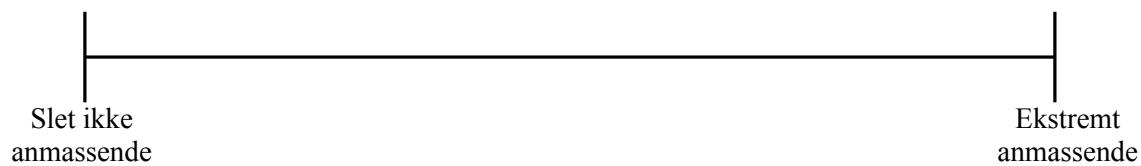
Figur 5.17. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er menneskelig*.

Årsagen til at Figur 5.17 præsenteres på en unipolær skala fremfor en bipolær skala er, at der ikke umiddelbart findes en naturlig og logisk modpart til menneskelig, hværfald i forhold til robotten. Der kan argumenteres for, at anvende ordet: *Umenneskelig*, hvor årsagen til at dette ord ikke anvendes er, at hvis en robot er umenneskelig er det så en

maskine eller et dødt objekt, og i så fald hvilket ord skal da anvendes som label. På baggrund af dette blev det besluttet at angive modparten som: *Slet ikke menneskelig*.

Ydermere sammensættes: *Jeg synes at R er anmassende* fra kategori: *R's væremåde*, med *Jeg synes at R er intimiderende* fra kategori: *Henvendelse*, samt *Er det let at undgå R* fra kategori: *Interagerer ikke med R*. Årsagen til at *Jeg synes at R er intimiderende* sammensættes med hvorhvidt robotten er anmassende er, at vurderes at hvis robotten perciperes som værende ekstremt anmassende vil den formentlig også blive perciperet som intimiderende, hvorhvis robotten ikke er anmassende så forventes det ligeledes at robotten heller ikke perciperes som intimiderende. Årsagen til at *Er det let at undgå R* indgår er forklaret tidligere. Den valgte skala fremgår på Figur 5.18.

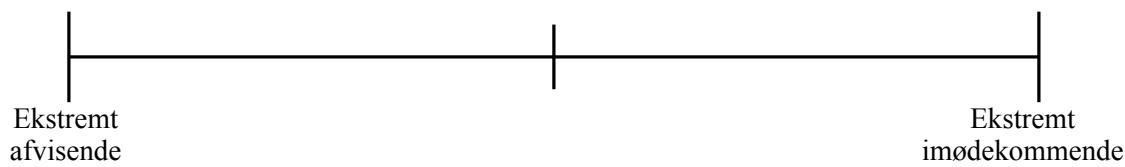
Jeg synes at robotten er anmassende



Figur 5.18. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er anmassende*.

Årsagen til at Figur 5.18 ikke præsenteres på en bipolær skala er, at det igen vurderes, at der ikke forekommer en naturlig og logisk modpart til anmassende. Af de fem potentielle skalaspørgsmål fra kategorien: *Henvendelse* er det kun *Jeg synes at R er intimiderende*, der er blevet sammensat med et andet parametre. De fire resterende potentielle skalaspørgsmål vælges derfor, hvilket fremgår af følgende figurer.

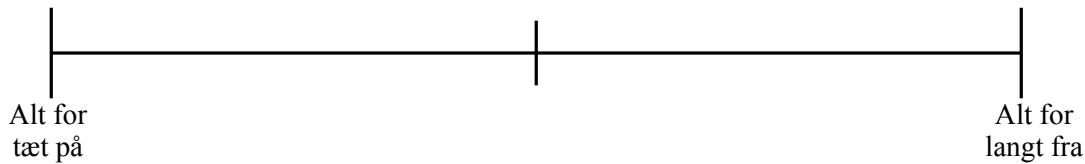
Jeg synes at robotten er imødekommen



Figur 5.19. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Jeg synes at robotten er imødekommen*.

Årsagen til at der ikke er et label på midtpunktet på Figur 5.19 er, at det vurderes at midtpunktet i sig selv naturligt og logisk indikerer midten. Derudover har det ikke været muligt at finde et pasende ord, som kunne anvendes som label til midtpunktet.

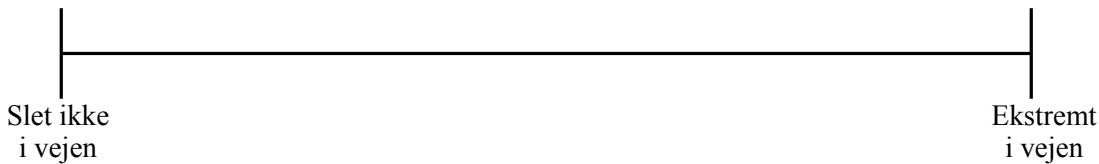
Jeg synes at robotten stoppede



Figur 5.20. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Jeg synes at robotten stoppede*.

De to labels på henholdsvis venstre og højre endepunkt, jævnfør Figur 5.20, afspejler testpersonernes egne udsagn, hvor årsagen til at der ikke er et label på midtpunktet er den samme som ved den foregående skala.

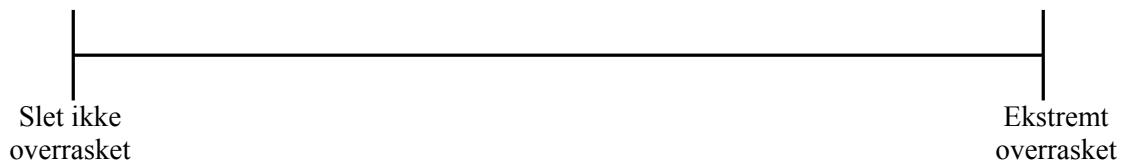
Jeg synes at robotten stod i vejen



Figur 5.21. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten stod i vejen*.

Årsagen til at skalaen på Figur 5.21 ikke præsenteres på en bipolær skala er, at det ikke har været muligt at angive en direkte modpart til at stå i vejen.

Jeg blev overrasket over robottens henvendelse

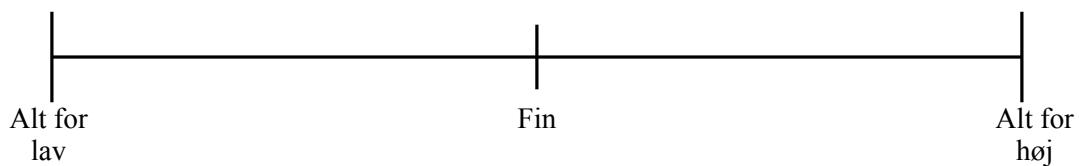


Figur 5.22. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg blev overrasket over robottens henvendelse*.

Af samme årsag som før præsenteres skalaen på Figur 5.22 på en unipolær skala.

I forhold til kategorien: *R's udseende* besluttes det, at ekskludere det potentielle skalaspørgsmål: *Jeg kan godt lide R's udseende*, fordi det formentlig er en form for overkategori til de resterende parametre. Det forventes at i tilfælde af, at en testperson vurderer robotten som værende elegant, fra samme kategori, sød, sjov og sej, som kommer fra kategori: *Positiv overfor R*, så vil en vurdering på hvor godt testpersonen kan lide robottens udseende ligeledes være positivt. Det antages derfor at hvorvidt danske rejsende kan lide robottens udseende måles indirekte ved de andre parametre, der omhandler robottens udseende.

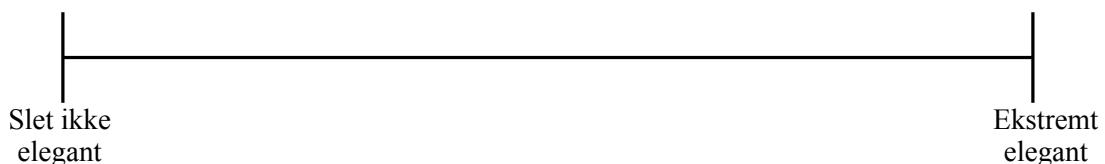
Jeg synes at robottens højde er



Figur 5.23. Bipolær VAS med lukkede endepunkter med navngivet midtpunkt til skalaspørgsmålet: *Jeg synes at robottens højde er*.

Årsagen til at robottens højde bliver evalueret på en bipolær skala, som illustreret på Figur 5.23, er, at høj og lav er modsætninger. At midtpunktet er navngivet: *Fin* er baseret på testpersonernes egen oplevelse af robottens højde. Derudover kan det være med til at kalibrere de næste testpersonernes evaluering af robottens højde i forhold til skalaen. En anden årsag til at *Fin* anvendes er, at der ikke nødvendigvis er en logisk højde lige akkurat mellem lav og høj. I dette tilfælde afspejler *Fin* en højde, der svarer til omkring albuehøjde, som er den højde testpersonerne gav udtryk for var fin.

Jeg synes at robotten er elegant

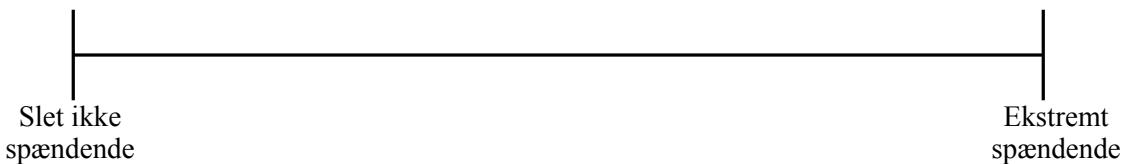


Figur 5.24. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er elegant*.

Af samme årsag som tidligere, har det ikke været muligt at finde en naturlig og logisk modpart til *elegant*, hvorfor denne skala præsenteres på en unipolær VAS, som fremgår af Figur 5.24.

Baseret på kategorien: *Interesse for R* fremgår der to potentielle skalaspørgsmål, som relaterer sig til hvorvidt robotten fangede ens opmærksomhed og om robotten er spændende, jævnfør Underafsnit 5.5.7 (*Interesse for R*). Det vælges at sammensætte de to potentielle skalaspørgsmål til ét fordi det vurderes, at de relaterer til den samme oplevelse, da de begge tilhører den samme kategori. På Figur 5.25 fremgår skalaen udviklet til skalaspørgsmålet: *Jeg synes at robotten er spændende*. Det antages derfor at i tilfælde, hvor en testperson synes, at robotten er spændende så vil robotten ligeledes have fangen testpersonens opmærksomhed, hvorvis testpersonen ikke synes, at robotten er spændende så er det ikke sikkert at den fanger ens opmærksomhed. Dertil kan der ligeledes opstå en situation, hvor robotten fanger en testpersons opmærksomhed uden, at testpersonen finder robotten spændende.

Jeg synes at robotten er spændende

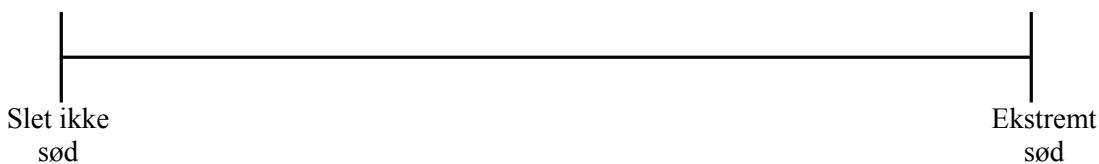


Figur 5.25. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er spændende*.

Igen har det hverken været muligt at finde en naturlig og logisk modpart til *spændende* eller finde et label til et potentielt midtpunkt i tilfælde af, at der fandtes en modpart.

Det vælges at medtage samtlige fire potentielle skalaspørgsmål fra kategorien: *Positiv overfor R*, jævnfør Underafsnit 5.5.8. Robotten evalueres derfor i forhold til hvor *sød*, Figur 5.26, *sjov*, Figur 5.27, *sej*, Figur 5.28, samt hvorvidt testpersonen kan lide at blive betjent af robotten, Figur 5.29. Årsagen til at de alle medtages er, at det forventes at de har en indflydelse på evalueringen af andre parametre. Derudover er *sød*, *sjov* og *sej* parametre, som går igen ved flere testpersoner.

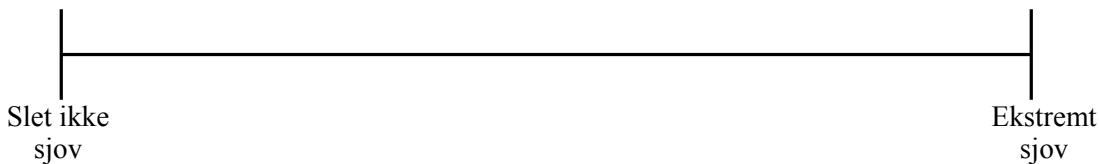
Jeg synes at robotten er sød



Figur 5.26. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er sød*.

Da det ikke med sikkerhed kan fastslåes, hvad testpersonerne har ment, da de gav udtryk for, at robotten er *sød*, har det ikke været mulig at finde en modpart. Det lader nemlig ikke til, at testpersonerne har brugt ordet *sød* om robotten, hvor modsætningen da har været *dum*.

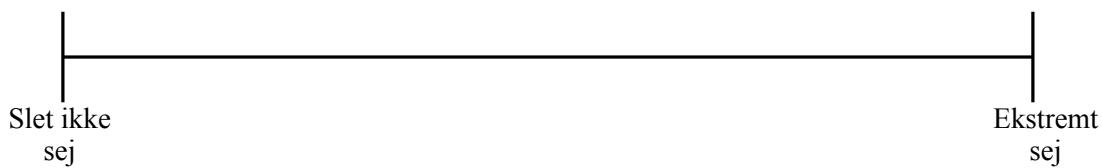
Jeg synes at robotten er sjov



Figur 5.27. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er sjov*.

Lignende er tilfældet med *sjov*, hvor det ikke vides med sikkerhed, hvad testpersonernes har ment, da de gav udtryk for, at robotten er *sjov*, hvorfor dette skalaspørgsmål ligeledes evalueres på en unipolær skala.

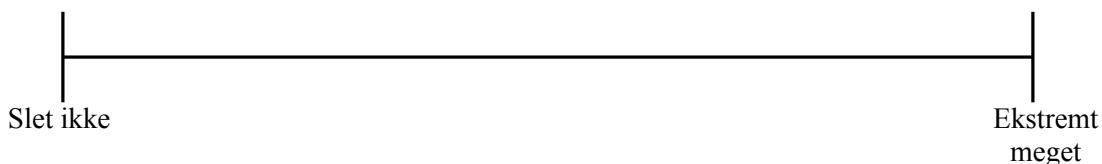
Jeg synes at robotten er dejlig



Figur 5.28. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg synes at robotten er dejlig*.

Tilsvarende gør sig gældende for hvorhvidt robotten er dejlig.

Jeg kan godt lide at blive betjent af robotten

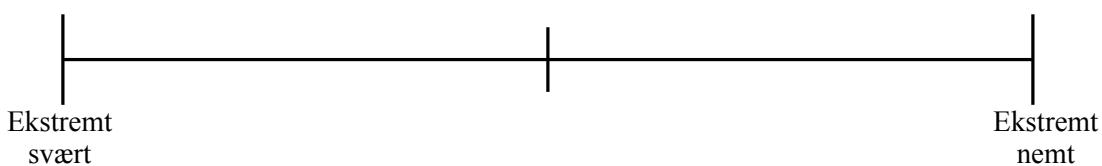


Figur 5.29. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Jeg kan godt lide at blive betjent af robotten*.

Det har ikke været muligt at formulere skalaspørgsmålet: *Jeg kan godt lide at blive betjent af robotten* på en måde, hvorpå det ville give mening at præsentere spørgsmålet på en bipolær skala, hvorfor det er præsenteret på en unipolær skala.

Det vælges at medtage de to potentielle skalaspørgsmål til kategorien: *Kendskab til teknologi*. Det første skalaspørgsmål vedrører hvordan det var at interagere med robotten, Figur 5.30, og det andet handler om ens kendskab til teknologi, Figur 5.9.

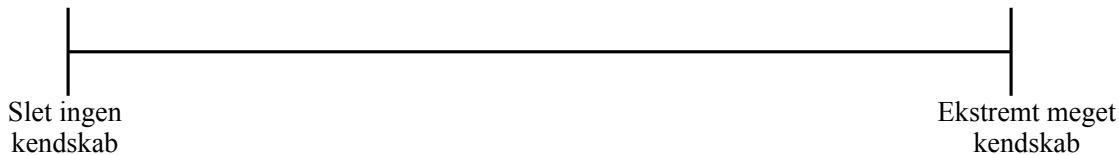
Hvordan var det at bruge robotten?



Figur 5.30. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Hvordan var det at bruge robotten?*

Det vurderes at *nemt* og *svært* er hinandens modpart, hvorfor de anvendes som labels på endepunkterne. Dog har det ikke været muligt, at finde et naturligt og logisk label til midtpunktet.

Hvor meget kendskab har du til teknologi/robotter?

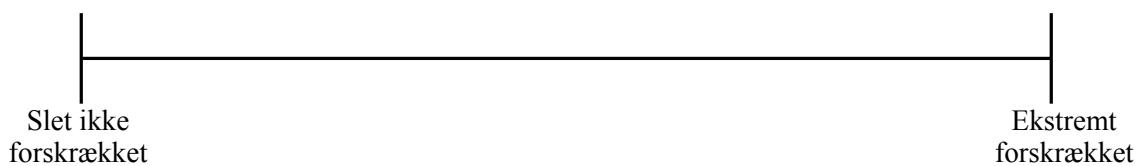


Figur 5.31. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Hvor meget kendskab har du til teknologi/robotter?*

Dette skalaspørgsmål medtages men det vil indgå som en del af demografi, hvorfor det ikke vil indgå samme med de andre skalaer. Årsagen til at det overhovedet inkluderes er, at det forventes, at ens kendskab til teknologi har indflydelse på hvordan HRI evalueres. Dog er det ikke besluttet om skalaspørgsmålet skal vedrører ens kendskab til teknologi eller ens kendskab til robotter.

De tre potentielle skalaspørgsmål fra kategorien: *Tillid til R* medtages. På den første skala skal testpersonen angive hvor forskräkket robotten gjorde en, jævnfør Figur 5.32, hvor på den næste evaluerer testpersonen hvorvidt de regnede med at robotten fulgte dem hen til det valgte sted, jævnfør Figur 5.33, og på den sidste skala evaluerer testpersonen hvor tryg de er ved robotten, jævnfør Figur 5.34.

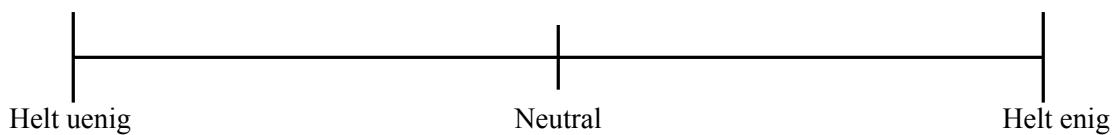
Robotten gjorde mig forskräkket



Figur 5.32. Unipolær VAS med lukkede endepunkter til skalaspørgsmålet: *Robotten gjorde mig forskräkket.*

Da der ikke kunne findes en naturlig og logisk modpart til forskräkket vælges det, at præsentere skalaspørgsmålet på en unipolær skala, jævnfør Figur 5.32.

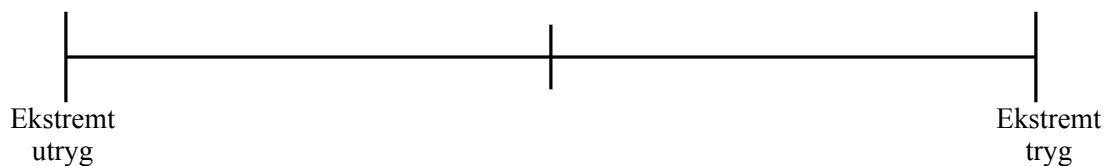
Jeg regnede med at robotten fulgte mig hen til det sted jeg valgte



Figur 5.33. Bipolær VAS med lukkede endepunkter og navngivet midtpunkt til skalaspørgsmålet: *Jeg regnede med at robotten fulgte mig hen til det sted jeg valgte.*

Da skalaspørgsmålet på Figur 5.33 er formuleret som det er, afspejler det i højere grad et udsagn en testperson kan erklærer sig enig eller uenig i, fremfor at testpersonerne angiver hvor meget de oplever en bestemt parameter. Derudover vurderes det at der findes et naturligt og logisk midtpunkt mellem at være *helt uenig* og *helt enig* nemlig *neutral*.

Jeg føler mig tryg ved robotten



Figur 5.34. Bipolær VAS med lukkede endepunkter og unavngivet midtpunkt til skalaspørgsmålet: *Jeg føler mig tryg ved robotten*.

Årsagen til at der ikke er et label på skalaens midtpunkt er, at der ikke forekommer et bestemt ord, som er en naturlig og logisk midte mellem *ekstremt utryg* og *ekstremt tryg*. Derudover vurderes det at markeringen af midtpunktet i sig selv bør være nok til at forstå skala.

Ud af de 30 potentielle skalaspørgsmål til de ti kategorier præsenteret i Afsnit 5.5 (*Resultater fra Affinity diagram*) er der nu udvalgt og opstillet 24 skalaer, som ud fra projektgruppens vurdering kan anvendes til at evaluere HRI. Hver skala afspejler en specifik parameter, som bygger på de danske rejsendes udtalelser omkring mødet med robotten i Aalborg Lufthavn.

5.7 Sammenligning mellem fundne parametre og relatede studier

Ud fra resultaterne præsenteret i Afsnit 5.5 (*Resultater fra Affinity diagram*) udledes de forskellige beskrevne parametre. Sammenholdes de fundne parametre med den beskrevne litteratur belyst i Afsnit 1.4 (*Interaktion med sociale robotter*), er nogle af parametrene gennemgående for HRI. Det gør sig gældende for følgende parametre:

- Kendskab til teknologi/robotter
- Hvor anmassende robotten er
- Hvor irriterende robotten er
- Hvor nemt/svært det var at bruge robotten
- Fornøjelse ved brug af robotten
- Brugbarheden af robotten
- Hvor sjov robotten er
- Højden af robotten
- Hastigheden af robotten
- Hvor rolige bevægelser robotten har
- Hvor menneskelig robotten percipieres
- Tryg ved robotten
- Hvor meget brugerden stoler på at robotten følger en på vej
- Skærmens reaktion
- Afstand til brugerden

Alle disse parametre har i højere eller mindre grad indflydelse på, hvordan interaktionen med robotten opleves og bør derfor tages i betragtning under designet af en social robot. Baseret på testpersonernes udsagn, jævnfør *affinity notes*, er der en tendens til at testpersonerne antropomorfiserer robotten, da de ofte tildeler robotten menneskelig karakteristikas blandt andet i forhold til dens bevægelser.

Udover de parametre, som går igen i projektets undersøgelse og litteraturen, så nævner testpersonerne i AAL andre aspekter, som adskiller sig fra den gennemgåede litteratur.

Under de forskellige kategorier bliver der, udover de nævnte parametre, lagt vægt på følgende:

- Om robotten er imødekommen
- Om robotten stod i vejen
- Om robottens hjælp perciperes som personlig
- Om robottens henvendelse overrasker
- Om robotten forskräkker
- Om robotten er elegant
- Om robotten er sød
- Om robotten er sej
- Om robotten er spændende

Disse parametre er ikke nødvendigvis nye, men kan ligge implicit i nogle af de parametre, der overlapper eller forefindes i litteraturen, uden at være nævnt med samme ord. Parametrene kommer til udtryk, når testen afvikles i AAL og har måske ikke betydning andre steder, der ikke minder om en lufthavnssituation. Dette gør sig blandt andet gældende for brugernes holdning til om robottens henvendelse overrasker eller forskräkker en, men også om hvorvidt robotten stod i vejen. I en lufthavn kan det forestilles, at rejsende har travlt eller er i stressede situationer, hvor interaktionen eller mødet med en robot kan virke overvældende, hvis brugerne godt ved hvor de er på vej hen. Det kan derfor være et vigtigt parameter at robotten ikke forskräkker og i hvert fald ikke står i vejen, da dette kan tilføje yderligere stress til situationen. Det er derfor muligt at parametrene kun gør sig gældende for en lufthavnssituation, men ikke desto mindre er det parametre, der kan måles, når brugerne interagerer med robotten, for at få en idé om helhedsindtrykket af robotten og hvor der kan foretages forbedringer.

Ifølge litteraturen er der også betydende parametre ved sociale robotter, som ikke er forefindes i undersøgelsen foretaget i AAL. Dette er blandt andet *image*, der vedrører brugerens overbevisning om at interaktionen med robotten øger ens status set med andres øjne, (Graaf og Allouch 2013, s. 1478), og i nogen grad *social influence*, hvor der kun er én testperson, der har givet udtryk for, hvad andre kunne tænke om situationen. Det kan muligvis skyldes at danske rejsende ikke interesserer sig specielt meget for hvad andre tænker om dem i den pågældende situation, hvilket kan adskille sig fra andre kulturer.

Ydermere nævner testpersonerne ikke direkte noget om *social distance*, hvilket kan skyldes, at robotten har overholdt de sociale normer, og at testpersonerne har følt, at de har været i kontrol over situationen. Ud fra teorien omkring *power distance* tyder det på, at testpersonerne har haft en følelse af at de var den dominante partner, hvor robotten var den underordnede. Da interaktionen har været baseret på at robotten skulle hjælpe den rejsende med at finde rundt i lufthavnen, hvilket kan betragtes som et samarbejde mellem robot og testperson, er det muligt at de rejsende har opført sig mere venligt, intimt og involveret end hvad der måske havde været tilfældet i en anden situation. Dette hænger sammen med *task distance*, (Kim og Mutlu 2014, s. 784).

Indgangsvinkel er ikke nævnt som et betydnende parameter blandt de danske rejsende. Dette står i modstrid til undersøgelsen foretaget af Dautenhahn, Walters m.fl. (2006), der konkluderer, at indgangsvinklen har betydning for HRI. Det kan have haft betydning, at testpersonerne i AAL har bevæget sig frit i test området efter sikkerhedskontrollen i lufthavnen. Det er blevet observeret, at rejsende ignorerer robotten, medmindre den står med front direkte mod én, hvorfor det ikke bemærkes, hvis robotten prøver at henvende sig fra den ene eller anden side.

En parameter, som ikke direkte er blevet udledt baseret på de danske rejsendes udtalelser er *companionship*, fortolkes der derimod på testpersonernes udtalelser tyder det på at denne parameter faktisk har en betydningen for HRI. Dette er dog primært rette mod at robotten kan hjælpe ældre og børn, der rejser alene, hvor det bliver nævnt at børn kan have nemmere ved at danne et forhold til robotter end voksne. Det vælges ikke at udvikle en decideret skala hvorpå det kan evalueres hvorvidt robotten betragtes som en følgesvend. Dog vurderes det, at dette parameter formentlig bliver dækket af skalaspørgsmålet: *Jeg oplever robottens hjælp som personlig.*

Selvom der fra litteraturen forekommer parametre, som har væsentlig betydning for HRI og som ikke er udledt i denne undersøgelse, vælges det kun at designe skalaer baseret på de fundne parametre.

Diskussion 6

I følgende afsnit vil feltundersøgelsen i AAL diskuteres. Da de forskellige valg løbende er blevet diskuteret, vil følgende indeholde diskussion af observationer gjort under feltundersøgelsen, resultater fra affinity, hvordan udvælgelsen af skalaer har foregået og hvad der kan gøres bedre til næste test.

6.0.1 Observationer og resultater fra feltundersøgelsen

Feltundersøgelsen i AAL var succesfuld, da flere danske rejsende havde lyst til at interagere med robotten og efterfølgende havde tid til at snakke om den med testlederen. Som forventet var der dog også mange rejsende, der ikke havde lyst til at interagere med robotten. Nogle rejsende så slet ikke robotten, hvilket kunne skyldes at den både var lav og ikke altid i bevægelse, men også at den stod med ryggen eller siden til og derfor ikke blev set som noget der kunne interageres med. Mange rejsende både så, snakkede om og grinede af robotten, men havde enten ikke lyst til at interagere med robotten eller trykkede nej til at få hjælp med at finde rundt i AAL. Dette kan både skyldes, at de bare ikke havde lyst til at interagere med robotten, men også at de ikke vidste at robotten henvendte sig til dem og at de havde mulighed for at interagere med den. Det var svært for robotstyreren altid at komme tæt på de rejsende, fordi udsynet til robotten ikke altid var lige godt og det ikke var ønsket at køre ind i nogen, hvilket kan have haft en betydning for om de rejsende følte den henvendte sig til dem eller ej. Ved udvikling af social intelligens er det derfor nødvendigt at brugertesten på hvordan en robot skal henvende sig til mennesker, for at de forstår denne henvendelse og reagerer på den.

Robotten kørte rundt i lufthavnen med en skærm hvorpå spørgsmålet: *Kan jeg hjælpe dig med at finde rundt i Aalborg Lufthavn?*, blev præsenteret. Nogle af de rejsende trykkede Nej på skærmen, hvilket kan skyldes at de ikke havde brug for hjælp til at finde rundt i lufthavnen, men betyder ikke nødvendigvis at de ikke i en anden situation ville have lyst til at interagere med og få hjælp af robotten. Spørgsmålet kan derfor være lidt misvisende og afskære nogle af de rejsende, som ellers ville have relevante og vigtige pointer i forhold til en social robot. Det blev diskuteret, hvorvidt det ville være en god idé at stoppe de rejsende, som trykker Nej, til næste test, for at få deres begrundelse for hvorfor de trykkede, som de gjorde. Dog blev det efterfølgende besluttet ikke at forsøge at indhente og interviewe de rejsende, der trykkede Nej, da det er svært at spørge ind til uden at lyde bebrejdende. Ydermere kan de rejsende have trykket Nej, fordi de ikke har tid til at interagere, hvorfor det vil være ubelejligt hvis projektgruppen stopper dem for at stille spørgsmål.

På trods af at der er flere, der ikke vil interagere med robotten, er der stadig rejsende, som gerne vil. Projektgruppen havde generelt ikke indflydelse på hvem, der blev rekrutteret. Det var kun robotstyreren, som kørte robotten hen til de rejsende, hvilket forsøges at gøre

på alle, der ikke stormer af sted eller virker optaget. Feltundersøgelsen er på den måde økologisk, da det kun er de rejsende, som vil interagere med en social robot i en lufthavn, der deltager i undersøgelsen. Dog er projektgruppen ikke helt usynlig i lufthavnen, da observatørerne sidder rundt omkring med en blok papir og flere af de rejsende smiler hen til testleder, robotstyrer eller observatører efter at have set robotten. Ydermere fortæller robotten, at den kommer fra Aalborg Universitet og at der er en undersøgelse i gang.

Selvom feltundersøgelsen er økologisk kan projektgruppen kun påvirke de rejsende med de ting, som robotten kan. Det kan være svært for testpersoner at påpege vigtige parametre, hvis helheden fungerer godt, hvorfor der kan være parametre, som ikke kommer frem, fordi robotten opfører sig som forventet.

Et parameter, der bliver nævnt flere gange er dog, at robottens hastighed har betydning. Der er flere af testpersonerne, der synes robotten bevæger sig alt for langsomt. Nogen gange så langsomt, at det er besværligt at gå bag ved den eller følge dens tempo, når testpersonerne går ved siden af den. Denne hastighed kan ikke ændres på *Double*-robotten, men ellers ville det være relevant at øge makshastigheden, så fremtidige brugere ikke er ved at falde over robotten eller bliver frustreret over at robotten er for langsom.

Et andet parameter, der bliver nævnt flere gang er, hvor godt touchskærmen på robotten virker. Det fremgår både gennem observationerne og når affinity diagrammet udvikles. Under interaktionen med robotten var det tydeligt, at flere rejsende blev irriteret enten over at skærmen ikke reagerede på deres berøring eller at robotten ikke ville acceptere et nej, når de prøvede at trykke på *Nej*. Dette havde betydning for, at flere af de rejsende stoppede interaktionen og gik videre, selvom de formentlig gerne ville have hjælp af robotten. Grunden til at touchskærmen ikke altid reagerede på berøring skyldes *wireframed* designet i *Marvel*. For det første kræver opsætningen en blid berøring på skærmen, hvilket ikke var det første de rejsende gjorde, da teknologien ikke virkede. For det andet eksekveres *wireframed* gennem *Doubles* applikation, hvorfor kommunikationen kan være langommere end hvad de rejsende er vant til. Til den kommende test kan det derfor være en fordel at overveje, hvordan de rejsende vil få nemmere ved at interagere med touchskærmen. Dette kan eventuelt gøres ved at implementere et stykke tekst, der fortæller at der skal trykkes blidt på skærmen.

Baseret på feltundersøgelsen, tyder det kraftigt på, at en social robot kan være en god idé, både i lufthavne såvel som på andre offentlige lokationer. I lufthavne er robotten god som vejviser, specielt til rejsende, der ikke forstår det talte sprog eller til børn, der rejser alene og skal følges til den rigtige gate. Specielt børn er meget begejstret for robotten, hvorfor det kan være eftertragtet at have en social robot, der kan hjælpe børn. Udover lufthavnssituationen blev der eksempelvis foreslægt, at robotten kan bruges på hospitaler, i storcentre eller på museer som vejviser, underholdning eller guide. Ud fra det opbyggede affinity diagram fremgår det, at kendskab til teknologi kan spille en rolle i hvorvidt den rejsende har lyst til at henvende sig eller interagere med en robot som *Double*. Flere testpersoner gav udtryk for at robotten gav mening i forhold til deres profession, nogle var interesseret i robotten fordi det var en kørende iPad og nogle viste interesse, fordi de selv arbejdede med teknologi eller robotter til dagligt. Generelt tyder resultaterne på, at brugere med et positivt forhold til teknologi har nemmere ved at acceptere en social robot

og mere lyst til at interagere med den.

6.0.2 Skalaer

Da skalaspørgsmålene i første omgang blev dannet, er det hovedsageligt gjort på baggrund af hvad testpersonerne selv har sagt og udtrykt som værende vigtige parametre. Efterfølgende er nogle af skalaspørgsmålene blevet slået sammen, fordi de dækkede over det samme. Da de tilhørende skalaer blev designet, blev der taget udgangspunkt i det enkelte skalaspørgsmål frem for at betragte helheden. Dog er der til alle spørgsmål designet en VAS, hvor det kun er ankerpunkter og labels, der varierer. Dette er gjort, for at udvikle den bedste skala til hver af de enkelte parametre. Modsat vil fordelen ved at gøre skalaerne mere konsistente være, at når testpersonerne præsenteres for dem og skal angive deres respons på dem, så skal de ikke forholde sig til flere forskellige typer af skalaer. Som skalaerne er bygget op stilles skalaspørgsmålene på forskellige måder med varierende endepunkter, hvilket kan skabe forvirring, misforståelser og i værste fald påvirke testpersonernes markering på skalaerne.

Skalaerne er blevet sat op med lukkede endepunkter. Dette er gjort fordi testpersonerne hverken bliver præsenteret for robotten på forskellige måder under interaktionen eller bliver præsenteret for den samme skala flere gang. Det vurderes derfor at være usandsynligt, at testpersonerne vil få brug for at kunne svare mere ekstremt end den tidligere besvarelse, netop fordi skalaspørgsmålene kun vil blive præsenteret én gang. Derudover vil lukkede endepunkter skabe nogle klare rammer for, hvor testpersonerne kan svare, og det forventes derfor at data bliver mere normalfordelt end hvis skalaerne havde åbne endepunkter. Ulempen ved at bruge lukkede endepunkter er, at testpersonerne kan blive kantforskrækket og svare længere væk fra endepunktet end de egentlig mener.

For at mindske misforståelser af skalaerne vælges det at præsentere de skalaer, der minder mest om hinanden sammen i den efterfølgende test.

Delkonklusion 7

For at besvare den første problemstilling: *Ud fra hvilke parametre beskriver danske rejsende interaktion med en social robot i en dansk lufthavn?*, blev der udført en feltundersøgelse i Aalborg Lufthavn på danske rejsende. Her sås det at flere rejsende havde lyst til at interagere og få hjælp af robotten. Til feltundersøgelsen blev der dels opstillet nogle samtaleemner, hvorfra testpersonerne fik mulighed for at give deres spontane respons relateret til interaktionen med robotten, og dels formuleret nogle specifikke spørgsmål, som omhandlede specifikke parametre, der har betydningen for HRI.

Kategori	Parameter
Skærmen virker ikke	Skærmens reaktion
Assisterer mennesker	R kan hjælpe Personlig hjælp
Væremåde	Bevægelse Hastighed Irriterende Menneskelig* Anmassende**
Henvendelse	Imødekommende Afstand I vejen Overrasket
Udseende	Højde Elegant
Interesse	Spændende
Postiv over for R	Sød Sjov Sej Betjent af R
Kendskab til teknologi	At bruge R*** Kendskab til teknologi/robotter
Tillid	Forskrækket Regnede med robotten Tryg

Tabel 7.1. Oversigt over de fundne og udvalgte parametre, samt hvilken kategori de stammer fra. Parameteren: *Menneskelig** stammer fra to kategorier: *Væremåde* og *Udseende*, *Anmassende*** stammer fra tre kategorier: *Væremåde*, *Henvendelse* og *Interagerer ikke med R*, hvor *At bruge R**** stammer fra: *Positiv over for R* og *Kendskab til teknologi*. *R* er en forkortelse for robot.

For at udlede hvilke parametre danske rejsende tilskriver interaktionen med en social robot, blev testpersonernes respons kategoriseret og analyseret igennem et *affinity diagram*,

jævnfør Afsnit 5.5 (*Resultater fra Affinity diagram*). Baseret på dette diagram blev der opstillet nogle potentielle skalaspørgsmål, hvorfra udvælgelsen og udviklingen af skalaerne er baseret. Der er i alt udvalgt 24 parametre, som der er blevet dannet skalaer ud fra, jævnfør Tabel 7.1. Sammenlignes de udvalgte parametre med de parametre, der er fundet til at have indflydelse på HRI i tidligere undersøgelser, fremgår det, at flere af parametrene går igen også for danske rejseende. Ydermere findes der parametre, som ikke er blevet nævnt i litteraturgennemgangen i Afsnit 5.5 (*Resultater fra Affinity diagram*), men som medtages, når skalaerne dannes.

Til hver af de udvalgte parametre er der udviklet en skala hvorpå parameteren kan evalueres, jævnfør Afsnit 5.6 (*Udvælgelse af skalaer*), hvilket er en del af at besvare den anden problemstilling: *Hvordan kan de fundne parametre anvendes til at evaluere HRI?*. Da disse skalaer, samt tilhørende skalaspørgsmål, er udviklet specifikt til én parameter er der ikke taget højde for det samlede helhedsindtryk, hvorfor det vurderes at den anden problemstilling ikke er besvaret tilfredstillende. Det er derfor nødvendigt at danne et helhedsindtryk samt videreudvikle de 24 skalaer og tilpasse dem i forhold til følgende: Formuleringen af skalaspørgsmålet, labels, skalatype, samt hvilken rækkefølge skalaerne skal præsenteres i, i den kommende undersøgelse.

Del III

Videreudvikling af skalaer til evaluering af HRI

For endeligt at kunne besvare den anden problemstilling: *Hvordan kan de fundne parametre anvendes til at evaluere HRI?*, vil de udvalgte skalaer blive yderligere tilpasset i forhold til skalaspørgsmål, labels og skalatype. Dernæst vil der udføres endnu en undersøgelse i Aalborg Lufthavn, hvor de rejsende skal evaluere HRI ved hjælp af de udviklede skalaer, som er baseret på de fundne parametre.

Tilpasning af skalaer 8

Som nævnt i Kapitel 7 (*Delkonklusion*) er skalaerne udviklet specifikt til ét enkelt parameter, hvorfor der ikke er taget højde for helhedsindtrykket. Det vil derfor i følgende afsnit fokuseres på at videreudvikle og tilpasse skalaerne, blandt andet i forhold til at opnå en form for konsistens relateret til formuleringen af skalaspørgsmål, labels samt skalatype. Målet er ikke at alle skalaer er ens, men at tilpasse skalaerne så de fremstår mere præsentable end hvad tilfældet er nu, samt at samle de skalaer, hvis skalaspørgsmål stilles på samme måde. Derudover tilstræbes det, at gruppere skalaerne efter det emne de evaluerer. For at undgå at testpersonerne angiver det samme punkt på skalaerne, fordi to parametre vedrører det samme emne, vil det tilstræbes ikke at placerer skalaerne lige efter hinanden, hvis de vedrører det samme emne. Det forventes, at ved at tilpasse og videreudvikle skalaerne vil det være lettere for testpersonerne at evaluere interaktionen med robotten. Der tages udgangspunkt i de udvalgte skalaer, der blev præsenteret i Afsnit 5.6 (*Udvælgelse af skalaer*).

Foruden at tilpasse og videreudvikle de udvalgte skalaer vil der ligeledes fokuseres på at vælge den rækkefølge, hvorved skalaerne efterfølgende skal præsenteres for testpersonerne.

8.1 Tilpasning og rækkefølge på skalaer

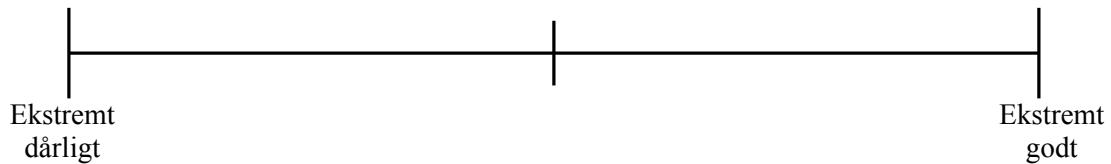
En af årsagerne til at skalaspørgsmål eller labels omformuleres og en anden skalatype vælges er, at risikoen for potentiel bias ved at lægge ord i munnen på testpersonerne mindskes. En anden årsag er, at der er situationer, hvor det valgte spørgsmål og de valgte labels ikke passer sammen i forhold til at fuldende en sætning og situationer, hvor spørgsmål og label modsiger hinanden. Da der tilstræbes en form for konsistens, i forhold til helhedsindtrykket, vil skalaerne ydermere blive vurderet i forhold til hvordan de passer sammen, for på den måde at gruppere skalaer, der er formulert og præsenteret ens.

Der vil foretages nogle grafiske ændringer på skalaerne, såsom at centrere skalaspørgsmålet over skalaen, i modsætning til skalaerne præsenteret i Afsnit 5.6 (*Udvælgelse af skalaer*), hvor skalaspørgsmålet stod til venstre. I tillæg vil grammatiske fejl blive rettet. Når skalaerne præsenteres for testpersonerne, vil testpersonerne maksimalt få præsenteret fire skalaer af gangen. Det skyldes, at skalaerne præsenteres på en *Microsoft Surface Pro (5)*, hvis skærm er 12.3" med en oplosning på 2736 x 1824. Det besluttes at der kun er plads til fire skalaer, da skalaerne ellers vil være for kompakte. I Afsnit 8.2 (*Program til skalaer*) vil den praktiske del af opsætningen af skalaerne blive uddybet.

Den første skala, der præsenteres for testpersonerne vedrører hvordan robottens skærm reagerede, jævnfør Figur 8.1. Det vælges at dette skalaspørgsmål skal præsenteres først, hver gang, da testpersonerne responderer på skalaerne på en anden skærm, som reagerer på

en anden måde. Ved at præsentere dette skalaspørgsmål først vil skærmens reaktionsevne på robotten være frisk i hukommelsen, og testpersonerne vil derfor have mindre at sammenligne med i forhold til reaktionsevne på skærmen, hvor skalaerne præsenteres. Da testpersonerne interagerer med to skærme; robottens og computeren hvorpå skalaerne præsenteres, vælges det at omformulere skalaspørgsmålet så det specificeres, hvilken skærm testpersonen skal evaluere.

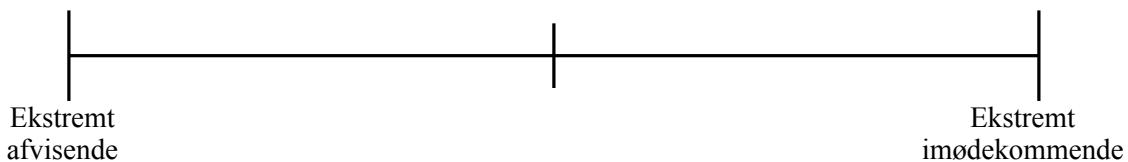
Hvordan synes du, at skærmen på robotten reagerede?



Figur 8.1. Tilpasset skala til: *Hvordan synes du skærmen på robotten reagerede?*

Det vælges at omformulere skalaspørgsmålet: *Jeg synes at robotten er imødekommen*de af to årsager: 1) At bruge ordet *imødekommen*de i den sætning er ledende, hvorfor det potentielt kan være en bias for testpersonerne, og 2) Formuleringen i spørgsmålet passer ikke med de valgte labels i forhold til at fuldende en sætning. Det giver ikke mening at stille spørgsmålet: *Jeg synes at robotten er imødekommen*de og så have en svarmulighed, der hedder: *Ekstremt afvisende*. Årsagen til at skalaen ikke præsenteres på samme måde som de skalaer, der tilhører: *Hvad synes du om robotten?* er, at fælles for de skalaer er, at de præsenteres på en unipolær skala. Hvis dette skalaspørgsmål skulle præsenteres på en unipolær skala, ville det være nødvendigt at danne to skalaer; én til *afvisende* og én til *imødekommen*de. Af de årsager vælges det at omformulere skalaspørgsmålet til: *Hvordan oplevede du robotten?*, hvor tidligere valgte labels bibeholdes, jævnfør Figur 8.2.

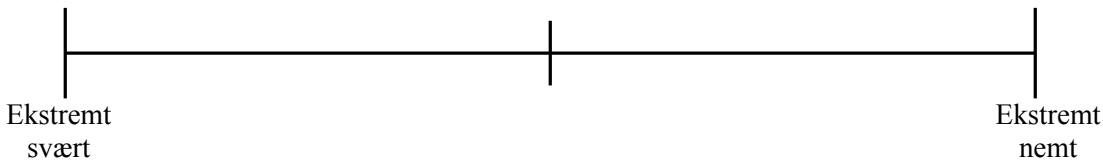
Hvordan oplevede du robotten?



Figur 8.2. Tilpasset skala til: *Hvordan oplevede du robotten?*

I henhold til skalaen gengivet på Figur 8.3, foretages der ingen ændringer i forhold til den udvalgte skala. Det vælges, at placere skalaen efter *Hvordan oplevede du robotten*, jævnfør Figur 8.2, fordi dette skalaspørgsmål dels er formulert på samme måde og dels fordi det hænger sammen med det første skalaspørgsmål: *Hvordan synes du, at skærmen på robotten reagerede?*.

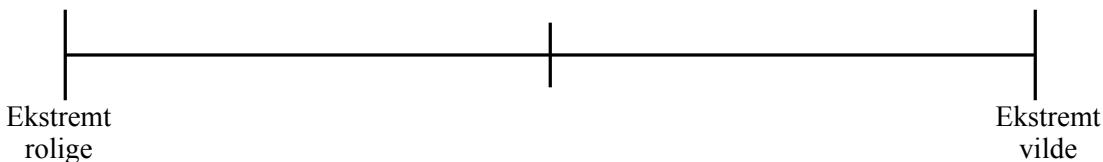
Hvordan var det at bruge robotten?



Figur 8.3. Tilpasset skala til: *Hvordan var det at bruge robotten?*.

Skalaspørgsmålet: *Jeg synes at robottens bevægelser er* omformuleres, fordi skalaens labels: *Ekstremt rolige* og *Ekstremt vilde*, ikke passer sammen med de labels, som er på skalerne, hvis skalaspørgsmål egentlig passer med *Jeg synes at robottens bevægelser er* såsom: *Jeg synes at robottens hastighed er*, hvis labels er: *Alt for langsom*, *Fin* og *Alt for hurtig*, for henholdsvis venstre, midtpunkt og højre label. Af den årsag vælges det er omformulere skalaspørgsmålet til: *Hvordan oplevede du robottens bevægelser?*, jævnfør Figur 8.4. Da denne formulering passer godt med: *Hvordan oplevede du robotten?* vælges det, at præsentere denne skala, som den fjerde skala.

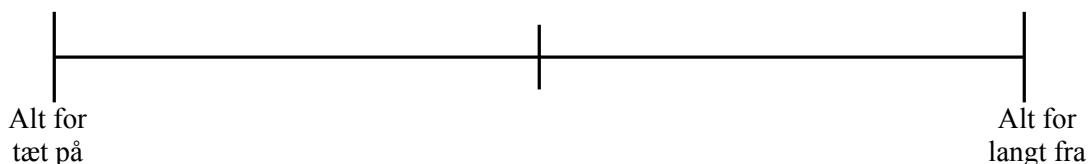
Hvordan oplevede du robottens bevægelser?



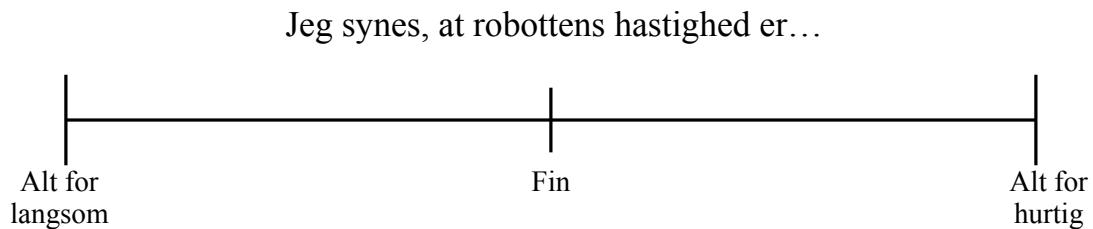
Figur 8.4. Tilpasset skala til: *Hvordan oplevede du robottens bevægelser?*.

De følgende tre skalaer for henholdsvis robottens afstand Figur 8.5, hastighed Figur 8.6, og højde Figur 8.7, vil blive præsenteret sammen, både fordi de minder om hinanden og fordi de alle tre er fysiske parametre. De tre skalaspørgsmål, labels og skalatype er stort set ens med den ene undtagelse, at der til skalaspørgsmålet: *Jeg synes, at robotten stoppede...*, ikke har et label på midtpunktet. De eneste ændringer, der er foretaget på de tre skalaer i forhold til dem, som blev præsenteret i Afsnit 5.6 (*Udvælgelse af skalaer*) er, at spørgsmålet er centreret, der er tilføjet tre punktummer efter spørgsmålet, og sat et komma efter *synes*.

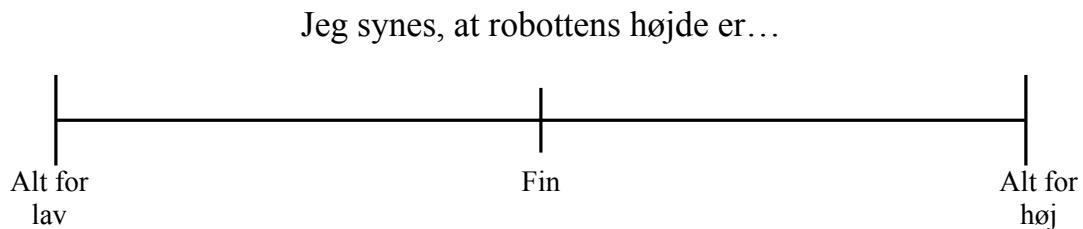
Jeg synes, at robotten stoppede...



Figur 8.5. Tilpasset skala til: *Jeg synes, at robotten stoppede....*

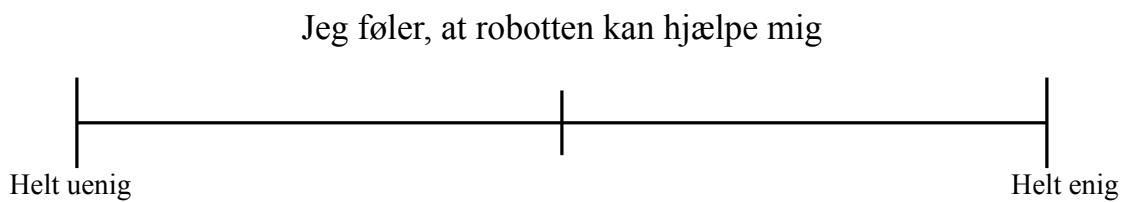


Figur 8.6. Tilpasset skala til: *Jeg synes, at robottens hastighed er....*



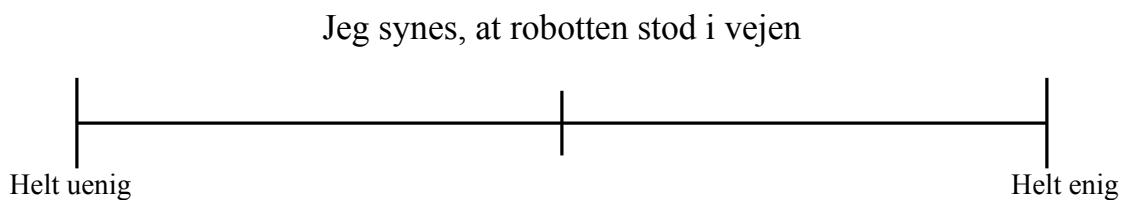
Figur 8.7. Tilpasset skala til: *Jeg synes, at robottens højde er....*

Da nogle af de udvalgte skalaer i forvejen er formuleret som udsagn testpersonerne kan erklære sig helt uenige eller helt enige i, vælges det at præsentere disse skalaer sammen og på ens skalaer. Den første skala, der vil blive præsenteret i det henseende er til skalaspørgsmålet: *Jeg føler, at robotten kan hjælpe mig*, jævnfør Figur 8.8. I forhold til udvælgelsen vil skalaen fremover blive præsenteret uden et navngivet label på midtpunktet.



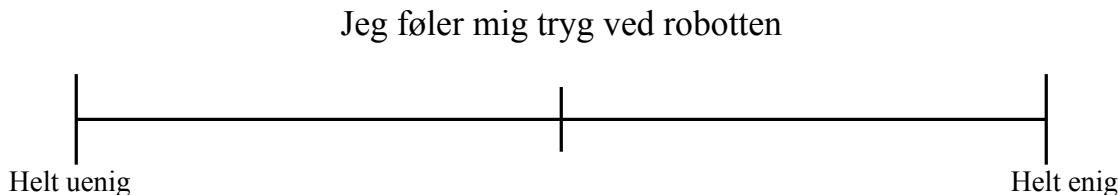
Figur 8.8. Tilpasset skala til: *Jeg føler, at robotten kan hjælpe mig.*

Skalaen hvorpå *Jeg synes, at robotten stod i vejen* evalueres får både tildelt nye labels og ny skalatype, jævnfør Figur 8.9. Årsagen til det er, at de labels, der var tildelt: *Slet ikke i vejen* og *Ekstremt i vejen* ikke passer særligt godt til at fuldende sætningen med skalaspørgsmålet og fordi det antages, at der kan opstå bias mellem spørgsmål og label, som de formuleret nu. Det vælges derfor at præsentere skalaen som en bipolær skala, hvor testpersonerne angiver, hvor helt uenige eller helt enige de er i, at robotten stod i vejen.



Figur 8.9. Tilpasset skala til: *Jeg synes, at robotten er i vejen.*

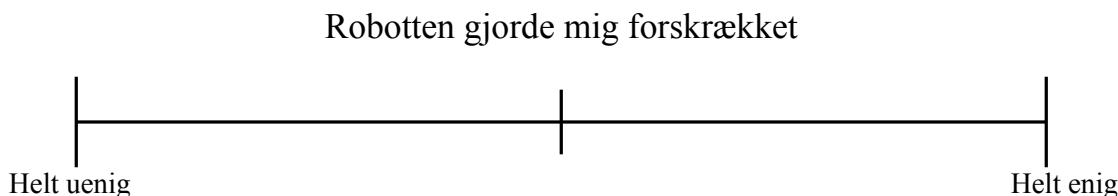
I henhold til de to labels: *Ekstremt utryg* og *Ekstremt tryg*, der er anvendt på skalaen til *Jeg føler mig utryg ved robotten*, er det klart, at der ikke forekommer særlig god overensstemmelse mellem label og spørgsmål, når sætningen fuldendes. Af den årsag og for at undgå at lægge ord i munden på testpersonerne, vælges det at omformulere de to labels, så skalaen går fra: *Helt uenig* til *Helt enig*, jævnfør Figur 8.10.



Figur 8.10. Tilpasset skala til: *Jeg føler mig tryg ved robotten*.

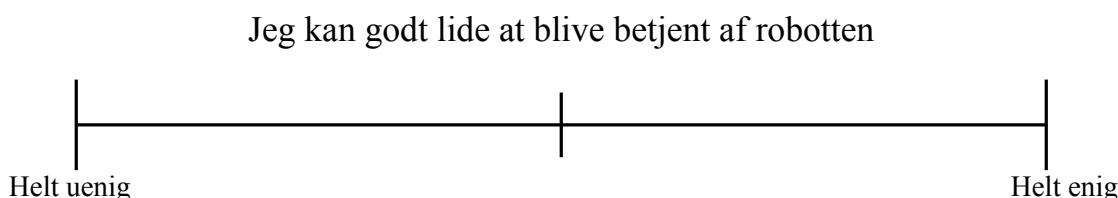
De følgende tre skalaer hænger egentlig sammen med de tre foregående skalaer, men da det er valgt maksimalt at præsenterer fire skalaer på en side, vælges det at præsentere de seks skalaer på to sider med tre skalaer på hver.

Igen opleves der uoverensstemmelse mellem labels: *Slet ikke forskrækket* og *Ekstremt forskrækket*, og skalaspørgsmålet: *Robotten gjorde mig forskrækket*. Det vælges derfor at ændre skalaen fra en unipolær til en bipolær skala, hvorpå testpersonerne kan angive hvor helt uenige eller helt enige de er i, at robotten forskrækkede dem, jævnfør Figur 8.11.



Figur 8.11. Tilpasset skala til: *Robotten gjorde mig forskrækket*.

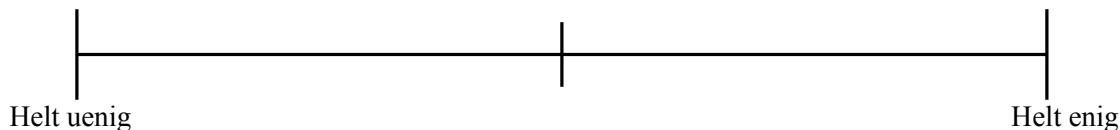
Tilsvarene uoverensstemmelse gør sig gældende for skalaspørgsmålet: *Jeg kan godt lide at blive betjent af robotten*, hvis endepunkter er navngivet med henholdsvis: *Slet ikke* og *Ekstremt meget*. Det vælges at bideholde formuleringen af skalaspørgsmålet og ændre skalatypen fra en unipolær skala til en bipolær skala med et midtpunkt. Dertil vælges det, at testpersonerne evaluerer skalaspørgsmålet i forhold til hvor helt uenige eller helt enige de er, jævnfør Figur 8.12.



Figur 8.12. Tilpasset skala til: *Jeg kan godt lide at blive betjent af robotten*.

De eneste ændringer, der er foretaget i forbindelse med skalaen til: *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte* er det grafiske med centreringen af skalaspørgsmålet, en kommaretteelse samt fjernelse af midtpunktets label: *Neutral*, jævnfør Figur 8.13.

Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte

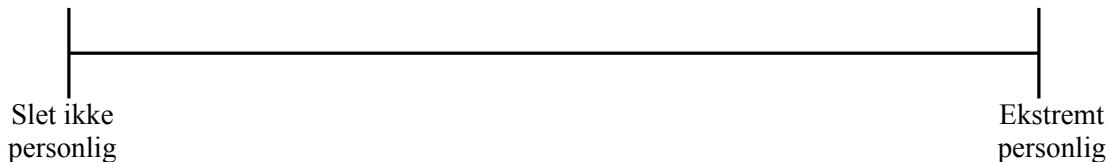


Figur 8.13. Tilpasset skala til: *Jeg regnede med at robotten fulgte mig hen til det sted jeg valgte*.

Selvom der er seks skalaer, hvorpå testpersonerne kan erklærer sig helt uenige eller helt enig i skalaspørgsmålet er det valgt, at præsentere de tre første på én side og de tre næste på den næste side. Rækkefølgen er valgt efter, at *Jeg føler, at robotten kan hjælpe mig* og *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte* minder om hinanden, hvorfor de to skalaer præsenteres på hver sin side. I tillæg minder *Jeg kan godt lide at blive betjent af robotten* i højere grad om *Jeg føler, at robotten kan hjælpe mig* end om *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte*, hvorfor det vurderes at dette skalaspørgsmål kan præsenteres på samme side som sidstnævnte. Derudover vælges det at præsentere: *Jeg føler mig tryg ved robotten* og *Robotten gjorde mig forskrækket* på hver sin side, da de muligvis hænger sammen.

Det vælges at omformulere skalaspørgsmålet: *Jeg oplever robottens hjælp som personlig*, dels for at undgå at lægge ord i munnen på testpersonerne og fordi det ved den formulering antages at testpersonerne har oplevet robottens hjælp som personlig. Derudover er det ikke muligt med de to labels: *Slet ikke personlig* og *Ekstremt personlig*, at fuldende sætningen efter spørgsmålet. Det vælges derfor at omformulere skalaspørgsmålet til: *Hvor personlig oplevede du robottens hjælp?* og bibe holde de to labels, jævnfør Figur 8.14.

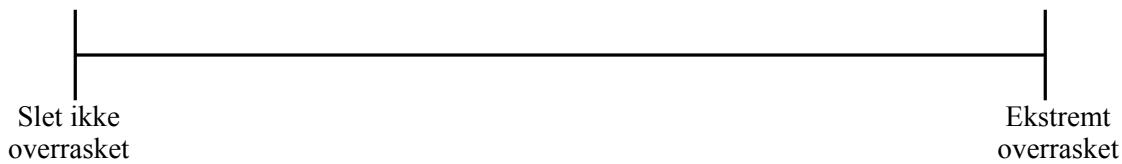
Hvor personlig oplevede du robottens hjælp?



Figur 8.14. Tilpasset skala til: *Hvor personlig oplevede du robottens hjælp?*

Af samme årsag omformuleres skalaspørgsmålet: *Jeg blev overrasket over robottens henvendelse til Hvor overrasket blev du over robottens henvendelse?*, hvor de to labels bibe holdes, jævnfør Figur 8.15.

Hvor overrasket blev du over robottens henvendelse?

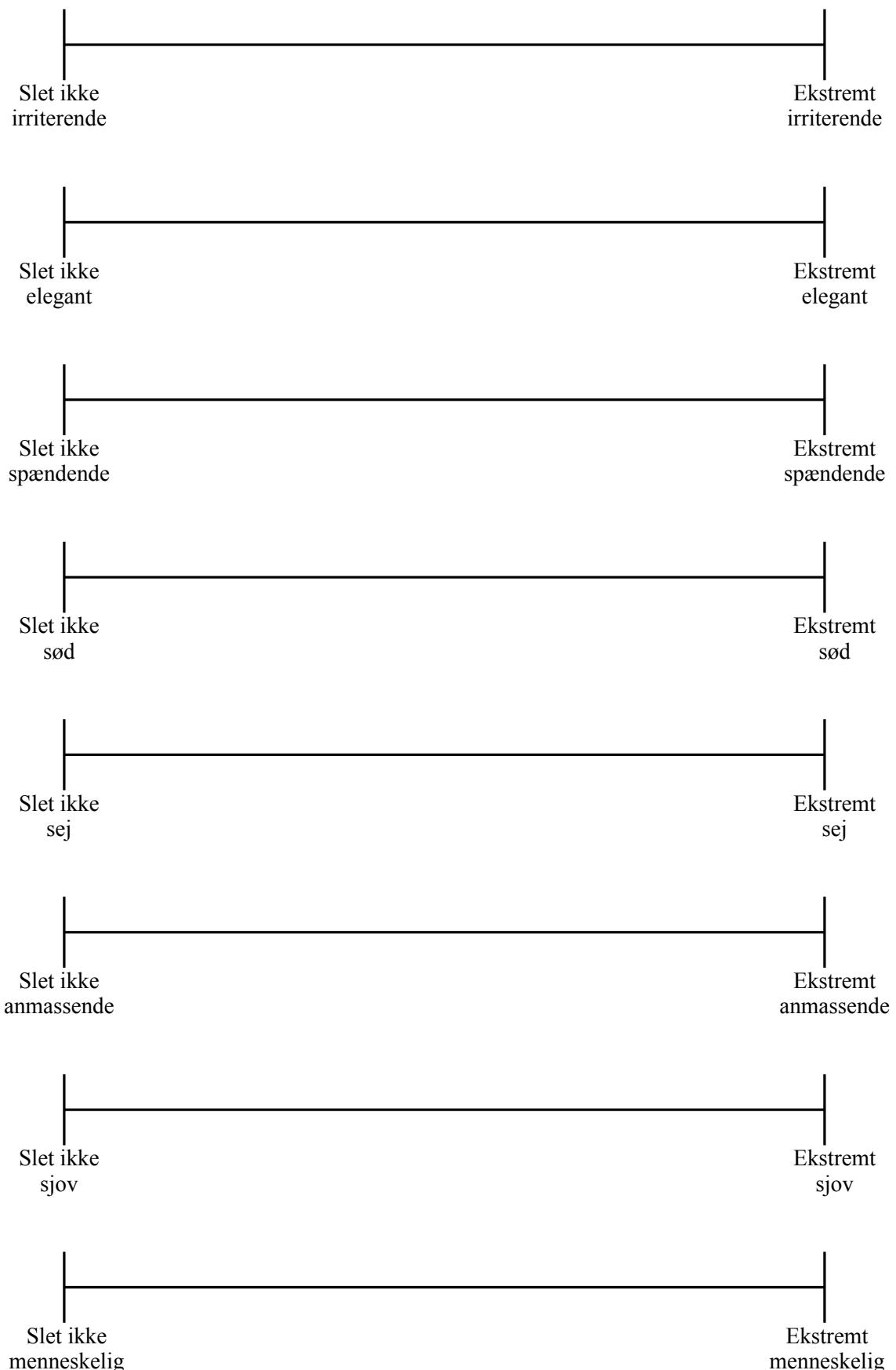


Figur 8.15. Tilpasset skala til: *Hvor overrasket blev du over robottens henvendelse?*

De resterende udvalgte skalaer, som har det til fælles, at deres skalaspørsgsmål er formulertet: *Jeg synes at robotten er* efterfulgt af en parameter, vil alle høre under ét samlet skalaspørsgsmål: *Hvad synes du om robotten?*, jævnfør Figur 8.16. Årsagen er blandt andet, at det for de enkelte skalaspørsgsmål ikke var muligt at fuldende sætningerne med de angivne labels, og derudover vurderes det, at hvis parameteren, som testpersonen skal evaluere, står i skalaspørsgsmålet er der risiko for at der opstår bias, når samme parameter står i begge labels. Grunden til at det overhovedet blev besluttet at sammensætte denne type skalaspørsgsmål til ét skalaspørsgsmål, med otte tilhørende skalaer, var skalaspørsgsmålet: *Jeg synes, at robotten er imødekommen*. Problemet med det skalaspørsgsmål i forhold til de resterende skalaer, vedrørerende testpersonernes mening om robotten, er, at skalaen er bipolær. Med en bipolær skala fra *Ekstremt avisende* til *Ekstremt imødekommen* kan det potentielt medføre, at det kun er den ene halvdel af skalaen, der bliver brugt så længe skalaspørsgsmålet indeholder ordet: *imødekommen*.

Da det er besluttet, at der maksimalt må præsenteres fire skala på skærmen af gangen er det nødvendigt, at fordele de otte skalaer i to. Ligesom tidligere vil to skalaer, der formentlig vedrører det samme emne, blive præsenteret på hver sin side. Med udgangspunkt på i det vil skalaerne med parameterne: *Irriterende* og *anmassende* blive præsenteret på hver sin side, hvor førstnævnte præsenteres på den første af de to sider. Det vælges at parameterne: *Elegant* og *sød* præsenteres på den første side, dog ikke efter hinanden, hvor *sej* og *sjov* præsenteres på den anden side. Grunden til at de fire parametre er fordelt på denne måde er, at det forventes, at særligt *sød*, *sjov* og *sej* formentlig måler det samme, hvor det i mindre grad forventes at *sød* og *elegant* måler det samme. Dertil er skalaen for parameteren *spændende* placeret mellem *elegant* og *sød* for at indikere, at der er forskel mellem de to parametre. Det samme er gældende for *sej* og *sjov*, som er adskilt med *anmassende*. Rækkefølgen på de otte skalaer fremgår af Figur 8.16.

Hvad synes du om robotten?



Figur 8.16. Tilpasset skalaer til: *Hvad synes du om robotten.*

Som tidligere nævnt vil skalaspørgsmålet: *Hvor meget kendskab har du til teknologi/robotter?* ikke blive behandlet på samme måde som de resterende parametre, da det vil indgå som en del af demografi indsamlingen, hvorfor det ikke vil blive beskrevet yderligere i dette afsnit.

8.2 Program til skalaer

Programmet, som blev brugt til at præsentere skalaerne for testpersonerne, blev udviklet i *Processing 3.3.6*, (Fry og Reas 2017). Det bygger på et widget-bibliotek udviklet af Søren Krarup Olesen m.fl. som hedder *PDPGUITest*. I programmet er der lavet klasser til blandt andet check bokse, radio knapper, sliders og lignende GUI elementer, som kan interageres med og som kan gemme data fra besvarelserne.

I *setup* konstrueres objekter fra klasserne: PDPGUI, PDPPushButton og PDPVASknapper. Heri defineres tekst, størrelse og placering af knapper. Det er også her skalaspørgsmål og labels på skalaerne defineres. Der laves også en tabel, med en navngivet kolonne og til hvert datapunkt.

I *draw*-loopet er der syv forskellige skærmbilleder, som præsenteres for testpersonerne. Skærmbillederne dannes ved at loade de widgets ind, som blev konstruerede i setuppet, jævnfør linje 154 og 160-164 på Figur 8.17. For de bipolare skalaer, som har en markering af midtpunktet er der tilføjet en linje oven på den pågældende widget, jævnfør linje 157-159. Det samme gælder i de tilfælde hvor der er et label på midtpunktet.

Hvert skærmbillede loades med variablen *screen*, som starter med at være 1. Når der trykkes på *Næste*, øges *screen* med 1, hvilket gør det næste *if-statement* sandt og dermed skifter til næste skærmbillede. Logikken bag tilbage-knappen er den samme, bortset fra at *screen* mindskes med 1 når der trykkes på *Tilbage*. Den del af koden, som er ansvarlig for navigationen mellem siderne kan fremgå på Figur 8.17, hvor linje 153 og 181 tjekker hvilket skærmbillede, der skal vises, hvor linje 165 og 173 øger *screen* så den går videre til næste skærmbillede og linje 176 og 177 mindske *screen*, så den går tilbage til forrige skærmbillede. *Næste*-knappen gør også at det indtastede data gemmes i datatabellen, hvilket fremgår af linje 166-172. Linje 174 og 178 fjerner de eksisterende widgets fra skærmbilledet, når der skiftes til henholdsvis næste eller forrige skærm.

```

153 if(screen==3){
154     gui.update();
155     progress="Side " + screen + " af 7";
156     text(progress, 1367,height-150);
157     line(1367, 210+150/2,1367,150+150/2);
158     line(1367, 560+150/2,1367,500+150/2);
159     line(1367, 910+150/2,1367,850+150/2);
160     gui.addWidget(vas4);
161     gui.addWidget(vas5);
162     gui.addWidget(vas6);
163     gui.addWidget(pbBack);
164     gui.addWidget(pbNext);
165     if (pbNext.isClicked()) {
166         Results[3]=vas4.getResult();
167         Results[4]=vas5.getResult();
168         Results[5]=vas6.getResult();
169         newRow.setFloat("Q4", Results[3]);
170         newRow.setFloat("Q5", Results[4]);
171         newRow.setFloat("Q6", Results[5]);
172         saveTable (saveResults, filePath + subjectID + ".csv");
173         screen = screen + 1;
174         gui.removeWidget(vas);
175     }
176     if (pbBack.isClicked()) {
177         screen=screen-1;
178         gui.removeWidget(vas);
179     }
180 }
181 if(screen==4){
182     gui.update();
183     progress="Side " + screen + " af 7";
184     text(progress, 1367,height-150);

```

Figur 8.17. Utpluk af koden, som viser hvordan et skærmbillede sættes op og hvordan der navigeres mellem forskellige skærmbilleder.

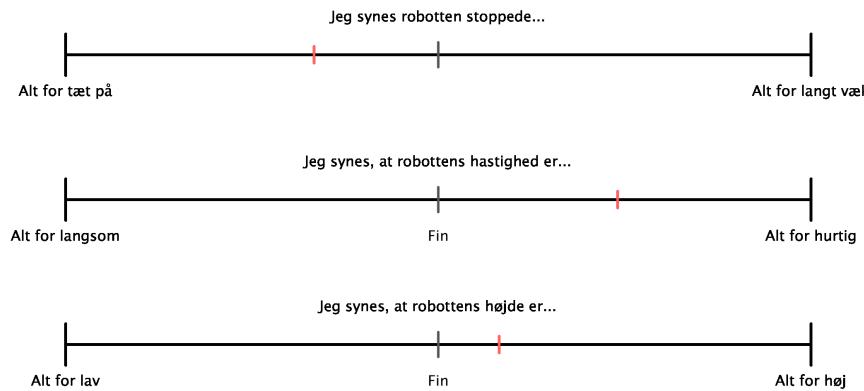
Samme fremgangsmåde er anvendt til alle skærmbilleder med skalaer. Når der trykkes *Næste* på det sidste skærmbillede med skalaer, skiftes der til en side hvor der står: *Tak for hjælpen*, jævnfør linje 304-308 på Figur 8.18.

```

304 if(screen==8){
305     background(255);
306     textSize(80);
307     text("Tak for hjælpen!",1367,912);
308 }
```

Figur 8.18. Utpluk af koden, som viser hvordan det sidste skærmbillede sættes op.

På Figur 8.19 forefindes et eksempel på én af de syv sider med skalaer testpersonerne præsenteres for, hvor de forskellige GUI elementer fremgår.

[Tilbage](#)

Side 2 af 7

[Næste](#)

Figur 8.19. Eksempel på et skærmbillede testpersonerne præsenteres for når de evaluere interaktionen med robotten. De røde markeringer afspejler testpersonen respons.

Programmet bærer kraftigt præg af, at det blev udviklet i sidste øjeblik, grundet tidspres. Det havde derfor nogle fejl og mangler, som med fordel kan udbedres til en anden gang. Herunder opsummeres de problemer som blev opdaget og eventuelle løsningsforslag til, hvordan det kan optimeres.

Det er muligt for testpersonerne at gå videre uden at angive en respons på de præsenterede skalaer. Det kan løses ved at tjekke om data er gemt for skalaerne på det nuværende skærmbillede og bruge det som en betingelse for at *screen* kan øges med 1.

Programmet skriver 0 når der ikke bliver angivet en respons. Ved at løse problemet med at testpersonerne gik videre uden at svare, forsvinder dette problem også, da der aldrig vil være ubesvarede skalaer. Alternativt kan det vælges at skrive *NaN*, hvis ikke testpersonen har angivet sin respons.

Programmet reagerer dårligt på tryk. Problemets skyldes sandsynligvis at programmet er en smule ineffektivt og bruger meget processorkraft på at læse kode, som ikke bruges på det pågældende skærmbillede. For at optimere det, kan der med fordel laves en PDPGUI for hvert skærmbillede, frem for at slette de gamle widgets og loade de nye hver gang. Endvidere burde *while-loops* anvendes for hvert skærmbillede, frem for *if-statements*, så computeren ikke skal gennem hele koden i hver iteration, men kun skal læse den kode, som bruges på det tilstede værende skærmbillede.

Det samlede program fremgår af Afsnit 13.3 (*Program til VAS*).

Testdesign 9

For dels at kunne besvare den anden problemstilling: *Hvordan kan de fundne parametre anvendes til at evaluere HRI?*, og dels for at få erfaring med samt for at evaluere de udviklede skalaer opstilles der en test, hvori skalaerne anvendes til evaluering af robotten samt HRI.

Testdesignet, der er gældende for denne test baseres på testdesignet for den allerede udførte feltundersøgelse, som er præsenteret i Afsnit 5.1 (*Ændringer af testdesign*). Der vil i dette kapitel kun blive præsenteret de elementer af testen, der er forskellig fra feltundersøgelsen. De elementer, der som udgangspunkt ikke ændres fra feltundersøgelsen til denne test er: Type af testpersoner samt rekrutteringen af dem, brugsscenerierne samt testlokation, som er beskrevet i henholdsvis: Afsnit 4.2, Afsnit 4.3 og Afsnit 4.6

9.1 Testens omfang

Med udgangspunkt i og for at besvare problemstillingen: *Hvordan kan de fundne parametre anvendes til at evaluere HRI?*, vil omfanget af denne test primært indebære at lade danske rejsende i Aalborg Lufthavn interagere med den social robot, for efterfølgende at evaluere robotten samt interaktionen med den. Testpersonernes evaluering vil foregå på de nu tilpassede 24 skalaer, der er udviklet på baggrund af den tidligere feltundersøgelse. De 24 skalaer, samt rækkefølgen hvorved de præsenteres fremgår af Afsnit 8.1 (*Tilpasning og rækkefølge på skalaer*). Modsat feltundersøgelsen, hvor der blev indsamlet kvalitativ data, vil der i denne test primært kun blive indsamlet kvantitativ data i form af responsen på skalaerne. Baseret på den indsamlede data vil det blive undersøgt, om der forekommer korrelation mellem parametrene. Fokus for denne test er derfor ikke at undersøge, hvordan og om testpersonerne er i stand til at anvende skalaerne eller om de forstår de angivne labels.

9.2 Brugsscenarier

Den eneste ændring, der er til brugsscenarierne, er når testpersonerne skal vælge destination, hvor antallet af afgange er ændret fra fire til fem, for at tilpasse destinationerne efter de reelle afgange, der er på den pågældende dag, hvor evalueringen udføres. De fem destinationer de rejsende kan vælge imellem fremgår af højre side på Figur 9.1, hvor venstre side gengiver skærmbilledet fra feltundersøgelsen. Undervejs som tiden går vil afgangene blive opdateret, så de passer med oversigtstavlerne.



Figur 9.1. Illustration af ændringerne foretaget fra feltundersøgelsen (venstre) til evalueringen med skalaerne (højre).

Udover ændringen i hvilke destinationer, der kan vælges, er informationen til den specifikke destination ligeledes tilpasset så det svarer til informationerne fundet på Aalborg Lufthavns hjemmeside gældende for den 01/12-2017.

9.3 Robottens bevægelse

Det vælges, at robotstyreren skal variere på robottens højde, afstand til testpersonen samt indgangsvinkelen, hvorved robotten henvender sig til en testperson. At netop de tre parametre er valgt skyldes, at det er muligt at ændre på dem ved brug af *Double*. Formålet med at ændre på de tre fysiske parametre er at kunne undersøge, om de har indflydelse på nogen af de andre parametre.

Den laveste højde robotten kan have er 118 cm, målt fra gulvet til det øverste punkt på rammen hvori iPad'en sidder, hvor robotten maksimalt kan indstilles til en højde på 151 cm, målt fra gulv til top. De to højder vil derfor agere som de to ekstremer: *Alt for lav* og *Alt for høj*. På den bipolare skala, hvor robottens højde evalueres, er midtpunktet navngivet med: *Fin*, som afspejler en højde svarende til omkring albuehøjde, jævnfør Afsnit 5.6 (*Udvælgelse af skalaer*). Denne label er med til at kalibrere skalaen, da der ikke findes en naturlig og logisk højde mellem de to ekstremer, hvorfor det vælges at én af højderne så vidt muligt skal afspejle, hvad testpersonerne forbinder med *Fin*. For at have et kvalificeret estimat af albuehøjden tages der udgangspunkt i den gennemsnitlige højde for danske mænd, som er

181.4 cm, og for danske kvinder, som er 167.2 cm, (Tolstrup 2016). Omregnes de to højder til en gennemsnitshøjde for danskere er højden 174.3 cm. Albuehøjden blev derfor målt ved, at en fra projektgruppen, som er ca. 174.3 cm høj, bukkede albueleddet i en vinkel på ca. 90° og interagerede med robotten i en højde, der føltes behageligt. Robottens højde blev målt til 129 cm når skærmen befandt sig i albuehøjde, hvilket svarer til midtpunktet på skalaen angivet med *Fin*. Denne højde blev bekræftet af flere fra PDP-uddannelsen, som blev spurgt om, hvad de synes om robottens højde, hvortil de svarede: *Fin*.

For at få yderligere variation i robottens højde og for at præge testpersonerne til at bruge mere af skalaen, vælges det er inkludere to ekstra højder. Den ene højde er midt imellem minimum højden og albuehøjden, hvilket giver en højde på 123.5 cm målt fra gulv til top. Den anden højde er midt imellem albuehøjden og maks højden, hvilket giver en højde på 140 cm. Det tilstræbes, at indstille robotten i disse højder, dog kan der forekomme en lille variation i forhold til den eksakte højde. At der kan forekomme en lille variation i højden vurderes at være acceptabelt, da formålet med at ændre højden på robotten er at undersøge, hvilken indflydelse det kan have på andre parametre.

I forhold til robottens afstand til testpersonerne vil tre afstande være gældende: Tæt på, tilpas og langt fra. Afstanden angives ikke i centimeter, da det ikke er muligt at sikre en specifik afstand, når robotten kører rundt i lufthavnen, i hvert fald ikke uden at forstyrre interaktionen mellem testperson og robot. Dog gør følgende retningslinjer sig gældende: Tæt på afspejler en afstand hvor robotten kommer så tæt på testpersonen, at testpersonen vil træde et skridt tilbage for at kunne interagere ordentlig på skærmen. En tilpas afstand afspejler en afstand, hvor testpersonen har mulighed for, at nå skærmen uden nødvendigvis at strække armen helt ud og uden at skulle træde et skridt tilbage. Afstanden langt fra afspejler at testpersonen med strakt arm ikke kan nå skærmen og derfor er nødsaget til at træde et skridt nærmere. Det forventes dog, at det ude i lufthavnen kan blive svært at overholde disse retningslinjer fuldt ud, da der vil forekomme situationer hvor de rejsende selv henvender sig til robotten og på den måde selv bestemmer en afstand.

Den sidste af de tre parametre, der ændres på er: Robottens indgangsvinkel når den henvender sig til en potentiel testperson. Det vil blive forsøgt at få robotten til at henvende sig forfra, fra den rejsendes højre side og fra den rejsendes venstre side. Der sættes ikke krav til, at indgangsvinklen skal være præcis en samme hver gang den kommer fra enten højre eller venstre. Ligesom med afstanden forventes det dog, at det kan blive svært fuldstændigt at kunne kontrollere indgangsvinklen, da det dels afhænger af hvor mange rejsende der er i test området på det pågældende tidspunkt, hvor den rejsende befinner sig i forhold til robotten og dels afhænger af, om den rejsende selv henvender sig til robotten, da det i så fald vil være den rejsende, der dikterer indgangsvinklen.

I Tabel 9.1 fremgår en plan for hvilken afstand robotten skal holde til testpersonen samt hvilken indgangsvinkel den skal henvende sig med når robotten er på sit laveste 118 cm. For at de fem højder alle præsenteres med de tre forskellige afstande og indgangsvinkler, resulterer det i 45 kombinationer. For de fire resterende højder vil rækkefølgen for afstand og indgangsvinkel være magen til den, der fremgår af Tabel 9.1.

Testperson	Højde	Afstand	Indgangsvinkel
1	118 cm	Tæt på	Forfra
2	118 cm	Tilpas	Forfra
3	118 cm	Langt fra	Forfra
4	118 cm	Tæt på	Højre side
5	118 cm	Tilpas	Højre side
6	118 cm	Langt fra	Højre side
7	118 cm	Tæt på	Venstre side
8	118 cm	Tilpas	Venstre side
9	118 cm	Langt fra	Venstre side

Tabel 9.1. Plan for hvilken afstand robotten skal have til testpersonen og hvilken indgangsvinkel den skal have ved henvendelse, når robotten er 118 cm høj. Tilsvarende er gældende for de fire resterende højder: 123.5 cm, 129 cm, 140 cm og 151 cm.

For at opnå de 45 kombinationer af højde, afstand og indgangsvinkel kræver det deltagelse fra minimum 45 testpersoner. Det tilstræbes, at have alle 45 testpersoner på én dag, men da der kun blev udført 17 interviews i feltundersøgelsen, jævnfør Afsnit 10.2 (*Testpersoner*), er det ikke sikkert at det er muligt. Er det ikke muligt, at indsamle data fra minimum 45 testpersoner på én dag vil testen foretages over flere dage til det ønskede antal testpersoner er opnået.

9.4 Rollefordeling

For at udføre testen er det nødvendigt at definere nogle roller, hvortil specifikke opgaver tildeles. Der opstilles i alt tre forskellige roller.

Robotstyrer

Som ved feltundersøgelsen vælges det ligeledes at fjernstyre *Double*-robotten. Robotstyren har til opgave at køre robotten rundt i lufthavnen og få den til at henvende sig til de rejsende i området. Hvordan robotten skal bevæge sig i forhold til afstanden til testpersonen og indgangsvinklen for henvendelse fremgår af Tabel 9.1.

Observatør af interaktion på skærmen

Observatøren skal holde øje med, hvad de rejsende trykker på skærmen. Hvis der trykkes *Næj* skal observatøren indikere til robotstyreren, at robotten skal køre væk. Dette indikeres ved, at observatøren vender tommelfingeren nedad. Hvis den rejsende, der interagerer med robotten, trykker *Ja*, skal observatøren holde øje med, hvornår der står: *Følg venligst efter mig* på skærmen. Dette indikeres ved, at observatøren vender tommelfingeren opad, hvorefter robotstyreren lader robotten køre mod shoppingområdet.

Observatør af robotten

Personen med denne rolle har til opgave at observere og notere, hvordan robotten bevæger sig i forhold til om det stemmer overens med robotstyrerens plan, jævnfør Tabel 9.1. De parametre, der er særlig vigtige at notere er:

- Højde
- Afstand
- Indgangsvinkel

I tilfælde af, at robotten bevæger sig anderledes end hvad den plejer skal dette ligeledes noteres af observatøren. Vurderer observatøren, at der er andre hændelser, der er vigtige at have til senere brug, noteres disse. Formålet med denne observatør er at indsamle data, der kan være relevant for hvordan robotten agerede ved den enkelte testperson i henhold til testpersonens evaluering.

Testleder

Testledererens opgave er at henvende sig til testpersonen, når robotten følger dem mod shoppingområdet, og spørger dem om de har tid til at svare på nogle spørgsmål. Spørgsmålene er skalaerne. Har testpersonen tid til at evaluere sin oplevelse på skalaerne følger testlederen dem hen til bordet, hvor de præsenteres for skalaerne på computeren. Testlederen har i den forbindelse til opgave at notere testpersonnummer i programmet, jævnfør Afsnit 8.2 (*Program til skalaer*), og efterfølgende eksekvere programmet, så testpersonen kan svare på skalaerne. Testlederen har desuden til opgave at observere og svare på spørgsmål vedrørende skalaerne, hvis testpersonen måtte have nogle. Dette vil testlederen sørge for at notere.

Det er vigtigt at testlederen, i tilfælde hvor testpersonen har spørgsmål til skalaerne, ikke forklarer projektgruppens forståelse af spørgsmålet, men i stedet lader det være op til den enkelte testpersonen at danne sin egen forståelse af spørgsmålet. Det kan eksempelvis gøres ved, at testlederen siger: *Det er ud fra din forståelse af det eller Hvad synes du det betyder?*. Efter testpersonen har besvaret alle skalaer og udfyldt demografien, takker testlederen testpersonen for sin deltagelse og ønsker dem en god rejse.

9.5 Demografi

Testpersonerne vil blive bedt om at udfylde demografi, som indeholder køn, alder, højde, hvor ofte testpersonen rejser og hvor meget kendskab de har til teknologi. Demografien fremgår af Figur 9.2.

Dit køn: Mand/dreng
 Kvinde/pige

Din alder: _____ år

Din højde: _____ cm

Hvor ofte rejser du i gennemsnit?

- Mindre end 1 gang om året
- 1-3 gange om året
- 4-11 gange om året
- 1 gang om måneden
- Flere gange om måneden

Hvad er din kendskab til teknologi?

- Ingen kendskab
- Lidt kendskab
- Nogen kendskab
- Meget kendskab
- Ekstremt meget kendskab

Figur 9.2. Demografi, som testpersonen skal udfylde efter skalaerne er besvaret.

Kønnet kan enten angives som *Mand/dreng* eller *Kvinde/pige* årsagen til, at dreng og pige indgår i demografien er, at i feltundersøgelsen var en af testpersonerne kun 8 år, hvorfor det forventes at have børn som testpersoner igen. Derudover er det ikke sikkert, at et barn identifierer sig som mand eller kvinde, hvorfor dreng og pige er noteret. Alder noteres for at vide hvor stort et spænd testpersonerne dækker over, da det tidligere er fundet, at der kan være aldersforskelle i forhold til interaktionen med sociale robotter. Højden noteres i centimeter for, at have sammenholde det med robottens højde i forhold til evalueringen, da det kan have indflydelse på nogle af parametrene.

I den foregående feltundersøgelse blev testpersonerne spurgt, hvor ofte de rejser. Det vælges at få testpersonerne til at notere, hvor ofte de rejser istedet for at spørger dem, da deres respons desværre var svær at fortolke og kategorisere. Nogen af testpersonerne svarede nemlig, at de rejser ofte, men det er uvist hvad ofte egentligt er.

Som tidligere nævnt vil spørgsmålet omkring testpersonernes kendskab til teknologi/robotter blive stillet i forbindelse med demografien. I det henseende er det valgt at formulere spørgsmålet som: *Hvad er din kendskab til teknologi* og istedet for en skala, besvares spørgsmålet på en *Likert-scale*, hvor graden af kendskab fremgår i fem niveauer. At

teknologi vælges frem for robotter skyldes, at teknologi er mere dækkende for har testpersonerne kendskab til robotter, så har de også kendskab til teknologi, men har de kendskab til teknologi er det ikke ens betydende med, at de har kendskab til robotter.

9.6 Testlokation og udstyr

De ændringer, der er til testlokation og udstyr i forhold til hvordan det er beskrevet i Afsnit 4.6 (*Testlokation og udstyr*), henvender sig blandt andet til hvad der var muligt at gøre i lufthavnen, og er beskrevet i Afsnit 5.1 (*Ændringer af testdesign*). Udover de ændringer, vil der blive anvendt en computer hvorpå skalaerne præsenteres, computeren er en: *Microsoft Surface Pro (5)*, hvis skærm er 12.3" med en oplosning på 2736 x 1824, dertil vil der være en trådløs mus til rådighed, som testpersonerne skal bruge til at markere det sted på skalaen, der er deres respons. I modsætning til feltundersøgelsen, hvor demografien indsamlet ved at stille testpersonerne spørgsmål, er der til denne test blevet udviklet et ark hvorpå testpersonerne selv kan notere informationerne.

Derudover vil testlederen ikke optage hvad testpersonen siger, da der som udgangspunkt ikke blive stilt spørgsmål. Stilles der spørgsmål vil det blive noteret af testlederen på papir.

9.7 Fremgangsmåde

Så snart en af de rejsende befinner sig i det aftalte område, vil robotten henvende sig ved at køre hen til den rejsende og spørge om den kan hjælpe dem med at finde rundt i Aalborg Lufthavn. Svarer testpersonen *Ja* vil der efterfølgende ganske kort stå, på skærmen, hvad formålet med testen er: At undersøge menneskers interaktion med robotter. Derefter har testpersonen mulighed for selv, at vælge et af de fire brugsscenerier. Uanset hvilket brugsscenario, der bliver valgt vil det ende med, at robotten opfordrer testpersonen til at følge efter. Kort efter testpersonen er begyndt at følge efter robotten vil testlederen stoppe dem, og komme med en kort introduktion, der lyder som følgende:

Hej, vi kommer fra Aalborg universitet. Vi er i gang med at undersøge menneskers interaktion med robotter. Har du fem minutter til at svare på nogle spørgsmål i forhold til robotten? Det foregår på computeren herovre. Hvis du har nogle spørgsmål undervejs, så bare sig til.

Efter introduktion viser testlederen testpersonen hen til bordet med computeren, hvorpå testpersonerne skal evaluere sin oplevelse på de udvalgte skalaer. Imens testpersonen svarer på skalaerne vil testlederen foretage sig noget andet, eksempelvis notere observationer, for at det ikke føles som om at testlederen holde øje med testpersonens besvarelser. Opstår der problemer vil testlederen selvfølgelig træde ind og forsøge at læse problemet. Når testpersonen er færdig med at besvare skalaerne på computeren vil de blive bedt om at udfylde noget demografi på et papir, som fremgår af Afsnit 9.5.

Afhængigt af hvor tidspresset testpersonen er vil testlederen forsøge at stille følgende spørgsmål:

- Var der noget, som du synes at der manglede ved vurderingen af robotten?
- Hvordan var det at besvare skalaerne?

Efterfølgende afsluttes testen ved at testlederen siger tak for testpersonens deltagelse og ønske testpersonen en god rejse.

9.8 Pilottest

Da der er foretaget ændringer i testdesignet i forhold til det testdesign der blev brugt i feltundersøgelsen vil der udføres nogle pilottests. Alle pilottest er udført på Aalborg Universitet Fredrik Bajers vej 7B i projektgruppens grupperum. Til hver pilottest er der både en robotstyrer, en observatør som holder øje med hvad testpersonerne trykker på, en testleder, som præsenterer skalaerne samt en observatør, der sørger for at notere pilottestens forløb.

Pilottest 1

Den første pilottest blev udført af en mand på 25 år, som læser Produkt- og Designpsykologi. Da programmet til præsentation af skalaer endnu ikke var færdigudviklet blev skalaerne præsenteret i udprintet form på A4 papir. Skalaerne blev præsenteret efter samme rækkefølge som hvis de blev præsenteret på computeren.

Robotten henvendte sig til testpersonen ude på gangen foran grupperummet og kørte tæt på testpersonen, som kommenterede at robotten kom lidt for tæt på, hvorfor testpersonen rykkede et skridt tilbage. Robottens højde var indstillet til albuehøjde, svarende til 129 cm.

Testlederen introducerede skalaerne med et meget formelt ordvalg, som for testpersoner, der ikke har samme faglige baggrund som projektgruppen, kan have svært ved at forstå. Derudover undskyldte testlederen flere gange ved at sige: *Jeg er ked af at sige det*, hvilket giver en lidt mærkelige stemning. Der skal ikke undskyldes for hvorfor testen er bygget op som den er, det virker useriøst og som om at der er problemer med hvordan testen afvikles. Derudover blev det bemærket at testlederen havde en tendens til at stå bag testpersonen, når testpersonen svarede på skalaerne, hvilket ikke hensigtsmæssigt, da det kan virke som om at testlederen overvåger testpersonens besvarelser.

Testpersonen var i tvivl om hvordan skalaspørgsmålet: *Hvordan synes du, at skærmen på robotten reagerede?* skulle forstås. Testpersonen var i tvivl om hvorvidt spørgsmålet handlede om hvordan det var at trykke på skærmen eller om det handlede om hvorvidt valgmulighederne var tilfredstillende. Årsagen til at tvivlen kan være opstået er formentlig, at testpersonen ikke oplevede problemer med at interagere på robottens skærm. Det vælges ikke, at omformulere spørgsmålet da det forventes, at en del danske rejsende vil opleve problemer med skærmen når testen afvikles i Aalborg Lufthavn.

En anden ting testpersonen var i tvivl om var hvad ordet *anmassende* betød. Testlederen fortæller, at det er testpersonens forståelse af ordet *anmassende*, der skal evalueres ud fra, hvorefter testpersonen begynder, at forklare ordets betydning. Baseret på den forklaring er det klart, at testpersonen har forstået hvad der menes med *anmassende*, nemlig at robotten kommer meget tæt på, og presse på ved at køre lidt ind i testpersonen. Dog forklarer testpersonen, at hvis robotten blev ved med at være tæt på ville det føles *anmassende*, men da robotten efterfølgende holdte afstand føltes det ikke *anmassende*.

Derudover kommenterer testpersonen på ordet *sjov* i forhold til det er meget subjektivt

hvor der forståes ved ordet *sjovt*. Testpersonen selv synes ikke, at robotten var *sjov*, hvor det virker som om den formulering afspejler noget, der får en til at grine.

Pilottest 2

Den anden pilottest blev udført af gruppens kvindelige vejleder. Da programmet endnu ikke er færdigudviklet og der ikke var tid til at printe et nyt sæt skalaer, foreslog vejlederen, at de udprintede skalaer fra den forrige pilottest bare kunne bruges igen.

Testpersonen har mange problemer med skærmen i begyndelsen af interaktionen, hvor den ikke reagerer på tryk. Det bør derfor overvejes om der bør tilføjes noget tekst på skærmen, hvorpå der står, at der skal trykkes på en bestemt måde. Ingen undskylder testlederen flere gange overfor testpersonen, hvilket ikke er hensigtsmæssigt og testlederen bruger stadig et meget formelt sprog. Et forslag til hvordan testlederen starter samtal'en med testpersonen er, at testlederen siger noget i stil med: *Du har lige snakket med den her robot, og jeg vil gerne vide lidt om din oplevelse*, på den måde undgåes det formelle ordvalg. Derudover bliver testpersonen fortalt, at der er 24 skalaer de skal svare på. Dette bør undgåes da 24 lyder af mange og det kan være svært at forholde sig til og så er det heller ikke meningen at testpersonen skal holde styr på hvor mange skalaer, der er besvaret og hvor mange der er igen.

Testpersonen er i tvivl om hvad der menes med ordet: *personlig* i forhold til skalaspørgsmålet: *Hvor personlig oplevede du robottens hjælp?*, om det er i forhold til testpersonen selv. Testlederen forklarer, at det er hvad testpersonen forstår ved ordet, der skal svares ud fra.

De otte skalaer, som hører til skalaspørgsmålet: *Hvad synes du om robotten?*, er ligeligt fordelt på to papirer, ligesom det på computeren vil være fordelt på to skærbilleder, hvorfor skalaspørgsmålet vil være på begge sider. Dette kommenterede testpersonen var forvirrende, da det kunne opleves som om, at der var opstået en fejl, så skalaerne skulle besvares igen. Testpersonen foreslår at lave en anden overskrift til de fire sidste skalaer, hvor overskrift henviser til skalaspørgsmål. Derudover kommenterer testpersonen at det er abstrakt og svært at forholde sig til spørgsmålet: *Hvad er din kendskab til teknologi?*, hvorfor dette bør revurderes. Efter en dialog med projektgruppen bliver det foreslået, at ændre spørgsmålet til: *Hvor glad er du for teknologi?*, da det egentlig er det, projektgruppen gerne vil vide. Der bliver også diskuteret om der skal spørges ind til om testpersonerne har let ved at bruge sin smartphone eller sin computer, og i det henseende hvornår de sidst købte en ny.

Pilottest 3

Tredje pilottest blev udført af en mand på 25 år, som læser Produkt- og Designpsykologi, hvor skalaerne besvares på computeren, igennem det udviklede program.

Igen nævner testlederen hvor mange skalaer, der er og i tillæg hvor langtid det kommer til at tage. Problemet med at nævne tiden i pilottesten er, at det er første gang programmet bliver brugt, så det vides ikke hvor lang tid det kan komme til at tage, og derudover afhænger det af, hvor hurtig testpersonen er til at besvare skalaerne. Testlederen står bag testpersonen hvilket for testpersonen kan virke overvågende, årsagen er dog, at programmet

ikke er testet før og for at sikre, at programmet fungerer står testlederen bag ved. Derudover oplever testpersonen problemer med at respondere på skalaerne på computeren.

9.9 Ændringer efter pilottest

Baseret på de tre pilottests vil der foretages ændringer til testen. Testlederen skal sørge for ikke, at bruge formelt sprog, ikke nævne antallet af skalaer eller undskynde, at testen er som den er. Så vidt muligt skal testlederen følge instruktionerne beskrevet i Afsnit 9.7 (*Fremgangsmåde*).

Da der både i feltundersøgelsen og i de før beskrevne pilottest gentagende gange er oplevet problemer med robottens skærm, vælges det er at tilføje en tekst på det første skærmbillede, hvor der står: *Tryk blidt på mig*, da det er fundet, at hvis der trykkes meget blidt på skærmen så reagerer den. På Figur 9.3 fremgår ændringen på det første skærmbillede, hvor teksten er tilføjet.



Figur 9.3. Ændringen til den første side, der præsenteres på robotten, hvor der tilføjet den grå tekst: *Tryk blidt på mig*.

Baseret på pilottest 2, hvor det blev besluttet at ændre spørgmålet: *Hvad er din kendskab til teknologi?* til: *Hvor glad er du for teknologi?*, i demografien. I det henseende vælges det at præsentere spørgsmålet som et skalaspørgsmål, hvor testpersonerne angiver hvor glade de er for teknologi på en unipolær VAS med lukkede endepunkter fra: *Slet ikke glad til Ekstremt glad*. Før ændringen skulle testpersonerne både angive hvor ofte de gennemsnitligt rejse og hvad deres kendskab til teknologi er på en Likert-skala med fem valgmuligheder. Da spørgsmålet omkring teknologi er ændret til et skalaspørgsmål med tilhørende skala, vælges det ligeledes at ændre både formuleringen og måden hvorpå *Hvor ofte rejser du i gennemsnit?* præsenteres. Spørgsmålet omformuleres til: *Angiv hvor mange gange du ca. flyver på et år*, hvilket skal besvares på samme måde som alder og højde. Ændringerne til demografi fremgår på Figur 9.4.

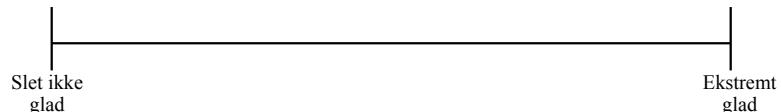
Dit køn: Mand/dreng
 Kvinde/pige

Din alder: _____ år

Din højde: _____ cm

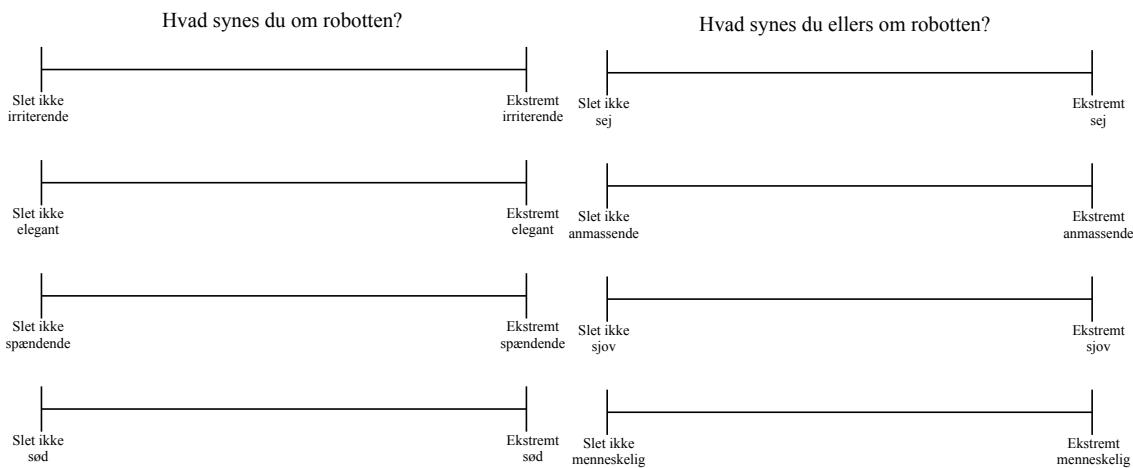
Angiv hvor mange gange du ca. flyver på et år: _____ gange
 (i hele tal)

Hvor glad er du for teknologi?



Figur 9.4. Ændringer til demografi.

Den sidste ændring, der er foretaget på baggrund af pilottestene er, at ændre skalaspørgsmålet: *Hvad synes du om robotten?*, på side to, så der ikke opstår forvirring. Skalaspørgsmålet på side to vil fremover være: *Hvad synes du ellers om robotten?*. Første side fremgår på Figur 9.5, hvor anden side fremgår på Figur 9.6.



Figur 9.5. Første side med de otte skalaer.

Figur 9.6. Anden side med de otte skalaer.

Databehandling 10

I følgende kapitel vil den indsamlede data blive behandlet og analyseret med forskellige metoder samt efter forskellige emner. Det samlede data forefindes i Afsnit 13.5 (*Rådata i Excel*). Det skal dog pointeres, at fordi denne undersøgelse er eksplorativ vil databehandlingen ligeledes være det. Ligesom ved den foregående feltundersøgelse, vil ændringer af testdesign blive beskrevet først. Dernæst hvilke testpersoner, der har deltaget i undersøgelsen, som er efterfulgt af de observationer, der er foretaget i AAL. Inden der foretages en decideret behandling og analyse af testpersonernes besvarelser, vil der blive beskrevet en frasortering af manglende datapunkter, som afspejler manglende besvarelser. Derefter vil der blive fokuseret på fordelingen af besvarelserne. For at undersøge hvordan de varierede fysiske parametre har indflydelse på testpersonernes respons og for at undersøge hvorvidt der forekommer korrelation mellem to eller flere parametre bliver der udført en *Principal Component Analysis* (PCA), først generelt for hele datasættet og derefter i forhold til de tre fysiske parametre: Højde, afstand og indgangsvinkel. Når PCA udføres kan det være en fordel, at have det tre dimensionelle *Bi*-plot ved siden af, der forefindes 3D *Bi*-plots for det samlede datasæt, højde og indgangsvinkel i Afsnit 13.8 (*3D Bi-plots*). Der er ikke et 3D *Bi*-plot for afstand, da der kun er to *Principal Components* (PCs). Derudover kan det være en god idé, at have oversigten over skalaspørgsmålene og skalaerne ved siden af, da disse vil blive noteret med SQ efterfulgt af nummer, oversigten forefindes i Afsnit 13.4 (*Skala Oversigt*). Derefter fokuseres der på, at sammenligne de korrelerede parametre, hvorefter der fokuseres på hvilken indflydelse testpersonernes demografi har.

10.1 Ændringer af testdesign

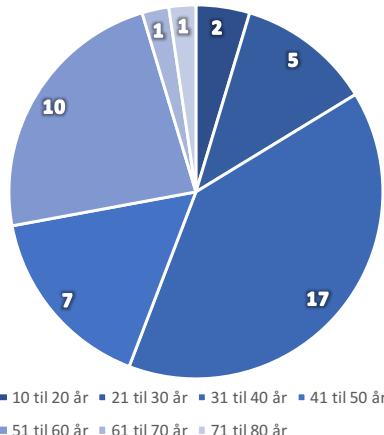
Efter at have afviklet testen i AAL den 01-12-2017 var der indsamlet data fra 22 testpersoner. Da det tilstræbes at have omkring 45 testpersoner blev det derfor besluttet at tage ud i lufthavnen og indsamle mere data den 05-12-2017. Information om afgangene blev derfor opdateret, så de passede til den tilsvarende dag testen blev kørt.

Ydermere blev det under testen observeret, at det var svært at styre hvor tæt på og fra hvilken vinkel robotten henvendte sig. Selvom det blev prøvet at følge skemaet, var der mange rejsende, som undveg robotten, når den kom for tæt på eller generelt ikke virkede til at have lyst til at interagere med den. Det blev derfor besluttet at prøve at variere de forskellige typer henvendelser, men uden at følge et bestemt skema. Da de rejsende også generelt virkede til at være mindre interesserede i robotten på de to testdage sammenlignet med de rejsende i feltundersøgelsen, blev det besluttet at det var vigtigere at få nogle testpersoner frem for at følge et bestemt skema. Højden på robotten blev i starten varieret efter skemaet, men som der blev mangel på testpersoner blev det besluttet at variere højden oftere, så alle højder blev testet.

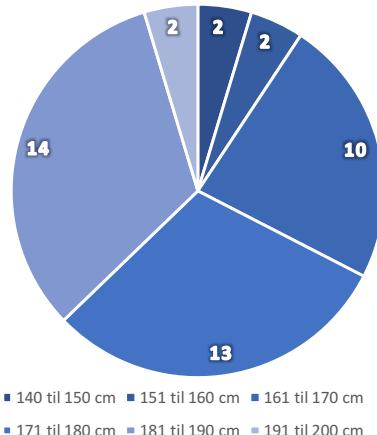
I lufthavnen var der en testperson, som gerne ville vide hvor toilettet var, men ikke ville følges derhen. Vedkommende virkede lidt forvirret da robotten efterfølgende spurgte hvad den så kunne hjælpe med, hvorefter vedkommende gik. Det valgtes derfor, at hvis der trykkes nej til gældende spørgsmål, så skifter skærmbilledet til: "Hav en god rejse" og robotten spørger derfor ikke hvad den skal hjælpe med endnu en gang.

10.2 Testpersoner

Der er i alt 43 testpersoner, som har interageret med robotten for derefter at svare på de forskellige skalaer. Testpersonerne udgør 16 kvinder og 27 mænd i et aldersspænd fra 10 til 72 år, ($M=40.1$, $SD=13.4$). En enkel testperson, TP2, nåede ikke at udfylde sin alder, hvorfor alderen er estimeret til at være 40 år. Observationer gjort for de enkelte testpersoner er beskrevet i Afsnit 10.3 (*Observationer i AAL*). Testpersonerne rejser mellem 1 og 100 gange om året ($M=15.3$, $SD=18.1$). Dette er udregnet efter testpersonernes nedskrevede besvarelser, hvor nogle testpersoner har svaret et interval, eksempelvis 1-2 gange om måneden eller 6-8 gange om året. Ved svar som 1-2 gange om måneden er det omregnet til antal gange om året, hvorefter det midterste tal er medtaget. Ved svar som 6-8 gange om året er det medtaget at testpersonen rejser 7 gange om året. Hvis testpersonen eksempelvis har svaret 2-3 gange om året er det højeste antal rejser medtaget.



Figur 10.1. Testpersonernes aldersfordeling, angivet i antal.



Figur 10.2. Testpersonernes højdefordeling, angivet i antal.

Aldersfordelingen fremgår af Figur 10.1, hvor det fremgår at størstedelen af testpersonerne har været mellem 31 år og 40 år, efterfulgt af 51 år til 60 år, med henholdsvis 17 og 10 testpersoner i hver aldersgruppe. Højdefordelingen fremgår af Figur 10.2, hvor det fremgår, at størstedelen af testpersonernes højde er mellem 161 cm til 190 cm. Mændene havde en gennemsnitshøjde på 182.1 cm ($SD=6.1$) og kvinderne havde en gennemsnitshøjde på 164.9 cm ($SD=10.8$).

10.3 Observationer i AAL

Projektgruppen var i lufthavnen fredag d. 01-12-2017 fra 07:00-15:15 og tirsdag d. 05-12-2017 fra 08:00-15:30. Det virkede meget sværere at få de rejsende til at interagere med

robotten, da de virkede til at have travlt. I løbet af de to dage blev højden på robotten, afstanden til de rejsende samt indgangsvinklen varieret for at få testpersonerne til at bruge mere af skalaerne. I de følgende afsnit vil observationer fra de to test dage blive beskrevet.

Problemet med at skærmen på *Double* reagerede dårligt blev ikke fuldstændig løst af at tilføje teksten: *Tryk blidt på mig* i wireframed. Der var stadig problemer med at skærmen ikke reagerede på testpersonernes berøring. Derudover var der også problemer med det udviklede skalaprogram. Ofte reagerede det ikke på første klik, hvorfor der skulle dobbeltklikkes. I nogle tilfælde skulle programmet genstartes, hvilket resulterede i at testpersonens besvarelser blev overskrevet, så de skulle angive en ny respons. Det blev observeret, at den nye respons i nogle tilfælde var meget anderledes end den oprindelige indskydelse.

10.3.1 Observationer fra d. 01-12-2017

Det virkede ikke som om, at robotten var spændende nok til at interagere med, når den kørte ret frem og tilbage. Med andre ord virkede det bedre, at styre robotten, så den kom skræt fra siden. Selvom mange rejsende virkede til at have travlt, eksisterede der stadig en generel interesse for robotten. Mange smilte, pegede og virkede overraskede selvom de ikke interagerede med robotten. Der var også mange, der virkede interesserede, men efter at have læst på skærmen gik videre eller trykkede nej til at robotten måtte hjælpe dem med at finde rundt. Årsagen til det kan være formuleringen af teksten. På skærmen står der om robotten må hjælpe med at finde rundt i lufthavnen. Aalborg Lufthavn er ret lille sammenlignet med mange andre lufthavne og det er derfor ikke svært at finde rundt, hvorfor det kan være årsagen til at de rejsende ikke har lyst til at deltage. Der var en dansk rejsende, der forsøgte at anvende gestikker til at stoppe robotten, ved at vifte hånden foran skærmen.

Der var en del, der stoppede og tog billeder eller optog video med deres telefoner. Flere lavede gestikker som om robotten var en hund, og klappede sig på lårerne og piftede for at få den til at følge efter. Vi mistede mange testpersoner igen som en konsekvens af, at *Double* skærmen reagerede dårligt. Hvis ikke skærmen reagerer ved første tryk, trykkes der hårdere og hurtigere, hvilket skærmen slet ikke reagerer på. Ofte tager de rejsende fat med fingrene bagom skærmen og trykker med tommelfingeren, muligvis for at kunne trykke hårdt og præcist samtidigt med at stabilisere robottens bevægelser.

Testpersoner (TP)

- **TP01** blev forstyrret under sin besvarelse af skalaerne, da programmet reagerede meget dårligt. Efter tre skalabesvarelser, blev programmet genstartet. Programmet reagerede efterfølgende hurtigere, men det blev observeret, at testpersonen svarede markant anderledes end sin oprindelige indskydelse på de første tre skalaer.
- **TP02** misforstod skalaens midtpunkt og antog, at det var en slider.
- **TP03** nævnte at skærmen på robotten reagerede bedre end skalaprogrammet.
- **TP04** besvare skalaerne sammen med sin kone. De konfererer omkring svarene, selvom det kun var manden, der trykkede på skærmen.
- **TP06** svarede på skalaerne, men det var både hende og sønnen der interagerede med robotten.

- **TP07** svarer på skalaerne, men det er sønnen der har interageret med robotten med hjælp fra faderen.
- **TP08** snakker først med robotstyrer og observatør om robotten, hvorefter han spørger om han må interagere med den.
- **TP09** bliver opfordret til at interagere med robotten.
- **TP10** er et ældre ægtepar, der trykkede for hårdt på skærmen, og blev afbrudt af en observatør, der assisterede dem igennem deres interaktion med robotten. Efterfølgende nævnte de også, at de ville have opgivet hurtigt, hvis ikke de havde fået hjælp. Kvinden nægtede at menneskeliggøre robotten ved at kalde den sød, som hun forbandt med en menneskelig egenskab.
- **TP11** interagerede først med robotten, men gik. Testpersonen kom dog tilbage omkring 5 minutter efter, hvorfor testpersonen ikke lige havde interageret med robotten, da skalaerne blev besvaret. Halvvejs inde i besvarelsen opdagede testpersonen, at robotten ikke kører af sig selv, hvilket måske kan have påvirket besvarelserne.
- **TP12** var muligvis ansat i lufthavnen, da testpersonen gik i civil sammen med personale.
- **TP13** er en far og søn, der begge interagerer med robotten. Det er kun faderen der svarer på skalaerne.
- **TP14** fik hjælp til at trykke på robottens skærm, da testpersonen trykkede for hårdt og holdte fingeren inde på skærmen. Testpersonen blev ikke fulgt afsted af robotten, da testlederen stod lige ved siden af. Testpersonen spørger ind til, hvad der menes med hastighed.
- **TP15** regnede med at kunne bruge midtpunktet som slider. Dobbeltklikkede på en skala under midtpunktet, da den ikke reagerede, flyttede musen over midten og satte sin markør der, før testpersonen gik videre.
- **TP16** forstod ikke skalaen, der handler om hvor langt fra/tæt på robotten stoppede. Udfylder skalaerne med sin kammerat ind til side 4 af 7. Derefter besvares skalaerne af en enkelt testperson.
- **TP17** havde problemer med skalaprogrammet fordi det reagerede dårligt. Spørger ind til midtpunktets label: *Fin* og om det også var fin alle andre steder. Testlederen forklarer, at det bygger på andre danske rejsendes udtalelser, de steder det er angivet.
- **TP18** havde ikke lige interageret med robotten da skalaerne blev besvaret, da testpersonen skulle på toilet først.
- **TP19** ventede et par minutter på at besvare skalaerne fordi TP18 stadig var igang. Kommenterede, at robotten var meget hyggelig; lidt som en hund.
- **TP20** interagerede med robotten, som havde en højde på 151 cm.
- **TP21** går først forbi robotten men vender sig om til sin mand og de interagerer sammen samt besvarer skalaerne sammen. De misforstår skalaspørgsmålet, der spørger ind til hvor langt fra/tæt på robotten stoppede. De troede det betød om robotten stoppede for langt fra/for tæt det sted de havde valgt den skulle følge dem hen til.
- **TP22** et par, der besvarer skalaerne sammen. De grinte rigtigt meget, da de fulgte efter robotten.

Der er ikke foretaget nogle særlige observationer for de testpersoner, som ikke er listet i ovenstående. Efter TP22 var der lang tid mellem diverse afgange, hvorfor det blev valgt at stoppe testen og tage ud i lufthavnen en anden dag for at få flere testpersoner.

10.3.2 Observationer fra d. 05-12-2017

Der var få ikke-danskthalende rejsende, der interagerede med robotten. En enkelt britte kommenterede, at robotten var ret imponerende og at han elskede den. En pige på 12 år virkede oprigtigt interesseret i robotten og ville interagere. Hendes mor sagde, at hun ikke kunne lide sådan noget og trykkede mange gange og ivrigt på nej. Igen var skærmen ofte et problem på både *Double* og computeren hvorpå skalaerne præsenteres.

Testpersoner (TP)

- **TP23** var en far og en mindre dreng. Faderen forsøgte først at trykke på skærmen, men trykkede ofte for hårdt. Han løftede drengen op (robotten var 151 cm) så drengen kunne trykke, hvorefter skærmen reagerede. Kommenterede at skalaprogrammet heller ikke virkede så godt. Kommenterer, at han ikke lagde mærke til robotten, før den pludselig var der og påpeger at den ikke skal være irriterende som en gadesælger.
- **TP24** var en familie på fire. Det var en pige, der trykkede på skærmen og udfyldte skalaerne med hjælp fra sin familie.
- **TP25** Både skærmen på robotten og skalaprogrammet reagerede rigtigt dårligt.
- **TP26** kommenterer: Hvis [Double] var pyntet op til jul, så var den ekstremt imødekommen.
- **TP27** blev ikke fulgt afsted. Testpersonen kom til at minimere *Double* applikationen, så iPad'ens baggrund viste sig.
- **TP28** var sammen med en kammerat. Testpersonen tastede ikke selv på *Double*. Testpersonen ændrer sin placering på skalaen, når den ikke reagerer første gang.
- **TP29** studerer Innovationsdesign på AAU.
- **TP31** var fire personer i alt. Testpersonen brugte meget sine rejsemakkere til at besvare skalaerne. Testpersonen fik forklaret hvad anmassende betød af sin rejsemakker. Testpersonen talte flydende dansk, men det var ikke hendes modersmål, hvorfor dette accepteres.
- **TP33** mener at robotten er for langsom.
- **TP34** var nordmand, men forstod udmærket dansk, hvorfor dette accepteres.
- **TP38** det hele gik uden problemer. Robottens skærm og skalaprogrammet reagerede hver gang, der blev trykket.
- **TP39** henvendte sig selv til robotten. Testpersonen reflekterede over, hvad der menes med hvor mange gange testpersonen flyver; er det per tur eller fysisk fly? Ender med med at tælle alle fly med som testpersonen satte sig ind i i løbet af et år.
- **TP40** reflekterede over ordet *sjov*. De synes robotten var sjov, men den fortalte ingen vittigheder, så den var ikke *sjov-sjov*.
- **TP41** mente det ville være nemmere at kunne tale til robotten. Brugte *Find whiskey* som et eksempel.
- **TP42** blev forvirret, da gateinformation om København ikke var blevet opdateret. Testpersonen forsøgte at swipe, da testpersonen bemærkede at tiden ikke passede og derefter trykkede på Aalborg Lufthavn logoet i bunden (måske for at komme tilbage eller opdatere siden.)

- **TP43** interagerede med robotten et andet sted end alle andre testpersoner, ved et pakkebord under informationsskærmen

Igen er der testpersoner, hvor der ikke er foretaget nogle særlige observationer hvorfor de ikke fremgår af ovenstående. Selvom det blev tilstræbet at følge skemaet over de planlagte højder, afstande og indgangsvinkler var det ikke muligt, at overholde det, hvorfor det ligeledes accepteres at der ikke er 45 testpersoner.

10.4 Skala Oversigt

<u>Side</u>	<u>SQ#</u>	<u>Scale question</u>	<u>Skala labels</u>
1	1	Hvordan synes du skærmen reagerede?	"Ekstremt dårligt", "Ekstremt godt"
	2	"Hvordan oplevede du robotten?"	"Ekstremt afvisende", "Ekstremt imødekommen"
	3	"Hvordan var det at bruge robotten?"	"Ekstremt svært", "Ekstremt nemt"
	4	"Hvordan oplevede du robottens bevægelser?"	"Ekstremt rolige", "Ekstremt vilde"
2	5	"Jeg synes robotten stoppede..."	"Alt for tæt på", "Alt for langt væk"
	6	"Jeg synes, at robottens hastighed er..."	"Alt for langsom", "Alt for hurtig"
	7	"Jeg synes, at robottens højde er..."	"Alt for lav", "Alt for høj"
3	8	"Jeg føler, at robotten kan hjælpe mig"	"Helt uenig", "Helt enig"
	9	"Jeg synes, at robotten stod i vejen"	"Helt uenig", "Helt enig"
	10	"Jeg føler mig tryg ved robotten"	"Helt uenig", "Helt enig"
4	11	"Robotten gjorde mig forskräkket"	"Helt uenig", "Helt enig"
	12	"Jeg kan godt lide at blive betjent af robotten"	"Helt uenig", "Helt enig"
	13	"Jeg regnede med, at robotten fulgte mig hen sted jeg valgte"	"Helt uenig", "Helt enig"
5	14	"Hvor personlig oplevede du robottens hjælp?"	"Slet ikke personlig", "Ekstremt personlig"
	15	"Hvor overrasket blev du over robottens hen vendelse?"	"Slet ikke overrasket", "Ekstremt overrasket"
6	16	"Hvad synes du om robotten?"	"Slet ikke irriterende", "Ekstremt irriterende"
	17		"Slet ikke elegant", "Ekstremt elegant"
	18		"Slet ikke spændende", "Ekstremt spændende"
	19		"Slet ikke sød", "Ekstremt sød"
7	20	"Hvad synes du ellers om robotten?"	"Slet ikke dej", "Ekstremt dej"
	21		"Slet ikke anmassende", "Ekstremt anmassende"
	22		"Slet ikke sjov", "Ekstremt sjov"
	23		"Slet ikke menneskelig", "Ekstremt menneskelig"

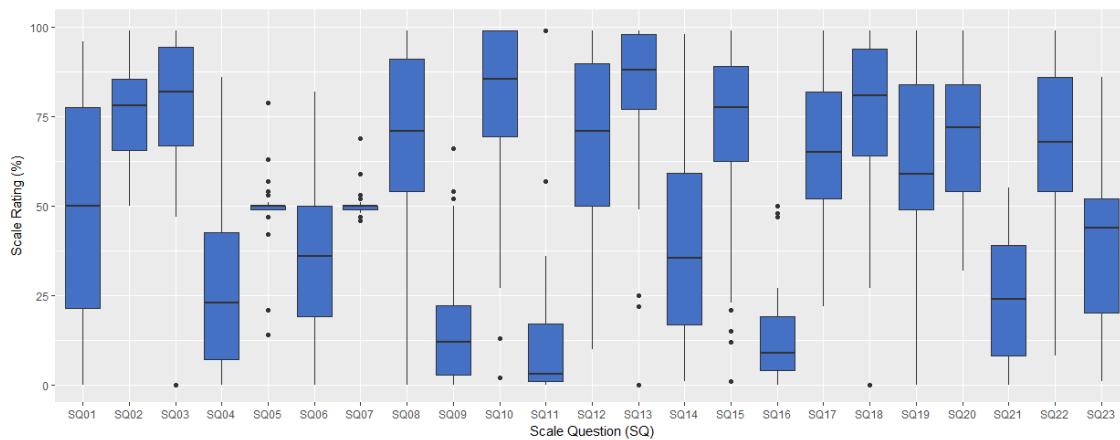
Figur 10.3. Oversigt over hvilke skalaspørgsmål (*Scale question*), der er hører til skala nummer (*SQ#*) samt tilhørende labels (*Skala labels*), der præsenteres på hver af de syv sider (*Side*). Forefindes også i Afsnit 13.4 (*Skala Oversigt*).

10.5 Frasortering af manglende data

Testpersonernes respons på skalaerne omregnes til procent, hvorfor det venstre endepunkt er 0 % og det højre endepunkt er 100 %. Ud fra datasættet fremgår det, at nogle skalaer har én eller flere besvarelser på 0 %, problemet med det er, at det ikke vides med sikkerhed om testpersonen har svaret 0 eller om testpersonen slet ikke har angivet en respons på den pågældende skala. At det ikke er muligt at adskille de to tilfælde fra hinanden skyldes, dels at programmet automatisk gemmer værdien 0 og ikke *Nan*, hvis en skala ikke er besvaret og dels at det har været muligt at trykke *Næste*, selvom der ikke er angivet en besvarelse. Det har derfor været nødvendigt at gennemgå datasættet og vurdere hver enkelte tilfælde i forhold til om det er en reel besvarelse eller om det er en manglende besvarelse. Dette afgøres på baggrund af to kriterier: 1) Hvis alle skalaerne, der er præsenteret på én af de syv sider, ikke er besvaret vil de alle være angivet med 0 i datasætte, hvorfor det kan fastslås at det ikke er en reel besvarelse men derimod en manglende besvarelse. Årsagen til at der forekommer tilfælde, hvor en testperson ikke har besvaret en eneste skala på en af de præsenterede sider skyldes formentlig, at der har været problemer med programmet. Disse problemer relaterer sig hovedsageligt til *Næste*-knappen i programmet, som reagerede dårligt og i nogle tilfælde kun ved dobbeltklik, så hvis en testperson klikkede flere gange på knappen er det muligt, at de sprang over en side og slet ikke blev præsenteret for de pågældende skalaer. 2) Hvis der ved en skala generelt er høje procentsatser og en besvarelse på 0 % virker usandsynlig vil det betragtes som en manglende besvarelse. Er det tilfældet vil 0'et blive behandlet som en outlier, der er tre standard afvigelser fra det resterende data.

I henhold til datasættet, jævnfør Afsnit 13.5 (*Rådata i Excel*), vil situationer, der opfylder kriterie 1) vedrørende manglende besvarelser, være markeret med en rød celle. Summeret forekommer det 16 gange, hvor største delen af fejlene er opstået på side syv, hvor hverken TP8, TP10, TP18, TP29, TP34 og TP36 har afgivet en besvarelse på hver af de fire præsenterede skalaer.

Der er 28 tilfælde hvor det er uvist hvorvidt der enten er afgivet en besvarelse på 0 % eller om 0 % afspejler en manglende besvarelse. Dette vil blive undersøgt nærmere ved hjælp af nedenstående Figur 10.4.



Figur 10.4. Boksplot for besvarelserne på de 23 skalaer. Outliers er markeret med sorte prikker.

Boksplottet er lavet i *RStudio*, hvor whiskers enten går fra den øverste eller nederste del af det interkvartile område til den største eller mindste værdi fundet 1.5 gang fra det interkvartile område, svarende til afstanden mellem første og tredje quartil. Forefindes der et punkt udenfor disse whiskers betragtes det som en outlier.

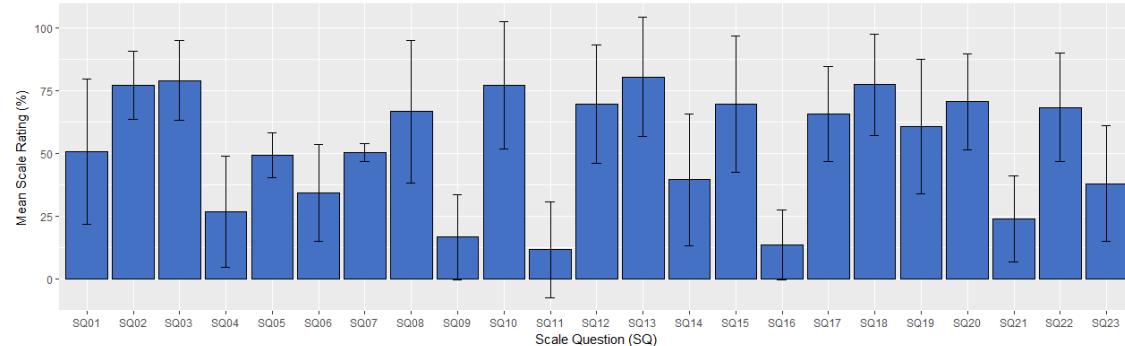
Der tages udelukkende udgangspunkt i de outliers, der har værdien 0, hvilket er tilfældet for SQ3, SQ13 samt SQ18. Disse outliers er baseret på respons fra to testpersoner: TP18 til både SQ3 og SQ18, og TP30 til SQ13. Årsagen til at outlieren til SQ3, som vedrører hvor let det er at interagere med robotten, fjernes er, at den største grund til robotten sommetider kan være svær at interagere med skyldes skærmens reaktionsevne. Ved dette skalaspørgsmål vurderede TP18, at skærmen reagerede 54 % og der blev ydermere ikke observeret nogle problemer med interaktionen, som kunne indikere, at testpersonen havde oplevelsen af, at robotten var svær at interagere med. Det vurderes derfor, at det er usandsynligt, at testpersonen aktivt har svaret 0, men nærmere har været resultat af, at scalaprogrammet reagerede dårligt på input. Derfor ekskluderes denne outlier. Den anden outlier, der forefindes i datasættet tilhørende TP18, er til SQ18, som vedrører hvor spændende robotten er. Baseret på observationer, hvor testpersonen virkede begejstret for mødet med robotten, samt besvarelserne til hvor glad testpersonen er for teknologi (91 %) virker det usandsynligt at testpersonen har oplevet robotten som værende slet ikke spændende. Det vurderes derfor, at det er usandsynligt at testpersonen aktivt har svaret 0, hvorfor denne outlier ekskluderes.

Den tredje outlier, som har værdien 0, forefindes ved SQ13, som vedrører hvorvidt testpersonen stolede på at robotten fulgte en hen til det valgte sted. Hvis testpersonerne ved, hvor de skal hen, og robotten kører en anden vej, er der stor sandsynlighed for, at de ikke tror på at robotten kører det rigtige sted hen. Flere testpersoner gav udtryk for, at robotten kørte den forkerte vej, hvis de for eksempel skulle på toilettet og robotten kørte mod DutyFree. Derudover forekommer der to andre outliers til dette skalaspørgsmål, jævnfør Figur 10.4, hvor testpersonerne aktivt har svaret. Det vurderes derfor at der ikke er nok belæg for at ekskludere denne outlier.

De resterende outliers, der er forskellig fra 0, medtages alle da det er besvarelser testpersonerne aktivt har foretaget og derfor afspejler deres oplevelse af interaktionen med robotten.

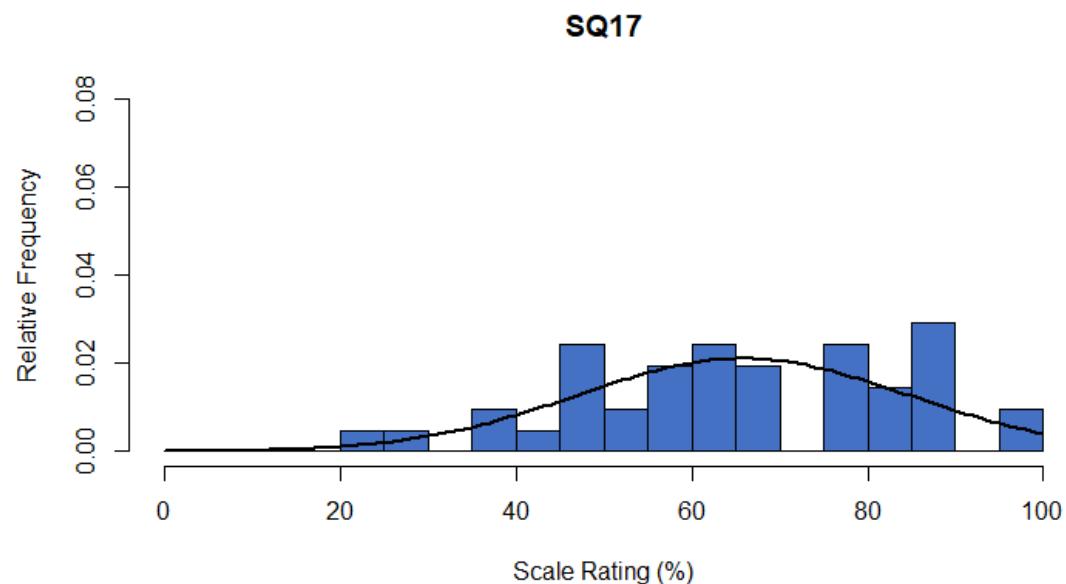
10.6 Fordeling af besvarelser

I følgende afsnit vil fordelingen af besvarelserne blive præsenteret. Dette dækker over den gennemsnitlige besvarelse til hver af de 23 skalaer samt eksempler på forskellige fordelinger, der er med til at illustrere, at der eksisterer stor varians på tværs af besvarelserne. Til sidst vil data blive præsenteret i et boksplot og diskuteret.



Figur 10.5. Søjlediagram over den gennemsnitlige besvarelse (%) til hvert skalaspørgsmål (SQ) samt barer, der repræsenterer standardafvigelsen for hver SQ.

På Figur 10.5 er det tydeligt, at der forekommer stor variation mellem besvarelserne til de enkelte skalaspørgsmål, dog med undtagelse af SQ7 vedrørende robottens højde og tildels SQ5 vedrørende robottens afstand. Der er udarbejdet et kombineret histogram og normalfordelingsplot for hver skala, for at få en mere nøjagtig indsigt i, hvordan besvarelserne er fordelt. Enkelte af dem vil blive beskrevet i dette afsnit, mens resten fremgår af Afsnit 13.6 (*Hisogram og normalfordelings plot*).

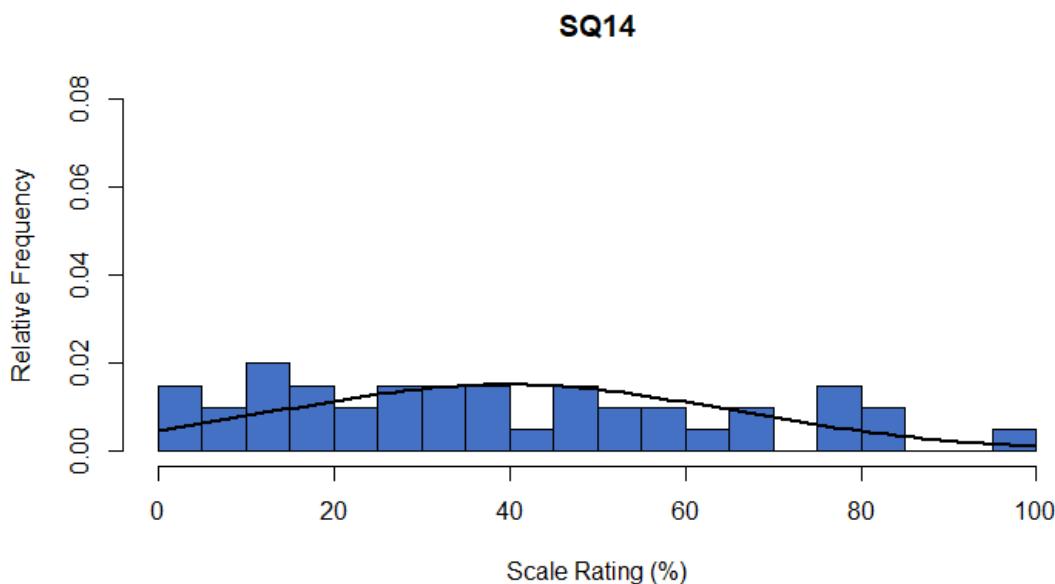


Figur 10.6. Fordelingerne af besvarelserne til SQ17, vedrørende hvor elegant robotten er. Histogrammet repræsenterer hyppigheden af besvarelser (y-aksen) målt som relativ frekvens (Antal/SampleSize), inden for de fastsatte intervaller (x-aksen), angivet i %. Den sorte kurve repræsenterer den underliggende normalfordelingskurve, som er baseret på middelværdien og standardafvigelsen.

Figur 10.6 repræsenterer histogrammet for SQ17, vedrørende hvor elegant robotten er, med tilhørende normalfordelingskurve og er et eksempel på hvordan resten af histogrammerne er visualiseret. Da en stor del af data befinner sig under den sorte kurve, tyder det på, at besvarelserne for SQ17 er tilnærmelsesvis normalfordelte. Dette undersøges dog ikke med en signifikantest, *Shapiro-Wilk Normality Test*, da formålet med gennemgangen blot er at få et overblik over fordelingerne.

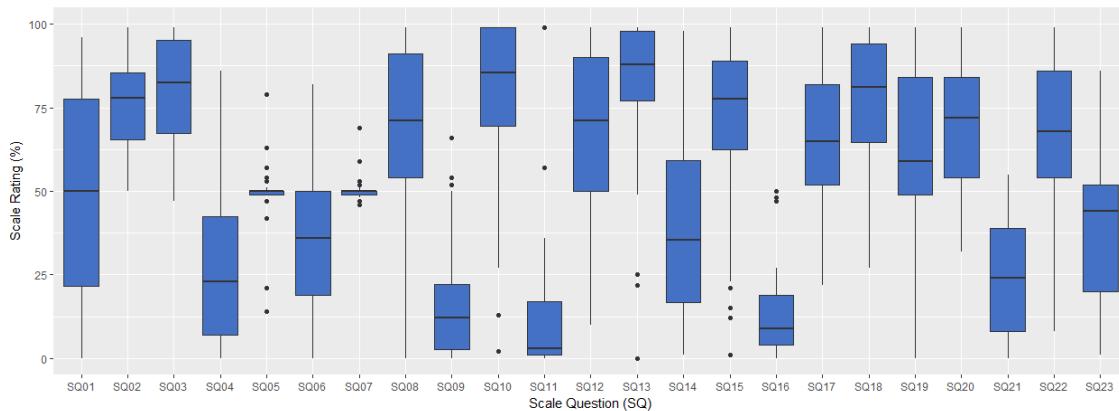
Normalfordelings kurverne i flere af plottene forekommer meget flade, hvilket skyldes, at akserne holdes konstante på tværs af plottene, for bedre at kunne sammenligne dem indbyrdes. Dette er dog med undtagelse af histogrammer og normalfordelinger for SQ5, SQ7 og SQ11, da data er meget centreret omkring midten og til venstre på disse histogrammer. Det kan derfor være svært at se fordelingerne på de resterende histogrammer, hvis akserne defineres ud fra SQ5, SQ7 og SQ11.

I mange af plottene er der en stor varians, hvilket indikerer at en stor del af skalaen er blevet anvendt. Her skal det tilføjes at besvarelserne er givet ud fra forskellige højder, afstande og indgangsvinkler og at variationen derfor ikke nødvendigvis er et udtryk for at testpersonerne er uenige, men snarere at de har vurderet forskellige stimuli. Et eksempel er givet på Figur 10.7.



Figur 10.7. Fordelingerne af besvarelserne til SQ14, vedrørende hvor personlig robottens hjælp opleves. Histogrammet repræsenterer hyppigheden af besvarelser (y-aksen) målt som relativ frekvens (Antal/SampleSize), inden for de fastsatte intervaller (x-aksen), angivet i %. Den sorte kurve repræsenterer den underliggende normalfordelingskurve, som er baseret på middelværdien og standardafvigelsen.

I disse histogram- og normalfordelingsplots fokuseres der på, hvordan besvarelserne er fordelt for hvert enkelt skalaspørgsmål, men for at få et samlet overblik, fokuseres der på Figur 10.8.



Figur 10.8. Boksplot over fordelinger af besvarelserne (angivet i % på y-aksen) for hvert skalaspørgsmål (angivet med SQ på x-aksen). De manglende besvarelser er ekskluderet fra dette boksplot, så det udelukkende bygger på reelle besvarelser.

Figur 10.8 deler alle besvarelserne til det pågældende skalaspørgsmål op i fire kvartiler. Medianen adskiller det interkvartile område, som repræsenterer 50 % af data, og kanterne på det interkvartile område adskiller disse fra de to yderste kvartiler. Outliers er ikke medregnet i kvartilerne for boksplottet, men er stadig plottet for at give en idé om hvordan de fordeler sig i forhold til resten af besvarelserne.

Ud fra Figur 10.8 fremgår det, at fordelingerne adskiller sig fra hinanden. Testpersonerne udnytter en stor del af nogle af skalaerne (SQ1, SQ4, SQ8, SQ12, SQ14, SQ19, SQ22, SQ23), hvorimod besvarelserne på andre skalaer fordeler sig over mindre områder, hvilket gør sig gældende for: SQ2, SQ3, SQ5, SQ7, SQ9, SQ11, SQ13 og SQ16. Nogle boksplot gengiver tilnærmelsesvis en normalfordeling, (SQ1, SQ2, SQ6, SQ17, SQ20, SQ21), mens andre er skævvredet den ene eller anden vej (SQ8, SQ9, SQ10, SQ11, SQ13, SQ18). Der forekommer også enkelte boksplots, hvor datapunkterne ophober sig omkring endepunkterne eller midtpunktet (SQ5, SQ7, SQ9, SQ10, SQ11, SQ13).

Den lave varians, som forekommer i nogle af boksplottene, kan skyldes forskellige ting. Det kan enten skyldes, at parameteren ikke er vigtig for oplevelsen af interaktionen med robotten og at testpersonerne derfor har svaret hvad de anser som værende neutralt. Det kan også skyldes, at det ikke er en parameter, som er blevet påvirket af de specifikke stimuli testpersonerne blev udsat for, men at den kan være vigtig i andre situationer. Effekten forstærkes især ved SQ5 og SQ7, som er bipolare skalaer. Midtpunkterne på disse skalaer kan fungere som et anker, hvor testpersonerne ofte vil centrere deres vurderinger omkring. Labelen *Fin*, som anvendes i SQ7 blev valgt på baggrund af testpersonernes udtalelser i feltundersøgelsen. Det kan dog være, at ordet *Fin* dækker over for bredt et spænd af skalaen, hvilket kan medføre at testpersonerne har angivet deres respons i midten, selvom robotten har været en lille smule for høj eller lav. Dette kan også være med til at centrere datapunkterne omkring midtpunktet og dermed mindske variationen i vurderingerne.

De lukkede endepunkter, som blev anvendt på skalaerne, kan ophøbe data omkring disse. Ophobningen af data omkring endepunkterne betyder også, at fordelingen bliver skævvredet, da den ofte stopper brat i den ende med endepunktet, men har en lang hale i den anden ende.

I nogle tilfælde bruges kun øverste halvdel, hvorpå andre skalaer er det kun nederste halvdel, som bruges. Det hænger formentlig sammen med formuleringerne på endepunkterne, da den positive label sommetider vil være svare til 100 % (Ekspemvelvis: *Ekstremt sød*) og i andre tilfælde ved 0 % (Ekspemvelvis: *Slet ikke amassende*).

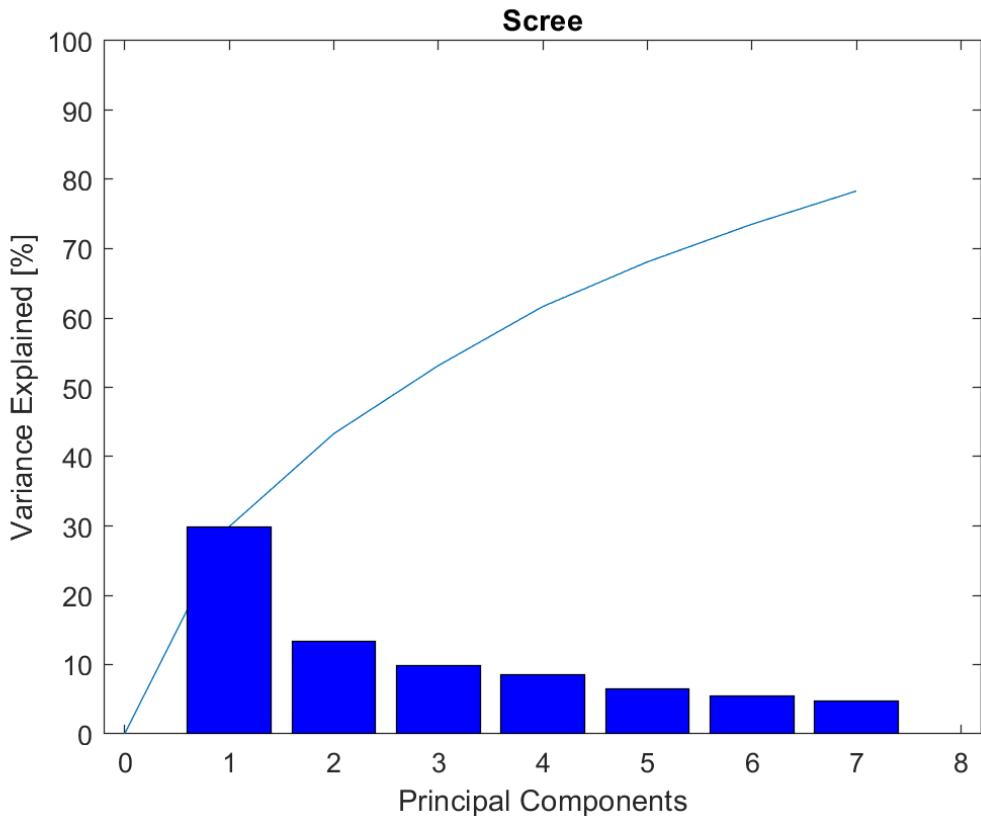
10.6.1 Standardisering af besvarelser

Når variansen ikke er ens for ens data kan det i nogle tilfælde vælges at udligne variansen når der skal udføres PCA. Hvis variansen skulle udlignes vil det kræve en standardisering af data. Formålet med at gøre dette er at data med lav varians i rå data får lige betydning i PCA som rå data med stor varians.

Det vælges ikke at standardisere data, da det kan give skævvridninger, som ikke er ønsket. Testpersonerne har reelt haft muligheden for at svare på hele skalaen, men har valgt at afgive besvarelser i nærheden af hinanden, hvilket tolkes som at de har været mere enige i dette tilfælde end i andre tilfælde hvor der er stor spredning.

10.7 Principal Component Analysis

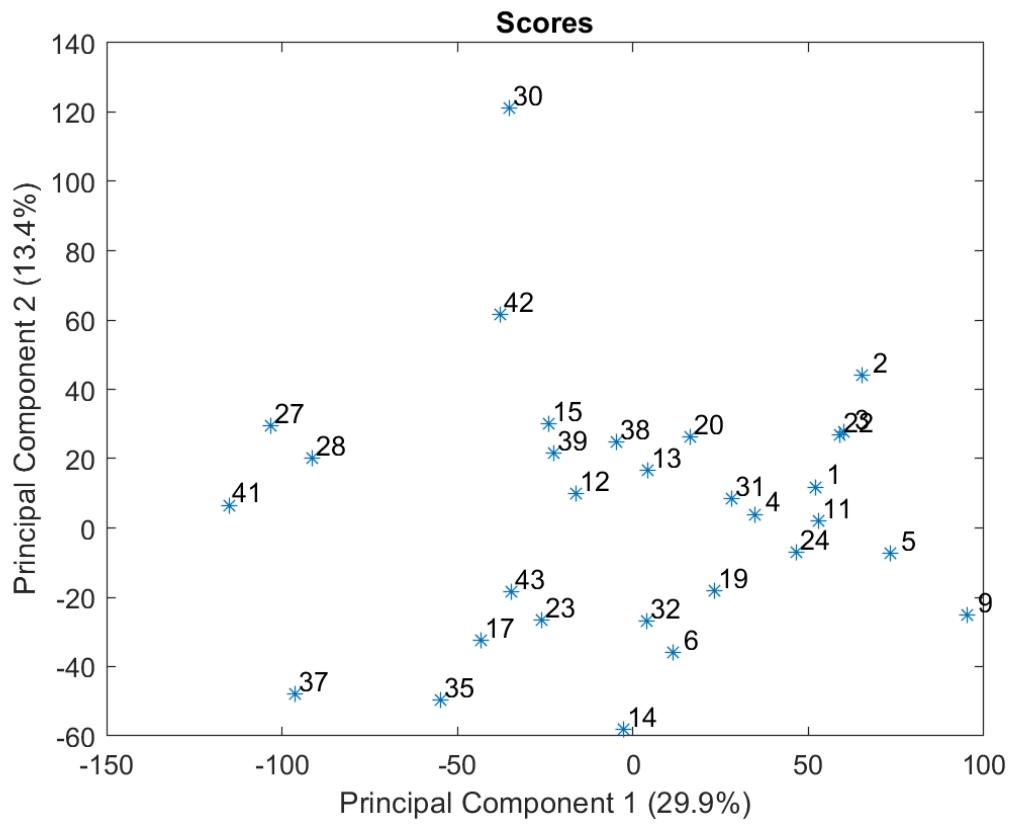
Først laves en *Principal Component Analysis* (PCA) analyse på det samlede datasæt, for at undersøge om der er korrelation mellem nogle af de 23 skalaspørgsmål. Analysen er udføres i *MatLab* og scriptet findes i Afsnit 13.7 (*MATLAB Scripts til PCA*). For at undersøge hvor meget af variansen, der kan forklares ud fra hver *Principal Components* (PC) udarbejdes der et *Scree*-plot, jævnfør Figur 10.9.



Figur 10.9. Scree-plot, hvorpå sammenhængen mellem antallet af *Principal Components* og *Variance Explained [%]* fremgår. Bidraget for hver PC er som følger: PC1 29.9 %, PC2 13.4 %, PC3 9.8 %, PC4 8.5 %, PC5 6.4 %, PC6 5.4 % og PC7 4.8 %.

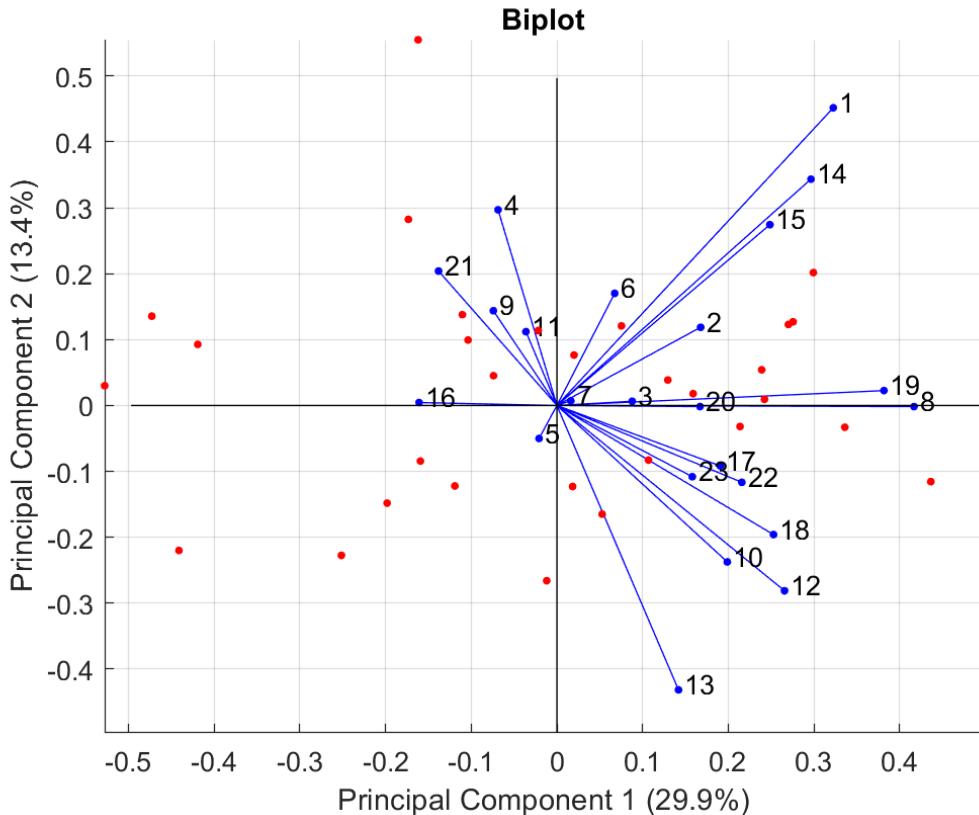
Ud fra *Scree*-plottet på Figur 10.9 fremgår det, at der ikke er en tydelig sammenhæng i variansen fra datasættet, da kun 29.9 % af variansen kan forklares af PC1 og 13.4 % af PC2. Selv hvis de tre første PCs medtages, er det kun 53.1 % af variansen, som forklares af modellen. Selvom det kun er de første syv PCs der fremgår på Figur 10.9 så er der reelt 23 PCs, men da de restrende PCs bidrager med meget lidt fremgår de ikke på *Scree*-plottet. Dette skyldes formentligt, at der er for mange variable, som bliver varieret på én gang og at det ikke har været muligt at kontrollere alle variable lige godt, og at der tilmed forekommer forskellige grupperinger i forhold til de variable, der tilnærmedesvis var mulige at kontrollere. På baggrund af det er det svært at vurdere hvilke parametre, der korrelerer.

Ud fra *Score*-plottet på Figur 10.10 fremgår det også at PC1 kun forklarer en lille del af variansen, da der er meget spredning i datapunkterne og spredningen ikke er systematisk. Hvis PC1 havde forklaret en større del af variansen, ville data være mere centreret omkring den horisontale akse.



Figur 10.10. Score-plot for PC1 og PC2.

For at undersøge hvilke parametre, der bidrager til hver PC opstilles et *Bi-plot*, hvori både *scores* og *loadings* fremgår. *Bi-plottet* på Figur 10.11 giver det bedste overblik, da det er nemmest at fortolke i to dimensioner, dog er det tredimensionelle plot vedlagt i Afsnit 13.8 (*3D Bi-plots*).



Figur 10.11. *Bi-plot* med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens højde.

Baseret på Figur 10.11 tyder det på, at der forekommer en positiv korrelation mellem SQ14, vedrørende hvor personlig robottens hjælp opleves, og SQ15, vedrørende hvor overrasket testpersonerne blev over robottens henvendelse. Lignende korrelation forefindes mellem SQ17, vedrørende hvor elegant robotten opleves, og SQ22, vedrørende hvor sjov robotten opleves. Derudover tyder det på, at der forekommer en positiv korrelation mellem SQ8, vedrørende hvorvidt testpersonerne føler, at robotten kan hjælpe dem, SQ19, vedrørende hvor sød robotten opleves, og SQ20, vedrørende hvor dejlig robotten opleves. De tre parametre tyder på at være negativ korrelerede med SQ16, vedrørende hvor irriterende robotten opleves. En negativ korrelation forefindes også mellem SQ4, vedrørende hvordan robottens bevægelser opleves, og SQ13, vedrørende hvorvidt testpersonerne regnede med at robotten fulgte dem hen til det valgte sted.

Baseret på Figur 10.11 fremgår det, at SQ7, vedrørende robottens højde, SQ5, vedrørende robottens afstand, og SQ3, vedrørende hvordan det var at bruge robotten, kun forklare en meget lille del af variationen. Tages der i 3D plottet udgangspunkt i SQ7, fremgår det ligeledes at denne parametre ikke bidrager med noget til de tre PCs. Sammenholdes det med hvad der er beskrevet i Afsnit 10.6 (*Fordeling af besvarelser*) på Figur 10.8, hvor det fremgår at SQ7 har en meget lille varians, så er det ikke overraskende,

at den ikke bidrager mere til de tre PCs end den gør. Dette afspejler at testpersonerne har svaret nogenlunde det samme til SQ7, det kan derfor tyde på at SQ7 ikke er særlig vigtig for den samlede brugeroplevelse. At SQ7 kun forklarer en meget lille del af variationen kan skyldes flere ting; enten så er robottens højde ikke en vigtig parameter eller så er testpersonerne ikke blevet eksponeret for en højde, der reelt påvirker deres oplevelse. Det kan også være fordi, at SQ7 bliver målt indirekte i andre parametre, som netop påvirkes af robottens højde, hvorfor dette vil blive undersøgt nærmere i det følgende afsnit.

Ud fra resultaterne fra denne PCA kan der ikke drages nogle særlig håndgribelige konklusioner i forhold til hvilke parametre korrelere positiv eller negativt, eller slet ikke. For at undersøge om der er nogle tendenser i forhold til de objektive parametre, som blev justeret undervejs i testen; højde, afstand samt robottens indgangsvinkel, vil hver af disse parametre blive undersøgt med en PCA. Dette gøres ved at fokusere på et af de tre objektive parametre ad gangen og dele data op i de grupper, som allerede findes.

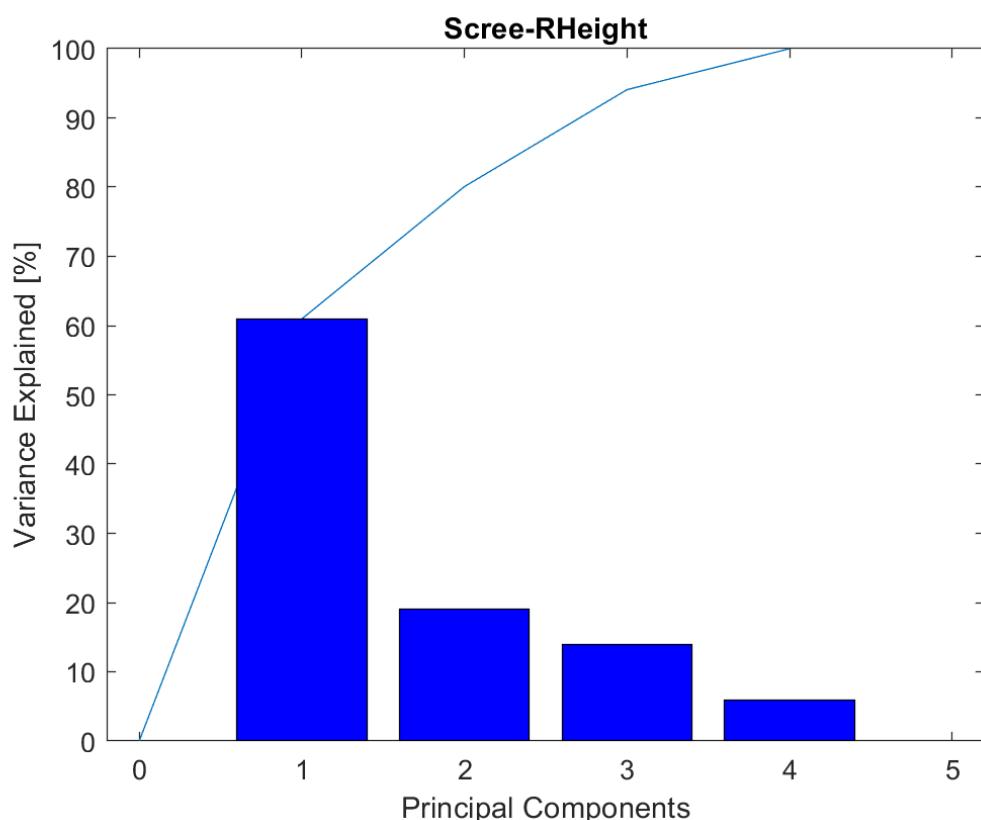
10.8 Principal Component Analysis: Robottens højde

I dette afsnit analyseres resultaterne med udgangspunkt i de fem forskellige højder robotten havde i testen. For hver højde er der mellem 7-10 testpersoner. De forskellige højder fremgår af Tabel 10.1, hvor antallet af testpersoner ligeledes er opgivet. Det undersøges, hvordan robottens højde påvirker de rejsendes oplevelse af robotten ved at undersøge, hvordan resultaterne ændrer sig afhængigt af de fem forskellige højder. Det er et *Between-subject* design, da hver testperson kun har oplevet robotten i én af højderne og svaret ud fra denne.

Højde (cm)	Antal testpersoner
118	9
123.5	10
129	9
140	7
151	8

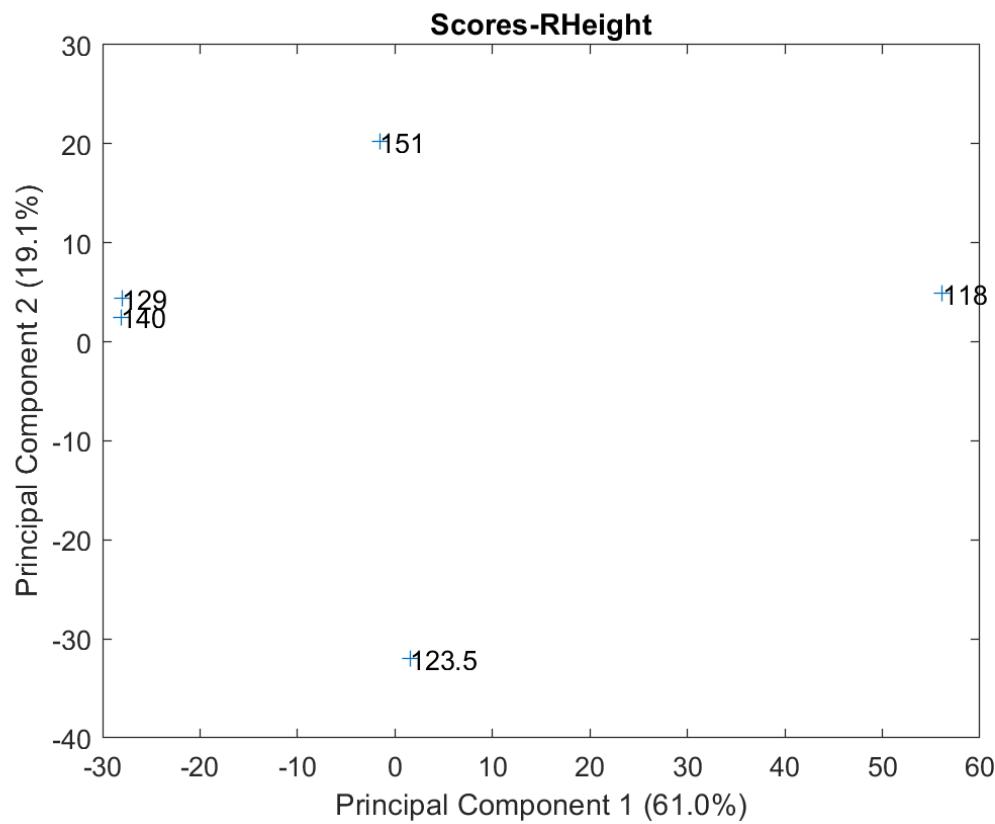
Tabel 10.1. Oversigt over antallet af testpersoner til hver af de fem højde robotten havde.

Ud fra *Scree*-plottet på Figur 10.12 fremgår det, at cirka 80 % af variansen kan forklares ud fra de to første *Principal Components* (PCs). Ved at tage PC3 med opnås cirka 94 % forklaret varians. Det er dog væsentligt simplere at fortolke resultaterne i to dimensioner og der fokuseres derfor som udgangspunkt kun på de to primære komponenter.



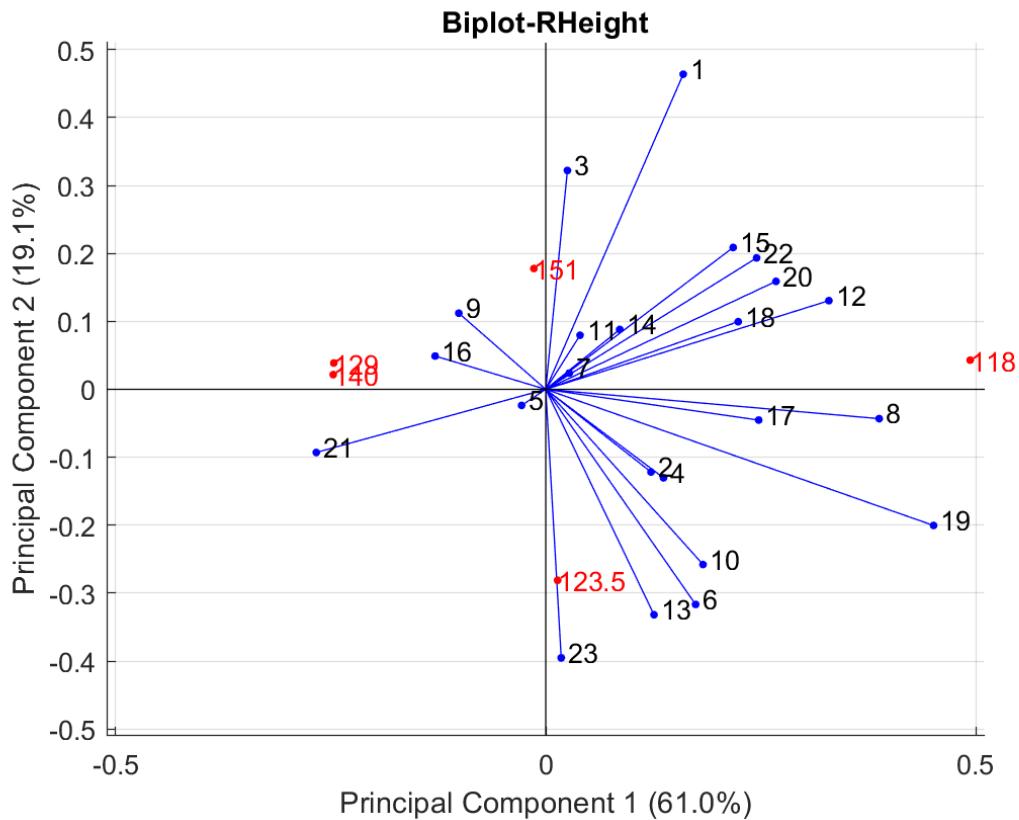
Figur 10.12. *Scree*-plot, hvorpå sammenhængen mellem antallet af *Principal Components* og *Variance Explained %* fremgår. Bidraget for hver PC er som følger: PC1 (61%), PC2 (19.1%), PC3(14%) og PC4(5.9%).

Ud fra *Score*-plottet på Figur 10.13 fremgår det, at der generelt er en del spredning mellem resultaterne ud fra de fem forskellige højder, men at 129 cm og 140 cm har lignende karakteristikas, da de ligger meget tæt i plottet.



Figur 10.13. Score-plot for PC1 og PC2 i forhold til robottens højde.

Ud fra *Bi*-plottet på Figur 10.14, hvorpå *loadings* og *scores* for hver parameter er præsenteret, fremgår det at SQ19 bidrager mest til PC1, SQ1 bidrager mest til PC2 og at SQ5, SQ7, SQ11 og SQ14 forklarer en meget lille del af variansen, hvorfor det ikke giver mening at kommentere på hvordan de indbyrdes ligger, medmindre de indgår i en korrelation med en mere betydende parameter.



Figur 10.14. *Bi*-plot med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens højde.

Når to eller flere parametre ligger tæt på hinanden, er det et udtryk for, at de er højt korrelerede. Dette gør sig gældende for SQ2 og SQ4, som henholdsvis dækker over hvordan testpersonerne oplevede robotten i forhold til om det er ekstremt afvisende eller ekstremt imødekommennde, og hvordan robottens bevægelser blev oplevet i forhold til om bevægelserne er ekstremt rolige eller ekstremt vilde. Lignende er gældende for SQ12 og SQ18, som henholdsvis dækker over hvorvidt testpersonerne kan lide, at blive betjent af robotten og om de synes, at robotten er spændende. SQ14 og SQ15 tyder også på at korrelere, da de nærmest ligger oven på hinanden. SQ14 og SQ15 vedrører henholdsvis hvor personlig robottens hjælp opleves og hvor overrasket testpersonen blev over robottens henvendelse. Lignende er gældende for SQ8 og SQ17, som også ligger relativt tæt. De to parametre vedrører henholdsvis, hvorvidt testpersonerne føler, at robotten kan hjælpe og hvor elegant robotten er.

Derudover forekommer det, at SQ21 er negativt korreleret med både SQ12 og SQ18, hvor SQ21 vedrører hvor anmassende robotten opleves. SQ9 er negativt korreleret med både SQ2 og SQ4, hvor SQ9, vedrører hvorvidt robotten stod i vejen. Der forekommer også negativ korrelation mellem SQ16, som vedrører hvor irriterende robotten er, og SQ19, som

vedrører hvor sød robotten er. Ydermere forekommer der en negativ korrelation mellem SQ3, som vedrører hvor nemt det var at bruge robotten, og SQ23, som vedrører hvor menneskelig robotten opleves.

I henhold til robottens højde, så ligger 118 cm i modsatte ende af PC1 i forhold til både 129 cm og 140 cm, som nærmest ligger oveni hinanden. Baseret på Figur 10.14 tyder det derfor på, at højderne 129 cm og 140 cm primært er domineret af SQ21, som vedrører hvor ammassende robotten opleves, og delvist af SQ16, som vedrører hvor irriterende robotten opleves, og SQ9, som vedrører hvorvidt robotten stod i vejen. For højde 118 cm tyder det derimod på, at den højde er domineret af SQ8, som vedrører hvorvidt testpersonen føler, at robotten kan hjælpe en, og delvist af SQ19, som vedrører hvor sød robotten opleves, og SQ12, som vedrører hvor godt testpersonen kan lide at blive betjent af robotten.

Lignede gør sig gældende for PC2, hvor højderne 123.5 cm og 151 cm adskiller sig. Baseret på Figur 10.14 fremgår det, at SQ23, som vedrører hvor menneskelig robotten opleves, primært dominerer denne højde. Dertil domineres højden også af SQ13, som vedrører hvorvidt testpersonen regnede med at robotten fulgte dem hen til det sted de havde valgt. Højden på 151 cm er derimod primært domineret af SQ3, som vedrører hvor nemt det var at bruge robotten, samt SQ1, som vedrører skærmens reaktion.

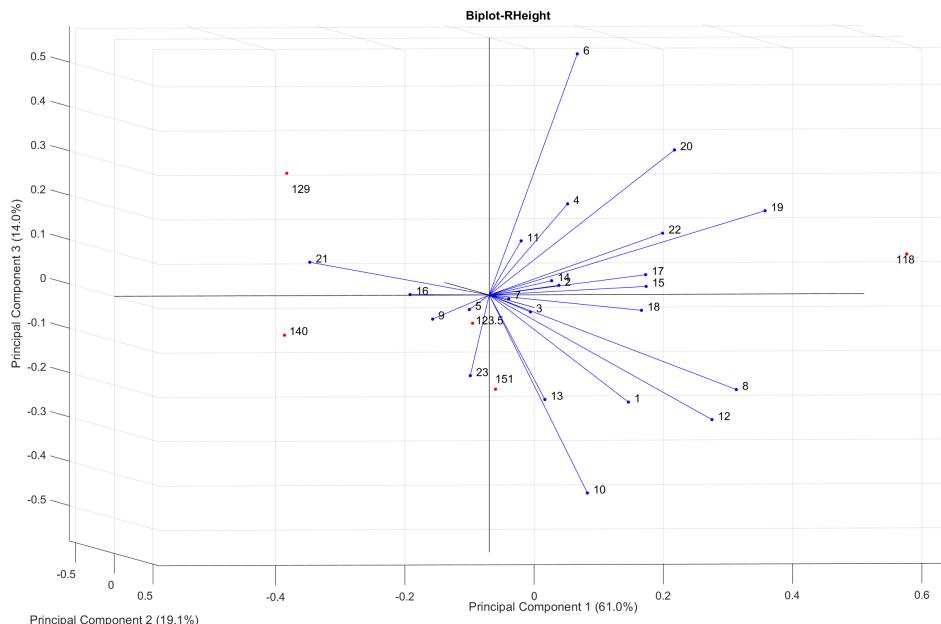
Hver *loading* angiver hvor meget variation hver parameter bidrager med i forhold til den specifikke PC. I Tabel 10.2 fremgår samtlige *loadings* for hver parameter til hver PC.

	PC1 (61%)	PC2 (19.1%)	PC3 (14%)	PC4 (5.9%)
SQ1	0.1598	0.4638	-0.2199	-0.0165
SQ2	0.1223	-0.1217	0.0152	-0.2757
SQ3	0.0252	0.3221	-0.0228	0.4574
SQ4	0.1369	-0.1303	0.1984	0.0372
SQ5	-0.028	-0.0237	-0.0339	-0.1285
SQ6	0.1742	-0.3167	0.5274	0.0873
SQ7	0.0273	0.0239	-0.0078	-0.0035
SQ8	0.3875	-0.0432	-0.2174	-0.3275
SQ9	-0.1010	0.1121	-0.0485	-0.1130
SQ10	0.1828	-0.2579	-0.4600	0.0328
SQ11	0.0399	0.0796	0.1256	0.1631
SQ12	0.3290	0.1305	-0.2765	-0.0732
SQ13	0.1258	-0.3318	-0.2525	0.5940
SQ14	0.0860	0.0881	0.0360	0.1441
SQ15	0.2178	0.2088	0.0282	-0.1953
SQ16	-0.1285	0.0490	0.0040	0.0326
SQ17	0.2475	-0.0453	0.0422	0.1889
SQ18	0.2236	0.0996	-0.0310	0.1844
SQ19	0.4507	-0.2004	0.1778	-0.0131
SQ20	0.2676	0.1590	0.3334	-0.0940
SQ21	-0.2666	-0.0929	0.0707	-0.1057
SQ22	0.2451	0.1935	0.1474	0.0239
SQ23	0.0180	-0.3953	-0.2013	-0.1809

Tabel 10.2. Oversigt over *loadings* til de fire PCs, for hver PC fremhæves den mest betydende parameter. *SQ* angiver skalaspørgsmålet.

Tages der derimod udgangspunkt i et tredimensionelt *Bi*-plot, jævnfør Figur 10.15, fremgår det, at de to højder: 129 cm og 140 cm adskiller sig en del i den tredje dimension. Derudover tyder det på, at SQ6, som vedrører testpersonernes oplevelse af robottens hastighed, samt SQ10, som vedrører hvor tryg testpersonen følte sig ved robotten, er de to parametre, som har størst indflydelse på PC3.

Derudover elimineres nogen af de fornævnte korrelationer mellem parametrene, når de præsenteres i det tredimensionelle plot, gengivet på Figur 10.15, eller vedlagt i Afsnit 13.8 (*3D Bi-plots*). Dette gælder blandt andet SQ2 og SQ4, som ikke længere ligger oven på hinanden, jævnfør ingen korrelation, samt SQ3 og SQ23, der ikke længere er negativt korrelerede. Derudover forefindes det at SQ10 og SQ13 samt SQ20 og SQ22 har en svag korrelation.

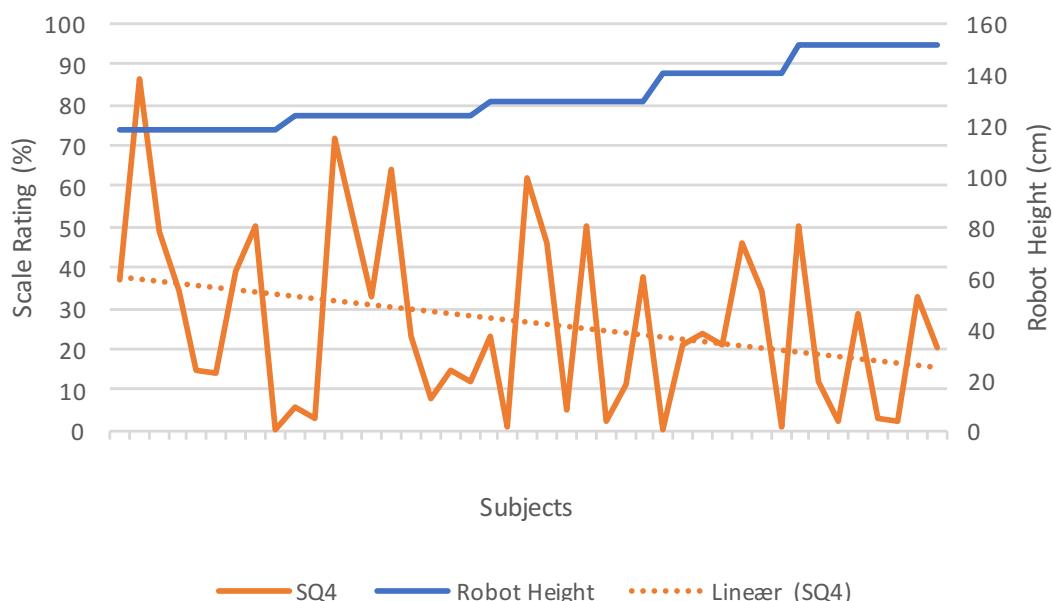


Figur 10.15. 3D *Bi*-plot med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens højde.

10.8.1 Tendenser i forhold til robottens højde

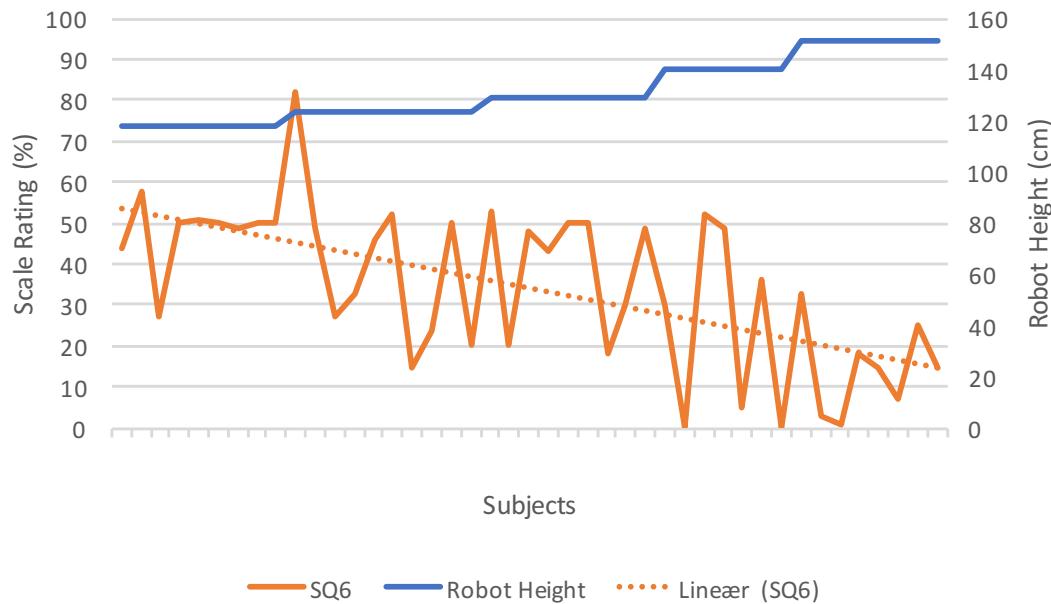
For at få et indtryk af hvilken indflydelse de fem forskellige højder, fra 118 cm til 151 cm, har på testpersonernes evaluering af interaktionen med robotten, vil der i det følgende afsnit undersøges hvordan højden påvirker parametrene. Undersøgelsen er foretaget på samtlige skalaer, men det vil kun være de skalaer, hvor der er en klar tendens til, at højden påvirker den pågældende parameter, der vil blive præsenteret. Samtlige grafer kan dog forefindes i Afsnit 13.9 (*Tendenser*). Dette vurderes blandt andet efter, at det højste og laveste punkt på tendenslinjen, den orange stippled linje, minimum skal variere omkring 10 % i forhold til y-aksen, som afspejler testpersonernes respons på den specifikke skala. Alle procentsatser er aflæst og ikke udregnet. Testpersonerne er ikke angivet kronologisk på x-aksen, da data er sorteret efter at robottens højde går fra den laveste til højeste, hvor testpersonernes vurdering ligeledes følger med. Det skal dog understreges, at tendenslinjerne kun er vejledende, da de er manipulerbare i forhold til hvordan datapunkterne struktureres, hvorfor tendenslinjerne skal sammenholdes med de kurver som de repræsenterer. På de følgende grafer er datapunkterne strukturet således, at testpersonerne står i kronologisk rækkefølge under den højde de har oplevet robotten i.

Den første skala hvorpå tendenslinjen ændrer sig med minimum 10 % er ved SQ4, som vedrører robottens bevægelse, hvor variationen aflæses til omkring 21 %, hvilket fremgår af Figur 10.16. Baseret på dette virker det som om, at der er en tendens til, at desto lavere robotten er, desto vildere opleves dens bevægelser, hvorimod bevægelserne opleves mere rolige desto højere robotten er.



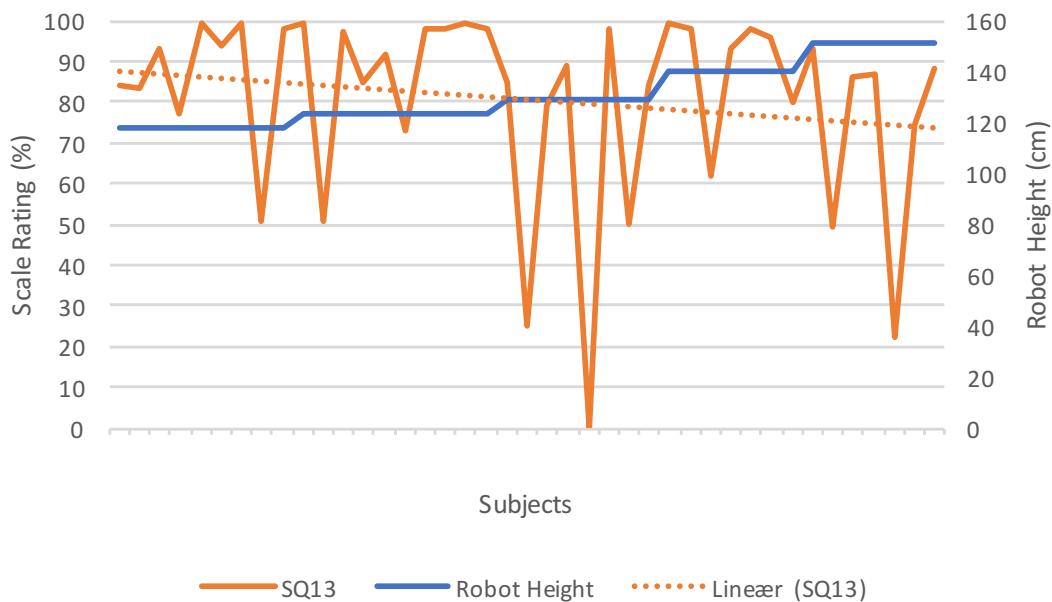
Figur 10.16. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ4: *Hvordan oplevede du robottens bevægelser?*. Denne graf bygger på samtlige 43 besvarelser.

Ud fra Figur 10.17 tyder det på, at der er en klar tendens til, at hvis robotten er på sit laveste (118 cm), så opleves robotten som værende omkring 40 % hurtigere end når robotten er på sit højeste (151 cm).



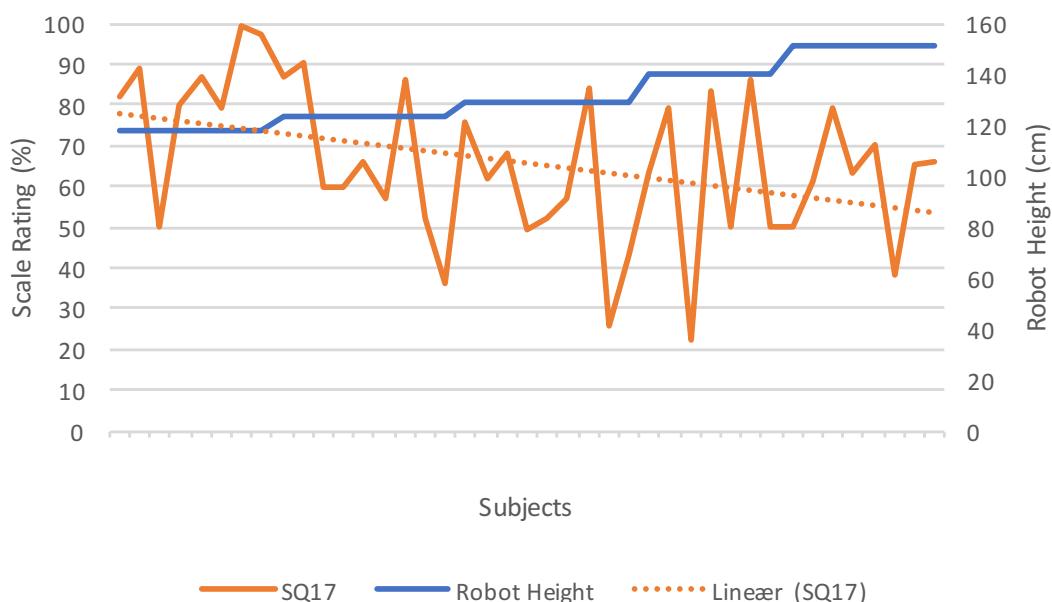
Figur 10.17. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ6: *Jeg synes, at robottens hastighed er...* Denne graf bygger på samtlige 43 besvarelser

På Figur 10.18 tyder det på, at når robotten er på sit laveste (118 cm) stoler testpersonerne omkring 18 % mere på, at robotten følger dem hen til det sted de har valgt, end når robotten er på sit højeste (151 cm). Det skal dog pointeres at for de to laveste højder er variationen i besvarelserne omrent det samme, hvorved 129 cm er der en meget stor variation i besvarelserne. Variationen i besvarelsen ved 140 cm er omrent det samme som for de to laveste højder, hvor det tyder på at besvarelserne ved 151 cm er lavere end ved 118 cm, 123.5 cm og 140 cm.



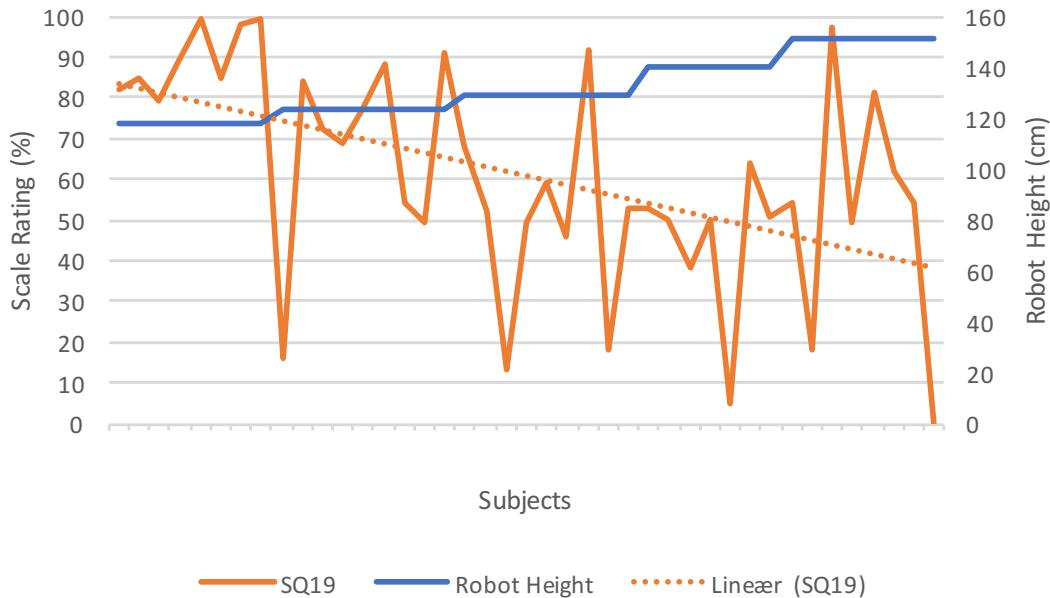
Figur 10.18. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ13: *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte*. Denne graf bygger på 41 besvarelser, da der manglede to.

På Figur 10.19 tyder det på, at når robotten er på sit laveste (118 cm) opleves den som værende omkring 27 % mere elegant end når robotten er på sit højeste (151 cm). Ingen skal der tages højde for den store variation særligt ved besvarelserne relateret til 129 cm og 140 cm.



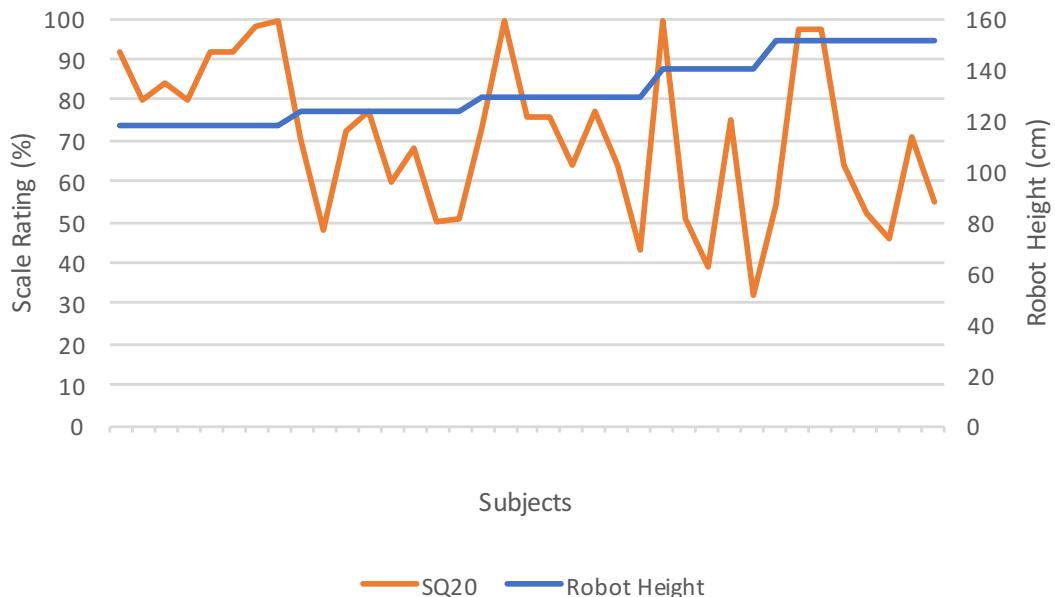
Figur 10.19. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ17: *Hvad synes du om robotten?*, i forhold til *elegant*. Denne graf bygger på 41 besvarelser, da der manglede to.

På Figur 10.20 er der en tendens til, at når robotten er på sit laveste (118 cm) opleves den omkring 35 % sjældere end når robotten er på sit højeste (151 cm). Dog forekommer der stor variation i besvarelserne, men det tyder på at besvarelserne til 140 cm er markant lavere end ved de resterende højder.



Figur 10.20. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ19: *Hvad synes du om robotten?*, i forhold til *sjæld*. Denne graf bygger på 41 besvarelser, da der manglede to.

Med forbehold for, at der mangler seks besvarelser til SQ20, omhandlende hvor sej robotten opleves, samt en stor variation mellem besvarelserne, hvorfor tendenslinjen er fjernet, tyder det på at når robotten er på sit laveste (118 cm) opleves den generelt som værende sejere end på de fire resterende højder, jævnfør Figur 10.21.



Figur 10.21. Sammenhæng mellem robottens fem højder (cm) og hvad testpersonerne angiver (%) på skalaen til SQ20: *Hvad synes du ellers om robotten?*, i forhold til *sej*. Denne graf bygger på 37 besvarelser, da der manglede seks.

10.9 Principal Component Analysis: Afstand

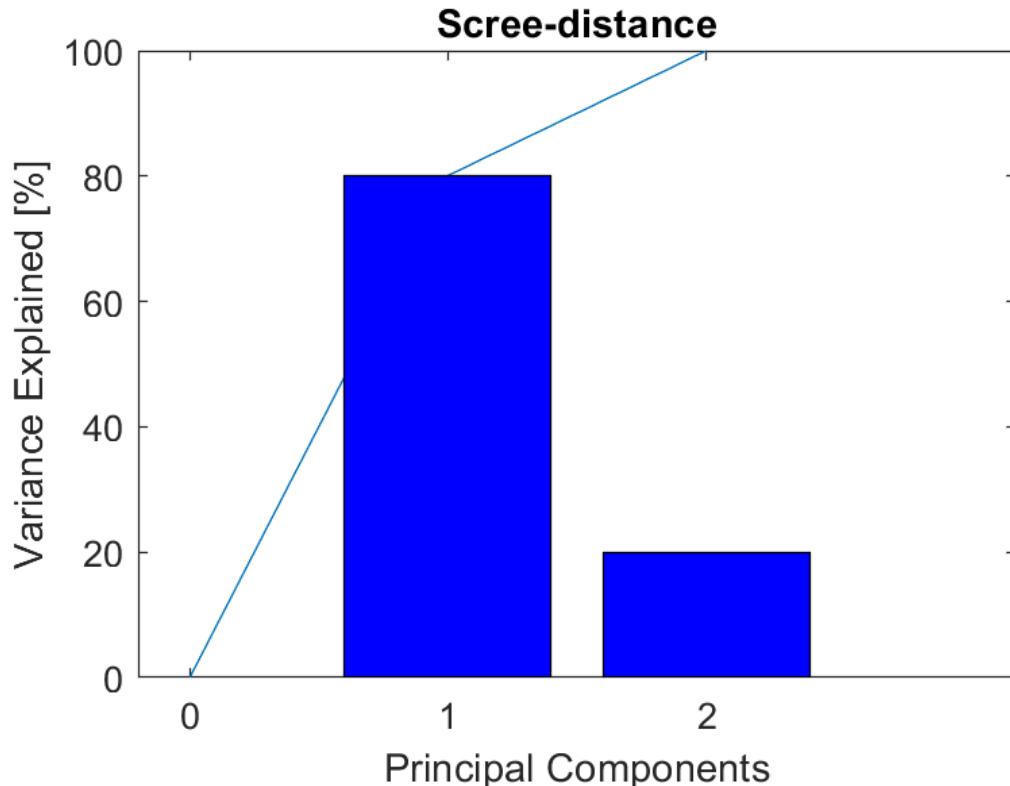
I dette afsnit analyseres resultaterne med udgangspunkt i de tre forskellige afstande robotten havde til testpersonerne i testen. De forskellige afstande fremgår af Tabel 10.3, hvor antallet af testpersoner ligeledes er opgivet.

Det undersøges, hvordan robottens afstand påvirker de rejsendes oplevelse af robotten ved at undersøge, hvordan resultaterne ændrer sig afhængigt af de tre forskellige afstande. Det er et *Between-subject* design, da hver testperson kun har oplevet robotten i én af afstandende og svaret ud fra denne.

Afstand	Antal testpersoner
Tæt på (1)	7
Tilpas (2)	28
Langt fra (3)	8

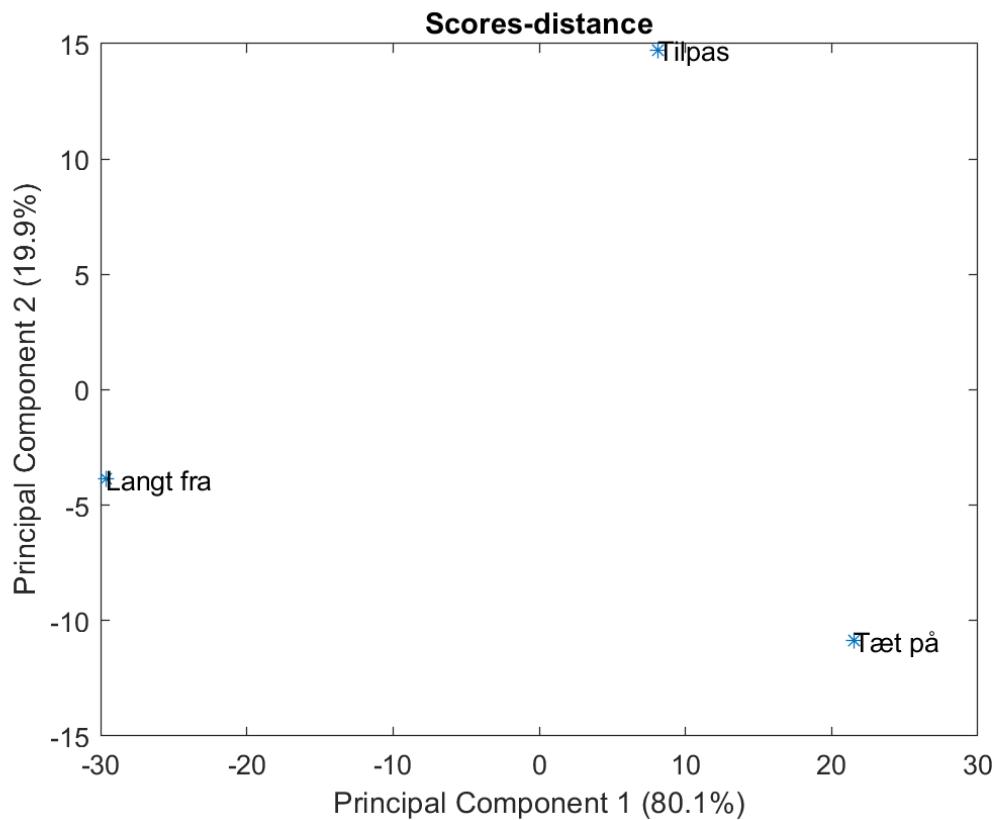
Tabel 10.3. Oversigt over antallet af testpersoner til hver af de tre afstande robotten havde.

Ud fra *Scree*-plottet på Figur 10.22 fremgår det at cirka 80 % af variansen kan forklares ud fra den første *Principal Component* (PC) og 100 % af variansen kan forklares ud fra to PCs.



Figur 10.22. *Scree*-plot, hvorpå sammenhængen mellem antallet af *Principal Components* og *Variance Explained [%]* fremgår. Bidraget for hver PC er som følger: PC1 (80.1%), PC2 (19.9%).

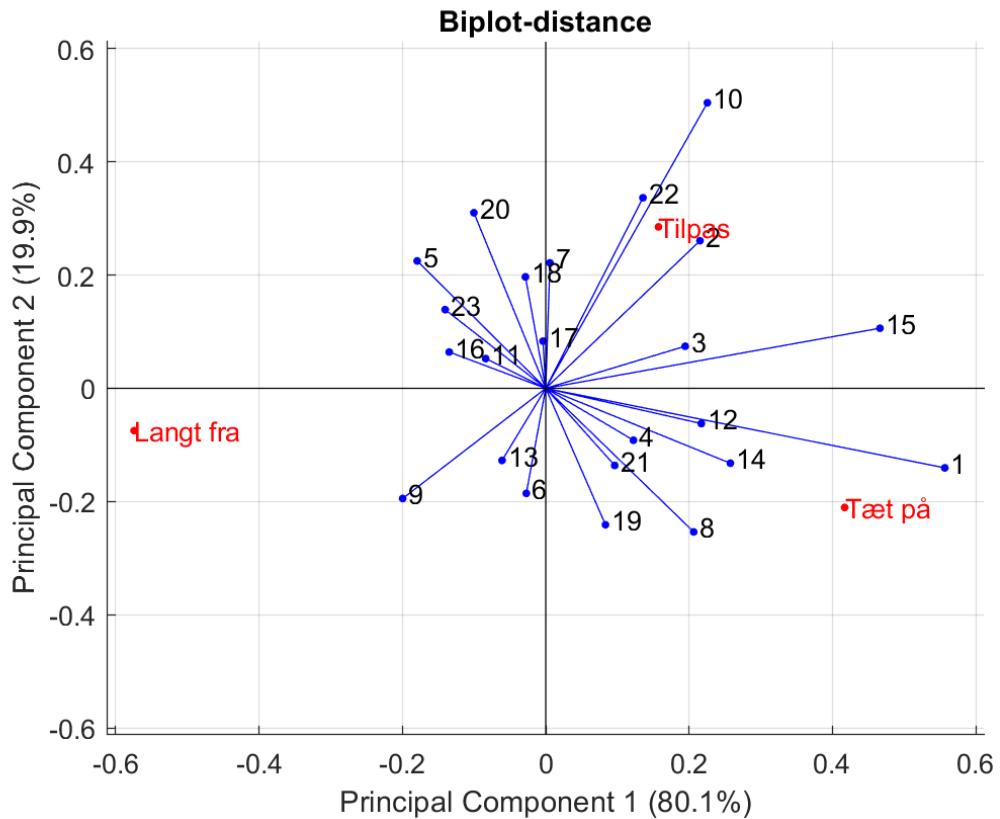
Ud fra *Score*-plottet på Figur 10.23 fremgår det, at der generelt er en del spredning mellem resultaterne ud fra de tre forskellige afstande, hvilket foreskriver at de tre afstande ikke umiddelbart har ens karakteristikas.



Figur 10.23. Score-plot for PC1 og PC2 i forhold til de tre afstande robotten holder til testpersonerne.

I henhold til de tre afstande tyder det ikke på, at afstanden *Langt fra* påvirkes særlig meget af nogen af parametrene, da ingen *loadings* peger direkte mod eller er tæt på denne afstand, jævnfør Figur 10.24. Det tyder derimod mod på, at afstanden *Tilpas* hovedsageligt er domineret af SQ2, SQ10 og SQ22, hvor afstanden *Tæt på* primært domineres af SQ1 og tildels SQ14.

Ud fra *Bi*-plottet på Figur 10.24, hvorpå *loadings* og *scores* for hver parameter er præsenteret, fremgår det at SQ1, omhandlende skærmens reaktion, og SQ15, vedrørende hvor overraskede testpersonerne blev ved robottens henvendelse, bidrager mest til PC1. Hvor for PC2 er det SQ10, vedrørende hvor tryg testpersonerne følte sig ved robotten, som har det største bidrag. Derudover tyder det på at både SQ11, vedrørende hvorvidt robotten gjorde testpersonerne forskrækket, og SQ17, vedrørende robottens elegance, kun beskriver en lille del af variansen, hvorfor det ikke giver mening at kommentere på, hvordan de indbyrdes ligger, medmindre de indgår i en korrelation med en mere betydende parameter.



Figur 10.24. Bi-plot med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens afstand.

Ud fra Figur 10.24 fremgår det, at SQ1, vedrørende skærmens reaktion, og SQ12, vedrørende hvor godt testpersonen kan lide at blive betjent af robotten, ligger tæt på hinanden, hvorfor de er korrelerede. Korrelationen skyldes ikke at de to parametre måler det samme, men formentlig påvirker hinanden, så når skærmen reagerer dårligt så vil det have en negativ effekt på hvor godt testpersonerne kan lide at blive betjent af robotten. Da det forventes at den færdigudviklede robot vil have en skærm, hvis reaktionsevne er væsentligt bedre end hvad der har været tilfældet i projektet, vurderes det at SQ1 formentlig kan være redundant.

Lignende gør sig gældende for SQ10, vedrørende hvor tryg testpersonerne følte sig ved robotten, og SQ22, vedrørende hvor sjov robotten er. Det giver mening, at de to parametre korrelerer da det antages, at hvis testpersonerne føler sig utrygge så vil de formentlig heller ikke opleve robotten som værende sjov. Det vurderes derfor, at de to parametre formentlig ikke måler det samme. For SQ8, vedrørende hvorvidt testpersonerne føler, at robotten kan hjælpe dem, og SQ21, vedrørende hvor anmassende robotten er, forekommer samme tendens; at de to parametre er positivt korreleret. Denne korrelation skyldes formentlig ikke, at de to parametre måler det samme, da anmassende også kan vedrøre at robotten kommer for tæt på, hvilket afspejles ved den negative korrelation med SQ5, vedrørende robottens afstand.

Udover de positive korrelationer forekommer der ligeledes negative korrelationer, hvilket blandt andet gør sig gældende for SQ2, vedrørende robottens bevægelser, og SQ9, vedrørende hvorvidt testpersonerne synes, at robotten stod i vejen. Lignende er gældende

for den positive korrelation mellem SQ10 og SQ22, som har en negativ korrelation med SQ13, vedrørende hvorvidt testpersonerne følte, at robotten fulgte dem hen til det valgte sted. Den negative korrelation, der forekommer mellem SQ10 og SQ13 blev fundet som værende positiv i foregående analyse omkring højden. Ydermere forekommer der negativ korrelation mellem SQ14, vedrørende hvor personlig robottens hjælp opleves, og SQ16, vedrørende hvor irriterende robotten opleves.

Derudover tyder det på, at der er en negativ korrelation mellem SQ19, vedrørende hvor sød robotten opleves, og SQ20, vedrørende hvor dejlig robotten opleves.

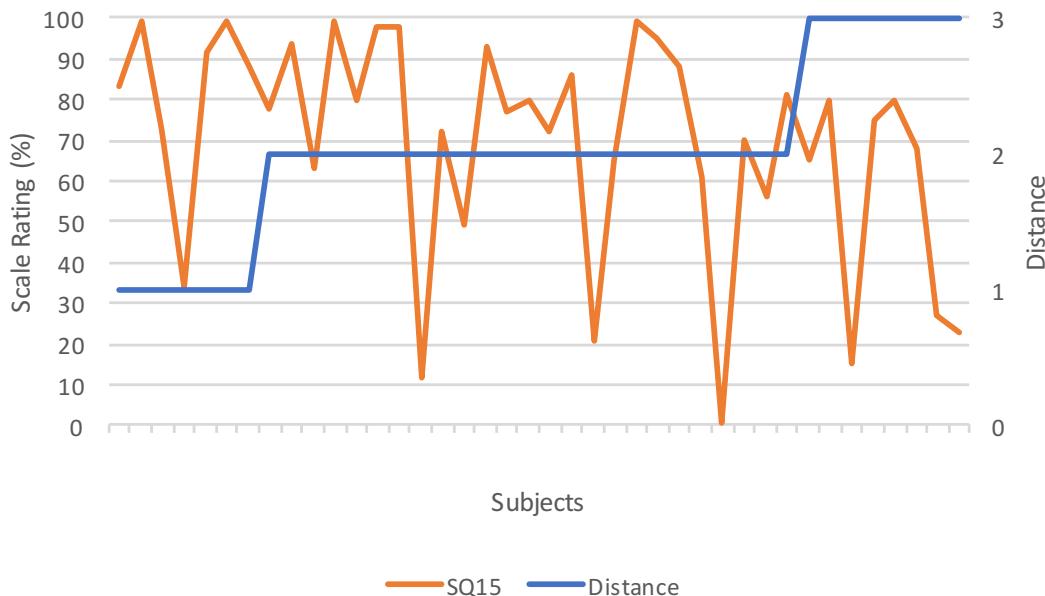
Hver *loading* angiver hvor meget variation hver parameter bidrager med i forhold til den specifikke PC. I Tabel 10.4 fremgår samtlige *loadings* for hver parameter til hver PC.

	PC1 (80.1%)	PC2 (19.9%)
SQ1	0.5567	-0.1405
SQ2	0.2152	0.2608
SQ3	0.1945	0.0744
SQ4	0.1224	-0.0917
SQ5	-0.1793	0.2251
SQ6	-0.0271	-0.1856
SQ7	0.0059	0.2217
SQ8	0.2064	-0.2537
SQ9	-0.1995	-0.1946
SQ10	0.2255	0.5045
SQ11	-0.0838	0.0526
SQ12	0.2171	-0.0623
SQ13	-0.0609	-0.1274
SQ14	0.2575	-0.1323
SQ15	0.4661	0.1062
SQ16	-0.1346	0.0641
SQ17	-0.0038	0.0834
SQ18	-0.0283	0.1969
SQ19	0.0834	-0.2412
SQ20	-0.1002	0.3102
SQ21	0.0962	-0.1363
SQ22	0.1356	0.3367
SQ23	-0.1403	0.1390

Tabel 10.4. Oversigt over *loadings* til de to PCs, for hver PC fremhæves den mest betydende parameter. *SQ* angiver skalaspørgsmålet.

10.9.1 Tendenser i forhold til robottens afstand

Da det ikke har været muligt at kontrollere afstanden mellem robotten og testpersonerne, dels fordi det ikke har været muligt at måle en decideret afstand og dels fordi det kan ændre sig løbende i interaktionen, er det begrænset hvad der kan udledes af tendenser, hvorfor de ikke fremgår her. Dog forefindes samtlige grafer i Afsnit 13.9 (*Tendenser*). Det har dog været muligt, at finde en tendens til at testpersonerne bliver mere overrasket når robotten er tæt på end når den er langt fra, jævnfør Figur 10.25. På figuren henviser afstand 1 til tæt på, afstand 2 til tilpas og afstand 3 til langt fra.



Figur 10.25. Sammenhæng mellem robottens tre afstande og hvad testpersonerne angiver (%) på skalaen til SQ15: *Hvor overrasket blev du over robottens henvendelse?*. Denne graf bygger på 40 besvarelser, da der manglede tre. Afstand 1 svarer til at robotten kommer tæt på, afstand 2 til en tilpas afstand og afstand 3 til at robotten stopper for langt fra.

10.10 Principal Component Analysis: Indgangsvinkel

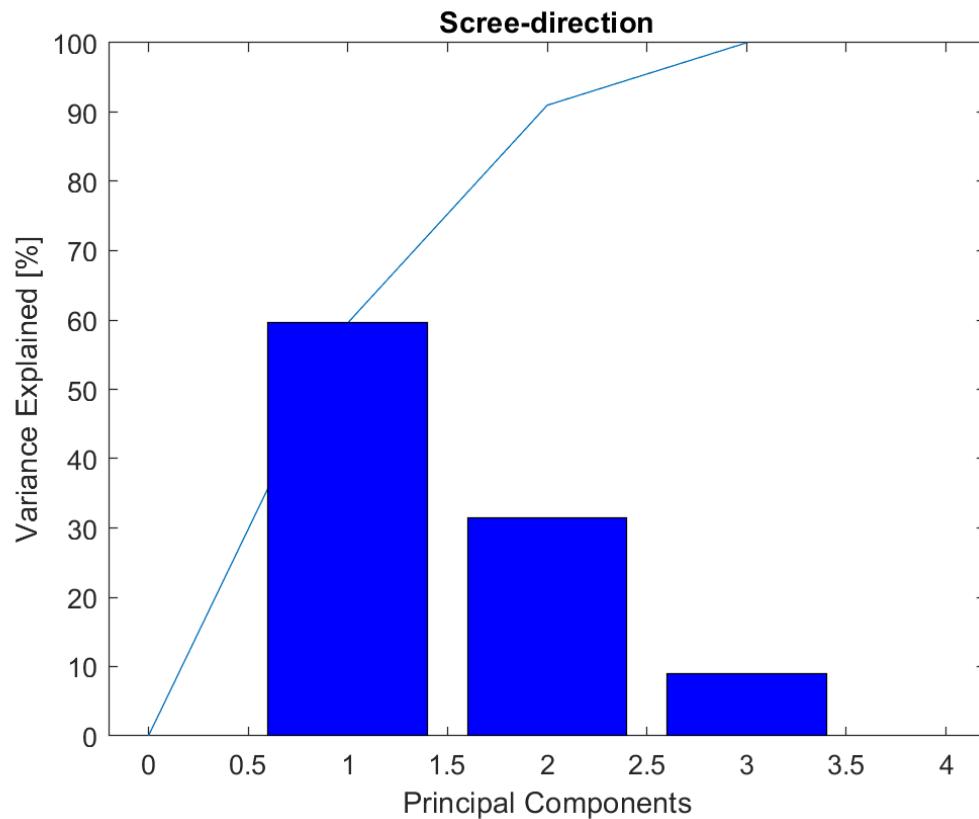
I dette afsnit analyseres resultaterne med udgangspunkt i de tre forskellige indgangsvinkler robotten havde til testpersonerne i testen og hvis testpersonerne selv henvendte sig. De forskellige indgangsvinkler fremgår af Tabel 10.3, hvor antallet af testpersoner ligeledes er opgivet.

Det undersøges, hvordan robottens indgangsvinkel for henvendelse påvirker de rejsendes oplevelse af robotten ved at undersøge, hvordan resultaterne ændrer sig afhængigt af de tre forskellige indgangsvinkler samt testpersonernes egen henvendelse. Det er et *Between-subject* design, da hver testperson kun har oplevet robotten i én af indgangsvinklerne eller selv henvendt sig, og svaret ud fra dette.

Indgangsvinkel	Antal testpersoner
Venstre (1)	4
Forfra (2)	18
Højre (3)	8
TP henvender sig (4)	13

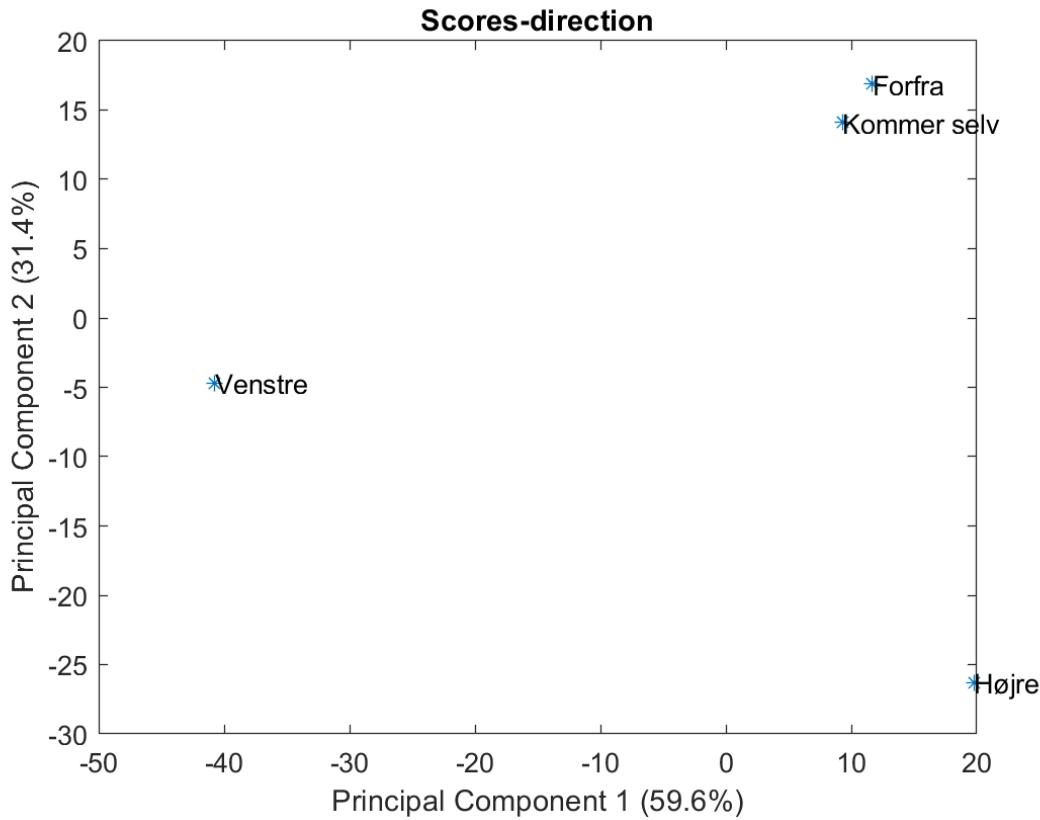
Tabel 10.5. Oversigt over antallet af testpersoner (TP) til hver af de tre indgangsvinkler robotten havde samt hvis testpersonerne selv henvendte sig.

Ud fra *Scree*-plottet på Figur 10.26 fremgår det, at cirka 60 % af variansen kan forklares ud fra den første *Principal Component* (PC) og 90 % af variansen kan forklares ud fra to PCs, hvor PC2 bidrager med cirka 30 %. Gøres der brug af alle tre PCs forklares 100 % af variationen.



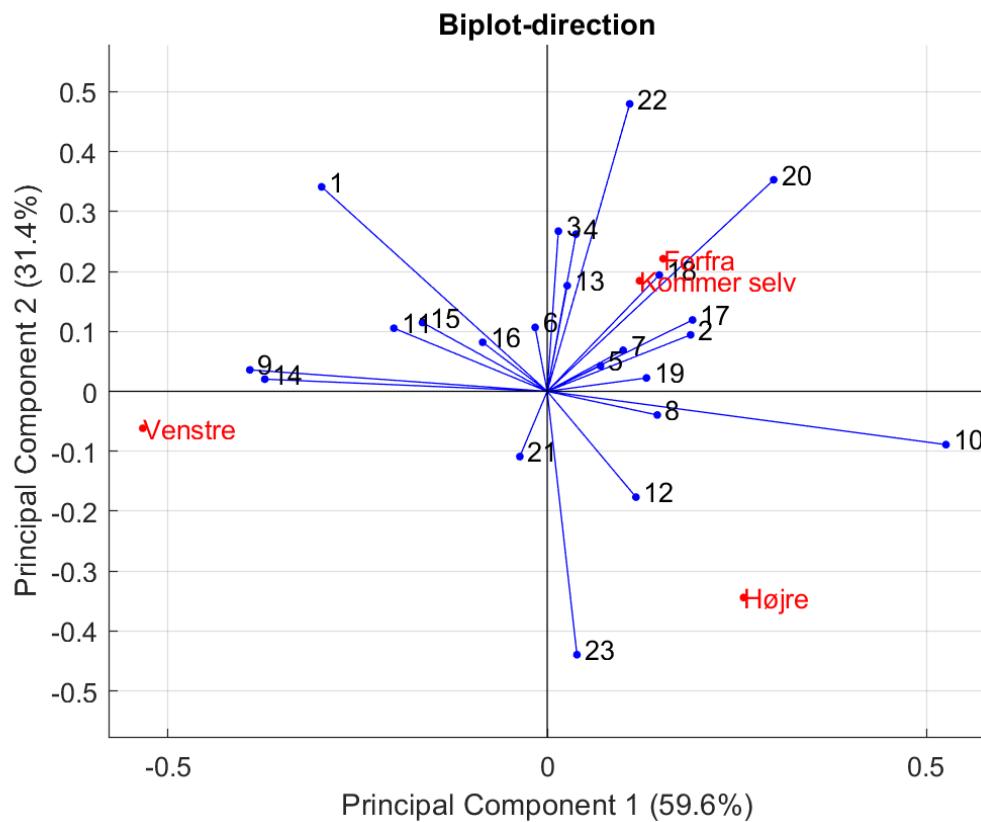
Figur 10.26. Scree-plot, hvorpå sammenhængen mellem antallet af *Principal Components* og *Variance Explained [%]* fremgår. Bidraget for hver PC er som følger: PC1 (59.6%), PC2 (31.4%) og PC3 (9.0%)

Ud fra *Score*-plottet på Figur 10.27 fremgår det, at der generelt er en del spredning mellem resultaterne ud fra de tre forskellige indgangsvinkler, hvilket foreskriver, at de tre indgangsvinkler ikke umiddelbart har ens karakteristikas. Dog ligger *Kommer selv*, svarende til at testpersonerne selv henvendte sig til robotten, meget tæt på *Forfra*, hvilket afspejler lignende karakteristikas for de to *scores*.



Figur 10.27. *Score*-plot for PC1 og PC2 i forhold til de tre indgangsvinkler robotten henvendte sig til testpersonerne med samt hvis testpersonerne selv henvendte sig.

Ud fra *Bi*-plottet på Figur 10.28, hvorpå *loadings* og *scores* for hver parameter er præsenteret, fremgår det at SQ10, vedrørende hvor tryg testpersonerne er ved robotten, forklarer den største varians for PC1. For PC2 fremgår det, at både SQ22, vedrørende hvor sjov robotten opleves samt SQ23, vedrørende hvor menneskelig robotten opleves, forklarer størstedelen af variansen. Da den forklarede varians fra SQ5, vedrørende robottens afstand, SQ6, vedrørende robottens hastighed, SQ7, vedrørende robottens højde, samt SQ16, vedrørende hvor irriterende robotten opleves, er relativt lille vil disse parametre ikke blive yderligere beskrevet, medmindre de indgår i en korrelation med en mere betydnende parameter.



Figur 10.28. *Bi*-plot med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens indgangsvinkel.

I forhold til korrelationen mellem parametrene, fremgår det, at SQ18, vedrørende hvor spændende robotten opleves, og SQ20, vedrørende hvor dejlig testpersonerne opleves, har en positiv korrelation. Den positive korrelation er gennemgående for flere af parametrene, SQ8, vedrørende hvorvidt testpersonerne føler, at robotten kan hjælpe dem, og SQ10, vedrørende hvor tryg testpersonerne er ved robotten, korrelerer ligeledes. Der forefindes en positiv korrelation mellem SQ9, vedrørende hvorvidt robotten stod i vejen, og SQ14, vedrørende hvor personlig testpersonerne oplevede robottens hjælp. SQ4, vedrørende robottens bevægelser, og SQ13, vedrørende hvorvidt testpersonerne følte, at robotten fulgte dem hen til det valgte sted, ligger oven på hinanden, hvorfor det vurderes at de korrelerer. Lignende gør sig gældende for SQ1 og SQ16, som også ligger tæt på Figur 10.28. Selvom SQ5 og SQ7 som tidligere nævnt forklarer en meget lille del af variansen, så korrelerer de begge med SQ17, vedrørende hvor elegant robotten opleves.

Udover de positive korrelationer forekommer der ligeledes negative korrelationer på Figur 10.28. Dette forekommer blandt andet mellem SQ1 og SQ12, vedrørende hvor godt testpersonerne kan lide at blive betjent af robotten. Den negative korrelation forekommer også mellem SQ21, vedrørende hvor anmassende robotten opleves, og SQ22, vedrørende hvor sjov robotten opleves. Den positive korrelation mellem SQ9 og SQ14 er negativ korrelerede med den positive korrelation mellem SQ8 og SQ10. Den sidste negative korrelation, der forefindes på Figur 10.28, er mellem SQ6, vedrørende robottens hastighed, og SQ23, vedrørende hvor menneskelig robotten opleves.

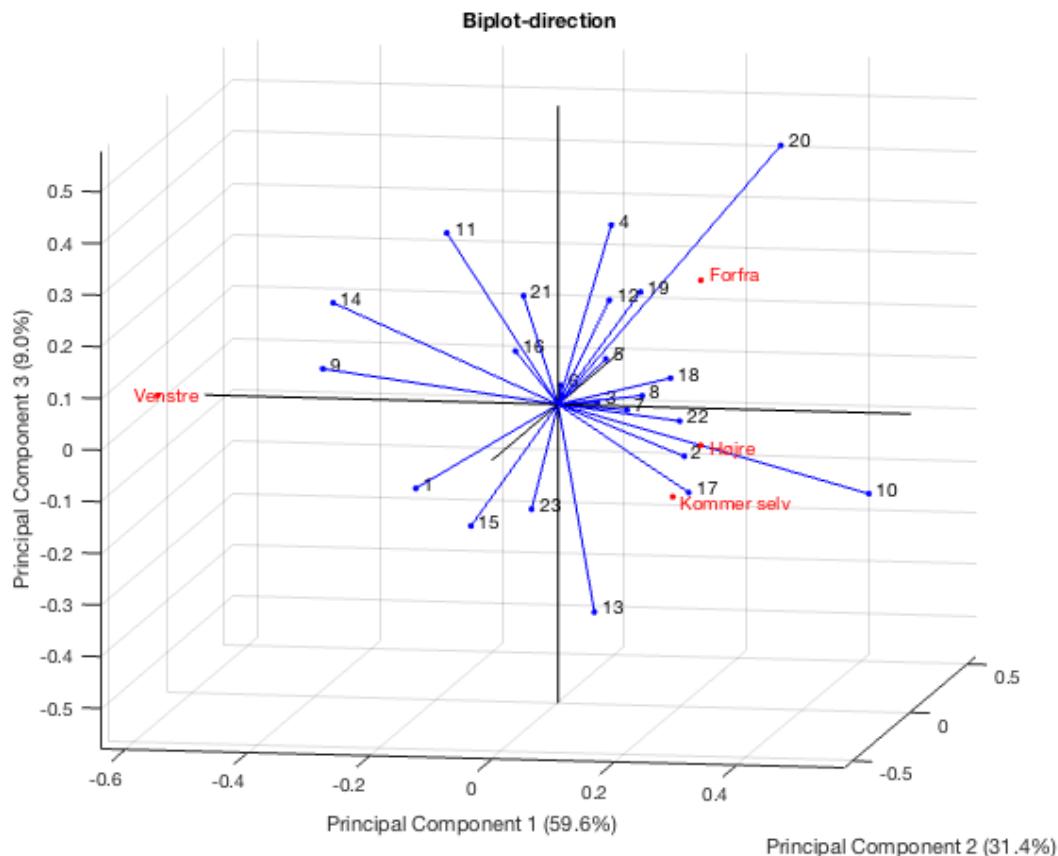
Hver *loading* angiver hvor meget variation hver parameter bidrager med i forhold til den specifikke PC. I Tabel 10.6 fremgår samtlige *loadings* for hver parameter til hver PC.

	PC1 (59.6%)	PC2 (31.4%)	PC3 (9%)
SQ1	-0.2974	0.3412	-0.2362
SQ2	0.1890	0.0942	-0.1122
SQ3	0.0147	0.2673	-0.0454
SQ4	0.0379	0.2624	0.2993
SQ5	0.0702	0.0427	0.0823
SQ6	-0.0159	0.1067	0.0165
SQ7	0.1001	0.0686	-0.0203
SQ8	0.1451	-0.0395	0.0291
SQ9	-0.3921	0.0358	0.0498
SQ10	0.5258	-0.0889	-0.1395
SQ11	-0.2022	0.1053	0.3059
SQ12	0.1170	-0.1765	0.2387
SQ13	0.0264	0.1764	-0.4342
SQ14	-0.3724	0.0203	0.1812
SQ15	-0.1644	0.1147	-0.2624
SQ16	-0.0851	0.0818	0.0851
SQ17	0.1916	0.1191	-0.1870
SQ18	0.1478	0.1942	0.0192
SQ19	0.1309	0.0223	0.2178
SQ20	0.2987	0.3530	0.4448
SQ21	-0.0361	-0.1089	0.2298
SQ22	0.1088	0.4797	-0.1187
SQ23	0.0392	-0.4395	-0.1192

Tabel 10.6. Oversigt over *loadings* til de tre PCs, for hver PC fremhæves den mest betydnende parameter. *SQ* angiver skalaspørgsmålet.

Tages der derimod udgangspunkt i et tredimensionelt *Bi*-plot, jævnfør Figur 10.29, fremgår det, at scoren for *Forfra* og *Kommer selv* ikke ligger lige så tæt på Figur 10.28, men derimod ligger *Kommer selv* tættere på *Højre*. Det tredimensionelle *Bi*-plot er vedlagt Afsnit 13.8 (*3D Bi-plots*). Derudover er mange af de fornævnte positive korrelationer ikke længere korrelerede. På det tredimensionelle *Bi*-plot, jævnfør Figur 10.29, fremgår det, at SQ1 og SQ16 i højere grad er ortogonale, hvilket gengiver at de ikke korrelerer. Lignende tendens forekommer ved SQ5 og SQ17, som er ortogonale fremfor korrelerede. I tillæg er SQ17 og SQ7 ikke længere placeret oven på hinanden, dog forekommer der en lille korrelation. Der forekommer ikke længere end positiv korrelation mellem SQ4 og SQ13

i det tredimensionelle *Bi*-plot, jævnfør Figur 10.29, korrelationen er nærmere negativ. Korrelationen mellem SQ18 og SQ20 er heller ikke lige så høj, som tidligere, men er stadig tilstede værende. På Figur 10.28 fremgår det at SQ21 er negativ korreleret med SQ22, dette er dog ikke tilfældet i Figur 10.29, hvor de to parametre i højere grad er ortogonale end negativ korrelerede. Dog forekommer der en negativ korrelation mellem SQ21 og SQ13.



Figur 10.29. 3D *Bi*-plot med både *loadings* (angivet med blå) og *scores* (angivet med rød) fremgår i forhold til robottens indgangsvinkel eller hvis testpersonerne selv henvendte sig.

10.10.1 Tendenser i forhold til robottens indgangsvinkel

Ligesom tilfældet med robottens afstand til testpersonerne, har det ikke været muligt at kontrollere indgangsvinklen, blandt andet fordi det afhænger af, hvor robotten er inden den henvender sig til en rejsende eller inden en rejsende henvender sig til den. Det har derfor ikke været muligt at spore nogle tendenser, da der har været meget variation blandt besvarelserne. Samtlige grafer forefindes i Afsnit 13.9 (*Tendenser*).

10.11 Sammenligning af korrelerede parametre

Som beskrevet udføres der en PCA analyse relativt til forskellige grupperinger. Ud fra disse PCA analyser forefindes der korrelation mellem nogen af de fundne og testede parametre. For yderligere at undersøge disse korrelationer opsættes relevante grafer, hvor forholdet mellem de forskellige parametre og skalaspørgsmål visualiseres. På alle grafer i det følgende afsnit vil x-aksen indikere testpersoner, men ikke nødvendigvis i kronologisk rækkefølge, hvorfor værdier på x-aksen er fjernet. Dertil er manglende besvarelser ekskluderet fra graferne, hvorfor antallet af testpersonerne varierer. Selvom dette afsnit er inddelt i tre underafsnit for henholdsvis: Højde, afstand og indgangsvinkel, hvori korrelationerne fra den foregående PCA præsenteres, så bygger graferne på det samlede datasæt og tager derfor ikke højde for hverken højde, afstand eller indgangsvinkel. Samtlige grafer forefindes i Afsnit 13.10 (*Korrelationsgrafer*).

10.11.1 Korrelerede parametre fra højde

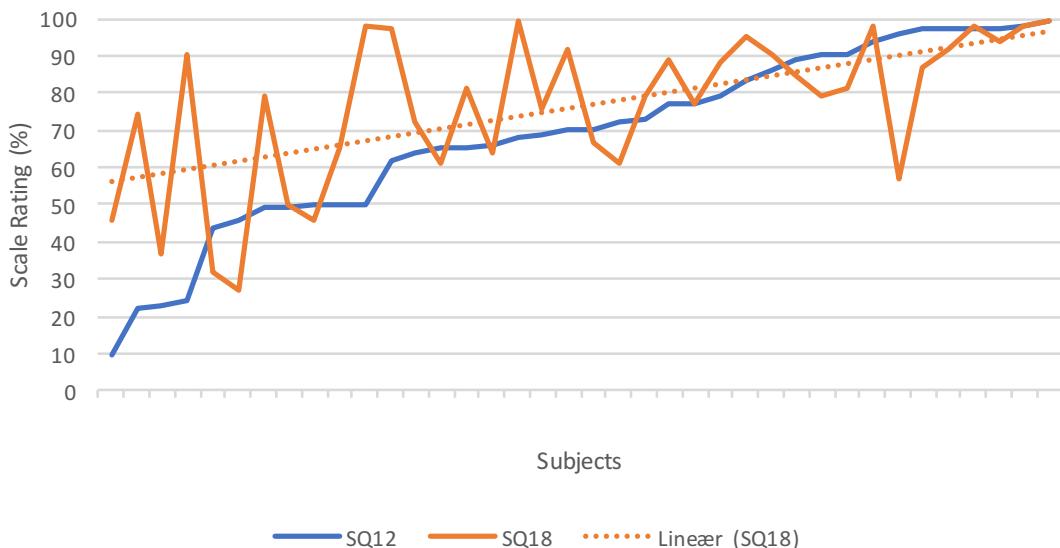
Når PCA analysen udføres relativt til højde, tyder det, som beskrevet i Afsnit 10.8 (*Principal Component Analysis: Robottens højde*), på, at en positiv korrelation mellem følgende parametre forekommer:

- SQ12 og SQ18
- SQ10 og SQ13
- SQ14 og SQ15
- SQ8 og SQ17
- SQ20 og SQ22

Derudover forefindes der negativ korrelation mellem følgende parametre:

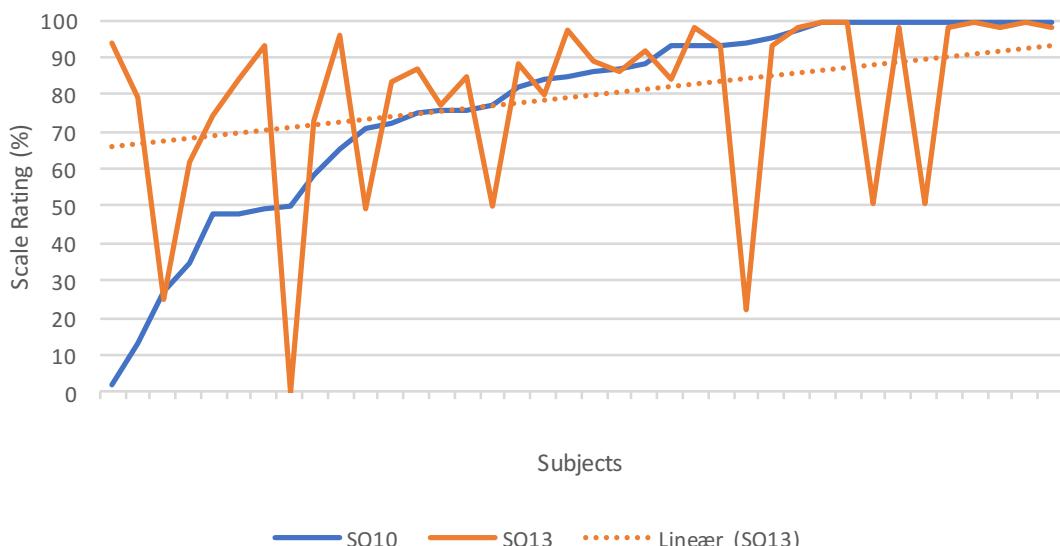
- SQ12 og S21
- SQ18 og SQ21
- SQ2 og SQ9
- SQ4 og SQ9
- SQ16 og SQ19

Korrelationen mellem SQ12 og SQ18 skyldes formentlig ikke, at de mäter det samme, men derimod, at når testpersonerne godt kan lide at blive betjent af robotten, så synes de også robotten er spændende, hvilket fremgår på Figur 10.30.



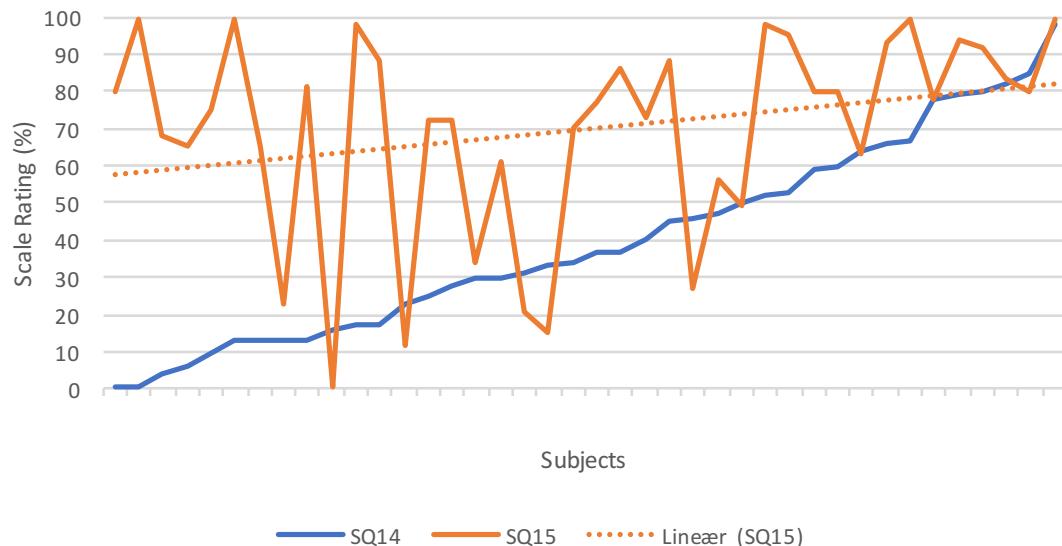
Figur 10.30. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ12: *Jeg kan godt lide at blive betjent af robotten* og SQ18: *Hvad synes du om robotten?*, i forhold til spændende. Denne graf bygger på 38 besvarelser, da der manglede fem.

Sammenholdes SQ10 med SQ13 forekommer der ud fra Figur 10.31, tiltrods for stor variation i besvarelserne til SQ13, en positiv korrelation. Dette indikerer derfor, at når testpersonerne regnede med at robotten fulgte dem hen til det valgte sted, så var de også mere trygge ved robotten.



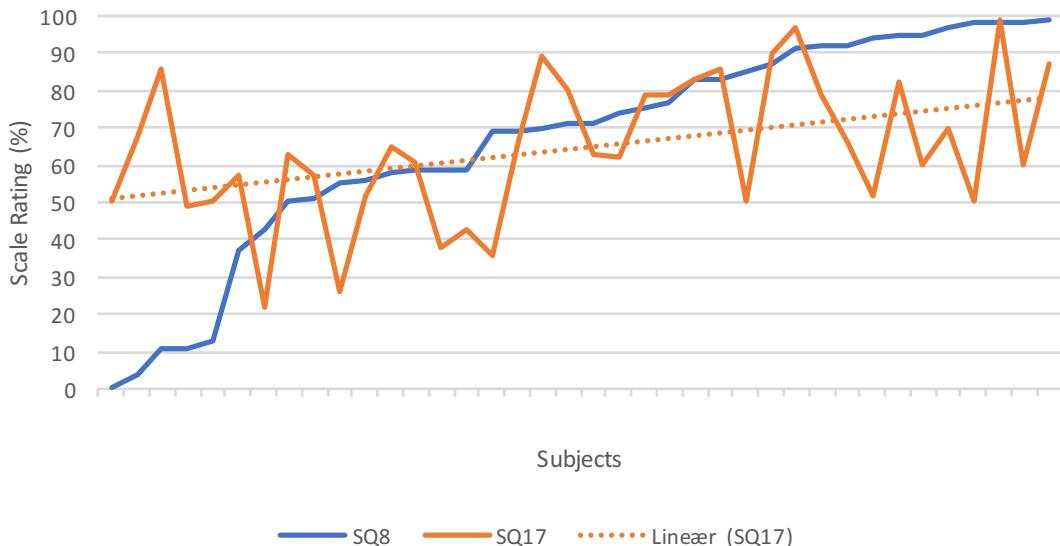
Figur 10.31. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ10: *Jeg føler mig tryg ved robotten* og SQ13: *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte*. Denne graf bygger på 38 besvarelser, da der manglede fem.

Korrelationen mellem SQ14 og SQ15 skyldes formentlig ikke, at de mäter det samme, og ud fra tendenslinjen på Figur 10.32 tyder det på, at der tilnærmelsesvis er en sammenhæng mellem de to parametre. Fokuseres der derimod på punkterne på graferne, virker det ikke som om, at der er en korrelation mellem hvor overrasket testpersonerne bliver og hvor personlig robottens hjælp opleves.



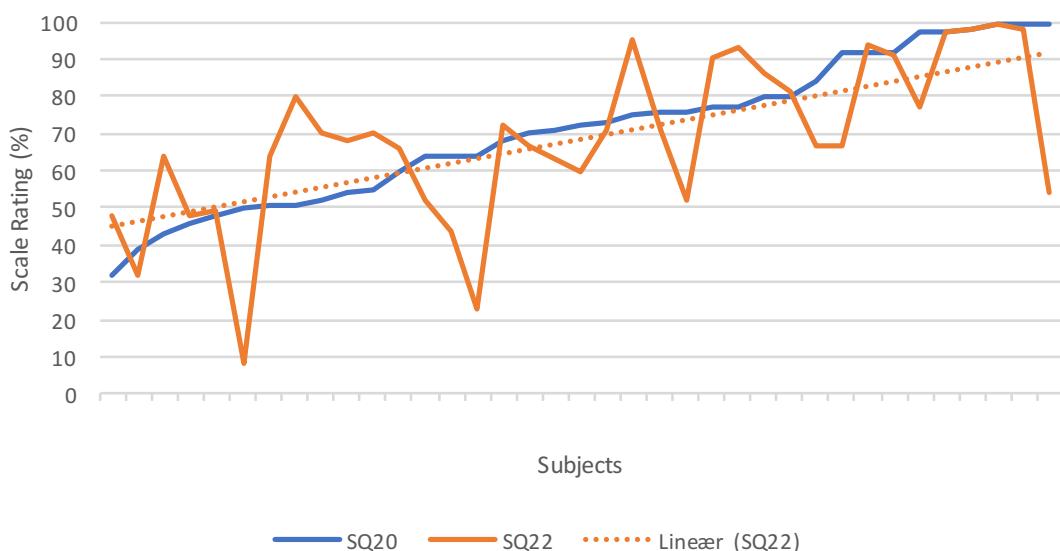
Figur 10.32. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ14: *Hvor personlig oplevede du robottens hjælp?* og SQ15: *Hvor overrasket blev du over robottens henvendelse?*. Denne graf bygger på 40 besvarelser, da der mangler tre.

Korrelationen mellem SQ8 og SQ17 skyldes formentlig ikke, at de mäter det samme, det tyder derimod på, at korrelationen forekommer når testpersonerne føler, at robotten kan hjælpe dem og samtidig vurderer at robotten er elegant. Dog er denne korrelation lille, fordi datapunkterne varierer en del, særligt i forhold til SQ17, jævnfør Figur 10.33.



Figur 10.33. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ8: *Jeg føler, at robotten kan hjælpe mig* og SQ17: *Hvad synes du om robotten?*, i forhold til elegant. Denne graf bygger på 38 besvarelser, da der manglede fem.

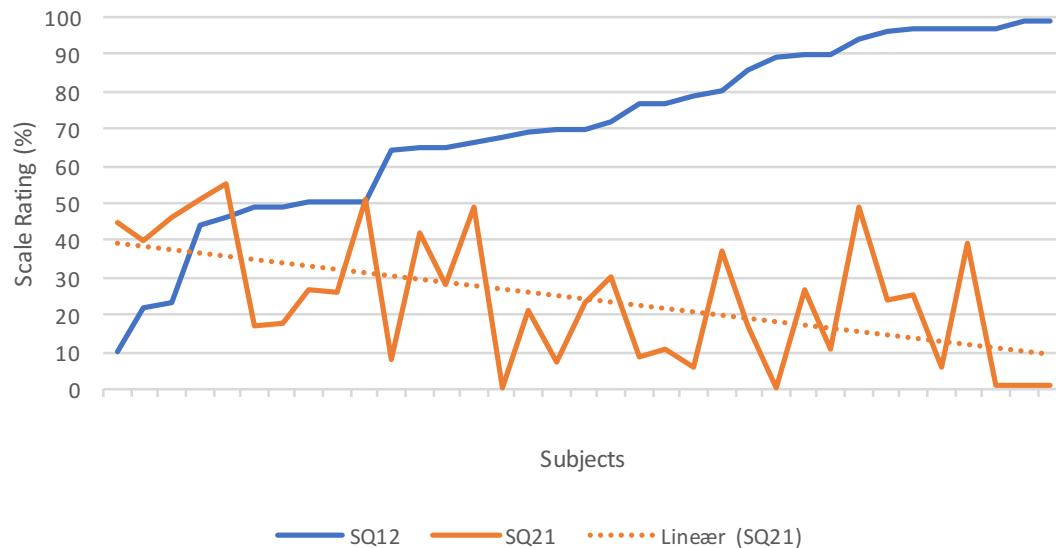
Korrelationen mellem SQ20, vedrørende hvor dej robotten opleves, og SQ22, vedrørende hvor sjov robotten er, kan tyde på, at parametrene mäter det samme, da testpersonerne synes robotten er mere sjov, desto mere dej den er. Datapunkterne varierer dog en del i forhold til SQ22, jævnfør Figur 10.34.



Figur 10.34. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ20 og SQ22, der begge har samme skalaspørgsmål: *Hvad synes du ellers om robotten?*, som vedrører henholdsvis dej og sjov. Denne graf bygger på samtlige 37 besvarelser, da der manglede seks.

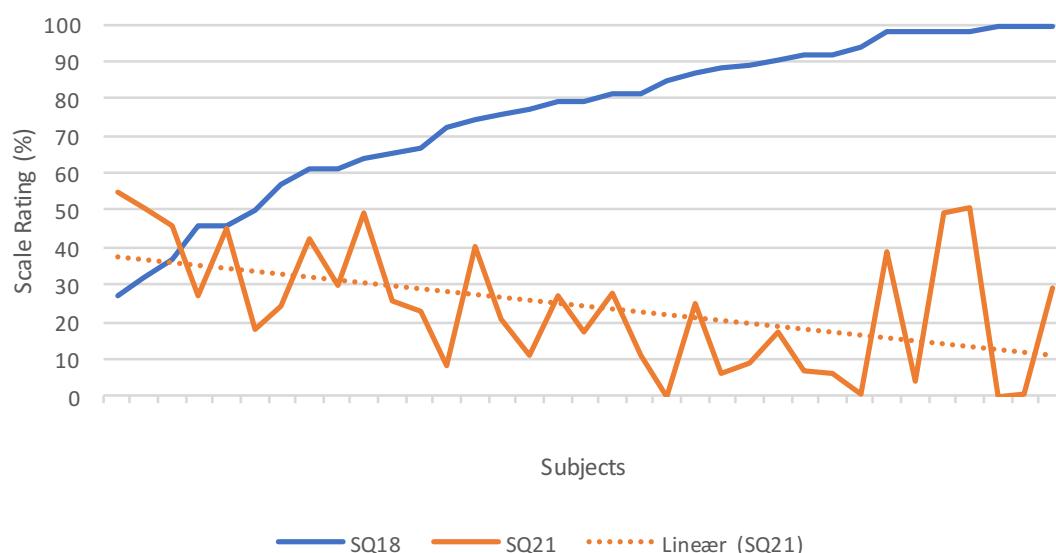
I forhold til den negative korrelation mellem SQ21 og SQ12, vedrørende hvor godt testpersonerne kan lide at blive betjent af robotten, fremgår det ud fra Figur 10.35, at

desto mindre anmassende robotten opleves, desto bedre kan testpersonerne lide at blive betjent af robotten, hvilket giver god mening.



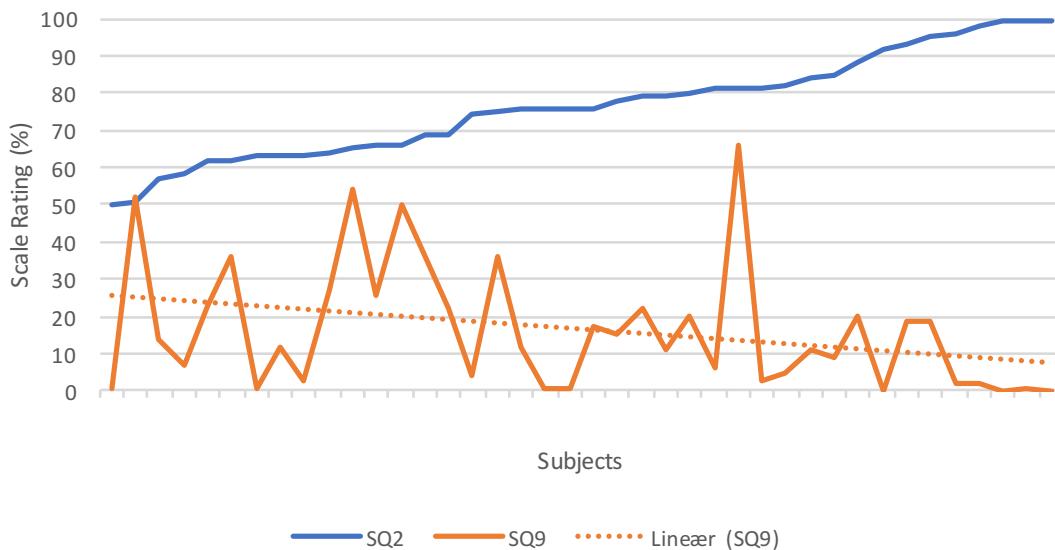
Figur 10.35. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ12: *Jeg kan godt lide at blive betjent af robotten* og SQ21: *Hvad synes du ellers om robotten?*, i forhold til *anmassende*. Denne graf bygger på 38 besvarelser, da der mangler fem.

I forhold til den negative korrelation mellem SQ18, vedrørende hvor spændende robotten opleves, og SQ21 fremgår det ud fra Figur 10.36, at desto mindre anmassende robotten opleves desto mere spændende vurderes robotten. Dog tyder det på, når SQ18 vurderes højest så stiger vurderingen på SQ21 ligeledes.



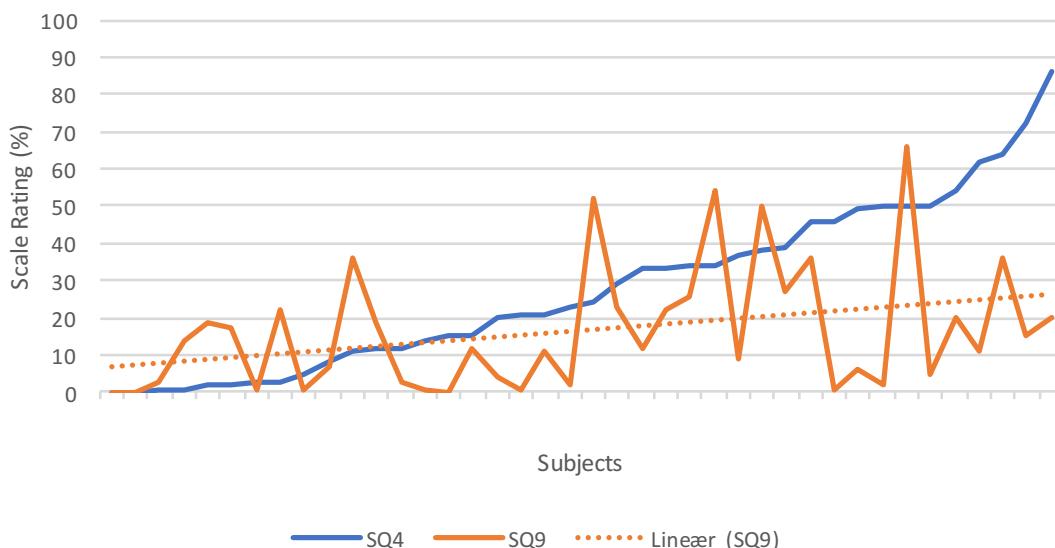
Figur 10.36. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ18: *Hvad synes du om robotten?*, i forhold til *spændende*, og SQ21: *Hvad synes du ellers om robotten?*, i forhold til *anmassende*. Denne graf bygger på 35 besvarelser, da der mangler otte.

I henhold til den negative korrelation mellem SQ2, vedrørende hvorvidt robotten er imødekommen eller afvisende, og SQ9 fremgår det, at desto mere imødekommen robotten opleves, desto mindre er den i vejen, jævnfør Figur 10.37.



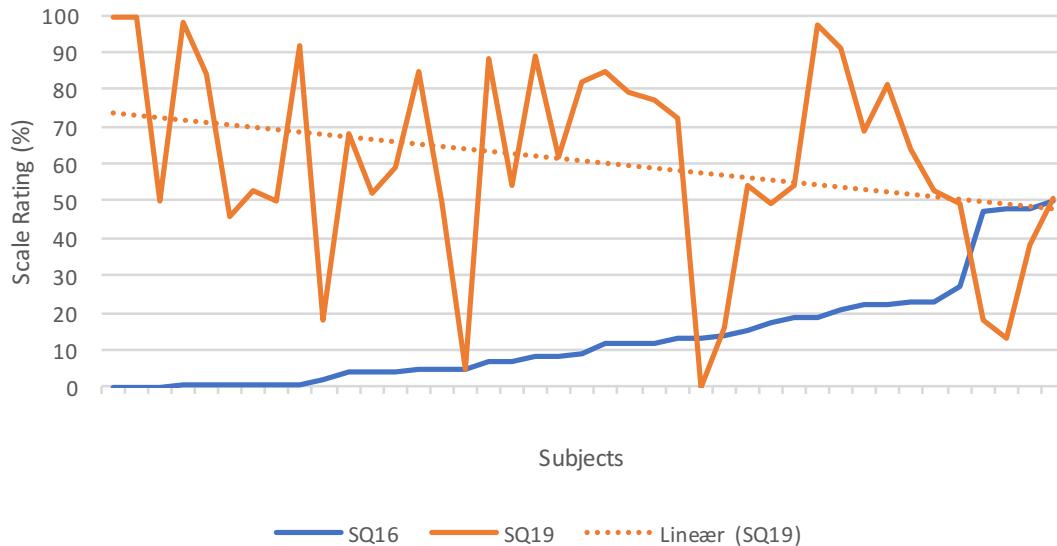
Figur 10.37. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ2: *Hvordan oplevede du robotten?*, i forhold til *afvisende* og *imødekommen*, og SQ9: *Jeg synes, at robotten stod i vejen*. Denne graf bygger på 40 besvarelser, da der mangler 3.

Selvom det på Figur 10.14 tyder på, at der er en negativ korrelation mellem SQ4, vedrørende robottens bevægelser, og SQ9 så fremgår denne korrelation ikke som værende negativ på Figur 10.15 eller på Figur 10.38. På Figur 10.38 tyder det nærmere på, at der er en svag positiv korrelation.



Figur 10.38. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ4: *Hvordan oplevede du robottens bevægelser?*, i forhold til *vilde* og *rolige* og SQ9: *Jeg synes, at robotten stod i vejen*. Denne graf bygger på 40 besvarelser, da der mangler 3.

Den negative korrelation mellem SQ16 og SQ19 forekommer både på Figur 10.14 og på Figur 10.15, men korrelationen er ikke lige så markant på Figur 10.39, da datapunkterne for SQ19 varierer en del. Det er derfor ikke muligt endegyldigt, at vurdere om der er en korrelation eller om det er en tilfældighed.



Figur 10.39. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ16 og SQ19, der begge har samme skalaspørgsmål: *Hvad synes du om robotten?*, men vedrører henholdsvis *irriterende* og *sød*. Denne graf bygger på samtlige 43 besvarelser.

10.11.2 Korrelerede parametre fra afstand

Når PCA analysen udføres relativt til afstand, tyder det, som beskrevet i Afsnit 10.9 (*Principal Component Analysis: Afstand*), på, at en positiv korrelation mellem følgende parametre forekommer:

- SQ1 og SQ12
- SQ7 og SQ17
- SQ10 og SQ22
- SQ8 og SQ21

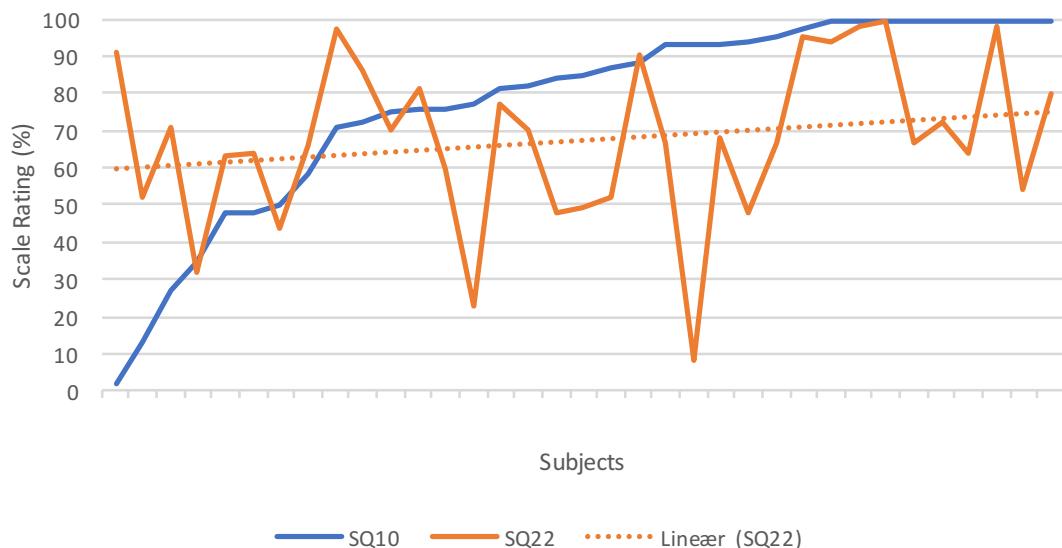
Derudover forekommer der negativ korrelation mellem følgende parametre:

- SQ2 og SQ9
- SQ5 og SQ8
- SQ5 og SQ21
- SQ10 og SQ13
- SQ13 og SQ22
- SQ14 og SQ16
- SQ19 og SQ20

Sammenholdes SQ1 og SQ12 tyder det ikke på, at der er en korrelation mellem dem, hvorfor grafen forefindes i Afsnit 13.10 (*Korrelationsgrafer*). Selvom SQ17 kun har en

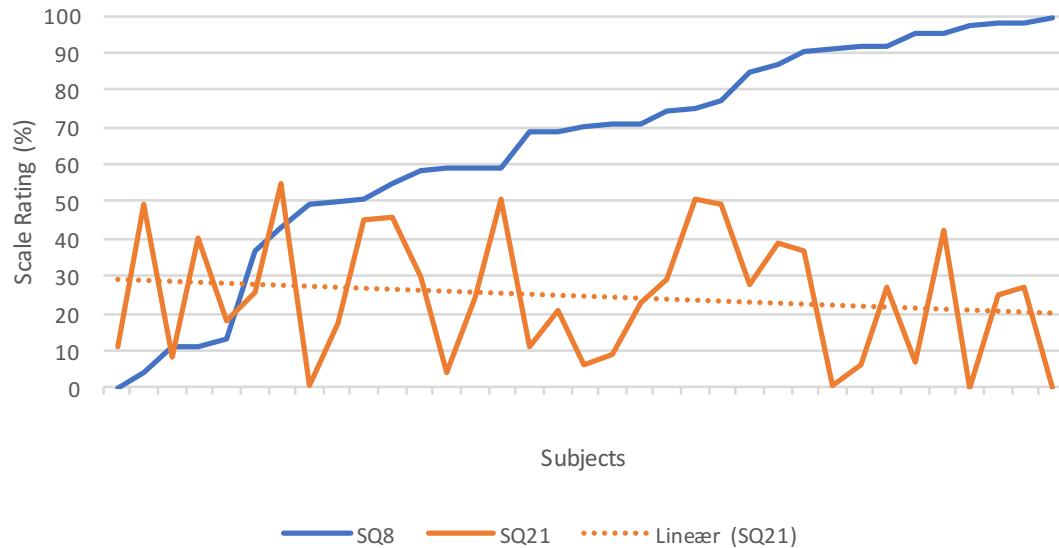
minimal betydningen, fremgår det på Figur 10.24, at SQ17 er højt korreleret med SQ7, vedrørende robottens højde. Sammenholdes de to skalaspørsgsmål tyder det ikke på, at der er en korrelation mellem dem, hvorfor grafen forefindes i Afsnit 13.10 (*Korrelationsgrafer*).

Sammenholdes SQ10 og SQ22 tyder det ikke på, at der er en korrelation mellem hvor tryg testpersonerne er ved robotten og hvor sjov robotten opleves, jævnfør Figur 10.40.



Figur 10.40. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ10: *Jeg føler mig tryg ved robotten* og SQ22: *Hvad synes du ellers om robotten?* i forhold til sjov. Denne graf bygger på 35 besvarelser, da der manglede otte.

Da der forefindes en positiv korrelation mellem SQ8 og SQ21 på Figur 10.24, forventes det, at tendenslinjen stiger, men når de to skalaspørsgsmål sammenholdes på Figur 10.41 tyder det ikke på, at det er tilfældet, jævnfør ingen korrelation.



Figur 10.41. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ8: *Jeg føler, at robotten kan hjælpe mig* og SQ21: *Hvad synes du ellers om robotten?* i forhold til anmassende. Denne graf bygger på 35 besvarelser, da der manglede otte.

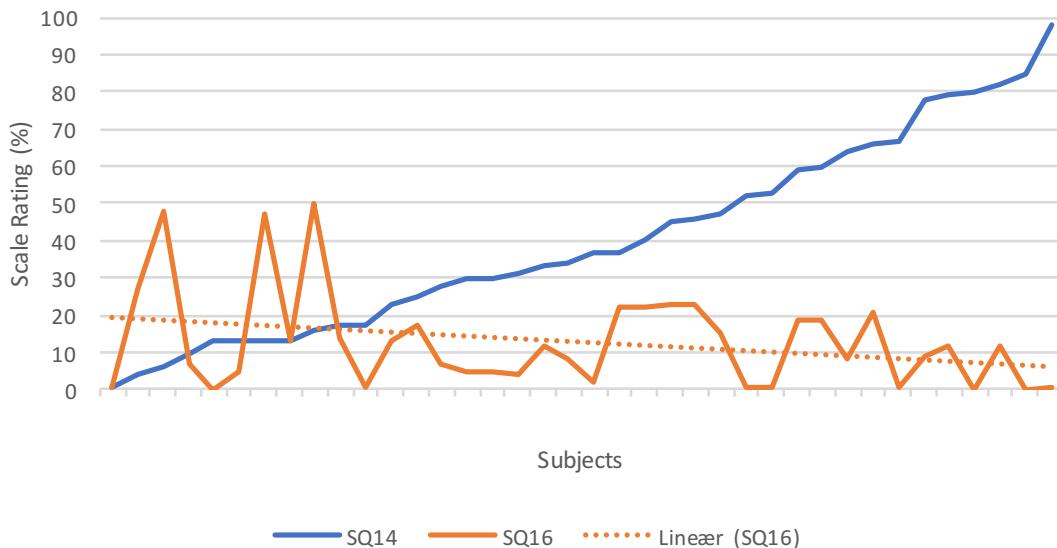
Den negative korrelation mellem SQ2 og SQ9 er præsenteret på Figur 10.37. Sammenlignes SQ5 og SQ8, jævnfør Figur 10.42, er det svært at udlede, hvorvidt der er en korrelation mellem de to parametre, da der forekommer stor variation i besvarelserne til SQ8. På trods af tidligere fundet negativ korrelation tyder det ikke på, at testpersonerne har en ændret oplevelse af om robotten kan hjælpe dem afhængigt af dens afstand. Det samme gør sig gældende, når SQ5 og SQ21 sammenlignes, hvorfor grafen forefindes i Afsnit 13.10 (*Korrelationsgrafer*).



Figur 10.42. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ5: *Jeg synes robotten stoppede...* og SQ8: *Jeg føler, at robotten kan hjælpe mig*. Denne graf bygger på 40 besvarelser, da der manglede tre.

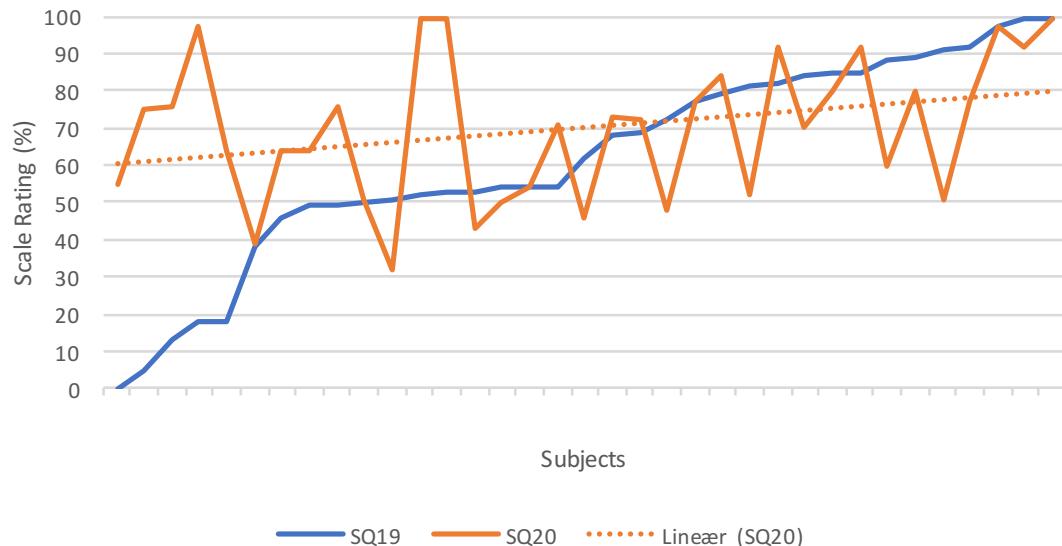
Lignende er gældende for korrelationen mellem SQ13 og SQ22, hvor der heller ikke forefindes en negativ korrelation, men derimod tyder det på, at når testpersonerne regnede med at robotten fulgte dem hen til det valgte sted, så var de mere tilbøjelige til at vurdere robotten som værende mere sjov. Dog forekommer der en del manglende besvarelser til SQ22, hvorfor grafen for dette forefindes i Afsnit 13.10 (*Korrelationsgraffer*).

Sammenholdes SQ14 med SQ16 tyder det ikke på, at der forekommer den før fundne negative korrelation, jævnfør Figur 10.43.



Figur 10.43. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ14: *Hvor personlig oplevede du robottens hjælp?* og SQ16: *Hvad synes du om robotten?*, i forhold til *irriterende*. Denne graf bygger på 38 besvarelser, da der manglede fem.

Sammenholdes SQ19 og SQ20 forefindes der ikke en negativ korrelation, men nærmere en postiv korrelation, jævnfør Figur 10.44.



Figur 10.44. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ19: *Hvad synes du om robotten?*, i forhold til *sød* og SQ20: *Hvad synes du ellers om robotten?*, i forhold til *sej*. Denne graf bygger på 35 besvarelser, da der manglede otte.

10.11.3 Korrelerede parametre fra indgangsvinkel

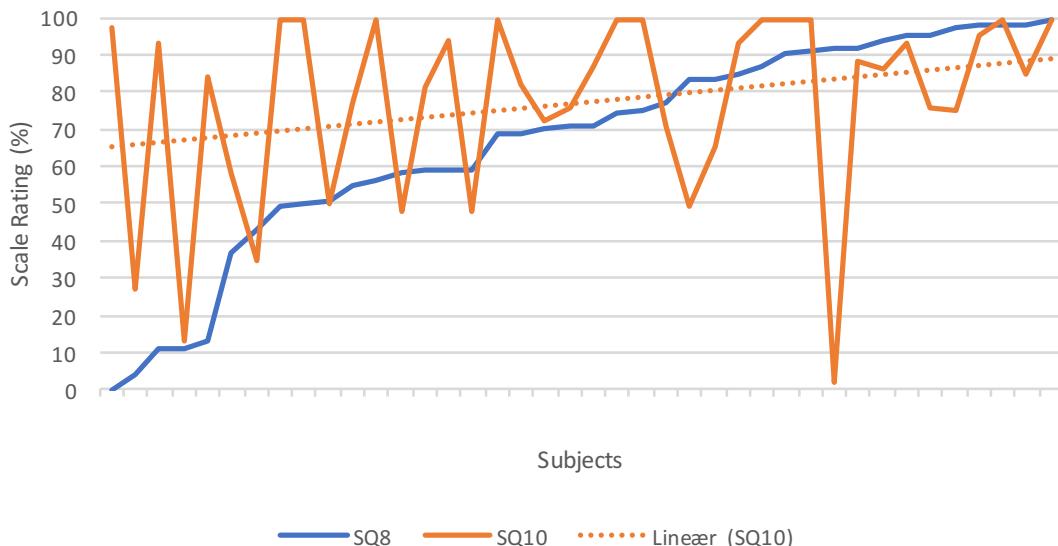
Når PCA analysen udføres relativt til indgangsvinkel, tyder det, som beskrevet i Afsnit 10.10 (*Principal Component Analysis: Indgangsvinkel*), på, at en positiv korrelation mellem følgende parametre forekommer:

- SQ8 og SQ10
- SQ18 og SQ20
- SQ9 og SQ14
- SQ5 og SQ7

Derudover forekommer der negativ korrelation mellem følgende parametre:

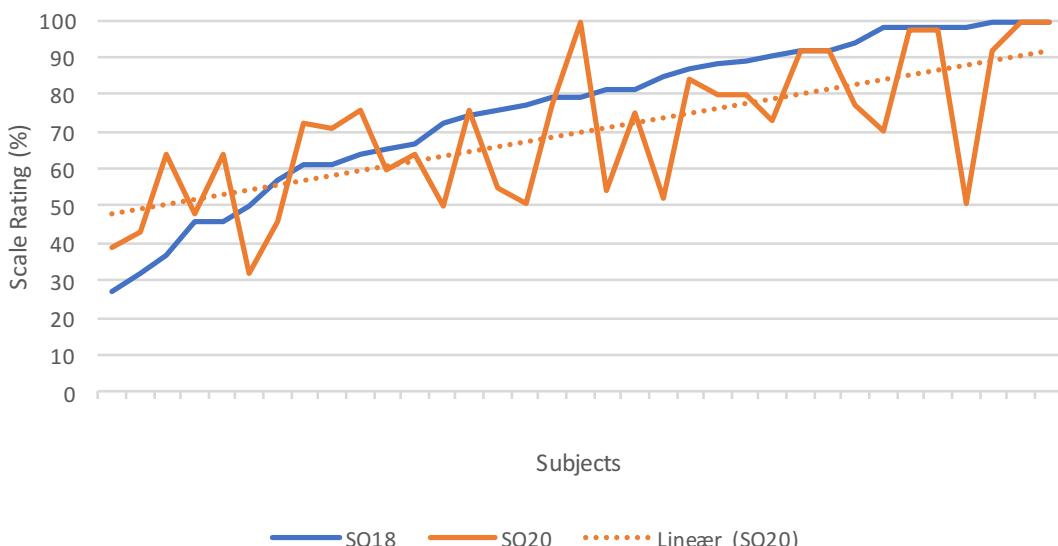
- SQ1 og SQ12
- SQ9 og SQ10
- SQ10 og SQ14
- SQ6 og SQ23
- SQ13 og SQ21

Sammenholdes SQ8 med SQ10 tyder det på, at der er en positiv korrelation mellem de to skalaspørgsmål, jævnfør Figur 10.45. Dette medfører, at desto mere testpersonerne føler, at robotten kan hjælpe dem, desto tryggere er de ved robotten.



Figur 10.45. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ8: *Jeg føler, at robotten kan hjælpe mig*, og SQ10: *Jeg føler mig tryg ved robotten*. Denne graf bygger på 40 besvarelser, da der manglede tre.

Sammenholdes SQ18 og SQ20, jævnfør Figur 10.46, tyder det også på, at det er en positiv korrelation mellem de to skalaspørgsmål. Dette medfører, at desto mere spændende testpersonerne finder robotten, desto mere dej synes de også den er.

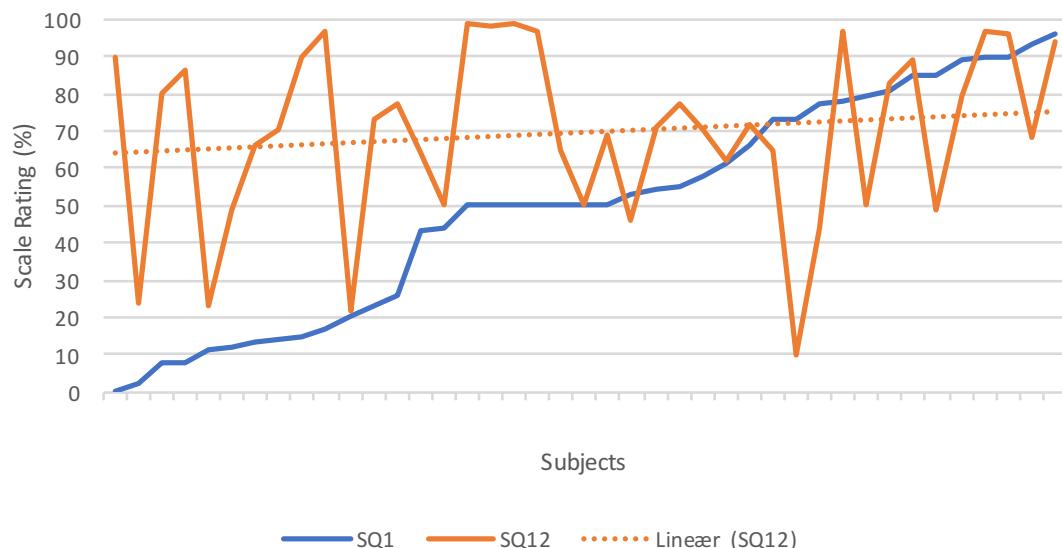


Figur 10.46. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ18: *Hvad synes du om robotten?*, i forhold til spændende og SQ20: *Hvad synes du ellers om robotten?*, i forhold til dej. Denne graf bygger på 35 besvarelser, da der manglede otte.

Selvom det ud fra Figur 10.28 tyder på, at SQ9 og SQ14 er positivt korreleerde, så tyder det ikke på, at der er en korrelation mellem de to parametre, hvorfor grafen forefindes i Afsnit 13.10 (*Korrelationsgrafer*). Det samme gør sig gældende for SQ5 og SQ7, hvor det ikke tyder på, at der forekommer en korrelation, jævnfør Afsnit 13.10 (*Korrelationsgrafer*).

For to af de negative korrelationer, tyder det ikke på, at der reelt er en korrelation. Dette gør sig gældende for SQ10 og SQ14 samt SQ6 og SQ23, jævnfør Afsnit 13.10 (*Korrelationsgrafer*).

Ud fra Figur 10.28 fremgår det, at SQ1 og SQ12 er negativ korreleret, men når de to skalaspørgsmål sammenholdes tyder det ikke på, at der overhovedet forekommer en korrelation, jævnfør Figur 10.47.



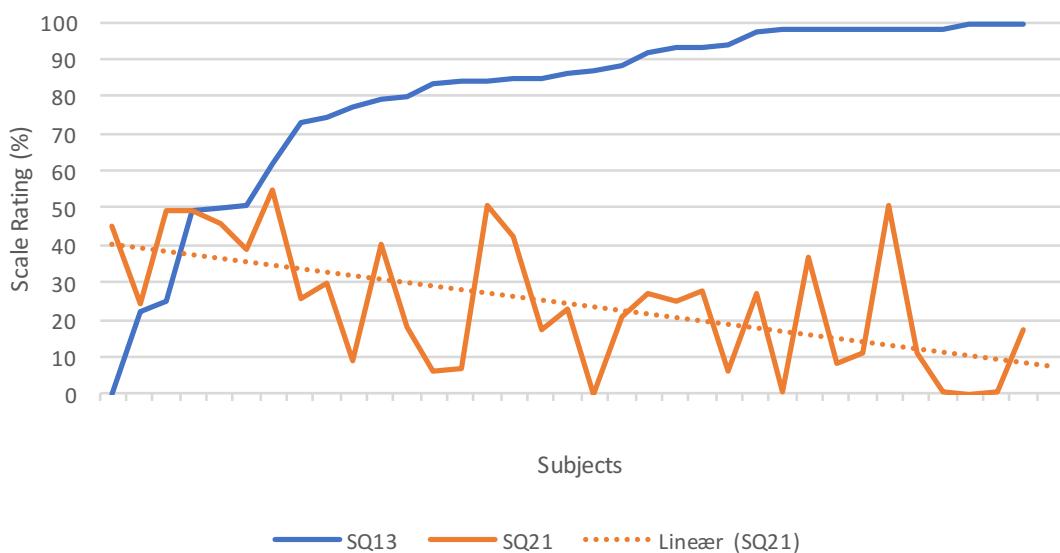
Figur 10.47. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ1: *Hvordan synes du skærmen reagerede*, og SQ12: *Jeg kan godt lide at blive betjent af robotten*. Denne graf bygger på 41 besvarelser, da der manglede to.

Sammenlignes SQ9 og SQ10, jævnfør Figur 10.48, forekommer der en lille korrelation relateret til at desto mere trygge testpersonerne er ved robotten, desto mindre synes de, at den stod i vejen.



Figur 10.48. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ9: *Jeg synes, at robotten stod i vejen* og SQ10: *Jeg føler mig tryg ved robotten*. Denne graf bygger på 40 besvarelser, da der manglede tre.

På Figur 10.49 sammenlignes SQ13 og SQ21, hvor den negative korrelation vurderes til at være lille. Korrelationen vedrører at testpersonerne oplever robotten som værende mindre anmassende, desto mere de regner med at robotten følger dem hen til det valgte sted.

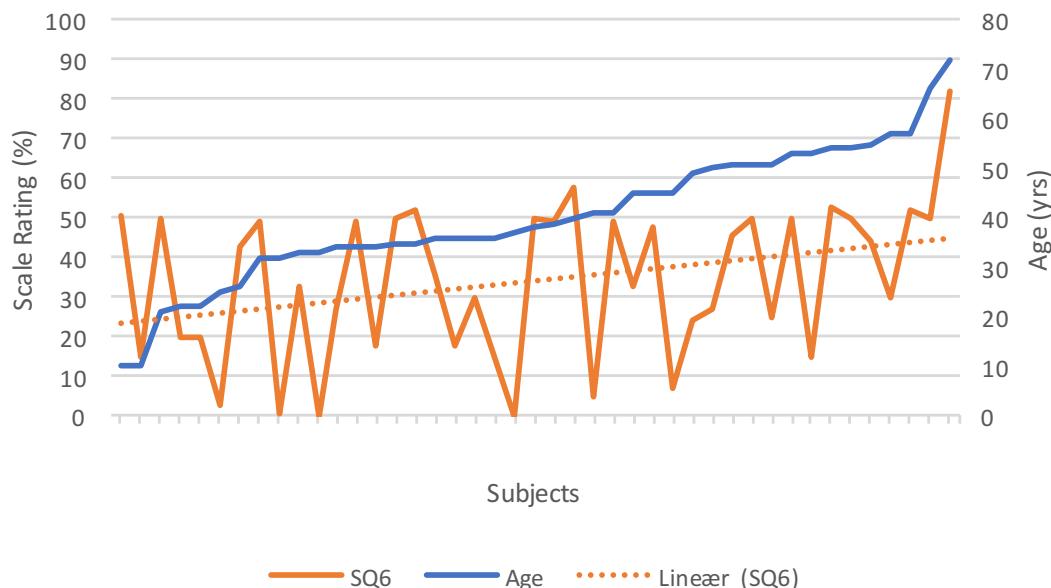


Figur 10.49. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ13: *Jeg regnede med, at robotten fulgte mig hen til det sted jeg valgte* og SQ21: *Hvad synes du ellers om robotten? i forhold til anmassende*. Denne graf bygger på 35 besvarelser, da der manglede otte.

10.12 Demografisk indflydelse

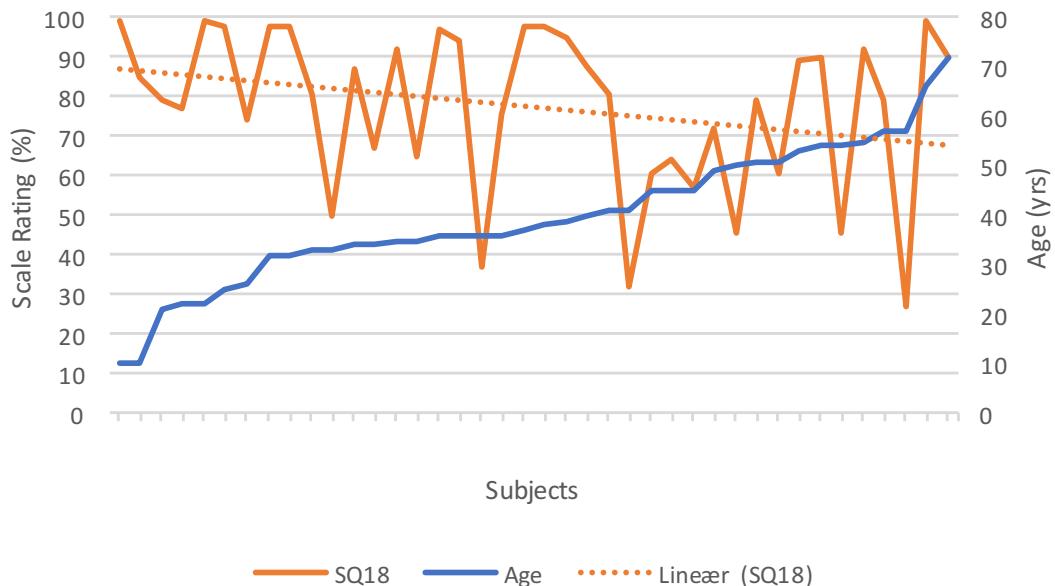
I dette afsnit beskrives, hvordan forskellene mellem testpersonerne kan være med til at påvirke vurderingen af robotten. Der fokuseres isoleret på, hvordan de demografiske faktorer påvirker hver enkelte skala. Det er valgt kun at medtage de grafer, hvor der forekommer en tendens for korrelation mellem alder og besvarelser til skalaerne. Samtlige grafer for undersøgelsen af testpersonernes besvarelser og alder, højdeforskel og hvor glade testpersonerne er for teknologi forefindes i Afsnit 13.9 (*Tendenser*). Dog er analysen foretaget på tværs af alle demografiske faktorer og samtlige skalaer, med undtagelse af hvor ofte testpersonerne rejser. Årsagen til at analysen ikke foretages i forhold til hvor ofte testpersonerne rejser skyldes hovedsageligt én ting, at AAL er relativt lille lufthavn, hvor der formentlig ikke er et lige så stort behov for robottens hjælp, som eksempelvis kan være tilfældet i Københavns Lufthavn. Med det i betragtning antages det, at for rejsende, som rejser privat flere gange årligt eller som rejser i forbindelse med arbejde, formentlig ikke adskiller sig specielt meget. Hvorimod hvis undersøgelsen blev foretaget i en større lufthavn, så kan det sagtens være at der er en forskel. Dette kan formentlig have særlig betydning for testpersonernes vurdering i forhold til hvorvidt de føler, at robotten kan hjælpe dem, om de stoler på, at den følger dem det rigtige sted hen, om den står i vejen og hvorvidt robotten opleves som værende anmassende.

I henhold til testpersonernes alder tyder det på, at der er en positiv korrelation mellem alder og SQ6, jævnfør Figur 10.50. Der er derfor en tendens til at desto ældre testpersonerne er, desto mere synes de, at robotten kører for hurtigt.



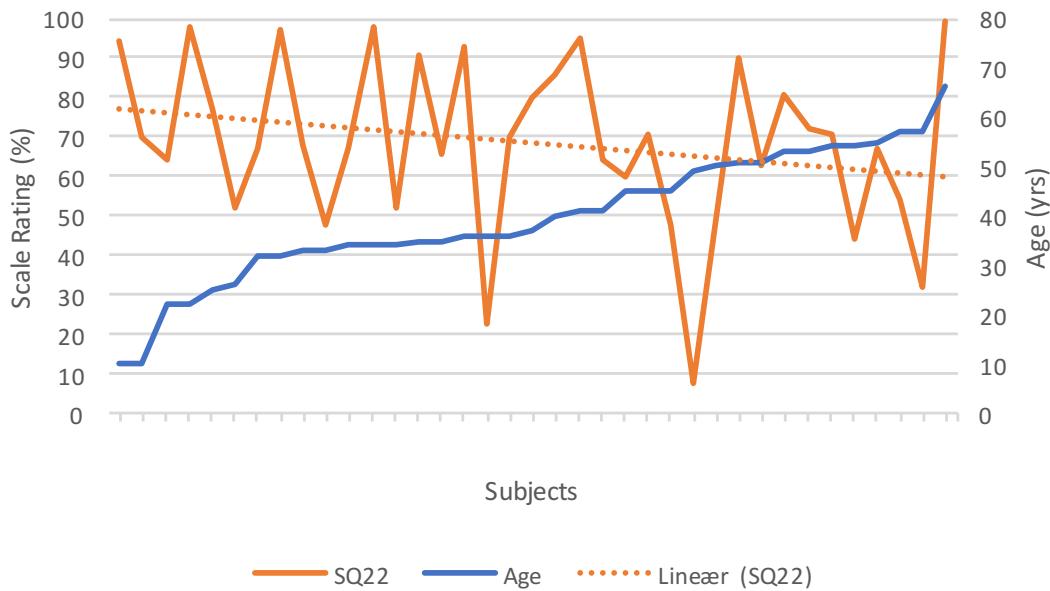
Figur 10.50. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ6: *Jeg synes, at robottens hastighed er...*, og deres alder. Denne graf bygger på samtlige besvarelser.

Derudover forekommer der en lille negativ korrelation mellem testpersonernes alder og hvor spændende de synes robotten er, jævnfør Figur 10.51. Der er derfor en tendens til at desto ældre testpersonerne, desto mindre spændende synes de, at robotten er. Derudover tyder det på at testpersoner mellem 35 år og 45 år finder robotten mindst spændende.



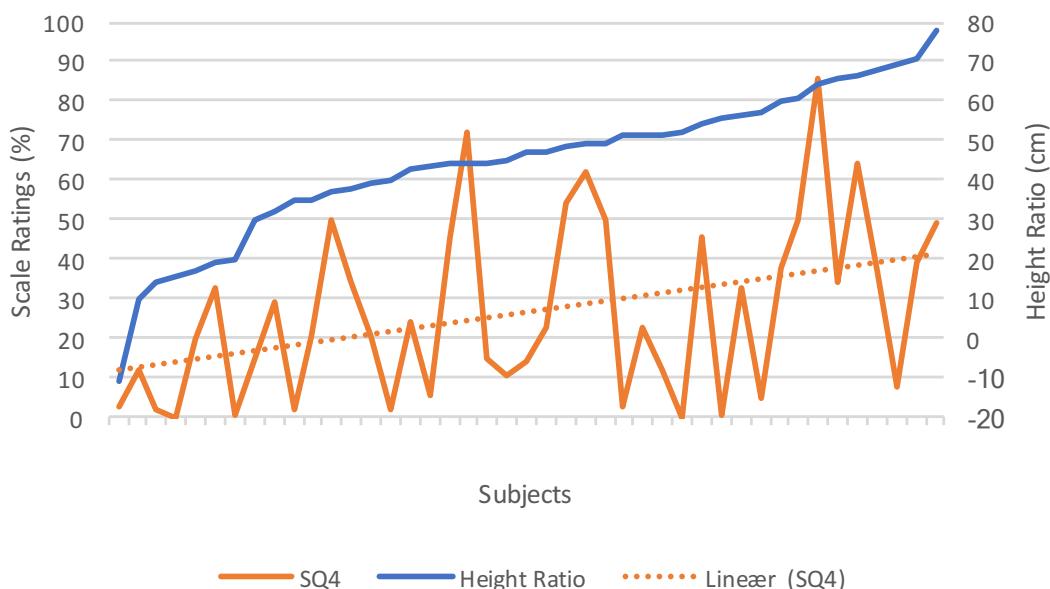
Figur 10.51. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ18: *Hvad synes du om robotten?*, i forhold til *spændende*, og deres alder. Denne graf bygger på 40 besvarelser, da der manglede tre.

Lignende forefindes mellem testpersonernes alder og SQ22, vedrørende hvor sjov robotten opleves, hvor der en lille tendens til en negativ korrelation, hvilket fremgår af Figur 10.52. Det tyder på at desto ældre testpersonerne er, desto mindre sjov synes de, at robotten er.



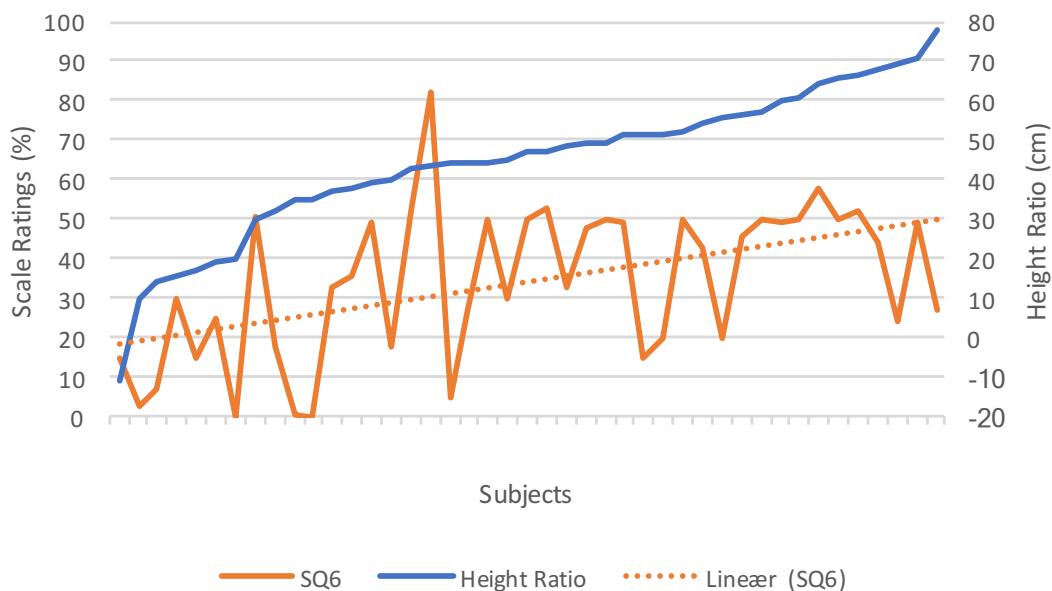
Figur 10.52. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ22: *Hvad synes du om robotten?*, i forhold til *sjov*, og deres alder. Denne graf bygger på 37 besvarelser, da der mangler seks.

Udover at analysere, hvordan robottens højde påvirker testpersonernes respons vælges det at analysere højdeforskellen mellem robot og testperson. Sammenholdes SQ4, vedrørende robottens bevægelser, med højdeforskellen forekommer der en lille positiv korrelation, jævnfør Figur 10.53, hvor det tyder på, at desto større højdeforskellen er mellem robot og testperson, desto mere vilde opleves robottens bevægelser. Dog skal der tages forbehold for, at der forekommer relativt stor variation mellem testpersonernes besvarelser.



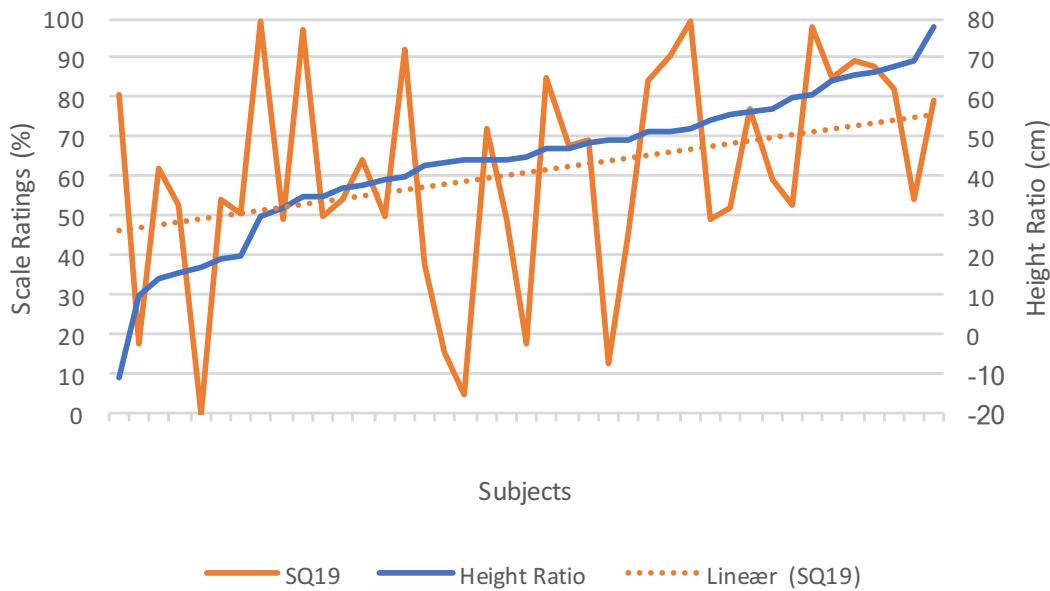
Figur 10.53. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ4: *Hvordan oplevede du robottens bevægelser?*, og højdeforskellen. Denne graf bygger på samtlige besvarelser.

Der forekommer en positiv korrelation mellem højdeforskellen og SQ6, vedrørende robottens hastighed, jævnfør Figur 10.54. Det tyder på, at desto større højdeforskellen er, desto hurtigere opleves robotten. Dette hænger formentlig sammen med, at når robotten er på sit laveste (118 cm), så kører den hurtigere end når den eksempelvis er på sit højeste (151 cm). Sammenholdes det med en stor højdeforskel kan det formentlig virke som om, at robotten er endnu hurtigere.



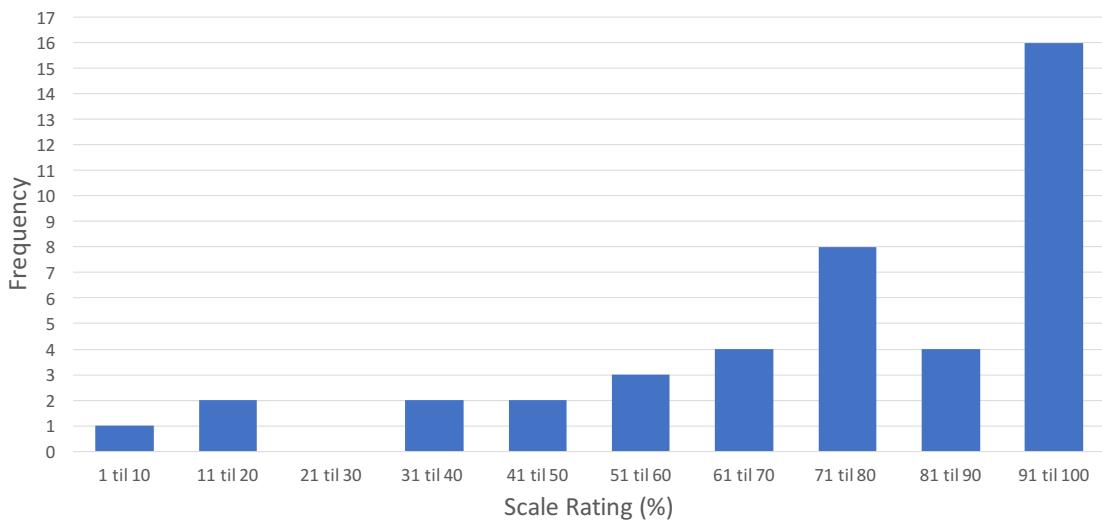
Figur 10.54. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ6: *Jeg synes, at robottens hastighed er...,* og højdeforskellen. Denne graf bygger på samtlige besvarelser.

Sammenholdes højdeforskellen med SQ19, vedrørende hvor sød robotten opleves, forekommer der en positiv korrelation, jævnfør Figur 10.55. Denne korrelation indikerer, at desto større højdeforskellen er, desto sødere opleves robotten.



Figur 10.55. Sammenhæng mellem hvad testpersonerne angiver (%) på skalaen til SQ19: *Hvad synes du om robotten?*, i forhold til *sød*, og højdeforskellen. Denne graf bygger på 41 besvarelser, da der manglerne to.

Baseret på Figur 10.56 fremgår det, at størstedelen af testpersonerne er glade for teknologi, hvor 16 ud af 43 testpersoner har angivet en respons mellem 91 % og 100 %. Derudover tyder det på, at mændende ($M=76.85$, $SD=22$) er mere glad for teknologi end kvinderne ($M=67.3$, $SD=30.2$), fælles for begge køn er dog, at standardafvigelsen er meget høj, hvorfør det ikke endeligt kan fastslås.



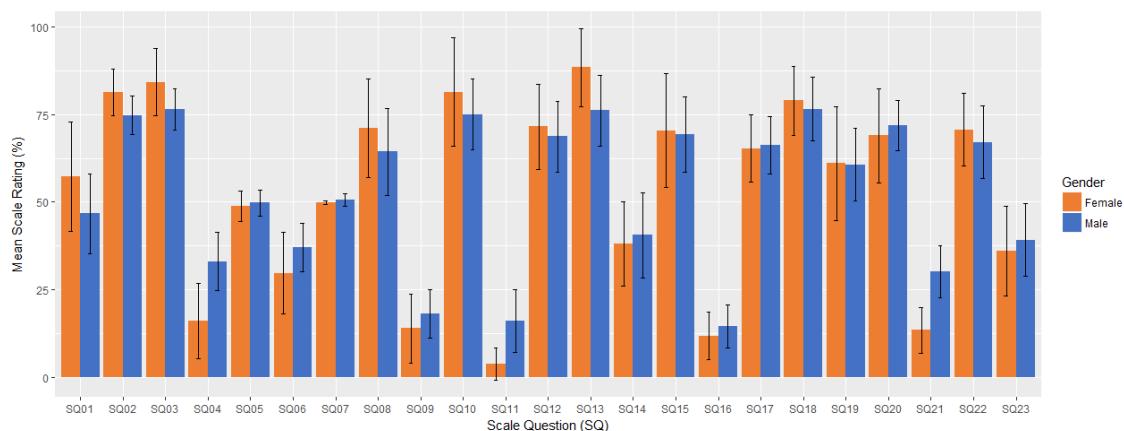
Figur 10.56. Histogram over besvarelserne (%) til skalaspørgsmålet: *Hvor glad er du for teknologi?*, som indgår i demografien.

I forhold til hvilken indflydelse det har, at testpersonerne er glade for teknologi eller ej, tyder det ikke på at der er forekommeth nogle korrelationer i forhold til specifikke skalaer. Selvom der ikke forekommer en korrelation mellem hvor godt skærmen reagerede og hvor glad testpersonerne var for teknologi, så er det ikke utænkeligt at de to parametre

måske har en indflydelse på hinanden. At det ikke vides skyldes hovedsageligt at det ikke har været muligt at kontrollere hvordan skærmen reagerede, hvorfor det kun er deres subjektive vurdering af det. Den vurdering kan være præget af hvor glad testpersonerne er for teknologi, hvor nogle formentlig vil være mere overbærerne når skærmen ikke reagerede første gang og andre vil blive irriteret eller frustreret over det. Det kan både gøre sig gældende for testpersonerne, som er meget glad for teknologi og for testpersoner, som ikke er.

10.12.1 Kønsforskelle

I forhold til testpersonernes køn er det valgt at opstille et søjlediagram med tilhørende error bars, jævnfør Figur 10.57. Søjlerne repræsenterer den gennemsnitlige besvarelse til hvert SQ fordelt på kvinder og mænd, hvor error bars repræsenterer 95 % konfidens intervallet (CI).



Figur 10.57. Søjlediagram over den gennemsnitlige besvarelse samt error bars (95 % CI) til hvert skalaspørgsmål (SQ) for henholdsvis kvinder (orange) og mænd (blå).

Ud fra Figur 10.57 undersøges det hvilke skalaspørgsmål hvortil besvarelserne forekommer forskellige mellem kvinder og mænd. Den tydeligste forskel forefindes ved SQ21: *Hvad synes du ellers om robotten i forhold til anmassende*, hvor konfidens intervallerne ikke overlapper hinanden, det forventes derfor at der er en signifikant forskel mellem kvinder og mænds besvarelser til dette SQ. Ved fortolkning af SQ21 indikerer det at mænd har oplevet robotten som værende mere anmassende end kvinderne har.

For SQ4: *Hvordan oplevede du robottens bevægelser* og SQ11: *Robotten gjorde mig forskrækket* tyder det på, at der kan være forskel mellem hvad kvinder og mænd har svarer. Dog overlapper konfidens intervallerne hinanden, hvorfor det ikke med sikkerhed kan vurderes om der reelt er en forskel eller ej. Ved antagelsen af at der er en forskel, har kvinder ved SQ4 vurderet robottens bevægelser som mere rolige end mændene og for SQ11 har robotten gjort mændene mere forskrækket end kvinderne.

For de resterende skalaspørgsmål er der så stort et overlap mellem konfidens intervallerne, at der ikke er en væsentlig forskel mellem kvinder og mænds besvarelser.

Diskussion 11

Denne diskussion vil tage udgangspunkt i den foregående databehandling og vedrører forskellige emner, som ikke allerede er diskuteret.

11.1 Skalaerne

Som det er beskrevet Afsnit 8.2 (*Program til skalaer*) opstod der flere problemer med programmet, hvori skalaerne blev præsenteret. For testpersonerne betød det, at programmet ikke altid reagerede, når de prøvede at angive deres respons, testlederen måtte derfor i flere tilfælde træde til og genstarte programmet. Dette forårsagede desværre, at det i flere tilfælde blev observeret, at testpersonerne ændrede deres respons anden gang de svarede på en skala. Dette er noteret for TP1 og TP28. Derudover var der flere testpersoner, som gav verbalt udtryk for, at programmet reagerede for langsomt, men også i deres kropssprog blev det observeret, at de var irriteret over, at programmet ikke reagerede og de som følge deraf måtte bruge længere tid på at besvare skalaerne.

Derudover opstod der situationer, hvor flere testpersoner fejlagtigt troede at midtpunktet var en slider, som de kunne bruge til at angive deres respons. Det kan muligvis skyldes den dårlige reaktion, så når programmet ikke reagerede første gang ved et musseklik på skalaen, så prøvede testpersonerne at tage fandt i midtpunktet for at rykke det. Dette blev noteret for TP2 og TP15.

TP17 er den eneste noteret testperson, som har kommenteret på det label, der er på to af de bipolare skalaer: *Fin*, hvor testpersonen spurgte ind til, om det var meningen, at det skulle stå alle steder, hvorfor testlederen måtte forklare, hvorfor det er valgt at anvende netop det label. Dog tyder det på, at både midtpunkt og labels har haft en indflydelse på besvarelserne, særligt ved SQ5, vedrørende robottens afstand, og SQ7, vedrørende robottens højde, hvor variationen er meget lille, hvilket gengiver at testpersonerne har svaret tæt på eller direkte på midtpunktet, jævnfør Figur 10.8. Selvom SQ5 ikke har haft labelen *Fin*, som de to andre skalaer præsenteret på side to, så kan det være at det har haft en indflydelse i form af, at testpersonerne har troet, at det skulle stå på den skala. Derudover så har testpersonerne selv kontrolleret afstanden til robotten, hvorfor det ligeledes er forståeligt, at deres respons samler sig omkring midtpunktet. At den lille variation forekommer til SQ7 kan formentlig skyldes, at robottens højde aldrig når en højde, som reelt ændrer testpersonernes oplevelses, men det kan også være, at testpersonerne bare ikke har en holdning til det, hvorfor de angiver en respons omkring midtpunktet *Fin*. Dette forudsætter, at der naturligt vil forekomme en del flere outliers, fordi der ikke skal særlig meget til for at afvige fra den resterende respons centreret omkring midtpunktet. Når variationen for de to skalaspørgsmål er så lille som den er på Figur 10.8 og at ingen af dem forklarer en særligt stor del af variationen i de foretaget PCA, indikerer det at

de formentlig ikke har særlig stor betydning for danske rejsende, hvert fald ikke som et selvstændigt skalaspørgsmål. Det skyldes formentlig, at de to parametre har indflydelse på andre parametre eller indirekte bliver målt ved andre parametre.

Foruden de to skalaspørgsmål forekommer der generelt en stor spredning på skalaerne, hvilket skyldes, at testpersonerne har haft forskellige oplevelser af interaktionen med robotten. Det betyder også at det er lykkedes at påvirke testpersonernes oplevelse ud fra de fysiske parametre, der blev ændret undervejs. Dog medfører den store varians, at det kan være svært at fastslå noget konkret ud fra de fundne korrelationer. Det er også derfor, at der skal tages forbehold for tendenslinjerne, som særligt påvirkes af variationen men også af den måde data sorteres på for at lave graferne, der blandt andet præsenteres i Afsnit 10.11 (*Sammenligning af korrelerede parametre*).

11.2 Testpersoner

Havde det været muligt at kontrollere robottens afstand til testpersonerne samt indgangsvinklen, ville der blive sigtet efter at få 45 testpersoner. Da det ikke var en mulighed og det tog væsentlig længere tid, at få nogle testpersoner, og testen derfor blev afviklet over to dage, resulterede det i, at der kun blev indsamlet data fra 43 testpersoner, hvilket er relativt få taget i betragtning af hvor mange timer der blev afsat til testen i lufthavnen, og hvor mange rejsende, der bare gik forbi robotten.

Selvom 43 testpersoner i nogle henseender kan være tilstrækkeligt, vurderes det, at det ikke har været tilfældet i denne test. Det skyldes hovedsagligt, at der bliver testet 23 forskellige skalaer, ændret på tre fysiske parametre; højde, afstand og indgangsvinkel, men også måden, hvorpå testpersonerne er inddelt i grupperne. Her referer grupperne til antallet af testpersoner repræsenteret ved de fem højder, tre afstande og fire indgangsvinkler. Med så mange forskellige scenarier burde der have været langt flere testpersoner for, at kunne fastslå noget endegyldigt.

På gruppen blev det diskuteret, hvorvidt rejsende, som ikke interagerede med robotten, men tydeligt registrerede den var der og rejsende, som starter en interaktion med robotten, men som enten svarer nej til om den må hjælpe dem med at finde rundt i AAL eller som af en eller anden grund stopper interaktionen undervejs, skulle stoppes og spørges om, hvorfor de ikke havde lyst til at interagere med robotten. Årsagen til at dette ikke blev gjort er, at det for de rejsende kan virke meget bebrejdende, at blive spurt ind til, hvorfor de ikke har interageret med robotten, og fordi det ikke ønskes, at forstyrre dem, hvis de nu har travlt. Derudover skal det også respekteres, hvis de rejsende bare ikke har lyst til at deltage.

Fordelen ved at have stoppet dem for at spørge ind til, hvorfor de valgte ikke at interagere med robotten eller valgte aktivt at trykke nej til hjælpen eller stoppede interaktionen undervejs, havde været at andre synspunkter kunne være kommet til udtryk. Det kan være, at nogen af de rejsende, som ikke interagerede med robotten, ikke føler sig trygge ved robotten eller ikke føler, at den kan hjælpe dem. Dette kommer ikke nødvendigvis til udtryk ved de indsamlede besvarelser, da det formentlig godt kan antages, at de rejsende, som har interageret med robotten og besvaret skalaerne, i højere grad har følt sig trygge ved robotten sammenlignet med nogen af dem, der ikke har. Ved at have stoppet de rejsende kan det have været med til, at give et mere nuanceret billede af hvordan

HRI opleves, da nogle af de mere negative parametre, såsom at robotten er anmassende og irriterende, formentlig ville være mere tydelige.

Da testpersonerne besvarede skalaerne, var der som regel en testleder i nærheden for dels at sikre, at programmet fungerede og dels for at besvare eventuelle spørgsmål. Dette kan potentelt have påvirket nogle af testpersonernes besvarelser, da de måske har været bevidste om, at testlederen kunne se besvarelserne, hvorfor de måske enten har vurderet nogle parametre lavere eller højere, end hvad de reelt mener, bare for at tilfredsstille projektgruppen. Derudover var der også nogle få testpersoner, eksempelvis TP10, som modtog hjælp i forhold til at trykke på robottens skærm under interaktionen, da de formentlig ellers havde opgivet og gået videre. Dette kan ligeledes have påvirket deres besvarelser, uden at det vides hvordan.

Andre ting, der kan have påvirket testpersonernes oplevelse, relaterer sig blandt andet til, at alle testpersoner ikke blev fulgt lige langt, enten fordi de vidste, hvor de rent faktisk skulle hen og robotten fulgte dem den forkerte vej, eller fordi testpersonerne startede interaktionen i nærheden af Duty Free, hvor robotten som regel stoppede. Dette gør sig blandt andet gældende for TP14 og TP27, som ikke blev fulgt afsted af robotten, da testlederen stod lige ved siden af dem. I relation til det, fremgår det fra observationerne, jævnfør Afsnit 10.3 (*Observationer i AAL*), at TP21 misforstod SQ5: *Jeg synes, at robotten stoppede*, hvor de to rejsende tror det handler om, hvorvidt robotten stoppede tæt på det sted, som de havde valgt at den skulle følge dem hen til. Det virker også til at TP16 ikke forstod SQ5.

Andre testpersoner har haft en samtale med testlederen samtidig med, at de evaluerede interaktionen med robotten på de præsenterede skalaer. Dette kan have haft en indflydelse på deres besvarelser, både fordi samtalens kan have hjulpet dem til at forstå skalaspørgsmålene, men også fordi det måske har påvirket dem til at svare mere positivt end hvad de reelt mener. Derudover har der også været situationer, hvor en familie enten skiftes til at trykke på robottens skærm, eller hvor det kun er én, der trykker på skærmen, men hvor det er et andet familiemedlem, som svarer på skalaerne. Derfor er der situationer, hvor en testperson har svaret på skalaerne, men ikke interagerede med robotten. Det er uvist hvilken indflydelse det har haft og om det overhovedet har haft en indflydelse.

11.3 Fysiske parametre

I Afsnit 9.3 (*Robottens bevægelse*) er der beskrevet og argumenteret for valget af de højder, som robotten havde under testen. Dog kan det være, at det har været unødvendigt at anvende alle fem højder, særligt 123.5 cm, som kun er 5.5 cm fra den laveste højde (118 cm) og midtpunktet (129 cm). Ved at have undladt denne højde, kunne antallet af testpersoner præsenteret for de fire andre højder være steget. Derudover forefindes der ikke en markant tendens, hvor testpersonerne, som interagerede med robotten på 123.5 cm, svarer anderledes end de testpersoner, som interagerede med robotten på 129 cm. Dog med forbehold for at der generelt forekommer stor variation i besvarelserne.

I forhold til robottens højde, blev det observeret, at det havde en indflydelse på hvor hurtigt robotten kørte. Når robotten var på sit laveste (118 cm) kørte den væsenligt hurtigere end når den var på sit højeste (151 cm). Problemets med at robottens hastighed afhænger af dens højde er, at det ikke vides, hvordan de hver især påvirker testpersonernes

oplevelse, og at der ikke er nogen kontrol over robottens hastighed. Det betyder også at testpersonerne, der har interagerede med robotten når den var 118 cm, har interageret med en hurtigere robot end dem, der interagerede med robotten, når den var 151 cm, hvor hastigheden var langsommere. Dette kommer blandt andet til udtryk på Figur 10.17, hvor det fremgår, at når robotten er 118 cm, bliver den vurderet hurtigere end når den er 151 cm.

I forhold til robottens afstand, var det ikke muligt at foruddefinere bestemte afstande. Dette havde formentlig heller ikke været hensigtsmæssigt, da testpersonerne så ikke selv kunne have indflydelse på afstanden, fordi robotten konstant ville tilpasse sin afstand, afhængigt af testpersonens placering. Dog havde fordelen med foruddefineret afstanden været, at det på samme måde som ved højde, havde været muligt at inddelte testpersonerne i mere specifikke grupper end hvad der har været tilfældet. Den manglende kontrol over afstanden kan være årsagen til, at SQ5 forklarer så lidt af variationen i de forskellige PCA'er.

Fordi det har været så svært at kontrollere afstanden til testpersonerne og notere hvorvidt afstanden var langt fra, tilpas eller tæt på, og fordi testpersonerne undervejs i interaktionen har mulighed for at ændre afstanden, er det ikke sikkert at en PCA er det bedste valg. Det samme gør sig gældende for indgangsvinkel, som ligeledes har været svært at kontrollere, specielt at få robotten til at henvende sig fra venstre i AAL grundet placeringen. Derudover har det ikke været muligt at kontrollere hvilken indgangsvinkel testperson, som selv henvendte sig til robotten, havde.

11.4 Redundans

Som beskrevet i Afsnit 10.11 (*Sammenligning af korrelerede parametre*) korrelerer nogle af parametrene. Det er derfor være relevant at undersøge, hvorvidt nogle af de korrelerede parametre dækker over det samme og derfor kan slås sammen under ét dækkende parameter. Det vælges kun at diskutere de parametre, hvor det giver mening at slå dem sammen.

Jævnfør Figur 10.30, hvor der forefindes en positiv korrelation mellem at kunne lide at blive betjent af robotten og hvor spændende den er, samt Figur 10.35, hvor der forefindes en negativ korrelation mellem at kunne lide at blive betjent af robotten og hvor anmassende den er, kan det overvejes at erstatte skalaspørgsmålet vedrørende at blive betjent af robotten med parametrene spændende og anmassende. Det kan være en tilfældighed, at der forefindes en korrelation, men hvis de rejsende ikke synes, at robotten er spændende, kan de formentlig heller ikke lide at blive betjent af den. Det samme gør sig gældende med anmassende. Hvis de rejsende kan lide at blive betjent af robotten, så synes de højest sandsynligt ikke at den er anmassende. På den måde kan det udledes, hvorvidt testpersonerne kan lide at blive betjent af robotten, ved at spørge dem om, om de synes at robotten er anmassende og spændende, uden at skulle spørge direkte om deres oplevelse af betjeningen.

Fokuseres der på Figur 10.34, for sjov og sjov, Figur 10.44, for sød og sjov samt Figur 10.46, for spændende og sjov, forefindes der positiv korrelation for de sammenlignede parametre. Under feltundersøgelsen kom det til udtryk, at sjov er et parameter, der dækker over mange forskellige ting, blandt andet at noget kan være humoristisk sjovt, men også underligt sjovt. Dette parameter er derfor svært at måle på, hvorfor det kan give mening at sammensætte

det med ej, da disse korrelerer. Ydermere kan ej sammensættes med enten sød eller spændende. Det virkede til, at børn oftest bedømte robotten som værende ej, hvorimod ældre rejsende ikke nødvendigvis kunne relatere til at robotten var ej. Da det bedømmes, at en robot ikke kan være ej uden at være spændende og da parametrene korrelerer positivt, kan det vælges at medtage parametrene sød og spændende dækende over både sød, spændende, ej og sjov.

Undersøges det hvor tryg testpersonerne var ved robotten i forhold til om de regnede med at den fulgte dem det rigtige sted hen og hvorvidt robotten kunne hjælpe dem, jævnfør Figur 10.31 og Figur 10.45, forefindes der en positiv korrelation herimellem. Det kan overvejes, hvorvidt tryghedsparameteret og at de rejsende regnede med robotten fulgte dem det rigtige sted hen kan slås sammen til et parameter omhandlende tillid til robotten. På den måde kan de tre parametre slås sammen til to; tillid til robotten og hvorvidt robotten kan hjælpe den rejsende. Det kan dog være problematisk at slå to skalaspørgsmål sammen til et nyt, for at danne en ny skala, da forståelsen af tillid til robotten ikke er testet på samme måde som de to andre er, og fordi tillid og tryghed ikke nødvendigvis er det samme.

Delkonklusion 12

Ud fra foregående afsnit, Kapitel 10 og Kapitel 11, er det muligt at konkludere på denne del af projektet og samtidig besvare den anden problemstilling: *Hvordan kan de fundne parametre anvendes til at evaluere HRI?*. De fundne parametre blev i Afsnit 5.6 (*Udvælgelse af skalaer*) gengivet i skalaer, som efterfølgende blev tilpasset for, at optimere formuleringerne og for at tilstræbe et helhedsindtryk, hvilket blev gjort i Kapitel 8 (*Tilpasning af skalaer*), hvorefter rækkefølgen for hvordan de præsenteres ligeledes blev besluttet. For at evaluere HRI blev de udviklede skalaer anvendt af danske rejsende i AAL efter de havde interageret med robotten.

Først og fremmest kan det konkluderes, at det tog betydeligt længere tid end forventet at afvikle testen og at der derfor kun deltog 43 testpersoner. Ydermere har det været svært at kontrollere fysiske parametre såsom afstand og indgangsvinkel, hvor de rejsende ofte gik udenom eller selv henvendte sig, hvis de ville interagere med robotten. Skal fysiske parametre som afstand og indgangsvinkel testes mere kontrolleret kan det derfor være fordelagtigt at opstille en test i et laboratorium eller tilføje en afstands måler til robotten, så det efterfølgende er muligt at vide hvilken afstand robotten havde.

Ligesom ved foregående feltundersøgelse reagerede robottens skærm ofte dårligt, og det virkede ikke til at hjælpeteksten: *Tryk blidt på mig*, havde den ønskede effekt. Dette har medvirket til at flere potentielle testpersoner opgav interaktionen og gik deres vej.

Af Afsnit 10.3 (*Observationer i AAL*) og Afsnit 11.1 (*Skalaerne*) fremgår det, at testpersonerne havde en forskellig forståelse af skalaerne. Nogle testpersoner troede at midtpunktet på de bipolare skalaer var en slider, som de kunne bruge til at angive deres respons.

Som beskrevet i Afsnit 10.7 (*Principal Component Analysis*) var det ikke fordelagtigt at udføre *Principal Component Analysis* på hele datasættet, hvorfor data grupperes efter henholdsvis robottens højde, afstand og indgangsvinkel. Fra PCA fremgår det hvilke parametre, der er positivt og negativt korreleret, hvilket er listet i Tabel 12.1.

PCA	Positive korrelationer	Negative korrelationer
Højde	SQ08 + SQ17	SQ02 + SQ09
	SQ10 + SQ13	SQ04 + SQ12
	SQ12 + SQ18	SQ12 + SQ21
	SQ14 + SQ15	SQ16 + SQ19
	SQ20 + SQ22	SQ18 + SQ21
Afstand	SQ01 + SQ12	SQ02 + SQ09
	SQ07 + SQ17	SQ05 + SQ21
	SQ08 + SQ21	SQ10 + SQ13
	SQ10 + SQ22	SQ13 + SQ22
		SQ14 + SQ16
Indgangsvinkel	SQ05 + SQ07	SQ01 + SQ12
	SQ08 + SQ10	SQ06 + SQ23
	SQ09 + SQ14	SQ09 + SQ10
	SQ18 + SQ20	SQ10 + SQ14
		SQ13 + SQ21

Tabel 12.1. Korrelationer fra PCA, *SQ* angiver skalaspørgsmål.

Ydermere er de parametre, hvor der er fundet en korrelation, blevet sammenlignet, hvilket fremgår af Tabel 12.2. Disse sammenligninger er foretaget på alt data, hvorfor der ikke er blevet taget højde for grupperinger.

Positive korrelationer	Negative korrelationer
SQ04 + SQ09	SQ02 + SQ09
SQ08 + SQ10	SQ09 + SQ10
SQ08 + SQ17	SQ12 + SQ21
SQ10 + SQ13	SQ13 + SQ21
SQ12 + SQ18	SQ18 + SQ21
SQ18 + SQ20	
SQ20 + SQ22	

Tabel 12.2. Korrelationer fra sammenligningsgrafer *SQ* angiver skalaspørgsmål.

En gennemgående tendens for de tre PCA'er for henholdsvis højde, afstand og indgangsvinkel er at både SQ5, vedrørende afstand, og SQ7, vedrørende højde, kun forklarer en meget lille del af variationen. Sammenholdes det med Figur 10.8, hvorpå det fremgår at spredningen i testpersonernes besvarelser til de to skalaspørgsmål er meget lille og centreret omkring midtpunktet. Dette kan betyde flere ting; 1) at de ikke er nogle parametre som har specielt stor betydning for danske rejsende, 2) at begge parametre indirekte måles ved et eller flere andre parametre, 3) at skalaen har påvirket testpersonerne til at afvige en respons omkring midtpunktet eller 4) at det simpelthen ikke har været muligt, at variere nok på afstanden og højden til at testpersonerne får en anderledes oplevelse. Det er dog stadig uvist hvilken årsag det rent faktisk er.

Udover de beskrevne korrelationer forekommer der også forskellige tendenser, når information fra demografien sammenlignes med besvarelserne på skalaerne. Det blev fundet, at når robotten bliver højere opleves bevægelserne roligere og hastigheden langsommere, hvilket formentlig hænger sammen med når *Double* er på sit laveste (118

cm) så kører den rent faktisk hurtigere end når den er på sit højeste (151 cm).

Derudover er det fundet, at når robotten er på sit laveste (118 cm), så stoler testpersonerne mere på, at den følger dem hen til det sted de har valgt, end når robotten er på sit højeste (151 cm), jævnfør Figur 10.18. Lignende er tilfældet for hvor elegant og sød robotten opleves, hvor den opleves som værende mere sød og elegant når den er lav end når den er høj, jævnfør Figur 10.19 og Figur 10.20. I forhold til robottens afstand oplever testpersonerne robotten som værende mere overraskende når den kommer tæt på end når den er langt fra. Sammenholdes testpersonernes alder og højdeforskell med deres vurdering af robottens hastighed fremgår det, at desto større højdeforskellen er og desto yngre testpersonerne er, desto langsommere oplever de robottens hastighed. Dog fremgår det at 75 % af testpersonerne besvarelser befinner sig mellem *Alt for langsom* og *Fin*, jævnfør Figur 10.8, hvorfor det formentlig kan konkluderes, at robottens hastighed godt kan øges. I tillæg fremgår det, at desto ældre testpersonerne er desto mindre spændende og sjov opleves robotten, jævnfør Figur 10.51 og Figur 10.52. Ydermere fremgår det af Figur 10.55, at desto større højdeforskellen er desto sødere opleves robotten. Af Figur 10.57 tyder det på, at den eneste kønsforskell der er fundet, hvor CI ikke overlapper, er i forhold til hvor anmassende robotten opleves, hvor mændende oplever robotten som værende mere anmassende end kvinderne.

Elektronisk bilag 13

Følgende er en oversigt over det elektroniske bilag, som er i den vedhæftede ZIP-fil: (/ElektroniskBilag).

13.1 Transskriberede data

Det transskriberede data er transskriberet ud fra lydoptagelserne foretaget i feltundersøgelsen, og forefindes i dokumentet: (/ElektroniskBilag/TransskriberedeData). Første side i dokumentet indeholder en oversigt over hvor transskriberingen til hver testperson findes.

13.2 Affinity notes

De udarbejde *affinity notes* fra feltundersøgelsen forefindes: (/ElektroniskBilag/Affinity-Notes), hvor *affinity notes* til hvert interview samt observationerne er skrevet ind.

13.3 Program til VAS

Det udarbejdede program til præsentation af skalaer forefindes i: (/ElektroniskBilag/VASProgram)

13.4 Skala Oversigt

En oversigt over hvilke skalaspørgsmål (*Scale question*), der er hører til skala nummer (*SQ#*) samt tilhørende labels (*Skala labels*), der præsenteres på hver af de syv sider (*Side*), forefindes i: (/ElektroniskBilag/SkalaOversigt.pdf)

13.5 Rådata i Excel

Rådata, der består af testpersonernes respons på de 23 skalaer samt demografi, forefindes i: (/ElektroniskBilag/RaaData).

13.6 Hisogram og normalfordelings plot

Til hver skalaspørgsmål, samt til besvarelserne ift teknologi, er der opstillet et histogram for besvarelserne hvor funktionen for normalfordelingen er plottet ovenpå. Skaleringen af akserne er ens for alle plots, med undtagelse af plottet for SQ5, SQ7 og SQ11. Alle plots findes i: (/ElektroniskBilag/HistogramNormalfordeling)

13.7 MATLAB Scripts til PCA

MATLAB scriptet der er anvendt i forbindelse med PCA samt script der anvendes til import af data i MATLAB forefindes i: (/ElektroniskBilag/MATLABScripts).

13.8 3D Bi-plots

Der er vedlagt tre 3D *Bi-plots* for henholdsvis: Det samlede data, højde og indgangsvinkel, som forefindes i: (/ElektroniskBilag/3DBiplots/), hvor det samlede data fremgår af: Biplot3D.fig, højde fremgår af: RHeight-3D.fig, og afstand fremgår af: Direction-3D.fig

13.9 Tendenser

For undersøgelsen af tendenser mellem testpersonernes besvarelser og robottens højde, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/Højde-skala.xlsx)

For undersøgelsen af tendenser mellem testpersonernes besvarelser og robottens afstand, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/Afstand-skala.xlsx). Afstanden angives i Excel-dokumentet som 1 (tæt på), 2 (tilpas) og 3 (langt fra).

For undersøgelsen af tendenser mellem testpersonernes besvarelser og robottens indgangsvinkel, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/Indgangsvinkel-skala.xlsx). Indgangsvinklen angives i Excel-dokumentet som 1 (venstre), 2 (forfra), 3 (højre) og 4 (kommer selv).

For undersøgelsen af tendenser mellem testpersonernes besvarelse og alder, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/Alder-skala.xlsx)

For undersøgelsen af tendenser mellem testpersonernes besvarelser og højdeforskellen mellem testpersonernes højde og robottens højde, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/Højdeforskelle-skala.xlsx)

For undersøgelsen af tendenser mellem testpersonernes besvarelser og hvor glade testpersonerne er for teknologi, forefindes samtlige grafer i: (/ElektroniskBilag/Tendenser/GladForTeknologi-skala.xlsx)

13.10 Korrelationsgrafer

Billederne: (/ElektroniskBilag/KorrelationsGrafer/SQ5+SQ7.pdf), dette er blot et eksempel på en af henvisningerne til en specifik graf. Henvises der til en specifik graf, der ikke er SQ5+SQ7, vil nummeret efter SQ ændres til de specifikke skalaer.

Samtlige grafer for sammenligning af korrelerede parametre fundet for højde forefindes i: (/ElektroniskBilag/KorrelationsGrafer/Korrelation-Højde.xlsx)

Samtlige grafer for sammenligning af korrelerede parametre fundet for afstand forefindes i: (/ElektroniskBilag/KorrelationsGrafer/Korrelation-Afstand.xlsx)

Samtlige grafer for sammenligning af korrelerede parametre fundet for indgangsvinkel forefindes i: (/ElektroniskBilag/KorrelationsGrafer/Korrelation-Indgangsvinkel.xlsx)

Bibliografi

- Abeele, V. V.; B. Zaman (2017):** “Laddering the User Experience!” I: *2009. INTERACT edition:12th IFIP TC13 Conference on Human-Computer Interaction* location:Uppsala, Sweden date:24-28 August 2009, s. 1–5. URL: <https://lirias.kuleuven.be/bitstream/123456789/267307/4/Laddering+the+User+Experience.pdf> (sidst set 27.Oktober.2017).
- Althaus, P. m.fl. (2004):** “Navigation for Human-Robot Interaction Tasks”. I: *IEEE Robotics and Automation*, s. 1894–1900. DOI: [10.1109/ROBOT.2004.1308100](https://doi.org/10.1109/ROBOT.2004.1308100).
- Arkin, R. C. m.fl. (2003):** “An ethological and emotional basis for human-robot interaction.” I: *Elsevier Science* (42): *Robotics and Autonomous Systems*, s. 191–201. DOI: [10.1016/S0921-8890\(02\)00375-5](https://doi.org/10.1016/S0921-8890(02)00375-5).
- Ball, A. m.fl. (2015):** “Group Vs. Individual Comfort When a Robot Approaches”. I: *Springer: Social Robotics*. Udg. af A. Tapus m.fl. 7th International Conference, ICSR 2015 Paris, France, October 26–30, 2015 Proceedings, s. 41–50. DOI: [10.1007/978-3-319-25554-5_5](https://doi.org/10.1007/978-3-319-25554-5_5).
- Bogicevic, V. m.fl. (2017):** “The impact of traveler-focused airport technology on traveler satisfaction.” I: *Elsevier* (123): *Technological Forecasting and Social Change*, s. 351–361. DOI: <http://dx.doi.org/10.1016/j.techfore.2017.03.038>.
- Breazeal, C. (2003):** “Toward sociable robots”. I: *Elsevier* (42): *Robotics and Autonomous Systems*, s. 167–175. DOI: [10.1016/S0921-8890\(02\)00373-1](https://doi.org/10.1016/S0921-8890(02)00373-1).
- Brida, J. G. m.fl. (2016):** “Customer perception of service quality: The role of Information and Communication Technologies (ITCs) at airport functional areas”. I: *Elsevier* (20): *Tourism Management Perspectives*, s. 209–216. DOI: <http://dx.doi.org/10.1016/j.tmp.2016.09.003>.
- Butler, J. T.; A. Agah (2001):** “Psychological Effects of Behavior Patterns of a Mobile Personal Robot”. I: *Elsevier. Autonomous Robots* 10 (2), s. 185–202.
- Chee, B. T. T. m.fl. (2012):** “Personality of social robots perceived through the appearance”. I: *Work*, s. 272–276. DOI: [10.3233/WOR-2012-0168-272](https://doi.org/10.3233/WOR-2012-0168-272).
- Dautenhahn, K.; A. Billard (2002):** *Universal Access and Assistive Technology. Games Children with Autism Can Play With Robota, a Humanoid Robotic Doll*. Udg. af K. S. m.fl. Springer. Kap. 18, s. 179–190. DOI: https://doi.org/10.1007/978-1-4471-3719-1_18.

Dautenhahn, K.; M. Walters m.fl. (2006): "How May I Serve You? A Robot Companion Approaching a Seated Person in a Helping Context". I: *ACM*, s. 172–179. DOI: 10.1145/1121241.1121272.

Eyssel, F. m.fl. (2011): "Effects of Anticipated Human-Robot Interaction and Predictability of Robots Behavoir on Perceptions of Anthropomorphism". I: *Human-Robot Interaction, HRI. Proceedings of the 6th international conference on Human-robot interaction*, s. 61–68. DOI: 10.1145/1957656.1957673.

Fry, B.; C. Reas (2017): *Processing. Download Processing*. URL: <https://processing.org/download/> (sidst set 11.12.2017).

Graaf, M. M. de; S. B. Allouch (2013): "Exploring influencing variables for the acceptance of social robots". I: *Elsevier* (61): *Robotics and Autonomous Systems*, s. 1476–1486. DOI: <http://dx.doi.org/10.1016/j.robot.2013.07.007>.

Graaf, M. M. de; S. B. Allouch; T. Klamer (2015): "Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot." I: *Elsevier* (43): *Computers in Human Behavior*, s. 1–14. DOI: <http://dx.doi.org/10.1016/j.chb.2014.10.030>.

Halpern, D.; J. E. Katz (2013): "Close But Not Stuck: Understanding Social Distance in Human-Robot Interaction Through a Computer Mediation Approach." I: *Intervalla : Platform for Intellectual Exchange* 1, s. 17–34. ISSN: 2296-3413.

Hansen, B. S. (2017): *Antropomorfisme*. Udg. af G. Den Store Danske. URL: http://denstoredanske.dk/Sprog,_religion_og_filosofi/Religion_og_mystik/Primitiv_religion/antropomorfisme (sidst set 16.Oktober.2017).

Heathrow (2017): *Facts and figures*. Udg. af L. A. Limited. URL: <http://www.heathrow.com/company/company-news-and-information/company-information/facts-and-figures#> (sidst set 31.Oktober.2017).

Heerink, M. (2010): "Assessing acceptance of assistive social robots by aging adults". PhD Thesis. Amsterdam: Informatics Institute (IVI). 183 s. URL: <http://hdl.handle.net/11245/1.327918>.

Holtzblatt, K.; H. Beyer (2015): *Contextual Design: Evolved. Synthesis Lectures On Human-Centered Informatics. Consolidation and Ideation: The Bridge to Design*. Udg. af J. M. Carroll. Morgan & Claypool Publishers. Kap. 4, s. 21–51. DOI: 10.2200/S00597ED1V01Y201409HCI024.

Holtzblatt, K. m.fl. (2005a): *Rapid Contextual Design. A How-To Guide to Key Techniques for User-Centered Design. Building an Affinity Diagram*. Udg. af S. Card m.fl. The Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers. Kap. 8, s. 159–179. 321 s. ISBN: 0-12-354051-8.

- (2005b): *Rapid Contextual Design. A How-To Guide to Key Techniques for User-Centered Design. Contextual Interview Interpretation Session.* Udg. af S. Card m.fl. The Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers. Kap. 5, s. 101–122. 321 s. ISBN: 0-12-354051-8.
- Höflich, J. R.; A. E. Bayed (2015): *Social Robots from a Human Perspective. Perception, Acceptance, and the Social Construction of Robots - Exploratory Studies.* Udg. af J. Vincent m.fl. Springer. Kap. 4, s. 39–51. DOI: 10.1007/978-3-319-15672-9_4.
- Joosse, M.; V. Evers (2017): “A Guide Robot at the Airport: First Impressions”. I: *HRI. Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, s. 149–150. DOI: <http://dx.doi.org/10.1145/3029798.3038389>.
- Kaerlein, T. (2015): *Social Robots from a Human Perspective. Minimizing the Human? Functional Reductions of Complexity in Social Robotics and Their Cybernetic Heritage.* Udg. af J. Vincent m.fl. Springer. Kap. 7, s. 77–88. DOI: 10.1007/978-3-319-15672-9_7.
- Kalakou, S. m.fl. (2015): “Future airport terminals: New technologies promise capacity gains.” I: Elsevier (42): *Air Transport Management*, s. 203–212. DOI: <http://dx.doi.org/10.1016/j.jairtraman.2014.10.005>.
- Karreman, D. m.fl. (2015): “Visiting Cultural Heritage with a Tour Guide Robot: A User Evaluation Study in-the-Wild”. I: Springer: *International Conference on Social Robotics. Social Robotics.* Udg. af T. A. m.fl. Part of the Lecture Notes in Computer Science book series (LNCS, volume 9388), s. 317–326.
- Katz, J. E. m.fl. (2015): *Social Robots from a Human Perspective. In the Company of Robots: Views of Acceptability of Robots in Social Settings.* Udg. af J. Vincent m.fl. Springer. Kap. 3, s. 25–38. DOI: 10.1007/978-3-319-15672-9_3.
- Kim, Y.; B. Mutlu (2014): “How social distance shapes human-robot interaction”. I: Elsevier. *Human-Computer Studies* 72 (72), s. 783–795. DOI: <http://dx.doi.org/10.1016/j.ijhcs.2014.05.005>.
- Lewis, J. R.; O. Erdinç (2017): “User Experience Rating Scales with 7, 11, og 101 Point: Does It Matter?” I: *Journal of Usability Studies*. Vol. 12, Issue 2, s. 73–91. URL: http://uxpajournal.org/wp-content/uploads/pdf/JUS_Lewis_Feb2017.pdf (sidst set 13.December.2017).
- Lufthavn, K. (2017a): *CPH og de regionale lufthavne vil styrke indenrigstrafikken.* Udg. af K. L. A/S. URL: <https://www.cph.dk/om-cph/investorer/trafikstatistik/2017/10/cph-trafiktal-september-2017/> (sidst set 30.Oktober.2017).
- (2017b): *Record growth in 2016.* Udg. af K. L. A/S. URL: <https://www.cph.dk/en/about-cph/investor/traffic-statistics/2017/1/copenhagen-airport-record-growth-in-2016/> (sidst set 31.Oktober.2017).

- Lufthavn, K. (2017c):** *Shops & Restaurants - Transit Area*. Udg. af K. L. A/S. URL: https://www.cph.dk/globalassets/3.-praktisk/91.-transfer-ny/transfer_map_2017.pdf (sidst set 30.Oktober.2017).
- Lufthavn, A. (2017a):** *Passagertal. Passagertal måned for måned*. Udg. af A. Lufthavn. URL: <http://www.aal.dk/om-aalborg-lufthavn/om-os/passagertal> (sidst set 30.Oktober.2017).
- (2017b): *Terminaloversigt. Oversigtskort over terminalen*. Udg. af A. Lufthavn. URL: <http://www.aal.dk/lufthavnen/terminaloversigt> (sidst set 30.Oktober.2017).
- Marshall, P. m.fl. (2011):** “Using F-formations to Analyse Spatial Patterns of Interaction in Physical Environments”. I: *ACM*, s. 445–454. DOI: [10.1145/1958824.1958893](https://doi.org/10.1145/1958824.1958893).
- Mori, M. (2012):** “The Uncanny Valley”. Oversat from Japanese af K. F. MacDorman; N. Kageki. I: *IEEE. Robotics and Automation Magazine*, s. 98–100. DOI: [10.1109/MRA.2012.2192811](https://doi.org/10.1109/MRA.2012.2192811).
- Pacchierotti, E. m.fl. (2005):** “Human-Robot Embodied Interaction in Hallway Settings: a Pilot User Study”. I: *IEEE. International Workshop on Robots and Human Interactive Communication*, s. 164–171. DOI: [10.1109/ROMAN.2005.1513774](https://doi.org/10.1109/ROMAN.2005.1513774).
- Robotics, D. (2017):** *Work from anywhere*. URL: <https://www.doublerobotics.com> (sidst set 14.Oktober.2017).
- Robotics, S. (2016a):** *For education and research*. URL: <https://www.ald.softbankrobotics.com/en/solutions/education-research> (sidst set 14.Oktober.2017).
- (2016b): *What is the ROMEO project?* URL: <https://www.ald.softbankrobotics.com/en/robots/romeo> (sidst set 14.Oktober.2017).
- Robots, P. (2014):** *PARO Therapeutic Robot*. URL: <http://www.parorobots.com> (sidst set 14.Oktober.2017).
- Severinson-Eklundh, K. m.fl. (2003):** “Social and collaborative aspects of interaction with a service robot”. I: *Elsevier. Robotics and Autonomous Systems* 42, s. 223–234. DOI: [10.1016/S0921-8890\(02\)00377-9](https://doi.org/10.1016/S0921-8890(02)00377-9).
- Shiomi, M. m.fl. (2013):** “Recommendation Effects of a Social Robot for Advertisement-Use Context in a Shopping Mall”. I: *International Journal of Social Robotics* 5 (2). Udg. af Springer, s. 251–262. DOI: [10.1007/s12369-013-0180-4](https://doi.org/10.1007/s12369-013-0180-4).
- SONY (2017):** *AIBOS HISTORY*. URL: <http://www.sony-aibo.com/aibos-history/> (sidst set 14.Oktober.2017).
- Taipale, S. m.fl. (2015):** *Social Robots from a Human Perspective. Robot Shift from Industrial Production to Social Reproduction*. Udg. af J. Vincent m.fl. Springer. Kap. 2, s. 11–24. DOI: [10.1007/978-3-319-15672-9_2](https://doi.org/10.1007/978-3-319-15672-9_2).

Tolstrup, P. J. (2016): *Danskerne er blevet 10 cm højere.* Udg. af S. I. for Folkesunhed. URL: http://www.si-folkesundhed.dk/Ugens%20tal%20for%20folkesundhed/Ugens%20tal/35_2016.aspx (sidst set 05.December.2017).

Triebel, R. m.fl. (2016): *Field and Service Robotics. Results of the 10th International Conference. SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports.* Udg. af D. S. Wettergreen; T. D. Barfoot. 113. udg. Tracts in Advanced Robotics. Springer, s. 607–622. DOI: 10.1007/978-3-319-27702-8_40.

Vincent, J. (2017): *LG's new airport robots will guide you to your gate and clean up your trash.* URL: <https://www.theverge.com/2017/7/21/16007680/lg-airport-robot-cleaning-guide-south-korea-incheon> (sidst set 25.Oktober.2017).

Vincent, J. (2015): *Social Robots from a Human Perspective. The Mobile Phone: An Emotionalised Social Robot.* Udg. af J. Vincent m.fl. Springer. Kap. 9, s. 105–115. DOI: DOI10.1007/978-3-319-15672-9_9.