

Research Article

Martin Westhoven* and Tim van der Grinten

Communicating Robotic Help Requests

Effects of Eye-Expressions, LED-Lights and Polite Language

<https://doi.org/10.1515/icom-2020-0013>

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 the used language 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. In this paper, the focus is placed on interaction effects and on the influence of covariates. Significant effects were found for the interaction of LED lighting and eye-expressions and for language and eye-expressions on help intention. The expression identification is significantly influenced by the interaction of LED lighting and eye-expressions. Several significant effects of the covariates were found, both direct and from interaction with independent variables. Especially the negative attitudes towards robots significantly influence help intention and perceived politeness. The results provide information on the effect of different design choices for help requesting robots.

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

1 Introduction

Robots are generally expected to act autonomously. Their role is typically thought of as a supporting one [4]. The increasing introduction of robots into more and more areas of human life requires an increased adaptability to dynamic environments [38]. However, equipping robots for all challenges posed by their possible environments

is a serious challenge [31]. 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 [42]. Technologically closing capability gaps can often be complex and time-consuming. Considering human assistance is an approach to compensate for this [32], which can significantly improve system performance [31]. While asking for help is not always successful [32], at least some factors possibly influencing the success of robotic help requests are known [8, 38, 7] and will be discussed in the related work section. Furthermore, knowledge of human-human-interaction can also be drawn upon [30, 44]. 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: (1) The knowledge of isolated factors influencing the success of robotic help requests is sparse. (2) There is lack of knowledge on how these factors interact with each other. And (3), they cannot be sure how this translates to their possibly unique robot features.

As an example, consider deploying an existing robot with voice and light outputs and an otherwise neutral robotic face. Even if there was perfect knowledge on how voice and lights will be perceived as isolated factors, we would still not know if the perception of a polite greeting could be influenced by the displayed light. And even if this interaction effect would be known from another robot with differently attached lights, we could not be sure if this translates well to our robot.

Often, and especially in the context of professional applications, the driving forces for the deployment of robots are mostly of economic nature. Nonetheless, 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

*Corresponding author: Martin Westhoven, Federal Institute for Occupational Safety and Health, Dortmund, Germany, e-mail: westhoven.martin@baua.bund.de

Tim van der Grinten, Ilmenau Technical University, Ilmenau, Germany, e-mail: tim.van-der-grinten@tu-ilmenau.de

of harmlessness of work also includes a psychological dimension (see e. g. [17]). In this regard, it is neither desirable to have humans being treated disrespectfully or being pressured by robots. Thus, when designing interactive robots, the goal should not only be an efficient and safe, but also a pleasant and socially appropriate interaction.

Consider a robot with dynamic facial expressions and voice output. One choice for using the face could be to passively display the robot's state while the voice output is used for direct interaction. From human-human interaction we know that facial expressions can radically change the intended as well as the perceived meaning of spoken words. Say the robot puts on a concentrated expression to signal being busy with a task. Then, a perfectly polite request could very well be perceived as being passive-aggressive.

Being faced with all the required knowledge on robot interaction design choices, we decided to perform a web- and video-based study with a help request from a robot to both add to existing knowledge, but also to explore and validate interaction 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 the 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. A first analysis with non-parametric methods was reported in Westhoven et al. [43]. Building upon this work, we re-analyzed the data with linear mixed models to better capture interaction effects and to include the effects from the covariates.

The remainder 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 [3] 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 [15]. They can transport emotions in human-robot interaction [5] and should therefore be present when social interaction is a core task 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.

Following the general understanding of the uncanny valley phenomenon, Mathur and Reichling [27] reported that staying on the mechanic-looks side of the mechano-humaneness 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 1, which was adapted from Mathur and Reichling [27]. Their findings are consistent with design guidelines for social robots by Duffy [12] and the suggestions of DiSalvo et al. [11], who evaluated the perceived humaneness 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 2) was decided to be aimed at the area of machine-likeness highlighted in figure 1.

As Marsh, Ambady, and Kleck [26] found, a sad or fearful look facilitates approach behaviours in perceivers, which is the exact thing a robotic system aims for in situations where it needs help. Lee, Šabanović, and Stolterman [22] 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. [24]. Those attributes are reported to relate to perceived warmth and fitness for spe-

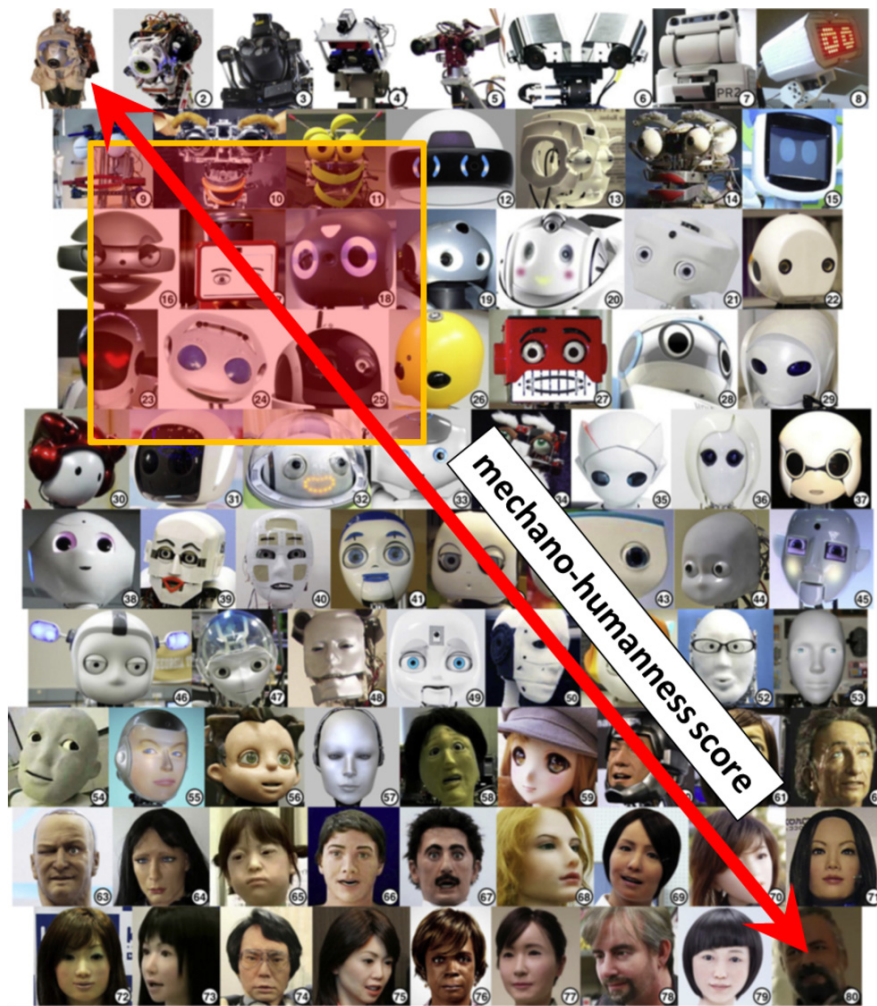


Figure 1: Robotic heads ordered from machine-like to human-like as compiled by Mathur and Reichling [27], and the target area of our design.

cific tasks. Lindberg et al. [23] 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.

Another known influencing factors in help requests is the used language. The use of polite language by robots

is motivated by general knowledge on effects of human-human use of politeness. An influential theory in this area is Brown and Levinson's Politeness Theory [6], which centers on the concept of avoidance of face loss towards interaction partners. 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. [19] for a robotic elderly assistant. Cameron et al. [8] 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 experiment, 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.

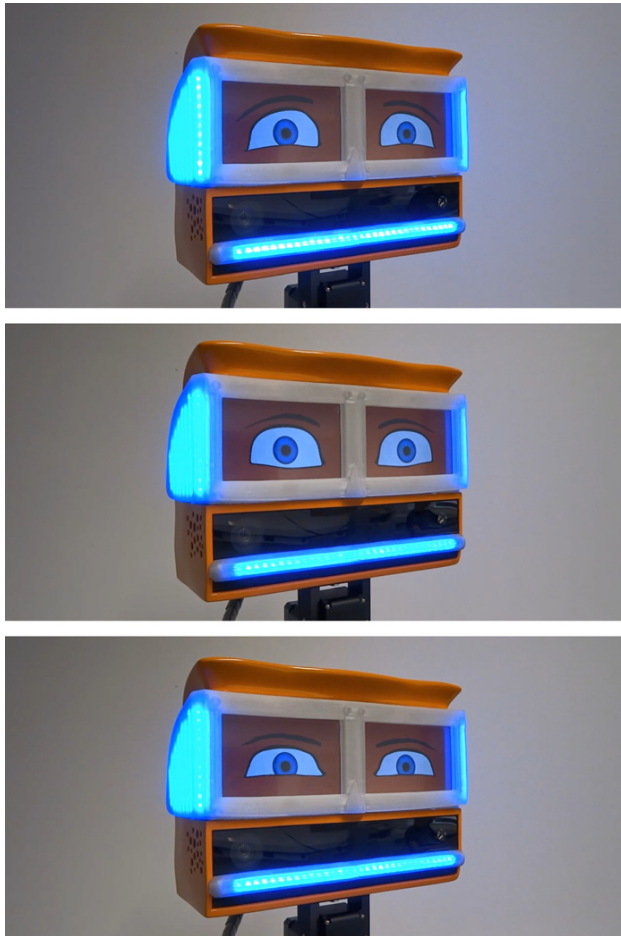


Figure 2: From top to bottom: Sad, neutral, and concentrated eye expression of the robot asking for help in our study.

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. [10] could show, video-based evaluation of robots is feasible, as long as no physical interaction is involved.

3.1 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 assess the effects of the variation of the independent variables.

3.1.1 Robot Head

The robot head used in our study was designed as a communication interface for a real social robot, which was built for the research project FRAME.¹ 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 its components. The case 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 eye-ball and a plane, which provides skin, eyelid, and eyebrow functionality. 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. 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. A detailed description of the robot's head can be found in van der Grinten et al. [41].

3.1.2 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 [6]. The direct variant was phrased for example as 'Please help me with ...'. In the polite vari-

¹ <http://www.frame-projekt.de/>



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

ant, 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 [9], 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 2 shows frames from the presented videos with the different eye-expressions and blue LED-lights.

3.1.3 Dependent Variables – Measurement

As in other work on robot design, one of the main performance 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 [3, 37]. The ninth alternative was added to include the concentrated expression we used in the experimental manipulation as well. In

addition, the participants could optionally mark one or 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 [14].

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. [34]. 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. [33] and one item on how much time the participant would be willing to invest in this situation to help the robot after Pavay et al. [29].

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. [28] was used to measure attitudes towards robots. A questionnaire on Prosopagnosia ('face blindness') by Kennerknecht et al. [21], and pictures of human facial expressions from the Cohn-Kanade dataset [20] in combination with the FEI-instrument [3, 37] (see above), were used to assess the participant's functionality of facial expression identification.

3.2 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 [16]. 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. surveycircle²) 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 Findings

In this Section, we report and discuss the findings of our experiment and point out practical implications.

The results are drawn from a follow-up analysis to Westhoven et al. [43] with linear mixed-effect models, a generalization of the general linear model. This, among others, allows to include significant interaction effects between covariates and independent variables into the regression model. In Westhoven et al. [43], non-parametric tests were 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. Also due to falling back to non-parametric tests, relations between the dependent variables could not be derived.

After introducing the recruited sample, we report all significant effects as well as the significant parameter estimates grouped by the dependent variables. We further provide more insightful graphs regarding the interaction effects. For improved readability, discussion and implications of the results are provided directly in each dependent variable's section. The main effects from Westhoven et al. [43] are not depicted again, since there was no deviation from the original results.

4.1 Sample Description

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.

4.2 Help Intention

As a requirement for the planned provision of help, help intention was measured in our web-based study to stand in for observable helping behaviour. In the following subsections, we provide results, discussion and implications.

4.2.1 Results for Help Intention

For help intention, the model did not need adjustment for interactions between covariates and independent variables. Significant main effects were found for LED-Lighting $F(6, 126) = 2.716$, $p = 0.016$, Eye-Expression $F(2, 672) = 32.842$, $p < 0.001$ and Language $F(1, 672) = 127.502$, $p < 0.001$. Also, significant interaction effects between Eyes and LEDs $F(12, 672) = 2.358$, $p = 0.006$ and between Eyes and Language $F(2, 672) = 3.645$, $p = 0.027$ were found. For the covariates, only the NARS S3 Subscale "Emotions" showed a significant effect $F(1, 126) = 8.437$, $p = 0.004$. The parameter estimates for significant levels of the reported terms are depicted in table 1. All results for LEDs and Eyes are tested in comparison with their respective neutral level.

The pairwise interactions of the independent variables LED, Eye-Expression and Language and their effects on Help Intention are depicted in the figures 4–6. The respective mean values are color coded to express the general level of Help Intention associated with the combinations of the independent variables' levels.

The interaction between LEDs and Eye-Expressions in figure 4 clearly shows the significant results for the blue LEDs and the concentrated expression. Although not significant in comparison to neutral levels of the independent variables, peak levels can be seen for the combination of constant and 1 Hz green LED lighting and a sad Eye-Expression.

² www.surveycircle.com

Table 1: Parameter estimates from the linear mixed-effect model for help intention.

Term	Coeff.	SE	Confidence Interval	
			L. Bound	U. Bound
Green 0 Hz	0.891	0.401	0.1	1.682
Green 1 Hz	0.856	0.414	0.04	1.672
Direct	−0.628	0.193	−1.006	−0.25
Conc. * Blue 0 Hz	−0.625	0.294	−1.202	−0.048
Conc. * Blue 0.5 Hz	−0.775	0.294	−1.352	−0.198
Sad * Blue 1 Hz	−0.725	0.294	0.148	1.302
Direct * Blue 1 Hz	−0.667	0.24	−1.138	−0.195
Direct * Green 0.5 Hz	−0.5	0.24	−0.971	−0.029
Conc. * Direct	0.388	0.158	0.079	0.698
Sad * Direct	0.345	0.158	0.036	0.655
NARS S3	−0.382	0.132	−0.642	−0.122

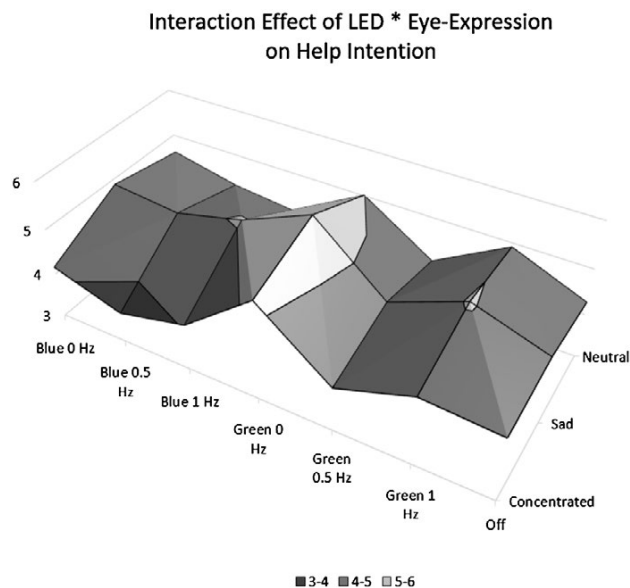
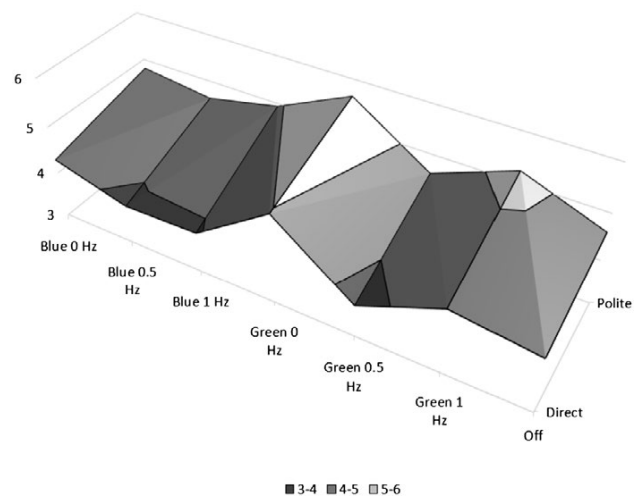
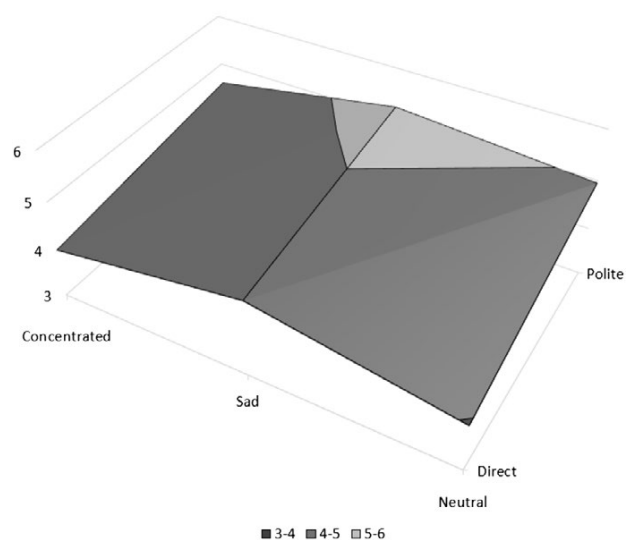
**Figure 4:** Help Intention under interaction of LED * Eye-Expression.

Figure 5 also clearly shows the significant results for blue and green LED Lighting and direct language. Peaks in Help Intention levels can again be seen at constant and 1 Hz Green LED Lighting when combined with polite language.

In figure 6 the minimum level of Help Intention lies at the combination of a neutral expression with direct language. The significant results are not as readily visible as for the interactions before, but the significantly different levels of Help Intention for the concentrated and sad expressions combined with direct language are still distinguishable when compared to the neutral levels of each independent variable. The peak level of Help Intention is

Interaction Effect of LED * Language on Help Intention**Figure 5:** Help Intention under interaction of LED * Language.**Interaction Effect of Eye-Expression * Language on Help Intention****Figure 6:** Help Intention under interaction of Eye-Expression * Language.

found for the combination of the sad Eye-Expression in combination with polite language.

As a short summary, the following levels of the independent variables and combinations of them stand out with high levels of Help Intention: Sad Eye-Expression, polite language and green LED lighting at constant or 1 Hz frequency.

4.2.2 Discussion and Implications for Help Intention

Significant main effects of all independent variables were found for the dependent variable of Help Intention, as well as interaction effects between Eye-Expression and LEDs and Eye-Expression and language. This confirms evidence from Westhoven et al. [43]. Slowly pulsing for both green and blue LED lighting seem to result in lower Help Intention scores with a minimum in the combination of slowly pulsing blue and the concentrated Eye-Expression. Peak scores result for the sad Eye-Expression, especially at constant and fast pulsing green LED lighting. Although not significant, the interaction between LEDs and language shows some very distinguishable properties: Constant and quickly pulsing green LED lighting in combination with polite language results in peaks for Help Intention, while the aforementioned slowly pulsing in combination with direct language result in the minima for this interaction. The interaction between Eye-Expression and language yields a high score for the sad Eye-Expression and polite language, whereas a neutral expression coupled with direct language results in the minimum score.

Negative attitudes towards emotion in human robot interaction negatively impact help intention, which is unsurprising, since decisions for helping are reported to be mediated by or at least coupled to our emotions [25].

First of all, this is in line with insights from human-human interaction that sad faces provoke sympathy and help intentions [13]. Also, the results are consistent with the literature in that politeness can increase successful help requests, which obviously requires an increased help intention of the helpers (see e. g. [38]).

Apart from adding up the main effects, the constant and slowly pulsing blue light in combination with the concentrated Eye-Expression as well as the sad expression together with fast pulsing blue LEDs unproportionally lower the Help Intention scores. This effect is hard to interpret and there exists no literature for reference. It may be that the effect is coupled to an altered identification of the intended expression or even that it is related to our emotions [40].

4.3 Politeness

As the connection of politeness and successful help requests is indicated by results from the literature and furthermore known to work in interhuman interaction, perceived politeness of the requests was measured. In the following subsections, we provide results, discussion and implications.

Table 2: Parameter estimates from the linear mixed-effect model for Politeness.

Term	Coeff.	SE	Confidence Interval	
			L. Bound	U. Bound
Sad	0.452	0.23	−0.001	0.904
Conc. * Blue 0 Hz	−0.481	0.223	−0.918	−0.043
Conc. * Blue 0.5 Hz	−0.598	0.222	−1.034	−0.162
Conc. * Green 1 Hz	−0.536	0.225	−0.977	−0.094
Sad * Direct	0.252	0.119	0.019	0.485
NARS S1	−0.334	0.096	−0.523	−0.146
Sad * sex	−0.29	0.122	−0.529	−0.051
Direct * Prosopagn.	0.384	0.123	0.142	0.625
Direct * NARS S1	−0.246	0.083	−0.409	−0.082
Direct * NARS S2	−0.182	0.09	−0.36	−0.005

4.3.1 Results for Politeness

Regarding Politeness, the model was adjusted because of significant interactions between LEDs and NARS S3 “Emotion”, Eyes and sex, Language and Prosopagnosia, Language and NARS S1 “Interaction Situations” and NARS S2 “Social Influence”. With the adjusted model, significant results were found for the Eye-Expressions $F(2, 667) = 15.814, p < 0.001$. Further significant effects were found for the covariate NARS S1 “Interaction Situations” $F(1, 120) = 28.24, p < 0.001$ and for the interactions of Eye-Expressions and sex $F(2, 667) = 6.867, p = 0.001$, Language and Prosopagnosia $F(1, 667) = 9.729, p = 0.002$, Language and NARS S1 “Interaction Situations” $F(1, 667) = 8.739, p = 0.003$ as well as Language and NARS S2 “Social Influence” $F(1, 667) = 4.064, p = 0.044$.

The parameter estimates for significant levels of the reported terms are depicted table 2. All results for LEDs and Eyes are tested in comparison with their respective neutral level.

The pairwise interactions of the independent variables LED, Eye-Expression and Language and their effects on perceived politeness are depicted in the figures 7–9. The respective mean values are color coded to express the general level of perceived politeness associated with the combinations of the independent variables’ levels.

Figure 7 clearly depicts the significant results for the concentrated Eye-Expression. Furthermore, low levels of perceived politeness are visible for all Eye-Expressions at LED lighting of 0.5 Hz blue. The sad Eye-Expression forms a ridge of high levels of perceived politeness for all levels of LED lighting.

The low level of perceived politeness at LED lighting of 0.5 Hz blue also shows for the interaction of LEDs and

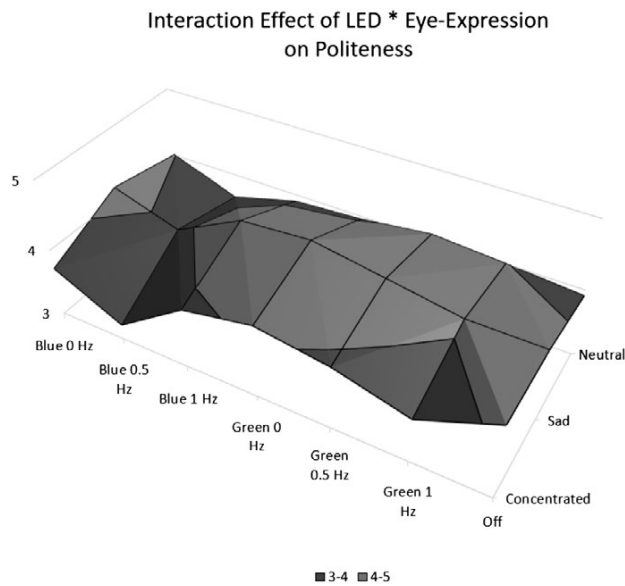


Figure 7: Politeness under interaction of LED * Eye-Expression.

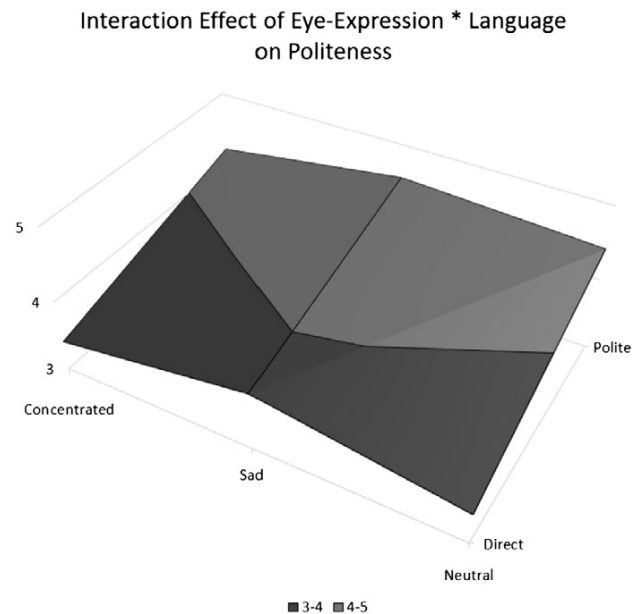


Figure 9: Politeness under interaction of Eye-Expression * Language.

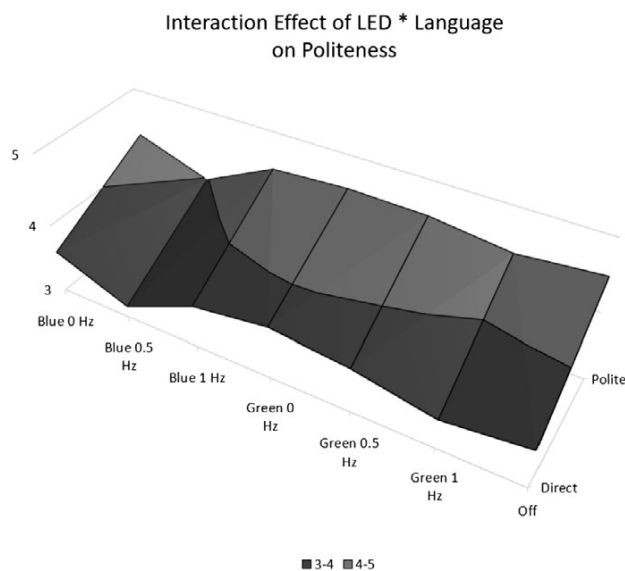


Figure 8: Politeness under interaction of LED * Language.

language in figure 8. Unsurprisingly, the polite language resulted in overall higher perceived politeness.

The expected effect of the use of polite language is also visible in figure 9, underlining the significant differences for the interaction of Eye-Expression and language. Also, slightly higher levels of perceived politeness are visible for the sad Eye-Expression.

In a short summary, the combinations with the highest levels of perceived politeness were found for polite language and the sad Eye-Expression. Regarding the LEDs,

only the 0.5 Hz blue stood out with especially low levels of perceived politeness.

4.3.2 Discussion and Implications for Politeness

For perceived politeness, only one significant main effect was found for the Eye-Expression. This is quite unexpected, because in Westhoven et al. [43], polite language was found to be the main contributor to perceived politeness. The model had to be adjusted for several interactions of covariates with independent variables, though, which was not included in the earlier non-parametric analyses of Westhoven et al. [43].

Additionally, negative attitude towards interaction situations with robots showed a significant effect. It seems that a negative attitude in this case affects the perception or judgement of what is polite and what is not.

Several interaction effects were found for covariates with different independent variables. The interaction between Eye-Expressions and gender implies a difference in what is judged to be an appropriate expression when asking for help. Language, which was thought to be the main driver of politeness perception, only significantly affects perceived politeness through interactions with negative attitudes towards interaction situations with robots and towards social influence of robots. As laid out above, interactions with attitudes are understandable.

There was however also an interaction between language and Prosopagnosia or face blindness. Since there is

Table 3: Parameter estimates from the linear mixed-effect model for Hedonic User Experience.

Term	Coeff.	SE	Confidence Interval	
			L. Bound	U. Bound
Green 1 Hz	1.094	0.444	0.217	1.972
Sad	0.411	0.182	0.055	0.768

no known connection between language perception and face blindness, this is a rather unexpected result. It could be connected to associated functional deficits, but these do not seem to follow a systematic pattern [35]. It may then also very well be something the Prosopagnosia questionnaire unintentionally assesses, apart from supporting a Prosopagnosia diagnosis itself. We assessed Prosopagnosia to control for difficulties in reading and perceiving faces, so we will leave it to future research on Prosopagnosia and politeness perception research to figure out this effect.

Not finding any interactions between the independent variables is in line with the findings of Westhoven et al. [43]. There, only spurious evidence for interaction effects was found, albeit in a regression model without including the covariates.

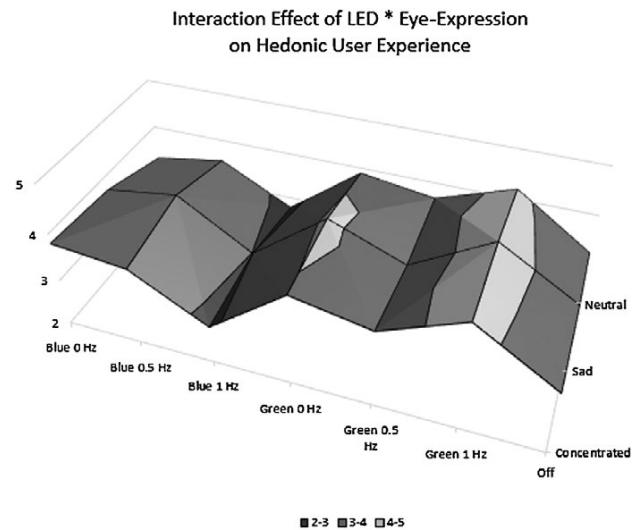
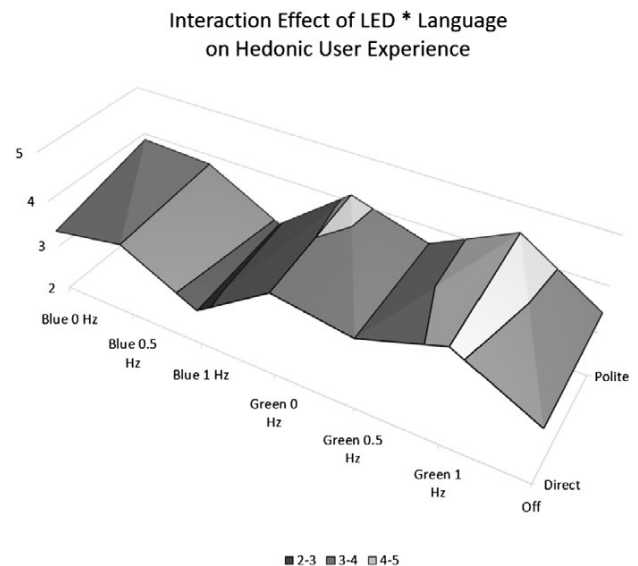
4.4 Hedonic User Experience

In the following subsections, we provide results, discussion and implications for Hedonic User Experience. As no direct interaction took place, we only measured the perception of the hedonic interaction aspects.

4.4.1 Results for Hedonic User Experience

Regarding hedonic User Experience, the model did not need adjustment for interactions between covariates and independent variables. Significant results were found for the LEDs $F(6, 126) = 2.53$, $p = 0.024$, Eye-Expressions $F(2, 672) = 6.606$, $p = 0.001$ and Language $F(1, 672) = 22.56$, $p < 0.001$. No significant interactions and no significant effects from covariates were found. The parameter estimates for significant levels of the reported terms are depicted table 3. All results for LEDs and Eyes are tested in comparison with their respective neutral level.

The pairwise interactions of the independent variables LED, Eye-Expression and Language and their effects on hedonic User Experience are depicted in the figures 10–12. The respective mean values are color coded to express the general level of hedonic User Experience as-

**Figure 10:** Hedonic User Experience under interaction of LED * Eye-Expression.**Figure 11:** Hedonic User Experience under interaction of LED * Language.

sociated with the combinations of the independent variables' levels. While no significant interaction effects were found, figure 10 clearly shows the effect of LED Lighting on hedonic User Experience, especially the high levels for constant or 1 Hz green light. For blue light at 1 Hz, very low levels of hedonic user experience can be seen. The visible effect of the different Eye-Expressions is very small, but shows a slight tendency for higher levels at the sad and neutral levels compared to the concentrated expression. The pattern for the effect of LED lighting can also be seen in figure 11. The effect of language is slightly

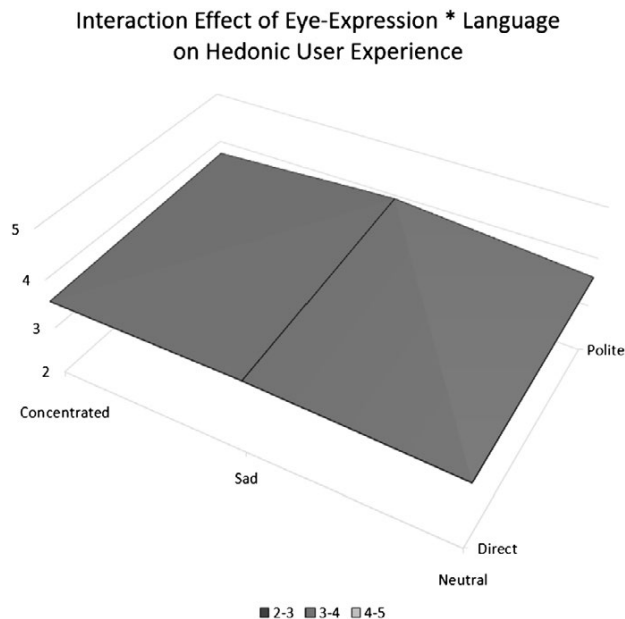


Figure 12: Hedonic User Experience under interaction of Eye-Expression * Language.

more pronounced and higher levels of hedonic User Experience are visible especially at the constant and 1 Hz green LED lightings. Only small differences can be seen for the levels of hedonic User Experience in figure 12 with very slightly elevated levels for polite language and the sad Eye-Expression.

Taken together, high levels of hedonic User Experience show mainly for constant and 1 Hz green LEDs while the other independent variables do not seem to exert a visible influence. Another significant result are the very low levels of hedonic User Experience at 1 Hz blue LED lighting.

4.4.2 Discussion and Implications for Hedonic User Experience

Significant main effects were found for all independent variables regarding hedonic UX. These effects seem to simply add up, since no interaction effects were found. Also, no main effects from covariates or interactions with them were found. While the results for expressions and language confirm the results from Westhoven et al. [43], the significant effect for LED lighting provides further design knowledge. The hedonic UX is increased for green LEDs, but with visible differences between the frequencies and with the highest scores at fast pulsing green. This suggests studying frequency and color of lighting as separate independent variables in future studies.

Table 4: Parameter estimates from the linear mixed-effect model for Facial Expression Identification.

Term	Coeff.	SE	Confidence Interval	
			L. Bound	U. Bound
Conc. * Blue 0.5 Hz	0.367	0.157	0.059	0.676
Sad * Blue 0.5 Hz	0.323	0.147	0.035	0.611
Sad * sex	−0.249	0.08	−0.406	−0.091

4.5 Facial Expression Identification

Finally, and mostly to control for a correct-as-intended identification of facial expressions, we measured the Facial Expression Identification. In the following subsections we report and discuss the results and their practical implications.

4.5.1 Results for Facial Expression Identification

Regarding Facial Expression Identification, the model was adjusted because of significant interactions between Eye-Expressions and sex. With the adjusted model, significant results were found for the interaction of Eye-Expressions and LEDs $F(12, 660.738) = 2.238, p = 0.009$ and for the interaction of Eye-Expressions and the covariate of sex $F(2, 683.962) = 4.794, p = 0.009$.

The parameter estimates for significant levels of the reported terms are depicted table 4. All results for LEDs and Eyes are tested in comparison with their respective neutral level.

The pairwise interactions of the independent variables LED, Eye-Expression and Language and their effects on Facial Expression Identification are depicted in the figures 13–15. The respective mean values are color coded to express the general level of Facial Expression Identification associated with the combinations of the independent variables' levels. In figure 13, the significant interaction effect between Eye-Expression and LED lighting can be seen at 0.5 Hz blue and at the LED's off-condition. While the off-condition yields the best overall identification scores, the sad expression also scores higher at 1 Hz green and at 0.5 Hz blue. The concentrated expression also scores high at constant or 0.5 Hz blue light. Figure 14 shows only small differences on the language axis, indicating a negligible effect of language on Facial Expression Identification.

The same is true for the interaction of the Eye-Expression and language depicted in figure 15.

Summarized, the Facial Expression Identification performance is mainly driven by the Eye-Expression. Some of

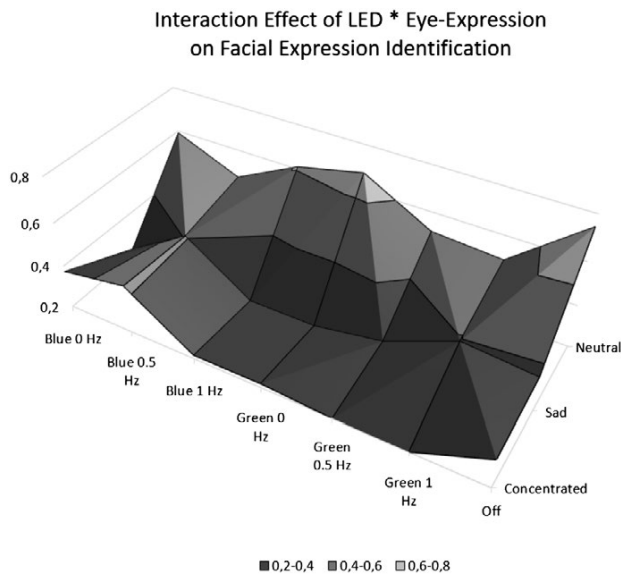


Figure 13: Facial Expression Identification under interaction of LED * Eye-Expression.

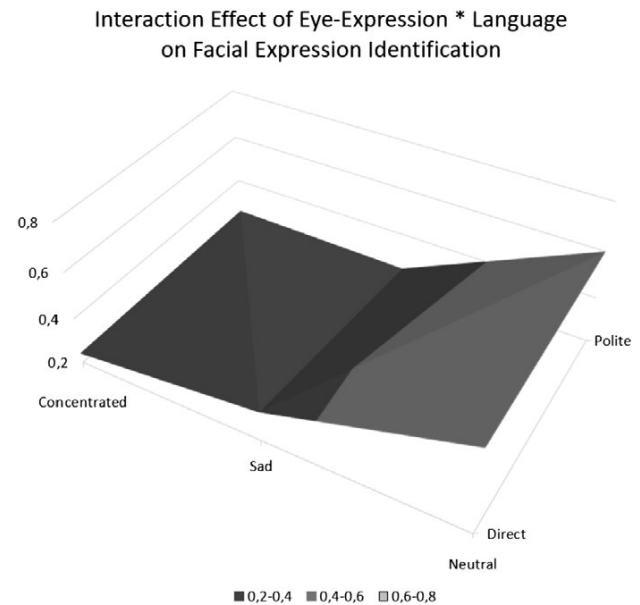


Figure 15: Facial Expression Identification under interaction of Eye-Expression * Language.

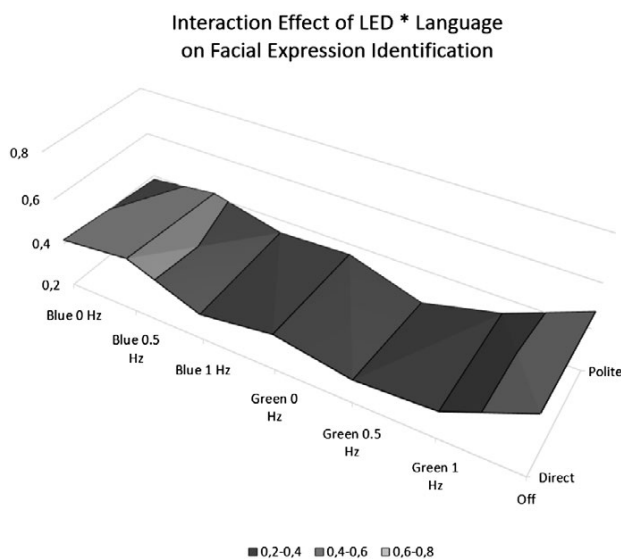


Figure 14: Facial Expression Identification under interaction of LED * Language.

the LED lighting conditions seem to influence the interpretation of the expressions, however.

4.5.2 Discussion and Implications for Facial Expression Identification

In Westhoven et al. [43], we reported the three Eye-Expressions differing in how well they were identified by the participants, with the neutral expression scor-

ing higher than both the sad and concentrated expressions.

With the adjusted model from the present work, only a significant interaction between Eye-Expression and LEDs was found. Especially the slowly pulsing blue LED lighting shows a visible difference for the concentrated and sad Eye-Expressions, where the lighting seems to support identification as intended. This frequency-color combination already showed a spurious finding regarding expression identification in Westhoven et al. [43]. Additionally, in figure 13 there is a visible dent for the neutral eye-expression at slowly pulsing blue as well, indicating a divergence from perceiving it as a neutral state.

The significant interaction of Eye-Expression and the covariate of gender manifests in the interpretation of the sad Eye-Expression. Here, females tend to identify the sad expression more readily. Looking at existing literature, this is unsurprising, since meta studies confirm an advantage for females in reading facial emotions [18].

The findings further indicate optimization potential for the eye-expressions. In Westhoven et al. [43], especially distinctiveness from non-intended interpretations such as fearfulness was identified. Also addressing the question, if acceptable identification performance can be achieved at all with only the eye-part of the face was suggested. With the current findings, the influence of lighting can be added to that list.

4.6 Discussion and Implications for the Research Method

In Westhoven et al. [43], non-parametric tests were used due to violations of assumptions for parametric tests. 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. In this more current paper, we presented a follow-up analysis using linear mixed models to better identify the interaction effects and to properly include the measured covariates. This increased effort for proper analysis of the experimental data may serve as a warning and/or reminder not to underestimate the analysis of such complex experimental designs.

Finally, it has to be noted 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.

5 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. We re-analyzed the data used in Westhoven et al. [43] with linear mixed models for inclusion of covariates and to better accommodate for interaction effects.

Using the sad eye-expression, polite language and constant green LED lights yields benefits regarding the hedonic UX, perceived politeness and help intention. We could show the interactions between these independent variables, which allows for better robot design choices. However, results regarding the expression identification performance suggest that there is further potential for optimization regarding the sad as well as the concentrated expression. Furthermore, a more thorough study of the influence of light color and frequency is strongly suggested as these affect identification of the expressions and in turn also their perception.

Especially the negative attitudes towards robots, but also other covariates, were shown to have significant effects either directly or in interaction with independent variables on help intention and perceived politeness, but not on hedonic UX and expression identification. This

calls for special consideration when designing robot interaction for specific target populations.

In summary, the results imply that to create a pleasant interaction with 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 [6], and LED lighting of constant green color. This combination yields significantly higher hedonic user experience, perceived politeness, and help intention than all other tested combinations.

6 Summary of Practical Implications

The results from this current paper and from Westhoven et al. [43] provide information on the design of help requesting robots. First of all, the main effects show in relative clarity that a sad expression, polite language and constant green LED lighting maximize help intention, perceived politeness and hedonic UX. This has to be taken with a grain of salt, as it may very well be that the true maximum is only reached with choices of colors, expressions or spoken words, which were not tested in this experiment.

The frequency effects of light are still not understood very well and as mentioned above, further studies with color and frequency as separate independent variables will be necessary to exactly pinpoint the occurring effects. As it stands now, the results indicate the use of constant green light and avoidance of the slowly pulsing blue light for optimal results.

The confirmed effects of negative attitudes towards robots suggest careful exploration of the deployment setting when thinking about help requesting robots. In case of a high prevalence of such negative attitudes, all design choices could be compromised due to the effects our attitudes have on our perception.

Acknowledgment: An earlier version of this work has been published at the 2019 Mensch und Computer conference [43], where it was invited to resubmit an extended version to ICOM.

Funding: 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] Asheber, W. T., Lin, C.-Y., and Yen, S. H. Humanoid head face mechanism with expandable facial expressions. *International Journal of Advanced Robotic Systems* 13, 1 (2016), 29.
- [2] Baraka, K., and Veloso, M. M. Mobile service robot state revealing through expressive lights: Formalism, design, and evaluation. *International Journal of Social Robotics* 10, 1 (2018), 65–92.
- [3] Bennett, C. C., and Šabanović, S. Deriving minimal features for human-like facial expressions in robotic faces. *International Journal of Social Robotics* 6, 3 (2014), 367–381.
- [4] Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, C. L., and Lee, D. The art of designing robot faces: Dimensions for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction* (New York, NY, USA, 2006), HRI'06, ACM, pp. 331–332.
- [5] Breazeal, C. Emotion and sociable humanoid robots. *International journal of human-computer studies* 59, 1-2 (2003), 119–155.
- [6] Brown, P., and Levinson, S. C. Universals in language usage: Politeness phenomena. In *Questions and politeness: Strategies in social interaction* (1978), Cambridge University Press, pp. 56–311.
- [7] Budde, V., Backhaus, N., Rosen, P. H., and Wischniewski, S. Needy robots-designing requests for help using insights from social psychology. In *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)* (2018), IEEE, pp. 48–53.
- [8] Cameron, D., Loh, E. J., Chua, A., Collins, E., Aitken, J. M., and Law, J. Robot-stated limitations but not intentions promote user assistance.
- [9] Crawford, A. The perception of light signals: The effect of mixing flashing and steady irrelevant lights. *Ergonomics* 6, 3 (1963), 287–294.
- [10] Dautenhahn, K., Walters, M., Woods, S., Koay, K. L., Nehaniv, C. L., Sisbot, A., Alami, R., and Siméon, T. 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* (2006), ACM, pp. 172–179.
- [11] DiSalvo, C. F., Gemperle, F., Forlizzi, J., and Kiesler, S. 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* (2002), ACM, pp. 321–326.
- [12] Duffy, B. R. Anthropomorphism and the social robot. *Robotics and autonomous systems* 42, 3-4 (2003), 177–190.
- [13] Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., and Reno, R. R. Relation of sympathy and personal distress to prosocial behavior: a multimethod study. *Journal of personality and social psychology* 57, 1 (1989), 55.
- [14] Ekman, P. Facial expression and emotion. *American psychologist* 48, 4 (1993), 384.
- [15] Ernest-Jones, M., Nettle, D., and Bateson, M. Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior* 32, 3 (2011), 172–178.
- [16] Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. Statistical power analyses using g* power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.
- [17] Hacker, W. *Psychologische Bewertung von Arbeitsgestaltungsmassnahmen: Ziele und Bewertungsmassstäbe*, vol. 1. Springer-Verlag, 2013.
- [18] Hall, J. A., Carter, J. D., and Horgan, T. G. Gender differences in nonverbal communication of emotion. In *Gender and emotion: Social psychological perspectives* (2000), pp. 97–117.
- [19] Hammer, S., Lugin, B., Bogomolov, S., Janowski, K., and André, E. Investigating politeness strategies and their persuasiveness for a robotic elderly assistant. In *International Conference on Persuasive Technology* (2016), Springer, pp. 315–326.
- [20] Kanade, T., Cohn, J. F., and Tian, Y. Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)* (2000), IEEE, pp. 46–53.
- [21] Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., and Grueter, M. First report of prevalence of non-syndromic hereditary prosopagnosia (hpa). *American Journal of Medical Genetics Part A* 140, 15 (2006), 1617–1622.
- [22] Lee, H. R., Šabanović, S., and Stolterman, E. How humanlike should a social robot be: A user-centered exploration. In *2016 AAAI Spring Symposium Series* (2016).
- [23] Lindberg, M., Sandberg, H., Liljenberg, M., Eriksson, M., Johansson, B., and Balkenius, C. The expression of mental states in a humanoid robot. In *International Conference on Intelligent Virtual Agents* (2017), Springer, pp. 247–250.
- [24] Lütkebohle, I., Hegel, F., Schulz, S., Hackel, M., Wrede, B., Wachsmuth, S., and Sagerer, G. The bielefeld anthropomorphic robot head ‘flobi’. In *2010 IEEE International Conference on Robotics and Automation* (2010), IEEE, pp. 3384–3391.
- [25] Mackay, N., and Barrowclough, C. Accident and emergency staff’s perceptions of deliberate self-harm: attributions, emotions and willingness to help. *British Journal of Clinical Psychology* 44, 2 (2005), 255–267.
- [26] Marsh, A. A., Ambady, N., and Kleck, R. E. The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion* 5, 1 (2005), 119.
- [27] Mathur, M. B., and Reichling, D. B. Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition* 146 (2016), 22–32.
- [28] Nomura, T., Kanda, T., Suzuki, T., and Kato, K. 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)* (2004), IEEE, pp. 35–40.
- [29] Pavey, L., Greitemeyer, T., and Sparks, P. ‘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.
- [30] Reeves, B., and Nass, C. I. *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press, 1996.
- [31] Rosenthal, S., Biswas, J., and Veloso, M. 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*

- (2010), International Foundation for Autonomous Agents and Multiagent Systems, pp. 915–922.
- [32] Rosenthal, S., Veloso, M., and Dey, A. K. 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* (2011).
- [33] Salem, M., Ziadde, M., and Sakr, M. 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)* (2014), IEEE, pp. 74–81.
- [34] Schrepp, M., Hinderks, A., and Thomaschewski, J. Design and evaluation of a short version of the user experience questionnaire (ueq-s). *IJIMAI* 4, 6 (2017), 103–108.
- [35] Sergeant, J., and Signoret, J.-L. Varieties of functional deficits in prosopagnosia. *Cerebral cortex* 2, 5 (1992), 375–388.
- [36] Song, S., and Yamada, S. 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* (2017), ACM, pp. 2–11.
- [37] Sosnowski, S., Bittermann, A., Kuhnlenz, K., and Buss, M. Design and evaluation of emotion-display eddie. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2006), IEEE, pp. 3113–3118.
- [38] Srinivasan, V., and Takayama, L. Help me please: Robot politeness strategies for soliciting help from humans. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (2016), ACM, pp. 4945–4955.
- [39] Tsiourti, C., Weiss, A., Wac, K., and Vincze, M. Designing emotionally expressive robots: A comparative study on the perception of communication modalities. In *Proceedings of the 5th International Conference on Human Agent Interaction* (New York, NY, USA, 2017), HAI'17, ACM, pp. 213–222.
- [40] Valdez, P., and Mehrabian, A. Effects of color on emotions. *Journal of experimental psychology: General* 123, 4 (1994), 394.
- [41] van der Grinten, T., Müller, S., Westhoven, M., Wischniewski, S., Scheidig, A., and Gross, H.-M. Designing an expressive head for a help requesting socially assistive robot. In *International Workshop on Human-Friendly Robotics* (2019), Springer, pp. 88–102.
- [42] Ventura, R. Two faces of human–robot interaction: Field and service robots. In *New Trends in Medical and Service Robots*. Springer, 2014, pp. 177–192.
- [43] Westhoven, M., van der Grinten, T., and Müller, S. Perceptions of a help-requesting robot – effects of eye-expressions, colored lights and politeness of speech. In *Proceedings of Mensch und Computer 2019* (2019), pp. 43–54.
- [44] Young, J. E., Hawkins, R., Sharlin, E., and Igarashi, T. Toward acceptable domestic robots: Applying insights from social psychology. *International Journal of Social Robotics* 1, 1 (2009), 95.

Bionotes



Martin Westhoven

Federal Institute for Occupational Safety and Health, 44149 Dortmund, Germany
westhoven.martin@baua.bund.de

Martin Westhoven is a researcher in Dr. Lars Adolph's workgroup at the Federal Institute for Occupational Safety and Health in Dortmund, Germany. His research until very recently focused on Human-Robot-Interaction, especially on safe interaction design and its effects on human perception of robots.



Tim van der Grinten

Technische Universität Ilmenau, 98693 Ilmenau, Germany
tim.van-der-grinten@tu-ilmenau.de

Tim van der Grinten studied computer science and systems engineering at Ilmenau University of Technology, Germany. He is now part of the Neuroinformatics and Cognitive Robotics Lab led by Prof. Horst-Michael Groß. His research focuses on environmental mapping and navigation algorithms for autonomous robots.