Perceptions of a Help-Requesting Robot - Effects of Eye-Expressions, Colored Lights and Politeness of Speech

Martin Westhoven
westhoven.martin@baua.bund.de
German Federal Institute for Occupational Safety and Health
Dortmund, Germany

Tim van der Grinten
tim.van-der-grinten@tu-ilmenau.de
Ilmenau University of Technology
Ilmenau, Germany

Steffen Mueller
steffen.mueller@tu-ilmenau.de
Ilmenau University of Technology
Ilmenau, Germany

ABSTRACT

In this paper we report results from a web- and video-based study on the perception of a request for help from a robot head. Colored lights, eye-expressions and politeness of speech were varied. We measured effects on expression identification, hedonic user experience, perceived politeness, and help intention. Additionally, sociodemographic data, a 'face blindness' questionnaire, and negative attitudes towards robots were collected to control for possible influences on the dependent variables. A total of n=139 participants were included in the analysis. Significant differences were found for the identification performance for our intended eye-expressions, for perceived politeness, help intentions and hedonic user experience. Especially for the negative attitudes towards robots, we found significant relationships with perceived politeness and help intentions.

CCS CONCEPTS

• **Human-centered computing** → *User studies*; *Empirical studies in HCI*; • **Computer systems organization** → **External interfaces for robotics**; • **Computing methodologies** → Perception.

KEYWORDS

human-robot-interaction, needy robots, politeness, user experience

ACM Reference Format:
Martin Westhoven, Tim van der Grinten, and Steffen Mueller. 2019. Perceptions of a Help-Requesting Robot - Effects of Eye-Expressions, Colored Lights and Politeness of Speech. In *Mensch und Computer 2019 (MuC '19), September 8–11, 2019, Hamburg, Germany.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3340764.3340783

MuC '19, September 8–11, 2019, Hamburg, Germany

ACM ISBN 978-1-4503-7198-8/19/09…$15.00
https://doi.org/10.1145/3340764.3340783

1 INTRODUCTION

Robots are generally expected to act autonomously. Their role is typically thought of as a supporting one [5]. The increasing introduction of robots into more and more areas of human life, however, requires an increased adaptability to dynamic environments [38]. But equipping robots for all challenges posed by their possible environments is a serious challenge [32]. Problems can e.g. arise from handling environmental factors the robot was not originally built to handle. An example for such a problem is the use of doors [40]. Technologically closing capability gaps can often be complex and time-consuming. Considering human assistance is an approach to compensate for this [33], which can significantly improve system performance [32]. While asking for help is not always successful [33], at least some factors possibly influencing the success of robotic help requests are known [8, 9, 38]. Furthermore, knowledge of human-human-interaction can also be drawn upon [31, 41]. Knowledge regarding interaction effects of the influencing factors is much more rare. Physical form, the robot's voice, speech content, mimics, lights and movement can all affect the outcome by themselves, but so can their interactions. In addition, existing robot heads vary greatly in their physical appearance (see e.g. [27] for a compilation) and consequently in their abilities to display information. Therefore, designers face several problems, when designing a robotic request for help. First, the knowledge of isolated factors influencing the success of robotic help requests is sparse. Second, there is lack of knowledge on how these factors interact with each other. And third, they cannot be sure how this translates to their possibly unique robot features.

Especially in context of professional applications, the driving forces for the deployment of robots are mostly of economic nature. But also human needs in regards to working

conditions, or generally in regards to well-being, have to be considered. The first concern with robots is often physical safety. But the aspect of harmlessness of work also includes a psychological dimension (see e.g. [19]). In this regard, it is neither desirable to have humans being treated disrespectfully by robots, nor having them perform only the remaining tasks, which were either too hard to automate or not automatable at all. Thus, when designing interactive robots, the goal should not only be an efficient and safe, but also a pleasant and socially appropriate interaction.

Being faced with all of these requirements on robot interaction design, we decided to perform a web- and video-based study to both add to existing knowledge but also to explore and validate variants to be used with a specific robot used in an ongoing research project. More specifically, we decided to vary the eye-expression, lighting and politeness of speech. We then measured the identification performance regarding our intended expressions, the perceived politeness, the hedonic user experience (UX), and the help intention of the participants. Sociodemographic data, attitude towards robots and the general ability to recognize human faces were additionally measured as control variables. The rest of the paper is organized as follows: We provide an overview of related work, describe the method of the study, report and then discuss the results, and conclude with a summary.

2 RELATED WORK

The literature provides results on the user reception of robot head design features mostly for isolated criteria. First among them is the correct identification of designer-intended meaning. There are also some hints at interaction effects.

Additionally, there are many reports on theory-guided design of specific robot heads alone, e.g. Asheber et al. [1]. They can provide insights into the design space for robot heads, but do not directly provide results in terms of user studies. Some studies also include the robot's body into the evaluation, with benefits for evaluating the perception of the whole robot, but making it generally more difficult to isolate the causes for different effects. Tsiourti et al. [39] report results from a study comparing two full body robots, while trying to provide also the singled-out effects caused by face, head, body, voice and locomotion.

Bennet and Šabanović [4] report results from a study on minimalist robot facial features for emotion expression. They use the upper and lower outline of the eyes as well as those of the mouth and achieved high accuracy in expression identification. Eyes can be a subtle cue for observation and increase cooperative behavior [16]. They can transport emotions in human-robot interaction [6] and should therefore be present when social interaction is a core part of a robot. Since our design includes only the eyes, albeit in a more detailed style, expression identification performance will have to be checked again. The design of the eye part of the robot's head orients itself on 'in the wild'-examples. A more comic-like style was chosen due to results regarding the uncanny valley, that is the lowering of trust and likeability if a robot gets more human-like, but not enough so. Following the general understanding of the uncanny valley phenomenon, Mathur and Reichling [27] reported that staying on the mechanic-looks side of the mechano-humanness score range also yields high values for trust and likeability as well as low response times. The range of mechanic and human looking robot heads is depicted in figure 2, which was adapted from Mathur and Reichling [27]). Their findings are consistent with design guidelines for social robots by Duffy [13] and the suggestions of DiSalvo et al. [12], who evaluated the perceived humanness of robot heads. They provide design suggestions for humanoid robot heads, which integrate human and robot features to underline the robotic nature of a robot's head. Taking into account all of the above aspects, considering also the trade-off between economic and performance-wise aspects, the robot's head we used (see figure 1) was decided to be aimed at the area of machine-likeness highlighted in figure 2.

A sad or fearful look was chosen for situations in which the robot needs help. As Marsh, Ambady, and Kleck [26] found, this facilitates approach behaviours in perceivers, which is the exact thing the robotic system aims for in this situation. Lee, Šabanović, and Stolterman [23] performed a qualitative study on social robot design. Their participants reported that eyes should not be too far apart and also not be too detailed. Also overly large eyes were reported to be intimidating and to be inducing a feeling of surveillance. A study regarding the perceived child-likeness, masculinity, and femininity of a robot head is reported by Lütkebohle et al. [25]. Those attributes are reported to relate to perceived warmth and fitness for specific tasks. Lindberg et al. [24] developed eyes where the pupils can be dilated and contracted. Different light colors of the eyes were used to further clarify the intended meaning. They report how well participants of a study were able to guess the intended meaning and how distinctive they found the different expressions.

Song and Yamada [36] studied colored lights, vibrations, and sounds on a real, but simplistic robot. This served to explore the effects in general, but also to verify the intended effects of their design decisions. As Baraka and Veloso [2] summarize, lights are seldom coupled with the state of a robot, although they are sometimes used to underline emotion. In a series of studies, they consequently showed how lighting patterns can be used to clarify a robot's state.

Mensch und Computer (MuC), Hamburg, Germany, pp. 43-54, ACM 2019

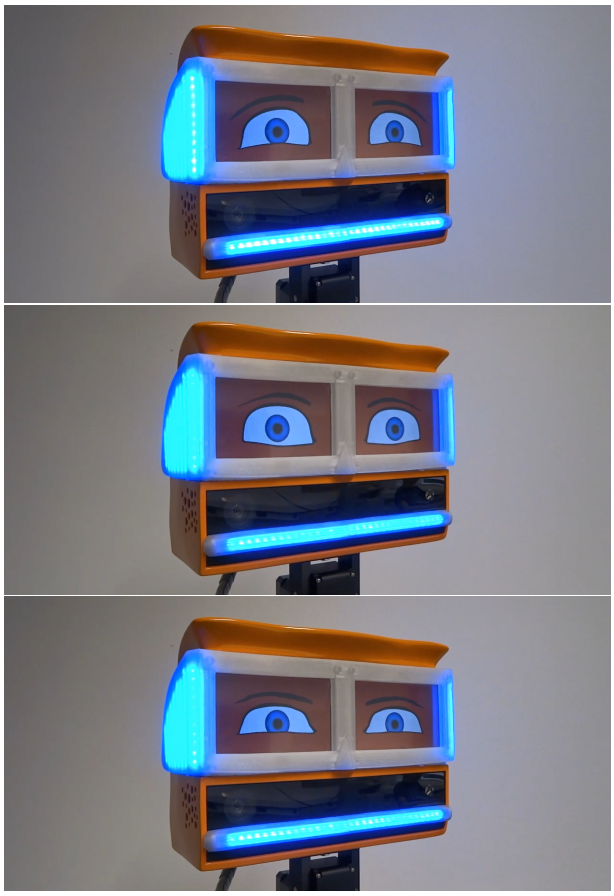


Figure 1: From top to bottom: Sad, neutral, and concentrated eye expression of the robot asking for help in our study.

The use of polite language by robots is motivated by general knowledge on human use of polite language. An influential theory in this area is Brown and Levinson's Politeness Theory [7]. Srinivasan et al. [38] could show, that polite language can increase a robot's success rate when asking for help. Positive effects of politeness strategies are also reported by Hammer et al. [20] for a robotic elderly assistant. Cameron et al. [9] add to this that stating limitations as the cause for a help request can increase a robot's success rate for help requests.

While all this prior work yields results for the effects of isolated criteria, there is a lack of work combining these results. In our work, the aim therefore was to study the combined effects of polite language, lighting and eye-expressions to gain further insights into how these variables can support or counter each others' effects.

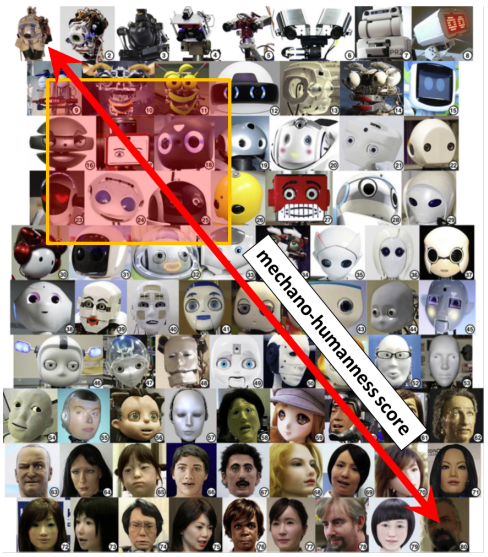


Figure 2: Robotic heads ordered from machine-like to human-like as compiled by Mathur and Reichling [27], and the target area of our design.

3 METHOD

A web-based questionnaire was used to compare variations of eye-expressions, colored lighting patterns and politeness of speech for the situation of being asked for help by a robot. Participants accessed the questionnaire online through devices of their own choice. As Dautenhahn et al. [11] could show, video-based evaluation of robots is feasible, as long as no physical interaction is involved. For statistical analysis, a mixed design MANOVA was planned, with non-parametric tests as a fallback option, should assumptions be violated.

Materials

This section provides information on the robot head as seen in the videos, on the actual contents of the videos and on the measures used to asses the effects of the variation of the independent variables.

Robot Head. The robot head used in our study was designed as a communication interface for a social robot. It is meant to provide a visual anchor to be used when interacting with the robot, to express internal states and to signal who is being addressed by eye-contact. The need to integrate a relatively large Kinect 2 sensor, consisting of cameras, infrared boosters, and a microphone array, constrained the design space regarding the minimum size. The maximum weight is constrained by the pan-tilt unit used. Since mechanical parts reduce the adaptability regarding future design choices and changes, the eyes were chosen to be represented by two displays. Figure 3 shows the head and it's components. The case

Figure 3: The robot's head with its components.

openings allow for the Kinect's built-in fan to provide ventilation. The sides of the head are equipped with two LED-arrays. Another single line of LEDs between the Kinect's cameras and microphones serves as the robot's mouth.

In terms of display content, the eyes are rendered using the OGRE 3D engine. They internally consist of an eyeball and a plane, which provides skin, eyelid, and eyebrow functionality. Different eye-expressions can be animated and also combined by using these objects. In a first step, the states angry, happy, sad, concentrated and neutral were implemented, of whom the latter three are used in this study. Additionally, eye blinking animations are possible, but were not used in this study.

Independent Variables - Video Contents. Participants were shown video-snippets of the robot asking for help. Each snippet was between ten and fifteen seconds long, depending on the politeness variant. A total of three variations for the eye-expressions were presented: A neutral, a sad and a concentrated expression. Politeness was varied between the two levels of direct and polite language, following the definitions from Brown and Levinson's Politeness Theory [7]. The direct variant was phrased for example as 'Please help me with ...'. In the polite variant, indirect speech was used to lessen the implied coercion resulting in the phrase 'It would be very nice, if you could help me with ...'. Since the perception of blinking lights is mainly studied in context of alarms and warnings, see e.g. Crawford [10], we decided to vary multiple blink-frequencies for exploratory analysis. Overall, LED-lighting was varied between two levels of color and three levels of blink-frequency against a control condition with all lights switched off. The colors shown were blue and green and each of them was presented at blink frequencies of 0 Hz, 0.5 Hz and 1 Hz. LED-lighting was designed as a between-groups measurement while eye-expression and politeness of speech were measured within-subjects. The resolution was 720p to also allow for participation over low-bandwidth connections. After each video-snippet, participants rated the situation they experienced with several questionnaire items. Figure 1 shows frames from the presented videos with the different eye-expressions and blue LED-lights.

Dependent Variables - Measurement. As in other work on robot design, one of the main criteria is the identification performance regarding the intended meaning. The other criteria focus on the pleasantness of the interaction, the perceived politeness, and on the resulting help intention. The questionnaire was based mostly on existing instruments. The first two questions asked the participants to answer, which facial expression they thought the robot showed. Nine alternatives were provided, of which eight were taken from the Facial Expression Identification instrument [4, 37]. The ninth alternative was added to include the concentrated expression as well. In addition, the participants could optionally mark one ore more further expressions they also thought could be fitting. Measuring facial expression recognition was deliberately chosen over emotion recognition, as facial expressions can also be fabricated to convey more universal non-verbal state information instead of just expressing emotional states [15]. The second part of the questionnaire consisted of the four items of the hedonic subscale of the short version of the User Experience Questionnaire (S-UEQ) by Schrepp et al. [35]. The items for the pragmatic subscale are not of primary relevance, since there is no real interaction between robot and human. To shorten the overall questionnaire length, they were thus left out. The third part of the questionnaire contained two items on the perceived politeness and rudeness of the robot's request after Salem et al. [34] and one item on how much time the participant would be willing to invest in this situation to help the robot after Pavey et al. [30].

Attitude towards robots and general functioning of facial expression identification were measured as well, to control for possible confounding effects. The Negative Attitudes towards Robots Scale (NARS) by Nomura et al. [29] was used to measure attitudes towards robots. A questionnaire on Prosopagnosia ('face blindness') by Kennerknecht et al. [22], and pictures of human facial expressions from the Cohn-Kanade dataset [21] in combination with the FEI-instrument [4, 37] (see above), were used to assess the participant's functionality of facial expression identification.

Hypotheses

We formulated the following hypotheses:

- **H1** The designed eye-expressions increase the identification of their intended meaning of 'sad', 'neutral' and 'concentrated'.
- **H2** The sad eye-expression increases help intention the most.
- **H3** Polite language increases the perceived politeness.
- **H4** Polite language increases the hedonic user experience.
- **H5** Polite language increases the help intention.

No hypotheses were formulated for the variation of LED-lighting, which was aimed to be subject to exploratory analysis.

Experiment Execution

The experiment was run over a period of four weeks. Typical experimental runs took about 20 minutes to complete. A target sample size of 139 was estimated for a within-between mixed design MANOVA using G*Power [17]. To correct for exclusions, a total of 157 participants had to be recruited to reach the planned 139 valid cases. The recruitment was mainly based on promotional platforms for web-based surveys (e.g. surveycircle[1]) and supplemented with recruitment from web-forums. No rewards were paid apart from those offered by the promotional platforms, typically a score to place own studies. All participants had the possibility to access these promotional rewards after finishing the survey. Participants were assigned to between-subjects levels randomly, but in a balanced manner, resulting in group sizes of 20 and one group (1 Hz green) of 19 participants. Exclusion was based on control items asking about the sincerity of answers and for one participant on being too young.

4 RESULTS

Of the recruited participants, 82 were female and 57 male (59% female). In each group, the share of females was between 50% to 70%. The age was distributed between 18 and 69 ($M = 27.48$, $SD = 6.92$). The age was evenly distributed, with some deviations due to the small share of older participants ($M = 25.45 - 30.00$, $SD = 2.783 - 9.975$). 98 participants reported being students and 30 being employees. 94 participants furthermore reported having a university or polytechnic degree and additional 36 reported having a high school degree. Non-parametric tests had to be used, because of the violation of several assumptions for parametric tests. The influence of the between-groups LED-lighting variation was analyzed with Kruskal-Wallis-Tests and post-hoc Mann-Whitney-U tests with Bonferroni-correction. The within-group variables eye-expression and politeness of speech were analyzed with Friedman's ANOVAs with post-hoc Wilcoxon signed-rank tests with Bonferroni-correction. Since chi square statistics with more than one degree of freedom are difficult to convert to an effect size [18, p. 555,565] and we are only really interested in focused comparisons, effect sizes are only reported for significant post-hoc tests. Also due to falling back to non-parametric tests, relations between the dependent variables All analyses were computed with IBM SPSS Statistics version 25.

[1] www.surveycircle.com

LED-lighting

A Kruskal-Wallis test with the independent variable LED-lighting regarding eye-expression and politeness of speech separately yielded significant results for help intention with neutral eye-expression $H(6) = 12.852$, $p = 0.045$, concentrated eye-expression $H(6) = 15.891$, $p = 0.014$, and direct language $H(6) = 12.770$, $p = 0.047$. For hedonic UX there was a significant effect with polite language $H(6) = 13.539$, $p = 0.035$. Post-hoc Mann-Whitney-U tests with LEDs off as a control group and with corrected significance level of $\alpha = 0.008\bar{3}$ were non-significant, however. A Kruskal-Wallis test with the independent variable LED-lighting regarding the combination of eye-expression and politeness of speech yielded significant results for the expression identification with concentrated eye-expression and polite language $H(6) = 15.006$, $p = 0.02$, for hedonic UX with concentrated eye-expression and polite language $H(6) = 15.631$, $p = 0.016$ and with sad eye-expression and polite language $H(6) = 12.795$, $p = 0.046$, and for help intention with concentrated eye-expression and direct language $H(6) = 15.604$, $p = 0.016$. Focused post-hoc Mann-Whitney-U tests with LEDs off as a control group and corrected significance level of $\alpha = 0.008\bar{3}$ were only significant for expression identification with concentrated eye-expression and polite language between the '0.5 Hz blue' pattern and 'LEDs off' $U = 130$, $p = 0.004$, $r = -0.4545$. With 'LEDs off' ($M = 0.35$, $SD = 0.489$), expression identification is higher compared to the '0.5 Hz blue' pattern ($M = 0$, $SD = 0$).

Eye-Expression

Friedman's ANOVAs for the three different eye-expressions yielded significant results for hedonic user experience $\chi^2(2) = 10.369$, $p = 0.006$, perceived politeness $\chi^2(2) = 42.792$, $p < 0.001$, facial expression identification $\chi^2(2) = 59.542$, $p < 0.001$ and help intention $\chi^2(2) = 53.043$, $p < 0.001$. Post-hoc Wilcoxon signed-rank tests used a corrected significance level of $\alpha = 0.01\bar{6}$.

The expression identification showed significant differences between sad and neutral, and between concentrated and neutral. All means and standard deviations are depicted in figure 4, test statistics for all significant post-hoc tests are shown in table 1.

For hedonic user experience, a significant difference was found between the sad and neutral expressions. Significant differences for perceived politeness were found between the sad and neutral, as well as between concentrated and sad expressions. The expression identification showed significant differences between the sad and neutral, and between the concentrated and neutral expressions. For help intention there were differences between all levels of eye-expressions.

Figure 4: Means and standard errors for the effect of eye-expression on expression identification, measured as the the share of interpretations consistent with our intended expression.

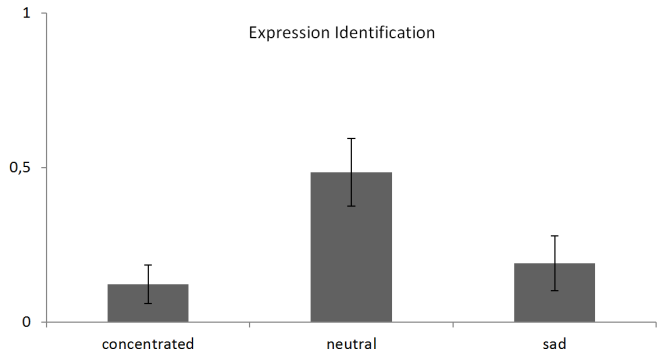


Table 1: Significant post-hoc Wilcoxon signed-rank tests for pair-wise comparisons of the eye-expressions. S = sad, N = neutral, C = concentrated.

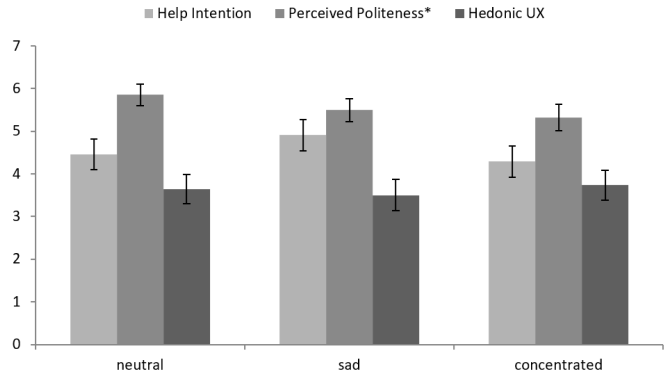
Dependent Variable	Comparison	T	p	r
Help Intention	S-N	771	< 0.001	-0.319
	S-C	686	< 0.001	-0.397
	N-C	1270	0.005	-0.167
Hedonic UX	S-N	1930.5	< 0.001	-0.222
Perceived Politeness	S-N	623.5	< 0.001	-0.29
	S-C	514.5	< 0.001	-0.337
Expression Identification	S-N	867	< 0.001	-0.331
	N-C	318.5	< 0.001	-0.417

Means and standard errors for help intention, perceived politeness and hedonic UX are depicted in figure 5.

Politeness of Speech

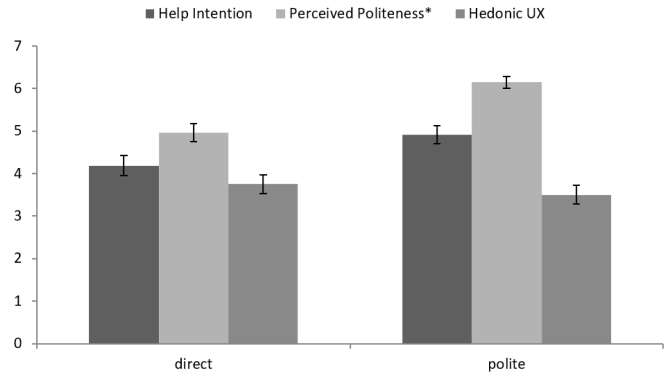
The effects of politeness of speech on the dependent variables were analyzed with Friedman's ANOVAs. For the independent variable politeness of speech, a significant effect for perceived politeness was found $\chi^2(1) = 84.746$, $p < 0.001$, $\phi = 0.522$. Polite language ($M = 4.391$, $SD = 0.588$) resulted in a higher politeness perception than the direct language ($M = 3.547$, $SD = 0.886$). Politeness of speech also had a significant effect on help intention $\chi^2(1) = 55.23$, $p < 0.001$, $\phi = 0.446$. Polite language ($M = 4.916$, $SD = 1.302$) resulted in a higher help intention than direct language ($M = 4.189$, $SD = 1.396$). Finally, there was a significant effect on the hedonic user experience $\chi^2(1) = 11.172$, $p < 0.001$, $\phi = 0.2$, where polite language ($M = 3.75$, $SD = 1.299$) resulted in a higher result on the hedonic subscale of the S-UEQ than

Figure 5: Means and standard errors for the effect of eye-expression on help intention, perceived politeness and the hedonic subscale of the S-UEQ. Perceived Politeness was upscaled from a five-point scale.



direct language ($M = 3.501$, $SD = 1.342$). All means and standard errors for the independent variable politeness of speech and the dependent variables perceived politeness, help intention, and hedonic UX are shown in figure 6.

Figure 6: Means and standard errors for the effect of polite language on help intention, perceived politeness and the hedonic subscale of the S-UEQ. Perceived Politeness was upscaled from a five-point scale.



Eye-Expression and Speech

For the combined effect of eye-expressions and politeness of speech, another set of Friedman's ANOVAs was performed. Significant effects were found for help intention $\chi^2(5) = 172.195$, $p < 0.001$, hedonic user experience $\chi^2(5) = 38.678$, $p < 0.001$, perceived politeness $\chi^2(5) = 248.539$, $p < 0.001$ and expression identification $\chi^2(5) = 86.721$, $p < 0.001$. Post-hoc Wilcoxon signed-rank tests used a corrected significance

level of $\alpha = 0.00\bar{3}$. The means and standard errors for help intention, perceived politeness and hedonic UX are shown in figure 7 and for expression identification in figure 8. For the sake of readability, the results from the significant post-hoc tests and corresponding effect sizes are depicted in the tables 2 (help intention, perceived politeness, hedonic UX) and 3 (expression identification).

Figure 7: Means and standard errors for the combined effect of politeness of speech and eye-expression on help intention, perceived politeness and the hedonic subscale of the S-UEQ. Perceived Politeness was up-scaled from a five-point scale.

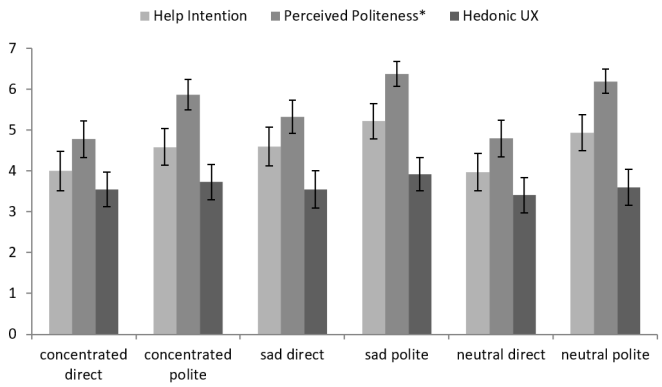
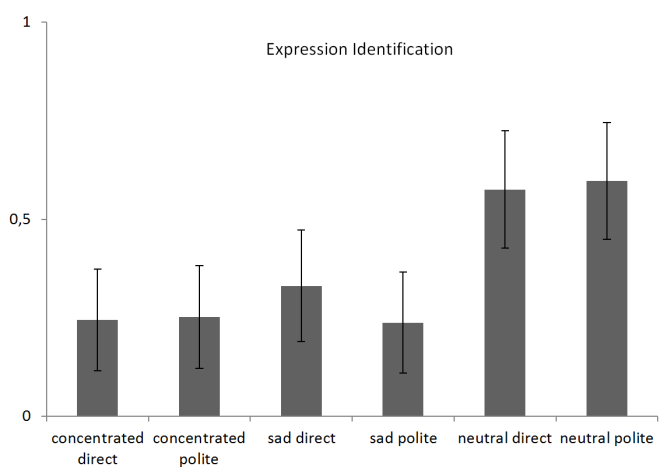


Figure 8: Means and standard errors for the combined effect of politeness of speech and eye-expression on expression identification, measured as the the share of interpretations consistent with the expression we intended.



Influence of Sociodemographics and Attitude

Since the assumptions for an analysis of covariance were violated, Pearson-correlations were computed to at least provide an overview over the relationships between covariates and dependent variables. There was a significant correlation between the subscales of the Negative Attitudes towards Robots Scale (NARS) and sex. The subscales of the NARS are as follows:

- **S1** Negative attitude regarding interaction situations with robots.
- **S2** Negative attitude regarding social influence of robots.
- **S3** Negative attitude regarding emotions in interactions with robots.

Females tended to higher values in negative attitudes towards robots (NARS S1: $r = -0.283$, $p = 0.001$; NARS S2: $r = -0.218$, $p = 0.01$; NARS S3: $r = -0.302$, $p < 0.001$). Sex is also related with perceived politeness for the concentrated eye-expression ($r = 0.184$, $p = 0.03$), where males rated politeness higher. Males showed a higher identification rate for the neutral eye-expression ($r = 0.173, p = 0.041$), as well as a generally higher identification rate for eye-expressions when polite language was used ($r = 0.174, p = 0.04$). There was a significant negative relationship between age and perceived politeness of the polite language variant ($r = -0.168$, $p = 0.049$).

The NARS S1 subscale showed significant negative relationships with help intention and perceived politeness for all within-subject variable levels. The S2 and S3 subscales also show significant negative relationships with help intention. For perceived politeness, S2 and S3 only show negative relationships for direct language and neutral eye-expression. S2 is also negatively related with perceived politeness for the concentrated eye-expression. All correlation coefficients for NARS and the dependent variables with two-tailed significance are shown in table 4.

There were no significant effects for the Prosopagnosia questionnaire or expression identification performance on pictures of human facial expressions.

Correlations between Dependent Variables

Non-parametric correlations using Kendall's Tau show significant relations for overall scores between the helping intention and hedonic UX ($r = 0.220$, $p < 0.001$), as well as between helping intention and perceived politeness ($r = 0.363$, $p < 0.001$). Highly significant correlations at similarly small to medium effect sizes also show for all variations of polite language and eye expressions (see table 5).

5 DISCUSSION

The discussion is structured into subsections on the methodology, the main effects of the independent variables, possible

Table 2: Significant post-hoc Wilcoxon signed-rank tests on help intention, perceived politeness and hedonic UX for the different combinations of eye-expressions and politeness of speech. Empty cells were non-significant. C = concentrated, N = neutral, S = sad, D = direct, P = polite.

Combination		Help Intention				Perceived Politeness				Hedonic UX			
		Z	Sig.	T	r	Z	Sig.	T	r	Z	Sig.	T	r
CD	CP	−4.510	< 0.001	796	−0.270	−6.910	< 0.001	314.5	−0.414				
	NP	−6.572	< 0.001	635	−0.338	−8.175	< 0.001	114	−0.490				
	SD	−5.370	< 0.001	577.5	−0.322	−4.440	< 0.001	383	−0.266				
	SP	−7.251	< 0.001	613	−0.435	−8.125	< 0.001	331.5	−0.487	−3.912	< 0.001	1457	−0.235
CP	ND	−4.825	< 0.001	1103.5	−0.289	−6.694	< 0.001	366.5	−0.401	−3.282	0.001	1399	−0.197
	NP	−4.023	< 0.001	465	−0.241	−3.083	0.002	268	−0.185				
	SD					−3.781	< 0.001	895.5	−0.227				
	SP	−5.628	< 0.001	500	−0.338	−4.173	< 0.001	304	−0.250				
ND	NP	−7.058	< 0.001	367.5	−0.423	−7.729	< 0.001	358.5	−0.464				
	SD	−5.353	< 0.001	602	−0.321	−4.486	< 0.001	470	−0.263				
	SP	−7.499	< 0.001	432	−0.450	−8.028	< 0.001	381.5	−0.481	−4.935	< 0.001	1093	−0.296
NP	SD	−3.747	< 0.001	957.5	−0.225	−6.566	< 0.001	315	−0.394				
	SP									−3.642	< 0.001	1538.5	−0.218
SD	SP	−5.191	< 0.001	802.5	−0.311	−6.788	< 0.001	313	−0.407	−3.848	< 0.001	1209.5	−0.231

Table 3: Significant post-hoc Wilcoxon signed-rank tests on expression identification for the different combinations of eye-expressions and politeness of speech. C = concentrated, N = neutral, S = sad, D = direct, P = polite.

Combination		Expression Identification			
		Z	Sig.	T	r
CD	ND	−5.208	< 0.001	632	−0.312
	NP	−5.513	< 0.001	600	−0.331
CP	ND	−5.341	< 0.001	468	−0.320
	NP	−5.237	< 0.001	765	−0.314
ND	SD	−4.064	< 0.001	639	−0.244
	SP	−5.356	< 0.001	585	−0.321
NP	SD	−4.331	< 0.001	666	−0.260
	SP	−5.455	< 0.001	722.5	−0.327

interaction effects between the independent variables and effects from covariates.

Method

Due to violations of assumptions for parametric tests, non-parametric tests were used as an alternative. The a priori sample size estimation based on parametric tests thus yielded a much smaller number of required participants than would have been optimally needed for post-hoc tests to reliably detect significant effects. Especially the significant results of the Kruskal-Wallis-H test for variations in LED-lighting could not be confirmed by post-hoc Mann-Whitney-U tests, most probably due to a lack of test power of the latter.

Eye-Expression

The three eye-expressions differed in how well they were identified by the participants. The neutral expression scored higher in this regard than both the sad and concentrated expressions. This indicates optimization potential for the latter. Table 6 highlights how the expressions were interpreted by the participants instead. Hypothesis H1 'The designed eye-expressions increase the identification of their intended meaning of 'sad', 'neutral' and 'concentrated'.' in conclusion only holds for the neutral eye-expression and has to be rejected. This knowledge will serve to further refine the respective expressions and to further distinguish them from the non-intended interpretations. Future research in this direction should also address the question, if acceptable identification performance can be achieved at all with only the eye-part of the face.

The hedonic UX was significantly higher for the sad eye-expression than for the neutral expression. The sad expression further increased the perceived politeness, compared to both the neutral and the concentrated expressions. Finally, help intention was highest for the sad eye-expression, followed by the neutral and the concentrated expressions. Hypothesis H2 ('The sad eye-expression increases help intention the most') is thus confirmed. This is in line with insights from human-human interaction that sad faces provoke sympathy and help intentions [14]. Furthermore, it's known that the expression of emotions by robots strengthens social acceptance [3], which could also affect the hedonic UX. Taking

Table 4: Significant correlations between the negative attitudes towards robots scale (NARS), and help intention and perceived politeness under different variations of within-subjects variables with corresponding p-values. Empty cells were non-significant.

NARS Subscale		Help Intention					Perceived Politeness				
		Eye Expression			Politeness of Speech		Eye Expression			Politeness of Speech	
		Conc.	Neutral	Sad	Direct	Polite	Conc.	Neutral	Sad	Direct	Polite
S1	Corr.	-0.372	-0.337	-0.327	-0.386	-0.306	-0.329	-0.378	-0.421	-0.453	-0.253
	Sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.003
S2	Corr.	-0.367	-0.338	-0.250	-0.381	-0.256	-0.189	-0.222		-0.297	
	Sig.	< 0.001	< 0.001	0.003	< 0.001	0.002	0.026	0.009		< 0.001	
S3	Corr.	-0.419	-0.400	-0.297	-0.425	-0.319		-0.178		-0.198	
	Sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001		0.037		0.019	

Table 5: Significant non-parametric correlations between the dependent variables of help intention, perceived politeness and hedonic user experience under different variations of within-subjects variables with corresponding p-values.

		Help Intention				
		Eye Expression			Politeness of Speech	
		Conc.	Neutral	Sad	Direct	Polite
Perceived Politeness	Corr.	0.411	0.343	0.381	0.429	0.336
	Sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Hedonic UX	Corr.	0.199	0.204	0.238	0.196	0.251
	Sig.	0.001	0.001	< 0.001	0.001	< 0.001

into consideration the large share of alternative interpretations as 'fearful' and 'neutral', future research should focus fearful expressions as well. Overall, the results confirm and add to the reported positive effects of sad expressions on onlookers by Marsh et al. [26].

LED-Lighting

Although the Kruskal-Wallis-H tests indicate significant differences between the varied LED-patterns, post-hoc tests confirmed only one very special case, where '0.5 Hz blue' in combination with using polite language seemed to inhibit the correct identification for the concentrated expression completely, whereas for 'LEDs off' this was significantly better. As can be seen in table 6, the concentrated expression suffers from a generally low correct identification rate. Nonetheless, this result hints at possibly distracting or even contradicting effects of lighting. As explained above, the post-hoc tests for LED-lighting had very low statistical power, warranting for further studies regarding this aspect. This is especially true since there seems to be no readily available literature on possibly interacting effects from light and other communication channels.

Politeness of Speech

For the polite language variant, the perceived politeness, hedonic UX, and help intention were significantly higher than for the direct variant. The hypotheses H3-H5 are thus confirmed. These results are consistent with the literature in that politeness can increase successful help requests, which obviously requires an increased help intention by helpers (see e.g. [38]).

Relations among Dependent Variables

Relations between the dependent variables were to be expected from literature, but also because the constructs are logically linked to each other. Significant correlations were found between the help intention and both hedonic user experience and perceived politeness. As to how exactly the hedonic user experience and perceived politeness are related to help intention and also to actual helping behavior, future research should connect to work on modeling the relationships of constructs surrounding helping behavior towards robots, as reported by e.g. Budde et al. [8].

Possible Interaction Effects

Due to the non-parametric tests used, interactions could not be analyzed directly. Instead, only the combined measures for variations of the within-variables can be drawn upon. The hedonic UX is significantly different only between some combinations of politeness of speech and eye-expressions. The significant differences are between the highest scoring combination of 'sad-polite' and all direct language variants

Table 6: Confusion matrix for the tested eye-expressions and answer options. All cells show the percentage of answers for the shown expression in the respective rows. The percentage of replies for the intended expression is underlined. Higher percentages of replies in non-intended categories than in the actually intended category are highlighted in gray.

Shown Expression	Answer Option								
	Angry	Happy	Sad	Fearful	Neutral	Surprised	Disgusted	Bored	Concelntrated
Sad	1.80	2.88	19.06	32.37	32.73	3.24	0.36	0.72	6.82
Neutral	6.14	3.25	0.72	19.49	48.74	12.64	0.00	1.44	7.58
Concentrated	31.29	2.16	2.16	3.96	42.45	2.88	1.44	1.44	12.23

plus the 'neutral-polite' combination. Also, there is a difference between 'concentrated-polite' and 'neutral-direct'. The results support the use of the 'sad-polite' combination to maximize the hedonic user experience regarding the studied combinations.

Only three of the fifteen pairwise comparisons for help intention are non-significant, of whom two differ only in the eye-expression component. As with the hedonic user experience, the 'sad-polite' combination maximizes the help intention among the studied combinations.

The perceived politeness does not show any obvious anomalies for the combined effect of politeness of speech and eye-expressions. The only two non-significant comparisons include concentrated and neutral with direct language, and sad and neutral with polite language. Both differ only in the eye-expression component, whose single effect on perceived politeness is generally smaller than the politeness of speech component.

For the expression identification, the combined effect of politeness of speech and eye-expressions does not seem to alter the results compared to eye-expression only: Both neutral-eyes variants show significantly higher correct identification rates than all non-neutral.

Overall, the polite language variant and the sad eye-expression yield the highest scores for hedonic UX and help intention when taken together. The results do not allow for any specific interpretation of a possible interaction effect. Further research will be required to better understand if both variables interact or if it is only their added effect showing here.

Effects from Covariates

Although a planned analysis of covariance could not be performed, significant and strong correlations between negative attitudes towards robots and help intention, and in parts also perceived politeness, indicate a possible effect. This would match with literature, as it is known that attitudes influence actual behavior when interacting with robots [28]. Considering prevalent attitudes in a target population is of increased significance when a target population for robotic deployment is expected to show higher than average negative attitudes, as the effect can compromise robot mission designs relying on human help. There are weaker significant correlations for sex and expression identification, as well as age and perceived politeness. Though these could be rooted in socialization effects, especially the age and politeness correlation, further research is required to isolate the causes.

6 CONCLUSION

We performed a web- and video-based study on the effects of colored lights, eye-expressions and politeness of speech on how a robot's request for help is perceived by humans. Using the sad eye-expression yields benefits regarding the hedonic UX, perceived politeness and help intention. However, results regarding the expression identification performance suggest that there is further potential for optimization regarding the sad as well as the concentrated expression. The sad expression was more often identified as fearful or neutral, and the concentrated expression more often as angry or neutral. Careful consideration of this fact is appropriate, until more precise follow-up studies can further clarify the results. Polite language increases the hedonic UX, perceived politeness, and help intention. In combination with the eye-expression, the 'sad-polite' combination again yields the best results. The results for the LED-lights indicate significant effects. Due to the attributes of the data set and the consequently performed non-parametric analyses, post-hoc analysis was not able to confirm these effects. Follow-up studies using a within-design will have to be performed to identify the possible effects. Negative attitudes towards robots were shown to have strong correlations to help intention and partially to perceived politeness. Age and sex show weaker correlations to expression identification and perceived politeness. Both, the negative attitudes and sociodemographic attributes, call for special consideration when designing robot interaction for specific target populations. It has to be noted though, that all of these effects show up for few and short interaction sequences, while prolonged exposure to help-requesting robots could result in different preferences, e.g. a briefer language style.

In summary, the results imply that a robot sporadically asking humans for help should be designed to use eye-expressions, which are interpreted as sad or fearful, in combination with

polite language, as defined by Brown and Levinson's Politeness Theory [7], to create a pleasant interaction. This combination yields significantly higher hedonic user experience, perceived politeness, and help intention than the other tested combinations.

ACKNOWLEDGMENTS

This research is funded by the German Federal Ministry of Education and Research (BMBF) within the Project FRAME (16SV7829K) and managed by the VDI-VDE IT. The authors are responsible for the contents of this publication.

REFERENCES

[1] Wagshum Techane Asheber, Chyi-Yeu Lin, and Shih Hsiang Yen. 2016. Humanoid head face mechanism with expandable facial expressions. *International Journal of Advanced Robotic Systems* 13, 1 (2016), 29. https://doi.org/10.5772/62181

[2] Kim Baraka and Manuela M Veloso. 2018. Mobile service robot state revealing through expressive lights: Formalism, design, and evaluation. *International Journal of Social Robotics* 10, 1 (2018), 65–92. https://doi.org/10.1007/s12369-017-0431-x

[3] Jenay M Beer, Cory-Ann Smarr, Arthur D Fisk, and Wendy A Rogers. 2015. Younger and older usersŒş recognition of virtual agent facial expressions. *International journal of human-computer studies* 75 (2015), 1–20.

[4] Casey C Bennett and Selma Šabanović. 2014. Deriving minimal features for human-like facial expressions in robotic faces. *International Journal of Social Robotics* 6, 3 (2014), 367–381. https://doi.org/10.1007/s12369-014-0237-z

[5] Mike Blow, Kerstin Dautenhahn, Andrew Appleby, Chrystopher L. Nehaniv, and David Lee. 2006. The Art of Designing Robot Faces: Dimensions for Human-robot Interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction (HRI '06)*. ACM, New York, NY, USA, 331–332. https://doi.org/10.1145/1121241.1121301

[6] Cynthia Breazeal. 2003. Emotion and sociable humanoid robots. *International journal of human-computer studies* 59, 1-2 (2003), 119–155. https://doi.org/10.1016/S1071-5819(03)00018-1

[7] Penelope Brown and Stephen C Levinson. 1978. Universals in language usage: Politeness phenomena. In *Questions and politeness: Strategies in social interaction*. Cambridge University Press, 56–311.

[8] Vanessa Budde, Nils Backhaus, Patricia H Rosen, and Sascha Wischniewski. 2018. Needy Robots-Designing Requests for Help Using Insights from Social Psychology. In *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. IEEE, 48–53. https://doi.org/10.1109/ARSO.2018.8625724

[9] David Cameron, Ee Jing Loh, Adriel Chua, Emily Collins, Jonathan M Aitken, and James Law. 2016. Robot-stated limitations but not intentions promote user assistance. (2016).

[10] A Crawford. 1963. The perception of light signals: The effect of mixing flashing and steady irrelevant lights. *Ergonomics* 6, 3 (1963), 287–294. https://doi.org/10.1080/00140136308930708

[11] Kerstin Dautenhahn, Michael Walters, Sarah Woods, Kheng Lee Koay, Chrystopher L Nehaniv, A Sisbot, Rachid Alami, and Thierry Siméon. 2006. How may I serve you?: a robot companion approaching a seated person in a helping context. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 172–179. https://doi.org/10.1145/1121241.1121272

[12] Carl F DiSalvo, Francine Gemperle, Jodi Forlizzi, and Sara Kiesler. 2002. All robots are not created equal: the design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM, 321–326. https://doi.org/10.1145/778712.778756

[13] Brian R Duffy. 2003. Anthropomorphism and the social robot. *Robotics and autonomous systems* 42, 3-4 (2003), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

[14] Nancy Eisenberg, Richard A Fabes, Paul A Miller, Jim Fultz, Rita Shell, Robin M Mathy, and Ray R Reno. 1989. Relation of sympathy and personal distress to prosocial behavior: a multimethod study. *Journal of personality and social psychology* 57, 1 (1989), 55.

[15] Paul Ekman. 1993. Facial expression and emotion. *American psychologist* 48, 4 (1993), 384.

[16] Max Ernest-Jones, Daniel Nettle, and Melissa Bateson. 2011. Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior* 32, 3 (2011), 172–178. https://doi.org/10.1016/j.evolhumbehav.2010.10.006

[17] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

[18] Andy Field. 2009. *Discovering statistics using SPSS*. Sage publications.

[19] Winfried Hacker. 2013. *Psychologische Bewertung von Arbeitsgestaltungsmassnahmen: Ziele und Bewertungsmassstäbe*. Vol. 1. Springer-Verlag.

[20] Stephan Hammer, Birgit Lugrin, Sergey Bogomolov, Kathrin Janowski, and Elisabeth André. 2016. Investigating politeness strategies and their persuasiveness for a robotic elderly assistant. In *International Conference on Persuasive Technology*. Springer, 315–326. https://doi.org/10.1007/978-3-319-31510-2_27

[21] Takeo Kanade, Jeffrey F Cohn, and Yingli Tian. 2000. Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. IEEE, 46–53. https://doi.org/10.1109/AFGR.2000.840611

[22] Ingo Kennerknecht, Thomas Grueter, Brigitte Welling, Sebastian Wentzek, Jürgen Horst, Steve Edwards, and Martina Grueter. 2006. First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *American Journal of Medical Genetics Part A* 140, 15 (2006), 1617–1622. https://doi.org/10.1002/ajmg.a.31343

[23] Hee Rin Lee, Selma Šabanović, and Erik Stolterman. 2016. How humanlike should a social robot be: A user-centered exploration. In *2016 AAAI Spring Symposium Series*.

[24] Markus Lindberg, Hannes Sandberg, Marcus Liljenberg, Max Eriksson, Birger Johansson, and Christian Balkenius. 2017. The Expression of Mental States in a Humanoid Robot. In *International Conference on Intelligent Virtual Agents*. Springer, 247–250. https://doi.org/10.1007/978-3-319-67401-8_32

[25] Ingo Lütkebohle, Frank Hegel, Simon Schulz, Matthias Hackel, Britta Wrede, Sven Wachsmuth, and Gerhard Sagerer. 2010. The bielefeld anthropomorphic robot head 'Flobi'. In *2010 IEEE International Conference on Robotics and Automation*. IEEE, 3384–3391. https://doi.org/10.1109/ROBOT.2010.5509173

[26] Abigail A Marsh, Nalini Ambady, and Robert E Kleck. 2005. The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion* 5, 1 (2005), 119. https://doi.org/10.1037/1528-3542.5.1.119

[27] Maya B Mathur and David B Reichling. 2016. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition* 146 (2016), 22–32. https://doi.org/10.1016/j.cognition.2015.09.008

[28] Tatsuya Nomura, Takayuki Kanda, and Tomohiro Suzuki. 2006. Experimental investigation into influence of negative attitudes toward robots

on human–robot interaction. *Ai & Society* 20, 2 (2006), 138–150.

[29] Tatsuya Nomura, Takayuki Kanda, Tomohiro Suzuki, and Kennsuke Kato. 2004. Psychology in human-robot communication: An attempt through investigation of negative attitudes and anxiety toward robots. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE, 35–40. https://doi.org/10.1109/ROMAN.2004.1374726

[30] Louisa Pavey, Tobias Greitemeyer, and Paul Sparks. 2012. 'I help because I want to, not because you tell me to' empathy increases autonomously motivated helping. *Personality and Social Psychology Bulletin* 38, 5 (2012), 681–689. https://doi.org/10.1177/0146167211435940

[31] Byron Reeves and Clifford Ivar Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge university press.

[32] Stephanie Rosenthal, Joydeep Biswas, and Manuela Veloso. 2010. An effective personal mobile robot agent through symbiotic human-robot interaction. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1-3*. International Foundation for Autonomous Agents and Multiagent Systems, 915–922. https://doi.org/10.1145/1838206.1838329

[33] Stephanie Rosenthal, Manuela Veloso, and Anind K Dey. 2011. Task behavior and interaction planning for a mobile service robot that occasionally requires help. In *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.

[34] Maha Salem, Micheline Ziadee, and Majd Sakr. 2014. Marhaba, how may I help you? Effects of politeness and culture on robot acceptance and anthropomorphization. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 74–81. https://doi.org/10.1145/2559636.2559683

[35] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *IJIMAI* 4, 6 (2017), 103–108. https://doi.org/10.9781/ijimai.2017.09.001

[36] Sichao Song and Seiji Yamada. 2017. Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2–11. https://doi.org/10.1145/2909824.3020239

[37] Stefan Sosnowski, Ansgar Bittermann, Kolja Kuhnlenz, and Martin Buss. 2006. Design and evaluation of emotion-display EDDIE. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3113–3118. https://doi.org/10.1109/IROS.2006.282330

[38] Vasant Srinivasan and Leila Takayama. 2016. Help me please: Robot politeness strategies for soliciting help from humans. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 4945–4955. https://doi.org/10.1145/2858036.2858217

[39] Christiana Tsiourti, Astrid Weiss, Katarzyna Wac, and Markus Vincze. 2017. Designing Emotionally Expressive Robots: A Comparative Study on the Perception of Communication Modalities. In *Proceedings of the 5th International Conference on Human Agent Interaction (HAI '17)*. ACM, New York, NY, USA, 213–222. https://doi.org/10.1145/3125739.3125744

[40] Rodrigo Ventura. 2014. Two Faces of Human–Robot Interaction: Field and Service Robots. In *New Trends in Medical and Service Robots*. Springer, 177–192. https://doi.org/10.1007/978-3-319-05431-5_12

[41] James E Young, Richard Hawkins, Ehud Sharlin, and Takeo Igarashi. 2009. Toward acceptable domestic robots: Applying insights from social psychology. *International Journal of Social Robotics* 1, 1 (2009), 95. https://doi.org/10.1007/s12369-008-0006-y